

**ESTIMATING THE HEALTH RISKS POSED BY INTERMITTENT WATER
SUPPLY USING QUANTITATIVE MICROBIAL RISK ASSESSMENT**

A Dissertation
Presented to
The Academic Faculty

By

Aaron William Bivins

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Environmental Engineering

Georgia Institute of Technology

December 2019

Copyright © Aaron William Bivins 2019

**ESTIMATING THE HEALTH RISKS POSED BY INTERMITTENT WATER
SUPPLY USING QUANTITATIVE MICROBIAL RISK ASSESSMENT**

Approved by:

Dr. Joe Brown
School of Civil & Environmental
Engineering
Georgia Institute of Technology

Dr. Heather Murphy
College of Public Health
Temple University

Dr. Kostas Konstantinidis
School of Civil & Environmental
Engineering
Georgia Institute of Technology

Dr. Mark Borchardt
Agricultural Research Service
United States Department of Agriculture

Dr. Anjali Bohlken
Sam Nunn School of International Affairs
Georgia Institute of Technology

Date approved: October 24, 2019

Without water we are nothing.

Even an emperor, denied water, would swiftly turn to dust.

Water is the real monarch and we are all its slaves.

- Rushdie, The Enchantress of Florence

To my family and friends,
who've always fought my corner.

ACKNOWLEDGEMENTS

If it takes a village to raise a child, then a city is required to raise a PhD. I am blessed with a wonderful city. First, to my advisor Dr. Joe Brown; thank you for taking a chance on me when I had little evidence to justify that you do so. I have enjoyed our conversations and travels. I hope for many more in the future. And I still owe you fly fishing lessons. Alongside Dr. Brown, I would like to thank each of my committee members, Dr. Kostas Konstantinidis, Dr. Anjali Bohlken, Dr. Heather Murphy, and Dr. Mark Borchardt, for providing thoughtful insight and guidance to this work. Thank you to Dr. Borchardt's team at USDA for welcoming me and showing me the ropes regarding the qPCR assays I used in this work. To Jackie Knee, thanks for showing me how to be a microbiologist, even if you did laugh out loud at my first pour plates. To David Taylor, thanks for all the stimulating conversations about IWS and India. You invited me to come to Delhi and collaborate at a point that was very dark for me in Nagpur. Nothing restores a sense of purpose like 4 am water sampling in Delhi. To Akanksha Singh, thank you for your friendship while I was in Nagpur. I learned so much from our conversations and enjoyed every single chai break. You kept me out of so much trouble that I was too ignorant to recognize. To Sarah Lowry, the undisputed ddPCR champ, your enthusiasm and optimism were just what a cynical old graduate student needed.

I would also like to thank the organizations that funded much of my PhD research. Many thanks to the United States Environmental Protection Agency which supported me through a Science to Achieve Results Graduate Research Fellowship. Thanks to the United States India Education Foundation which funded nine months of my field work in India where I

was a Fulbright-Nehru Scholar. And thanks for the American Water Works Association which generously supported my last semester of dissertating through the Larson Aquatic Research Scholarship.

While living and working in India I depended on the generosity and support of many colleagues. At the National Environmental Engineering Research Institute, Dr. Pawan Labhasetwar was the one who suggested I apply for a Fulbright. Thank you for welcoming me into your research group and providing generous support to my work. Many thanks to Rajashree Hajare for orienting me to the microbiology lab and logistics of environmental microbiology at NEERI. Thank you to my research assistant Sonal Wankhede who helped me with field sampling in 120°F heat and monsoon rain.

In Nagpur, my field sampling was supported by collaboration with Orange City Water. Thank you to Mr. Sanjoy Roy, CEO, for granting the support of his organization to my work and Mr. Praveen Sharon, GM, for ensuring that we received assistance in the field. Thank you to the OCW personnel that went above and beyond to make sure my work was successful – Dr. Dilip Thakre, Sandeep Patel, Sandeep Khumbhare, Savit Walde, and Aniket Godekar. I will always remember saoji poha and chai tea from the best stall in town.

In Jaipur, my field sampling was supported by Mr. HS Devinda with the Public Health Engineering Department and Dr. Kailash Sharma with the Ground Water Department of the Government of Rajasthan. Thank you sharing your time and city with us.

Finally, I would like to thank my family. To my mom, thanks for reading my papers and asking me about them. To my dad and Mrs. Laurie, thanks for checking in on me and inviting me to go fishing. To Pop, I wish you could be here for this. You always told me

that it wasn't where you started; it was where you finished. I started my first semester as an undergraduate at Georgia Tech with a 1.69 GPA. Now, I finish with a PhD.

And last among my city, but first in my corner is my wife Katherine. She has seen and known every single moment of pain and satisfaction along this journey and never once wavered. You are my rock.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	v
LIST OF TABLES	xiii
LIST OF FIGURES	xvii
LIST OF ABBREVIATIONS	xxiv
SUMMARY	xxix
INTRODUCTION	1
Intermittent Water Supply	3
Microbial Water Quality in IWS	5
Limitations of Epidemiology for Quantifying GII associated with IWS	7
CHAPTER 2: ESTIMATING THE GLOBAL BURDEN OF DIARRHEAL DISEASE ASSOCIATED WITH INTERMITTENT WATER SUPPLY USING QMRA	10
ABSTRACT	10
INTRODUCTION	11
MATERIALS AND METHODS	12
Hazard Identification	13
Dose-Response	14
Exposure Assessment	15
Risk Characterization	18
Population Served by IWS Estimate	19
Burden of Disease Calculations	20

RESULTS AND DISCUSSION	21
Point Estimates of Infection Risks	21
Monte Carlo Estimates of Infection Risks	22
Model Sensitivity	23
Global Population Served by IWS	23
Diarrheal Burden of Disease Calculations	24
Uncertainties and Limitations	26
Data Gaps	30
Policy Implications.....	31
 CHAPTER 3: LEVERAGING DEUF AND DDPCR TO DETECT WATERBORNE PATHOGENS AT LEVELS RELEVANT TO RISK-BASED STANDARDS: A CASE STUDY IN JAIPUR, INDIA	 33
ABSTRACT.....	33
INTRODUCTION	34
MATERIALS AND METHODS.....	40
Physicochemical Measurements	41
Fecal Indicator Bacteria by Culture	41
Dead-End Ultrafiltration	42
PEG Precipitation and Ultracentrifugation	43
Nucleic Acid Extraction	44
Reverse Transcription	45
Microbes of Interest and Molecular Assays for ddPCR.....	45

Adapting qPCR Assays to ddPCR	48
Interpreting ddPCR Results.....	50
RESULTS AND DISCUSSION.....	51
Physicochemical Results.....	51
Fecal indicators by Culture Results.....	53
ddPCR LODs.....	55
Interpreting ddPCR Results.....	56
Jaipur Drinking Water Quality as Assessed by ddPCR	60
Implications for Drinking Water Safety in Jaipur.....	63
Improvements in Interpreting ddPCR Results for DEUF Concentrate.....	66
Difficulty of Making Risk-Relevant Measurements of Waterborne Pathogens	67
CHAPTER 4: MICROBIAL WATER QUALITY IN INTERMITTENT VERSUS CONTINUOUS WATER SUPPLY IN NAGPUR, INDIA	69
ABSTRACT.....	69
INTRODUCTION	70
MATERIALS AND METHODS.....	73
Study Design	74
Physicochemical Measurements	75
Culture Methods.....	76
Dead-End Ultrafiltration	77
PEG Precipitation and Ultracentrifugation	78
Nucleic Acid Extraction.....	79

Reverse Transcription	79
ddPCR	80
RESULTS AND DISCUSSION	81
Service Pressure	82
Free and Total Chlorine.....	82
Turbidity.....	83
Fecal Indicators by Culture	84
Biochemical Testing of Isolates	88
ddPCR Analysis of DEUF Samples	89
Drinking Water Quality in IWS vs CWS	94
Waterborne Pathogens in IWS vs CWS.....	95
Heterogeneity in Water Quality in IWS.....	95
Limitations	96
Policy Implications for IWS in India	98
CHAPTER 5: RISKS OF INFECTION WITH WATERBORNE PATHOGENS ATTRIBUTABLE TO INTERMITTENT WATER SUPPLY IN INDIA.....	101
ABSTRACT.....	101
INTRODUCTION	102
MATERIALS AND METHODS.....	104
Hazard Identification.....	105
Protozoan Pathogens.....	106
Viral Pathogens.....	107

Bacterial Pathogens	108
Exposure Assessment.....	109
Dose Harmonization.....	111
Dose-Response	112
Risk Characterization	114
RESULTS AND DISCUSSION.....	117
Daily Probability of Infection	117
Sensitivity Analysis.....	118
Annual Probability of Infection.....	119
Comparison with Previous Risk Assessment of IWS	120
Infections Attributable to IWS in India.....	122
Limitations and Uncertainties	122
CONCLUSION.....	127
APPENDIX A.....	133
APPENDIX B	136
APPENDIX C	148
APPENDIX D.....	160
APPENDIX E	172
REFERENCES	189
VITA.....	215

LIST OF TABLES

TABLE 1: Initial estimate of the global population served by IWS based on the 2015 JMP Update and IWS prevalence reported in WHO reports from 2000 and 2001.	5
TABLE 2: Descriptive statistics of the probability density functions used to model each stochastic parameter in the Monte Carlo simulation of the IWS QMRA to estimate the global burden of diarrheal disease.	18
TABLE 3: Median, 10th percentile, and 90th percentile daily probabilities of infection for each reference pathogen assuming consumption of fecally contaminated tap water from an IWS as estimated using Monte Carlo simulation.	23
TABLE 4: Annual infections, diarrheal cases, DALYs, and deaths attributable to IWS as calculated using the median daily probability of infection and its associated 95% confidence interval for <i>Campylobacter</i> , <i>Cryptosporidium</i> , and rotavirus assuming consumption of fecally contaminated tap water from an IWS.	25
TABLE 5: Concentration of reference pathogens at which the annual infection risk is equal to the EPA level of acceptable risk for drinking water, and annual infection risk of 1 in 10 persons, and the WHO acceptable risk level for drinking water as calculated assuming a daily consumption of 1 liter of drinking water and the mean dose-response parameters as reported in the previous chapter.	36
TABLE 6: Primer and probe sequence information for qPCR assays as adapted to ddPCR for interrogation of DEUF concentrate from drinking water samples collected in Jaipur, India.	49
TABLE 7: ddPCR analysis performance metrics tabulated for manual and Umbrella thresholding stratified by molecular assay as observed in ddPCR assays of DEUF concentrate collected from the municipal drinking water supply in Jaipur, India.	58
TABLE 8: Proportion of samples positive and concentration ranges for gene targets in groundwater (GW), surface water (SW) and mixed source (M) samples collected from the Jaipur municipal water supply and assayed by ddPCR. Reported results are stratified by strength of the ddPCR evidence.	62

TABLE 9: A summary of the stochastic variables as implemented in our risk assessment model to estimate the probability of infection associated with IWS tap water in India.	117
TABLE 10: Measures of central tendency of the daily risk of infection (diarrhea for ETEC) and the 10th and 90th percentiles as estimated by bootstrapping of the Monte Carlo model of risk associated with ingesting IWS tap water in India.	119
TABLE 11: The annual risk of infection for each pathogen attributable to IWS as estimated by subsampling of the daily probability of infection.	121
TABLE 12: The annual probability of infection for each pathogen attributable to IWS in India as calculated with the daily probability of infection fixed at the denoted value.	121
TABLE A1: A summary of observed proportions of samples positive and summary statistics for measures of various fecal indicators in IWS networks throughout the globe.	134
TABLE A2: A summary of epidemiological evidence linking IWS and diarrheal disease.	136
TABLE B1: Point estimates of the daily and annual probability of infection, and annual disease burden assuming the median <i>E. coli</i> concentration and raw sewage as the source of fecal contamination in an IWS. These calculations demonstrate the mathematical framing and calculations associated with the QMRA model. Annual disease burdens highlighted in bold exceed the normative burden of disease threshold of 10^{-6} DALYs per person per year.	139
TABLE B2: Point estimates of the daily and annual probability of infection, and annual disease burden assuming the mean <i>E. coli</i> concentration and raw sewage as the source of fecal contamination in an IWS. Annual disease burdens highlighted in bold exceed the normative burden of disease threshold of 10^{-6} DALYs per person per year.	139
TABLE B3: Point estimates of the daily and annual probability of infection, and annual disease burden assuming the median <i>E. coli</i> concentration and raw sewage as the source of fecal contamination in an IWS, as characterized in the Guidelines for Drinking Water Quality 4th Edition Table 7.6. Annual disease burdens highlighted in bold exceed the normative burden of disease threshold of 10^{-6} DALYs per person per year.	140
TABLE B4: Results of the tornado analysis and rank correlation to assess the sensitivity of the estimated risk of infection for <i>Campylobacter</i> to input parameters when considering consumption of fecally contaminated tap water by an IWS user.	144

TABLE B5: Results of the tornado analysis testing the sensitivity of the estimated risk of infection for <i>Cryptosporidium</i> to input parameters when considering consumption of fecally contaminated tap water by an IWS user.	144
TABLE B6: Results of the tornado analysis testing the sensitivity of the estimated risk of infection for rotavirus to input parameters when considering consumption of fecally contaminated tap water by an IWS user.	144
TABLE B7: Estimate of the global population served by IWS by projecting IBNET reported percentages of IWS population onto JMP 2015 Update “piped on premise” water supply populations.	145
TABLE B8: The annual burden of diarrheal disease attributable to IWS based on the median daily probabilities of infection for <i>Campylobacter</i> , <i>Cryptosporidium</i> and rotavirus assuming consumption of fecally contaminated tap water from an IWS. The rotavirus burden is calculated using metrics for LMICs.	146
TABLE B9: The annual burden of diarrheal disease attributable to IWS based on the median daily probabilities of infection for <i>Campylobacter</i> , <i>Cryptosporidium</i> and rotavirus assuming consumption of fecally contaminated tap water from an IWS. The rotavirus burden is calculated using metrics for high-income countries.	148
TABLE C1. Primer and probe sequences for probe-based qPCR assays adapted to ddPCR and used as controls in interrogation of DEUF concentrate samples from drinking water.	151
TABLE C2: Optimized thermal cycling conditions and 95% LODs for qPCR assays as adapted to ddPCR.	152
TABLE C3: Optimized thermal cycling conditions for qPCR control assays as adapted to ddPCR.	153
TABLE C4: Concentration factor calculations for the DEUF and ddPCR workflow as used to analyze concentrated water samples from Jaipur, India.	154
TABLE D1: Performance of ddPCR assays as observed during the interrogation of DEUF concentrate from samples collected from the municipal water supply in Nagpur, India.	169
TABLE D2: Number of DEUF samples positive by ddPCR for each genetic target and an aggregate of all targets associated with pathogens (All Path) above the 95% LOD observed at each sampling point within IWS and CWS areas of Nagpur.	171

TABLE D3: Number of DEUF samples positive by ddPCR for each genetic target and an aggregate of all targets associated with pathogens (All Path) above the 95% LOD as observed in IWS versus CWS areas of Nagpur.	172
TABLE E1: Exposure assessment equations as implemented in the IWS in India QMRA.	173
TABLE E2: Dose harmonization equations as implemented in the IWS in India QMRA.	175
TABLE E3: Dose-response equations as implemented in the IWS in India QMRA.	177

LIST OF FIGURES

FIGURE 1: The multiple-barrier approach to ensure drinking water safety during water distribution consists of physical integrity of the pipeline, hydraulic integrity through positive internal pressure, and chemical integrity in the form of a residual disinfectant.	2
FIGURE 2: Reported measures of central tendency (mean or median) and range for studies quantifying fecal indicator bacteria in intermittent water supplies. (EC – <i>E. coli</i> ; FC – fecal coliforms; HPC – heterotrophic plate count; TC – total coliforms; TTC – thermotolerant coliforms).	6
FIGURE 3: A schematic of the Monte Carlo framework used to estimate the daily probability of infection for <i>Campylobacter</i> , <i>Cryptosporidium</i> , and rotavirus assuming the consumption of contaminated tap water from an IWS.	13
FIGURE 4: Sample collection locations from the cross-sectional sampling of the Jaipur municipal water supply with groundwater source samples shown in blue, surface source samples in green, and mixed source in yellow.	41
FIGURE 5: Box and whisker plots of free and total chlorine measures in groundwater samples from tubewells (n=9), surface source (n=8), and mixed source (n=4) drinking water samples from the municipal water supply in Jaipur.	52
FIGURE 6: Box and whisker plots of pH and conductivity measures in groundwater samples from tubewells (n=10), surface source (n=9), and mixed source (n=4) drinking water samples from the distribution network in Jaipur.	53
FIGURE 7: Frequency distribution of total coliform and <i>E. coli</i> counts as observed in grab samples from groundwater (n=9), surface source (n=7), and mixed source (n=4) drinking water samples from the distribution network in Jaipur.	54
FIGURE 8: Frequency distribution of total coliform and <i>E. coli</i> counts as observed in DEUF samples from groundwater (n=10), surface source (n=9), and mixed source (n=4) drinking water samples from the distribution network in Jaipur.	55

FIGURE 9: A schematic of the study design and water quality sampling as executed during the 2017 sampling period. We collected samples along the drinking water supply chain for an IWS command area (Khamla - green) and a CWS command area (Laxmi Nagar Old - purple).	75
FIGURE 10: Boxplots of free chlorine concentration as observed in grab samples collected from taps served by IWS versus taps served by CWS in Nagpur.	83
FIGURE 11: Turbidity measures as observed at household taps in IWS (Khamla) and CWS (Laxmi Nagar Old) command areas in Nagpur.	84
FIGURE 12: Counts of thermotolerant coliforms as observed in grab samples from household taps in an IWS versus a CWS service area in Nagpur.	86
FIGURE 13: TTC counts as observed via DEUF at ESRs and household taps in IWS (Khamla) and CWS (Laxmi Nagar Old) service areas in Nagpur.	87
FIGURE 14: <i>E. coli</i> counts as observed in grab samples collected from household taps in IWS (n=28 countable) versus CWS service areas (n=1 countable) in Nagpur.	88
FIGURE 15: <i>E. coli</i> counts as observed via DEUF at ESRs and household taps in IWS (Khamla) and CWS (Laxmi Nagar Old) service areas in Nagpur.	89
FIGURE 16: Box plots of observed gene copies of ybbW per liter of drinking water at sampling points along the IWS and CWS drinking water delivery chain in DEUF concentrate collected from the Nagpur water supply in 2017.	92
FIGURE 17: The proportion of DEUF samples positive for ybbW, a gene associated with <i>E. coli</i> , versus proportion of samples in which <i>E. coli</i> was detected by culture.	92
FIGURE 18: The proportion of DEUF samples positive for each genetic target as observed in taps served by IWS versus those served by CWS in Nagpur. The “All Pathogen” category is the aggregate of all the detections of a gene target associated with a pathogen across all IWS versus CWS taps.	94
FIGURE 19: The proportion of DEUF samples positive for any gene associated with a pathogen and the associated Wilson Score Interval as observed in samples from an IWS ESR, IWS Taps, a CWS ESR, and CWS taps in Nagpur.	95
FIGURE 20: Schematic of the mathematical framework developed to estimate the probability of infection for each of six waterborne pathogens detected at household taps served by IWS in India.	107

FIGURE B1: Boxplots of the field observed log <i>E. coli</i> counts and the modeled log <i>E. coli</i> counts from the normal distribution as parameterized using maximum likelihood estimation (MLE).	137
FIGURE B2: Frequency distributions of field-observed and maximum likelihood estimated <i>E. coli</i> counts at taps in an IWS.	137
FIGURE B3: Cumulative distribution of <i>E. coli</i> counts in an IWS as observed in the field and modeled using a lognormal distribution parameterized using maximum likelihood estimation (MLE).	138
FIGURE B4: Record screening flow chart for IBNET records used to estimate the global population served by IWS.	138
FIGURE B5: Cumulative distributions of the daily probability of infection with <i>Campylobacter</i> and the associated annual burden of disease assuming consumption of fecally contaminated tap water from an IWS.	141
FIGURE B6: Cumulative distributions of the daily probability of infection with <i>Cryptosporidium</i> and the associated annual burden of disease assuming consumption of fecally contaminated tap water from an IWS.	142
FIGURE B7: Cumulative distributions of the daily probability of infection with rotavirus and the associated annual burden of disease assuming consumption of fecally contaminated tap water from an IWS.	143
FIGURE B8: Map of the geographic range and estimated magnitude of the population served by IWS by country.	145
FIGURE B9: Cumulative distributions, by etiology and in total, of the annual diarrheal burden of disease and deaths attributable to consumption of fecally contaminated tap water among the 925 million global users of IWS.	148
FIGURE C1: Attributable fraction of mild-to-severe diarrhea in 24- to 59-month olds as measured by molecular methods at three southeast Asian study sites during GEMS. Adapted from (131).	149
FIGURE C2: Attributable fraction of mild-to-severe diarrhea in 24- to 59-month olds as measured by molecular methods are three southeast Asian study sites during GEMS with <i>Shigella</i> omitted to allow closer examination of other etiological agents. Adapted from (131).	149
FIGURE C3: Attributable fraction of diarrhea in 12- to 24-month olds as measured by molecular methods are four southeast Asian study sites during MAL-ED. Adapted from (67).	150

FIGURE C4: Example output from manual thresholding in QuantaSoft (Version 1.7.4.0917; Bio-Rad, Hercules, CA) during quantification of process control BRSV.	155
FIGURE C5: Umbrella output from Well A01, a sample spiked with 5 uL of BRSV. The y-axis reports the probability of a partition being negative for the target based on the fluorescence amplitude distributions observed in the NTC and NEC given the observed fluorescence value noted on the x-axis.	156
FIGURE C6: Umbrella results for well H03 an NTC from the BRSV ddPCR experiment. The histogram of the fluorescence values reflects the dual banded negative cluster observed from the QuantaSoft output in Figure C4. For this well all 12,543 droplets have been estimated to have less than a 20% probability of being false negatives and are therefore as negative and are shown in blue.	157
FIGURE C7: Correlation between the number of positive droplets by manual thresholding and Umbrella thresholding as observed for BRSV process control spiked into extracted DEUF concentrates from drinking water samples collected in Jaipur, negative extraction control, and no-template-control (n=24).	158
FIGURE C8: Variance in the fluorescence amplitude in the negative clusters as observed in the NEC (Well E04), two NTCs (Well F04 and G04), and a positive control (Well H04) for the hexon assay experiment.	159
FIGURE C9: Proposed strength-of-evidence paradigm for reporting molecular evidence of waterborne pathogens in Jaipur drinking water as determined by ddPCR assay.	160
FIGURE C10: Linear model fit to the log of ybbW gc per PCR reaction versus the log of CFU per reaction as observed during drinking water sampling in Jaipur.	160
FIGURE D1: Overview of sample collection locations in Nagpur during 2015 and 2017 sampling. Yellow and red markers denote sample collections in Reshim Bagh (IWS) and Karve Nagar (IWS), respectively, during 2015. Purple markers in the upper left are samples collected from the Pench 2 WTP and green and blue markers are samples collected in the Laxmi Nagar Old (CWS) and Khamla (IWS) command areas, respectively, in 2017.	161
FIGURE D2: Sample collection locations in Karve Nagar (IWS) in 2015 (red markers), Khamla (IWS) in 2017 (blue markers), and Laxmi Nagar Old (CWS) in 2017 (green markers).	162
FIGURE D3: Sample collection locations in Reshim Bagh (IWS) in 2015 (yellow markers).	163

FIGURE D4: Boxplots of observed water pressure at household service taps in an IWS area (Khamla) and a CWS area (Laxmi Nagar Old) in Nagpur.	163
FIGURE D5: Boxplots of observed turbidity at Khamla ESR and household taps (IWS) and Laxmi Nagar Old ESR and household taps (CWS) in Nagpur.	164
FIGURE D6: Heterotrophic plate counts (HPC) as observed in grab samples from an IWS and CWS service area during sampling in Nagpur in 2017.	164
FIGURE D7: Proportion of DEUF backflush samples positive for thermotolerant coliforms along the CWS versus the IWS drinking water supply chain as observed during sampling in 2017.	165
FIGURE D8: Mean TTC counts and associated 95% confidence intervals as observed in the IWS and CWS drinking water supply chains via DEUF in Nagpur.	165
FIGURE D9: The proportion of grab samples collected from household taps positive for <i>E. coli</i> in three IWS service areas (Karve Nagar, Reshim Bagh, Khamla) versus a CWS service area (Laxmi Nagar Old).	166
FIGURE D10: <i>E. coli</i> counts at household taps as observed in grab samples collected from Karve Nagar (IWS, n=8 countable), Reshim Bagh (IWS, n=9 countable), Khamla (IWS, n=11 countable), and Laxmi Nagar Old (CWS, n=1 countable).	166
FIGURE D11: Proportion of DEUF backflush samples positive for <i>E. coli</i> along the CWS versus the IWS drinking water supply chain as observed during sampling in 2017.	167
FIGURE D12: Mean <i>E. coli</i> counts and associated 95% confidence intervals as observed in the IWS and CWS drinking water supply chains via DEUF in Nagpur.	167
FIGURE D13: Quantification of process control BRSV in ddPCR reaction mix by manual thresholding and Umbrella thresholding demonstrated perfect linear correlation between the two thresholding methods.	168
FIGURE D14: Linear model fit to the log of ybbW gc per PCR reaction versus the log of CFU per reaction as observed in DEUF samples from the municipal water supply in Nagpur.	168
FIGURE E1: Lognormal model fit to observed concentrations of beta giardin (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.	178

FIGURE E2: Lognormal model fit to observed concentrations of <i>Crypto. 18S rRNA</i> (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.	179
FIGURE E3: Lognormal model fit to observed concentrations of norovirus GI and GII ORF1-2 (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.	179
FIGURE E4: Lognormal model fit to observed concentrations of hexon (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.	180
FIGURE E5: Lognormal model fit to observed concentrations of STh (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.	180
FIGURE E6: Lognormal model fit to observed concentrations of ipaH (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.	181
FIGURE E7: Lognormal model fit to observed concentrations of ybbW (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.	181
FIGURE E8: Lognormal model fit to observed counts of <i>E. coli</i> (CFU/100 mL) at IWS taps as enumerated via DEUF and membrane filtration.	182
FIGURE E9: The cumulative distribution of the daily risk of infection for <i>Giardia</i> associated with ingestion of tap water supplied by an IWS in India.	182
FIGURE E10: The cumulative distribution of the daily risk of infection for <i>Cryptosporidium</i> associated with ingestion of tap water supplied by an IWS in India.	183
FIGURE E11: The cumulative distribution of the daily risk of infection for norovirus associated with ingestion of tap water supplied by an IWS in India as estimated by the high-risk dose-response model.	183
FIGURE E12: The cumulative distribution of the daily risk of infection for norovirus associated with ingestion of tap water supplied by an IWS in India as estimated by the low-risk dose-response model.	184
FIGURE E13: The cumulative distribution of the daily risk of infection for adenovirus associated with ingestion of tap water supplied by an IWS in India.	184
FIGURE E14: The cumulative distribution of the daily risk of diarrhea from infection with ETEC associated with ingestion of tap water supplied by an IWS in India.	185

FIGURE E15: The cumulative distribution of the daily risk of infection with <i>Shigella</i> associated with ingestion of tap water supplied by an IWS in India.	185
FIGURE E16: The sensitivity of the estimated risk of infection with <i>Giardia</i> to model input parameters as assessed by rank order correlation.	186
FIGURE E17: The sensitivity of the estimated risk of infection with <i>Cryptosporidium</i> to model input parameters as assessed by rank order correlation.	186
FIGURE E18: The sensitivity of the estimated risk of infection with norovirus (high-risk model) to input parameters as assessed by rank order correlation.	187
FIGURE E19: The sensitivity of the estimated risk of infection with norovirus (low-risk model) to input parameters as assessed by rank order correlation.	187
FIGURE E20: The sensitivity of the estimated risk of infection with adenovirus to input parameters as assessed by rank order correlation.	188
FIGURE E21: The sensitivity of the estimated risk of infection with ETEC to input parameters as assessed by rank order correlation.	188
FIGURE E22: The sensitivity of the estimated risk of infection with <i>Shigella</i> to input parameters as assessed by rank order correlation.	189

LIST OF ABBREVIATIONS

95% confidence interval	95% CI
acute gastrointestinal illness	AGI
base pairs	bp
Basic Local Alignment Search Tool nucleotide	BLASTn
bovine herpes virus	BoHV
bovine respiratory syncytial virus	BRSV
Centers for Disease Control and Prevention	CDC
centigrade	C
colony forming unit	CFU
complementary deoxyribonucleic acid	cDNA
continuous water supply	CWS
crossing threshold	CT
dead-end ultrafiltration	DEUF
deoxyribonucleic acid	DNA
disability-adjusted life year	DALY
droplet digital polymerase chain reaction	ddPCR
elevated storage reservoir	ESR
enteroaggregative <i>E. coli</i>	EAggEC
enterohemorrhagic <i>E. coli</i>	EHEC
enteroinvasive <i>E. coli</i>	EIEC
enteropathogenic <i>E. coli</i>	EPEC

enterotoxigenic <i>E. coli</i>	ETEC or ST-ETEC
environmental enteric dysfunction	EED
environmental enteropathy	EE
<i>Escherichia coli</i>	<i>E. coli</i> or EC
fecal coliform	FC
fecal indicator bacteria	FIB
fecal streptococci	FS
gastrointestinal illness	GII
gene copies	gc
Global Enteric Multicenter Study	GEMS
Guidelines for Drinking Water Quality	GDWQ
heat stable toxin	STh
heterotrophic plate count	HPC
Indian Rupee	INR
intermittent water supply	IWS
International Benchmarking Network	IBNET
Joint Monitoring Program	JMP
kilometer	km
limit of detection	LOD
liter	L
low- and middle-income countries	LMIC
male-specific coliphage	MS2
Malnutrition and Enteric Disease Study	MAL-ED

master balancing reservoir	MBR
maximum likelihood estimation	MLE
median tissue culture infective dose	TCID50
meter	m
method reporting limit	MRL
microliter	uL
milligrams per liter	mg/L
milliliter	mL
million liters per day	MLD
molecular weight cut off	MWCO
most probable number	MPN
Nagpur Municipal Corporation	NMC
nanomolar	nM
National Center for Biotechnology Information	NCBI
National Environmental Engineering Research Institute	NEERI
negative extraction control	NEC
nephelometric turbidity unit	NTU
no-template control	NTC
odds ratio	OR
open reading frame	ORF
Orange City Water	OCW
phosphate-buffered saline	PBS
polyethylene glycol	PEG

positive control	PC
pounds force per square inch	psi
probability density function	PDF
proportion positive	PP
quantitative microbial risk assessment	QMRA
real-time quantitative polymerase chain reaction	qPCR
reverse transcriptase real-time PCR	RT-qPCR
reverse transcription	RT
ribonucleic acid	RNA
risk ratio	RR
Shiga-toxin-producing <i>E. coli</i>	STEC
small subunit ribosomal RNA	18S rRNA
sodium chloride	NaCl
sodium polyphosphate	NaPP
tangential-flow ultrafiltration	TFUF
thermotolerant coliforms	TTC
too numerous to count	TNTC
total coliform	TC
total dissolved solids	TDS
unaccounted for water	UFW
United States Dollar	USD
United States Environmental Protection Agency	US EPA or EPA
United States of America	US or USA

viable but not culturable	VBNC
water safety plan	WSP
water treatment plant	WTP
World Health Organization	WHO

SUMMARY

From 2000 to 2015, an additional 1.2 billion people gained access to a piped-on-premise drinking water source. While piped water supplies allow for the convenient delivery of water in greater quantity and with improved quality compared to other unimproved sources, evidence indicates that piped supplies are still frequently contaminated with fecal pollution. This contamination is not without effect. Epidemiology trials have documented increased risks of gastrointestinal illness among people receiving their drinking water from deficient piped distribution networks.

One prevalent deficiency is intermittent water supply (IWS). In many settings, water distribution systems are intentionally pressurized to distribute water intermittently as a response to scarcity of water or other resources. However, this mode of operation exposes such supplies to increased likelihood of microbial contamination through stagnation, scouring, intrusion, and household-level water handling in response to inconsistent water delivery. Increased prevalence of fecal contamination in intermittent compared to continuous water supplies (CWS) has been documented. But, the levels of such contamination vary both within and between water supplies. Epidemiology trials have documented increased risks of typhoid fever, cholera, diarrhea, and gut inflammation among children exposed to IWS in certain contexts.

In the current work, we leveraged quantitative microbial risk assessment (QMRA) to estimate the global burden of diarrheal disease associated with IWS. We also used dead-end ultrafiltration (DEUF) and droplet digital PCR (ddPCR) along with traditional culture-based methods to assess the microbial water quality at household taps served by IWS in

Jaipur and Nagpur, India. Lastly, we leveraged these datasets in another QMRA to estimate the risks of infection among those exposed to IWS in India.

In our initial risk assessment using *E. coli* counts observed at IWS taps and pathogen to *E. coli* ratios in sewage, we estimated that IWS could account for 17.2 million infections causing 4.52 million cases of diarrhea, 109,000 DALYs, and 1,560 deaths among the 925 million exposed to IWS globally. However, without empirical evidence concerning the presence of waterborne pathogens in IWS networks, we relied upon a reference pathogen paradigm and pathogen to indicator ratios to quantify human exposure to waterborne pathogens via IWS drinking water.

In response to these limitations, we combined DEUF of large volumes of drinking water and ddPCR to achieve sensitive detection and absolute quantification of gene targets associated with waterborne pathogens from two IWSs in India. To quantify gene targets using ddPCR we adapted existing qPCR assays to droplet digital format and determined the 95% limits of detection (LOD) for each one. During our sampling in Jaipur, we detected gene targets associated with *Cryptosporidium* spp., *Giardia duodenalis*, and enterotoxigenic *E. coli* (ETEC) concurrently with culturable *E. coli* in groundwater samples from tube wells. This groundwater is injected directly into the Jaipur water distribution network without treatment and our results indicate it likely poses a significant risk to public health. Importantly, we found that the detection limits of our experimental workflow were such that detecting and quantifying waterborne pathogens at levels corresponding to conservative risk-based standards via DEUF and ddPCR is unlikely because it would require the filtration of hundreds of thousands to millions of liters of drinking water.

In Nagpur, we leveraged the ongoing transition of the municipal water supply from IWS to CWS to observe differences in the microbial water quality in intermittent versus continuous water supply zones. Physicochemical and fecal indicator bacteria results suggest that water quality is more degraded as water is distributed from the municipal storage reservoir to the household taps in IWS versus CWS zones; but, many of these differences are not significant. We did, however, observe a significant increase in the proportion of samples positive for culturable *E. coli* and gene targets associated with waterborne pathogens at household taps served by IWS compared to those served by CWS. At household taps served by IWS we detected genes associated with ETEC, *Shigella* spp., norovirus GI and GII, adenovirus, *Cryptosporidium* spp., and *Giardia duodenalis*. While at household taps served by CWS we only detected genes associated with *Cryptosporidium* spp. in one sample. Our results indicate IWS users are more likely to be exposed to waterborne pathogens than CWS users in the same context.

Finally, we used the resulting dataset from Nagpur and QMRA to estimate the risks of infection with six fecal-oral pathogens among the urban population of India where no city yet provides continuous water supply. Our model estimates that the daily risks of infection for *Giardia*, *Cryptosporidium*, norovirus, adenovirus, and *Shigella* exceed the US EPA acceptable annual threshold of 1 in 10,000 at the 10th percentile. Thus, even a single day of ingesting tap water from an IWS poses considerable risk. At the 10th percentile of daily risk, IWS could account for up to 11 million *Giardia* infections, 60 million *Cryptosporidium* infections, and 2.17 million *Shigella* infections annually among the 460 million urban-dwelling Indians served by IWS. Our estimate of the risk is sensitive to

assumptions about the quantitative relationship between genetic elements and infectious pathogens.

Collectively, the results of our work indicate, even given large uncertainty and variability, the public health risks associated with IWS likely exceed acceptable levels established by the WHO and US EPA. Our environmental microbiology data also indicate that this risk is significantly elevated among users of IWS versus CWS in the same context. Our findings also indicate that detecting waterborne pathogens in drinking water at levels corresponding to risk-based thresholds is unlikely via DEUF and ddPCR since currently formulated infection risks would require detecting 1 to 10 pathogenic organisms in hundreds of thousands to millions of liters of drinking water. Given this limitation, risk assessment and management will likely continue rely on culture-based enumerations with large uncertainty and unclear relevance to risk-based standards.

INTRODUCTION

Since the Roman empire piped water supplies have been revered as the most reliable means of delivering drinking water to households (1). During the Millennium Development Goals era, the status of “piped-on-premise” water delivery was bolstered by its classification as the highest level of access to drinking water (2). Between 2000 and 2015, an estimated 1.2 billion people gained access to a piped-on-premise water source – the equivalent of connecting the entire population of the United States with a new piped-on-premise water service every 5 years for 15 years (3). This tremendous expansion brought the total number of people receiving drinking water via a piped supply to 4.7 billion or 64% of the world’s population (3).

Piped-on-premise water delivery is a reliable form of drinking water delivery because it conforms with a multiple-barrier paradigm, as illustrated in Figure 1, to protect drinking water safety (4). In the ideal scenario, treated drinking water is distributed via structurally sound pipelines that prevent contaminants from entering the water. Additionally, the drinking water within the pipeline is maintained at a positive pressure so that if there were cracks in the pipes, drinking water would leak out of the pipe rather than contaminants leaking in. And finally, a residual disinfectant, such as free chlorine, is maintained in the drinking water so that if microbial contaminants do enter the pipeline, they are inactivated by the disinfectant. However, in the event of a failure in one or more of these barriers or in the treatment prior to distribution, piped networks can become efficient transmitters of waterborne disease as exemplified during the 1993 *Cryptosporidium* outbreak in Milwaukee, Wisconsin in which 403,000 were estimated to have been sickened (5).

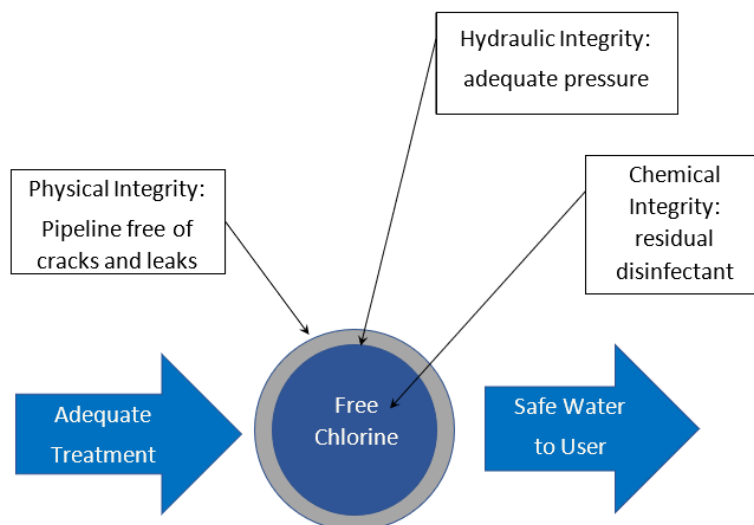


FIGURE 1: The multiple-barrier approach to ensure drinking water safety during water distribution consists of physical integrity of the pipeline, hydraulic integrity through positive internal pressure, and chemical integrity in the form of a residual disinfectant.

Increasing evidence indicates the multiple barriers afforded by piped networks do not guarantee water that is microbiologically safe for human consumption. A study of water quality from “improved” sources found that piped water supplies in Cambodia (n=142) and Vietnam (n=553) delivered water with median counts of *E. coli*, a bacteria whose presence indicates fecal pollution, of 10 and just under 100 CFU or MPN per 100 mL, respectively (6). A more comprehensive systematic review of fecal contamination in water supplies found that although piped supplies were associated with decreased odds of fecal contamination compared to other types of water sources, piped supplies were still frequently subject to high levels of fecal contamination with a median of 25% of samples positive for fecal indicator bacteria (FIB) and more than half of samples positive for FIB in approximately 30% of the studies included (n=119) (7).

To quantify the health impacts of such fecal contamination, several epidemiological trials have examined the incidence of gastrointestinal illness (GII) among a population consuming water as delivered to their home via piped supply compared with a population using supplemental point-of-use treatment (8–12). Drinking water-attributable GII in these studies ranged from 0.08% to 34%; however, these trials suffer from many limitations including a lack of blinding, high dropout rates, and small sample sizes, which ultimately result in low statistical power after accounting for random variation (13). A systematic review of such trials in distribution systems with known deficiencies in one or more of the barriers, the risk of GII among a population drinking tap water without supplemental treatment was 1.34 times greater (RR=1.34 95% CI: 1.00-1.79) than a population using supplemental treatment (14). The same review found that interruptions in continuous water supply (CWS) were associated with a 3.26 times greater risk of GII among consumers exposed to such events (95% CI: 1.48 – 7.19). These findings highlight the increased risk of disease among users of piped water supplies that are malfunctioning and especially piped water supplies that are interrupted during service.

Intermittent Water Supply

While most piped water supplies in high-income countries are only rarely interrupted, in certain resource-constrained settings water supply is intentionally interrupted. An intermittent water supply (IWS) is a piped water supply that delivers water to end-users on a discontinuous basis, with days or hours of planned or unplanned interruption, due to operational constraints including inadequate access to water or energy, distribution system deficiencies, pipe breakages, poor governance or other issues (15). IWS is prevalent in many low and middle-income countries (LMICs) (16). From 2004 to 2013, the

International Benchmarking Network (IBNET), documented water supply lasting less than 24 hours per day in 44 of the 102 countries included in the database (17). In the early 2000s, the World Health Organization (WHO) estimated that 60% of the population served by piped water in Latin America and the Caribbean were served by IWS and that at least one in three urban water supplies in Africa and one in two in Asia were operated intermittently (18,19). As shown in Table 1, based on these prevalence values and the Joint Monitoring Program (JMP) reported urban population receiving piped-on-premise water supply in each region, the population served by IWS could be as many as 1 billion people. Further, the ongoing rapid development of piped water supplies in LMICs, especially in rural and peri-urban areas, climate change, and urbanization, together exert increasing pressure on the resources required to maintain piped water supply functionality, and suggests that the population served by IWS could increase significantly in the coming years (20,21).

TABLE 1: Initial estimate of the global population served by IWS based on the 2015 JMP Update and IWS prevalence reported in WHO reports from 2000 and 2001.

REGION	URBAN POPULATION WITH PIPED-ON-PREMISE WATER SUPPLY	IWS PREVALENCE	POPULATION SERVED BY IWS
Asia	1,427,000,000	50%	713,500,000
Africa	215,100,000	33%	70,983,000
Latin America & Caribbean	473,400,000	60%	284,040,000
Global	2,115,500,000	--	1,068,523,000

Microbial Water Quality in IWS

Given the intentional, and often prolonged, periods of low pressure associated with IWS, such supplies are subject to increased microbial contamination through the intrusion of environmental water from outside the pipeline during low-pressure events, microbial regrowth during stagnant periods, biofilm scouring during re-pressurization, and household storage in response to unreliable supply (16,22,23). As summarized in Table A1, the available evidence suggests large variability in the prevalence of fecal contamination in IWS networks with the proportion of samples positive for fecal coliforms ranging from 0.9% to 76% and *E. coli* from 0.2% to 32%. Quantitative studies of fecal indicators also suggest high variability in measures of both central tendency and range with differences often spanning several orders of magnitude -- *E. coli* from 0.5 MPN/100 mL to 520 CFU/100 mL and fecal coliform from 4 CFU/100 mL to 175 CFU/100 mL as shown in Figure 2. In the only study documenting *E. coli* counts in an IWS compared with a CWS, 31.7% of samples in the IWS were positive for *E. coli* while only 0.7% of samples were positive in the CWS (24).

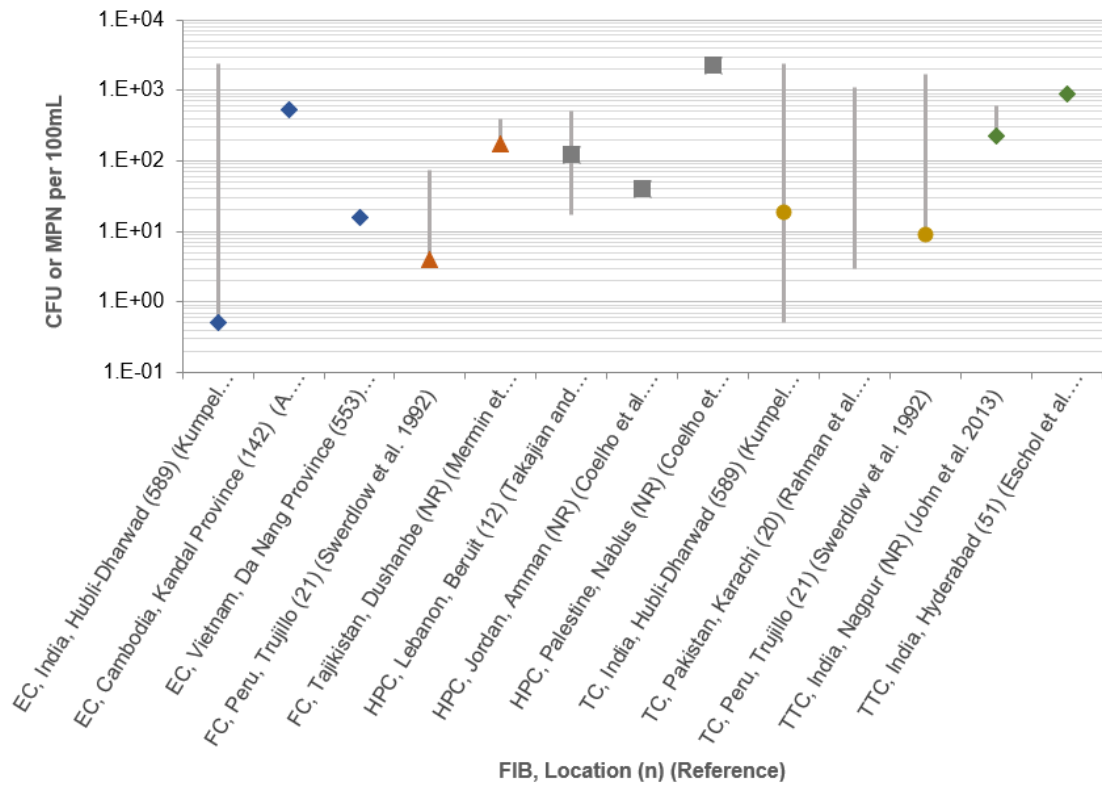


FIGURE 2: Reported measures of central tendency (mean or median) and range for studies quantifying fecal indicator bacteria in intermittent water supplies. (EC – *E. coli*; FC – fecal coliforms; HPC – heterotrophic plate count; TC – total coliforms; TTC – thermotolerant coliforms).

Many of the studies documenting microbial contamination in IWS networks are cross-sectional and include a small sample size and, therefore, fail to adequately document the temporal and spatial variability of microbial water quality in an IWS. Nonetheless, the best available data indicate that fecal contamination is frequently detected in IWS tap water and that contamination prevalence is likely to be much greater in an IWS than a CWS. Given the observed prevalence of microbial contamination in IWS, maintaining an adequate

disinfectant residual is essential to protect water safety. Yet, low disinfectant residuals are often observed in water supplies in LMICs, which potentially increases the risk of waterborne disease associated with IWS (25).

It is not surprising then, that fecal contamination in an IWS has been associated with epidemics of typhoid in Tajikistan and cholera in Peru (26,27). However, endemic GII associated with IWS has proven harder to detect via epidemiologic methods. In the previously mentioned meta-analysis, Ercumen concluded that users of IWS had 1.61 times greater odds of GII compared to those receiving uninterrupted supply (OR=1.61, 95% CI: 1.26-2.07) (14). But, larger and more recent trials of IWS and GII have yielded mixed results. An analysis of time series data of suspected cholera cases and water supply interruptions in the Democratic Republic of the Congo found that in the 12 days following a day without water supply the incidence of suspected cholera cases increased 1.55-fold (28). Meanwhile, a longitudinal cohort study of IWS and self-reported diarrhea in Hubli-Dharwad, India found no association between IWS and diarrhea, but did note a higher incidence of typhoid fever among children under 5 exposed to IWS compared to CWS (29). This epidemiological evidence, summarized in Table A2, suggests that IWS has been associated with epidemic transmission of waterborne diseases such as cholera and typhoid, but statistically meaningful associations between IWS and endemic GII are more difficult to establish.

Limitations of Epidemiology for Quantifying GII associated with IWS

As can be observed from the numerous studies cited, historically, epidemiology has been the tool used to link drinking water exposure and diarrheal diseases. But there are many limitations of epidemiology, especially for measuring the association between disease and

civil infrastructure, such as piped water supply. First, the design and construction of piped drinking water systems makes randomization, an important strategy to isolate causal relationships between interventions and outcomes, logistically, economically, and ethically difficult. Second, epidemiology studies connecting drinking water and disease are generally not mechanistic in nature and often rely on stratification by various risk factors and potential confounders to isolate disease cases attributable to the exposures of interest. Often these exposures are not directly measured but are quantified through proxies or indicators. Outcomes are also often recorded via subjective measures such as self-reported diarrhea and often within unblinded trials. Finally, epidemiological studies often require a large sample size to achieve adequate statistical power for low frequency outcomes, such as diarrhea in high-income countries and, as a result, are expensive. These design limitations make it difficult to interpret and generalize associations between drinking water and GII and make robust hypothesis testing via epidemiological trials cost prohibitive.

Quantitative microbial risk assessment (QMRA) offers an alternative means of quantifying associations between drinking water and disease by using stochastic methods and an emphasis on characterizing exposures to pathogens to predict risks to human health (30). Since its first application to drinking water in 1983, there has been a steady increase in applications of QMRA growing from an average of approximately 1 publication per year from 1983 to 1993 to an average of 15 publications per year in the last 10 years. QMRA has been used to make risk estimates for a variety of waterborne pathogens in drinking water including *Giardia*, *Cryptosporidium*, various viruses, and bacteria, and is now advocated in the WHO's Guidelines for Drinking-water Quality and for water safety management (31–38). QMRA has been used in the United States (US) to formulate

drinking water regulations, including the Surface Water Treatment Rule (39–41). The technique has also been applied to assess risks in a broad range of drinking water exposure scenarios including distributed water, surface water, rain water, and groundwater (42–48).

CHAPTER 2: ESTIMATING THE GLOBAL BURDEN OF DIARRHEAL DISEASE ASSOCIATED WITH INTERMITTENT WATER SUPPLY USING QMRA

Citation for the published manuscript:

Bivins, A. W., Sumner, T., Kumpel, E., Howard, G., Cumming, O., Ross, I., ... Brown, J. (2017). Estimating Infection Risks and the Global Burden of Diarrheal Disease Attributable to Intermittent Water Supply Using QMRA. *Environmental Science and Technology*, 51(13), 7542–7551. <https://doi.org/10.1021/acs.est.7b01014>

ABSTRACT

Intermittent water supply (IWS) is prevalent throughout low and middle-income countries. IWS is associated with increased microbial contamination and potentially elevated risk of waterborne illness. We used existing datasets to estimate the population exposed to IWS, assess the probability of infection using quantitative microbial risk assessment, and calculate the subsequent burden of diarrheal disease attributable to consuming fecally contaminated tap water from an IWS. We used reference pathogens *Campylobacter*, *Cryptosporidium*, and rotavirus as conservative risk proxies for infections via bacteria, protozoa, and viruses, respectively. Results indicate that the median daily risk of infection is an estimated 1 in 23,500 for *Campylobacter*, 1 in 5,050,000 for *Cryptosporidium*, and 1 in 118,000 for rotavirus. Based on these risks, IWS may account for 17.2 million infections causing 4.52 million cases of diarrhea, 109,000 diarrheal DALYs, and 1,560 deaths each year. The burden of diarrheal disease associated with IWS likely exceeds the WHO health-based normative guideline for drinking water of 10^{-6} DALYs per person per year. Our results underscore the importance water safety management in water supplies and the potential benefits of point-of-use treatment to mitigate risks.

INTRODUCTION

Given the global prevalence of IWS, the observed fecal contamination in such supplies, and the absence of unambiguous epidemiological evidence concerning the endemic health risks associated with IWS, QMRA offers a potentially useful tool for characterizing the risk of infection for fecal-oral pathogens associated with IWS and the attributable burden of diarrheal disease (49). QMRA can make use of relevant microbiological datasets alongside mathematical models to estimate the health effects of human exposures to pathogens (30). QMRA has been used to estimate the health risks associated with drinking water for a number of waterborne pathogens including viruses, bacteria, and protozoa, and for a variety of exposure scenarios, including intrusion of groundwater, surface water, and sewage (42,44,50,51). The application of QMRA in LMICs has been limited by scarcity of the data required to populate models. However, QMRA approaches have been used to estimate public health risks attributable to piped water supplies in Kampala, Uganda, and Accra, Ghana (52,53). Such studies demonstrate the viability of the approach and its importance in risk management in resource limited settings such as those where IWS is prevalent. In this chapter, we use QMRA to estimate the global burden of infection, morbidity, and mortality associated with IWS.

MATERIALS AND METHODS

Briefly, we used Monte Carlo techniques to estimate the risks of infection associated with human exposures to three reference pathogens (*Campylobacter*, *Cryptosporidium*, and rotavirus) through the consumption of contaminated tap water delivered by an IWS. We made use of three existing datasets: *E. coli* measurements in IWS tap water samples, measured pathogen to *E. coli* ratios in sewage, and published dose-response models to estimate the risk of infection. We fit probability distributions to each input dataset and executed Monte Carlo simulations in Oracle Crystal Ball software (Oracle Crystal Ball, Release 11.1.2.4.600 32-bit; Redwood Shores, CA). We then used the predicted median annual risk of infection for each reference pathogen and an estimate of the number of IWS users to quantify a global burden of diarrheal disease, including disability adjusted life years (DALYs) and deaths, associated with IWS. This manuscript is organized using the conventional QMRA framework consisting of hazard identification, exposure assessment, dose-response and risk characterization (54). The framework for the risk assessment model as implemented is illustrated in Figure 3.

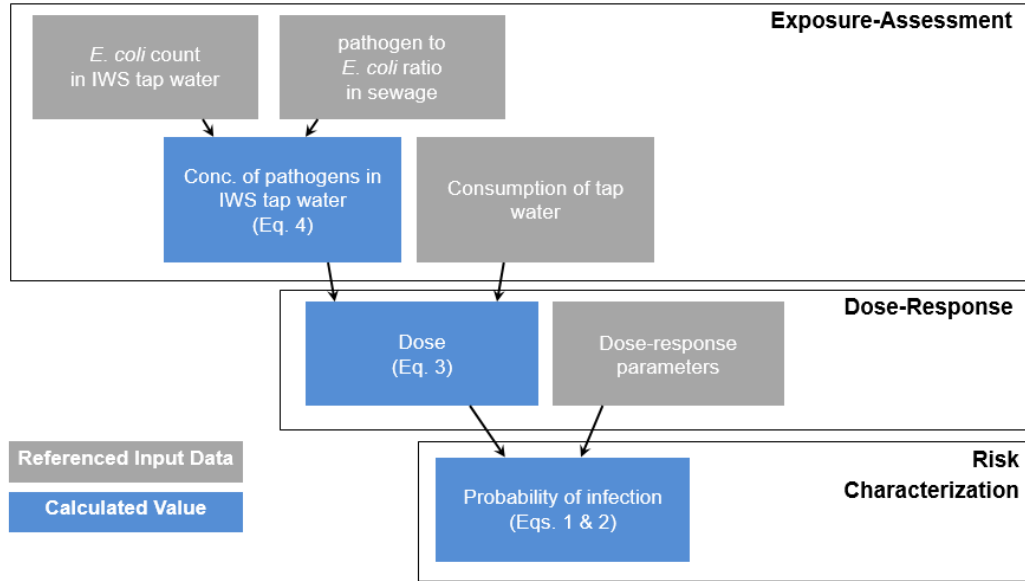


FIGURE 3: A schematic of the Monte Carlo framework used to estimate the daily probability of infection for *Campylobacter*, *Cryptosporidium*, and rotavirus assuming the consumption of contaminated tap water from an IWS.

Hazard Identification

In the absence of published measurements of waterborne pathogens in an IWS, we utilized a reference pathogen approach (36). I selected *Campylobacter jejuni*, *Cryptosporidium parvum*, and rotavirus as reference pathogens, following the model development guidance articulated in the WHO Guidelines for Drinking-water Quality (GDWQ) and supporting documentation (36,37,55,56). While these reference pathogens may not represent the greatest microbial drinking water exposure risks globally, they can be used as conservative proxies for each of the major waterborne pathogen classes in risk estimation. They also have well-characterized dose-response relationships, moderate to long persistence in water supplies, high infectivity, and moderate to high resistance to chlorine, making them suitable as proxies in risk estimation for waterborne pathogens (36).

Campylobacter is a pathogenic bacterium that has caused disease outbreaks associated with contaminated drinking water supplies (57,58). It has a low infectious dose with symptoms including diarrhea, fever, nausea, and vomiting, with rare sequelae (Guillain-Barré syndrome) (59,60). *Cryptosporidium* is a protozoan parasite that has caused large outbreaks of disease through transmission in piped water supplies (5). The infectious dose of *Cryptosporidium* has been estimated to be as low as 1 to 10 oocysts with most infections leading to acute diarrhea, with increased risks of serious illness and death among immunocompromised individuals (60,61). Although commonly associated with hygiene-related transmission, rotavirus has caused significant waterborne disease outbreaks in Rio de Janeiro, Colorado, and China (62–64). One rotavirus particle is capable of initiating an infection leading to fever, vomiting, and acute diarrhea and, in low income settings, presents a significant risk of death among children (60,65). The selection of *Campylobacter*, *Cryptosporidium*, and rotavirus as reference pathogens is supported by findings from the Global Enteric Multicenter (GEMS) Study and a multisite birth cohort study (MAL-ED) that identified each of them as important etiological agents of moderate-to-severe cases of diarrhea among children under 5 in LMICs (66–68).

Dose-Response

The probability of infection following ingestion of a dose of *Campylobacter* or rotavirus is best fit by an approximate beta-Poisson function, Equation 1, characterized by the median infectious dose, N_{50} , the beta distribution parameter alpha, α , and the dose, d (59,65,69). Probability of infection for ingesting *Cryptosporidium* is best characterized by an exponential dose-response function, Equation 2, described by parameters k , and the ingested dose, d (70). For each reference pathogen, we used the dose response parameters

from previously published dose-response fittings and modeled them using lognormal probability density functions (PDF) as described in Table 2 (71).

$$P_{inf}(d) = 1 - \left[1 + \frac{d}{N_{50}} \left(2^{\frac{1}{\alpha}} - 1\right)\right]^{-\alpha} \quad (1)$$

$$P_{inf}(d) = 1 - e^{(-k*d)} \quad (2)$$

Exposure Assessment

In an IWS, periods of low-pressure allow contamination from sewage, groundwater, surface water, or other environmental waters to intrude into the pipelines through holes and cracks (72). When the system is re-pressurized to deliver water to consumers, contaminated water is transported to the taps where it is either used upon delivery or stored for later use. Due to a lack of robust datasets on water quality following household storage due to IWS, our analysis considers the risks of infection posed by IWS if the drinking water were consumed the moment it arrives at the tap (i.e., point-of-entry), without considering re-growth, inactivation, recontamination in storage via unsafe handling practices, or point-of-use water treatment and further storage (6). Quantifying the dose of pathogen ingested at the moment of exposure as shown in Equation 3 is termed exposure assessment.

$$Dose (d) = C_{pathogen, IWS} \left(\frac{N}{mL}\right) * V_{water\ consumed, IWS} (mL) \quad (3)$$

We modeled water consumption in milliliters ($V_{water\ consumed, IWS}$) as a uniform PDF with a minimum of one thousand per day and maximum of two thousand per day based on

the use of one liter per day in WHO risk estimates and two liters per day for adult drinking water consumption in the United States (36,73). To estimate the PDF for the concentration of each reference pathogen ($C_{\text{pathogen, IWS}}$), in the absence of direct measurements of pathogens in IWS tap water, we used a previously developed method of quantifying waterborne pathogens in water distribution networks using pathogen to *E. coli* or thermotolerant coliform ratios (42,51). In this approach, the number of pathogens per volume of drinking water is calculated by multiplying the concentration of *E. coli* measured in IWS tap water by the observed ratio of pathogen to *E. coli* in a potential source of fecal contamination, in this scenario sewage, as shown in Equation 4.

$$C_{\text{pathogen, IWS}} \left(\frac{N}{100\text{mL}} \right) = C_{E. coli, IWS} \left(\frac{N}{100 \text{ mL}} \right) * \left[\frac{C_{\text{pathogen}} \left(\frac{N}{100\text{mL}} \right)}{C_{E. coli} \left(\frac{N}{100\text{mL}} \right)} \right]_{\text{raw sewage}} \quad (4)$$

We developed a PDF of the *E. coli* count in IWS tap water using data from three studies of fecal contamination in IWS systems in three locations: Kandal Province, Cambodia; Da Nang Province, Vietnam; and Hubli-Dharwad, India (24,74,75). These studies were selected because of their large sample size and use of robust methods to quantify *E. coli*. We log transformed the *E. coli* counts and used maximum likelihood techniques to parameterize the normal distribution that maximized the likelihood of obtaining the observed values. For values below and above detection limits, we used the value of the cumulative normal distribution function to incorporate these censored measures into the maximum likelihood estimation (MLE) per previously described methods (76). We estimated the log transformed *E. coli* counts to be normally distributed with mean of 0.17 and standard deviation of 1.57 as shown in Table 1. Boxplots of log *E. coli* counts from each study, the pooled dataset, and the MLE model (Figure B1) show that the quartiles,

median, and mean of the underlying data compare well with the modeled distribution. The frequency and cumulative distributions (Figures B2 and B3) indicate that the MLE model of the *E. coli* count is comparable to the underlying field observed *E. coli* distributions.

For the second term of Equation 4, we developed PDFs of the ratio of each reference pathogen to *E. coli* in raw sewage using paired measurements from sewage. Paired measures from sewage sources specific to locations where IWS is prevalent could not be found in the literature, so we used observations from a sewage treatment plant in the Netherlands (ratio of *Cryptosporidium* and enterovirus to thermotolerant coliforms) and German sewer systems (ratio of *Campylobacter* to *E. coli*) (77,78). Since robust measurements of thermotolerant coliform measurements in IWSs were unavailable in the literature, we assumed that 95% of thermotolerant coliforms in the measured ratios were *E. coli*. Additionally, we substituted rotavirus for enterovirus in the observed ratio. We used the previously described MLE technique on the log transformations of the observed ratios to parameterize the normal distribution that maximized the likelihood of observing the documented measures. The probability distributions and parameters for the reference pathogen to *E. coli* ratios in are summarized in Table 2.

TABLE 2: Descriptive statistics of the probability density functions used to model each stochastic parameter in the Monte Carlo simulation of the IWS QMRA to estimate the global burden of diarrheal disease.

DOSE-RESPONSE PARAMETERS				
Pathogen	Dose-Response Parameter	Distribution	Distribution Description	Reference
<i>Campylobacter</i>	α	Lognormal	Mean: 1.51×10^{-1} Std. Dev.: 5.90×10^{-2}	(59)
	N_{50}	Lognormal	Mean: 1.69×10^3 Std. Dev.: 2.78×10^3	(69)
<i>Cryptosporidium</i>	k	Lognormal	Mean: 3.44×10^{-1} Std. Dev.: 2.02	(70)
rotavirus	α	Lognormal	Mean: 2.48×10^{-1} Std. Dev.: 1.46×10^{-1}	(65)
	N_{50}	Lognormal	Mean: 8.16 Std. Dev.: 6.65	
EXPOSURE ASSESSMENT PARAMETERS				
Parameter		Distribution	Distribution Description	Reference
Tap Water Consumption		Uniform	Min: 1 L Max: 2 L	(36,73)
Log <i>E. coli</i> count in IWS tap water		Normal	Mean: 0.17 Std. Dev.: 1.57	(24,74,75)
<i>Campylobacter</i> to <i>E. coli</i> ratio in sewage		Lognormal	Mean: 8.89×10^{-3} Std. Dev.: 1.33	(78)
<i>Cryptosporidium</i> to fecal coliform ratio in sewage		Lognormal	Mean: 1.13×10^{-6} Std. Dev.: 9.26×10^{-6}	(77)
rotavirus to fecal coliform ratio in sewage		Lognormal	Mean: 8.79×10^{-7} Std. Dev.: 1.77×10^{-6}	(77)

Risk Characterization

To test the mathematical framework and plausibility of the proposed model, we first made point estimates of the daily and annual risk of infection, and the subsequent diarrheal burden of disease. After we reviewed the point estimates, we entered each stochastic variable using the PDFs as described and conducted Monte Carlo simulations in Crystal Ball. Each variable was drawn 10,000 times per the PDF that describes it and each individual input was propagated through the described equations to produce a distribution of the daily probability of infection. We estimated the median, mean, their associated confidence intervals, and percentiles of the probability of infection by bootstrapping the

model with 200 samples of 1,000 trials each. We evaluated the sensitivity of the estimated risks of infection to changes in the input variables by means of tornado analysis and rank correlation. In the tornado analysis, we varied each input from its 10th to 90th percentile and measured the associated variability in the predicted risk of infection while holding all other inputs constant. Rank correlation was determined using Spearman's rank correlation between each input variable and the predicted risk of infection.

Population Served by IWS Estimate

We made a robust estimate of the population served by IWS by projecting the IBNET reported prevalence of intermittent service onto JMP measures of access to piped-on-premise water supplies (79). The IBNET database contains more than 22,000 records from 119 countries dating from 1995 to 2014 (80). Each record consists of a single utility's self-reported performance data for a single year. For this analysis, we used only the most recent record from any single utility that contained both the number of hours the utility supplied water per day and the number of people it supplied. To exclude supply interruptions for repairs and maintenance associated with normal operations in a CWS, we defined an IWS as a utility reporting less than an average of 23 hours per day of service. We further limited our analysis to utilities reporting from countries defined as LMICs by the World Bank. After we removed records that were incomplete, outdated, or from high income countries, 2,591 records pertaining to utilities serving over 773 million people in 91 LMICs were included in the analysis (Figure B4). After screening, we stratified utilities reporting IWS into WHO regions and calculated an average percentage of utilities in that region that were such. We then bootstrapped this average percentage using 10,000 iterations to estimate 95% confidence intervals for each region. We then calculated the average and 95%

confidence interval for the global estimate similarly. To calculate the magnitude or persons served by IWS for each WHO region and globally, we multiplied the estimated percentages and confidence intervals by the number of persons receiving their drinking water from a piped-on-premise supply for each WHO region per the 2015 JMP Update.

Burden of Disease Calculations

We combined the probabilities of infection for each reference pathogen with the estimated number of IWS users by region to calculate the total number of infections, cases of diarrhea, diarrheal disability-adjusted life years (DALYs), and deaths attributable to the consumption of fecally contaminated tap water from an IWS. Following previously articulated methods, it was assumed that the probability of a case of diarrheal illness given infection with *Campylobacter* was 30% with 100% of the population susceptible, *Cryptosporidium* was 70% with 100% of the population susceptible, and rotavirus was 50% with 13% of the population susceptible (36,81). The DALY weighting used in the burden of disease calculations for *Campylobacter* was 4.6×10^{-3} DALYs per case, *Cryptosporidium* was 1.47×10^{-3} DALYs per case, and for rotavirus in low-income countries was 0.482 DALYs per case (81). We calculated deaths attributable to infection with each reference pathogen assuming probability of mortality for *Cryptosporidium* of 10^{-5} per case of diarrhea, probability of mortality due to gastroenteritis associated with *Campylobacter* of 10^{-4} per case of diarrhea and probability of mortality associated with rotavirus of 0.6% per case of diarrheal illness (81). We also assumed that 2.3% of *Campylobacter* cases develop Guillain-Barré syndrome with an associated probability of mortality of 2×10^{-4} (81). We compared the estimated annual burdens of diarrheal disease to the level of acceptable risk from drinking water of 10^{-6} DALYs per person per year as

proposed by the WHO (36). This threshold represents an excess risk of 1 in 100,000 and equates to everyone experiencing one mild self-limiting case of diarrhea every 10 years due to the consumption of unsafe water.

RESULTS AND DISCUSSION

Point Estimates of Infection Risks

We made point estimates of the daily and annual risk of infection, and the annual burden of diarrheal disease, for each reference pathogen using median values of the observed *E. coli* concentration in IWS tap water (1.3 CFU/100 mL) along with median values of the ratio of reference pathogen to *E. coli* in sewage, tap water consumption, and dose-response parameters. These point estimates indicate that, of the pathogens considered, *Campylobacter* poses the greatest risk of infection, possibly due to the greater ratio of *Campylobacter* to *E. coli* observed in sewage from Germany (78). At the median *E. coli* value in IWS tap water, the annual burden of diarrheal disease for *Campylobacter* and rotavirus both exceed the WHO threshold value of 10^{-6} DALYs per person per year (Table B1). When the mean *E. coli* concentration observed in IWS tap water is used, the annual burden of diarrheal disease for each reference pathogen exceeds this threshold (Table B2). For comparison, point estimates of infection risks and burden of diarrheal disease were also calculated for each pathogen using pathogen to *E. coli* ratios in untreated wastewater as documented in the Table 7.6 of the GDWQ (36). As shown in Table B3, the ranking of pathogens by risk of infection remains consistent between the GDWQ pathogen to *E. coli* ratios and the pathogen to *E. coli* ratios used in the model.

Monte Carlo Estimates of Infection Risks

The median daily probabilities of infection predicted by the Monte Carlo simulations, summarized in Table 3, are consistent with the point estimates with the highest risk associated with *Campylobacter* (4.26×10^{-5} 95% CI: $1.92 \times 10^{-5} - 7.89 \times 10^{-5}$) followed by rotavirus (8.47×10^{-6} 95% CI: $3.77 \times 10^{-6} - 1.77 \times 10^{-5}$) and *Cryptosporidium* (1.98×10^{-7} 95% CI: $8.31 \times 10^{-8} - 3.71 \times 10^{-7}$). These translate to median annual probabilities of infection of 1.54% for *Campylobacter*, 0.309% for rotavirus, and 0.007% for *Cryptosporidium*. The upper bounds of the daily probability of infection, as defined by the 90th percentile and shown in Table 3, were 25% for *Campylobacter*, 0.34% for *Cryptosporidium*, and 7.3% for rotavirus. The cumulative distributions of the daily probability of infection for each reference pathogen, shown in Figures B5, B6, and B7, illustrate that the mean daily risk of infection for each reference pathogen was greater than the 80th percentile. For this reason, we used the median risks of infection and their associated confidence intervals to make a conservative calculation of the diarrheal burden of disease associated with the consumption of fecally contaminated tap water delivered by an IWS.

TABLE 3: Median, 10th percentile, and 90th percentile daily probabilities of infection for each reference pathogen assuming consumption of fecally contaminated tap water from an IWS as estimated using Monte Carlo simulation.

PATHOGEN	10 th PERCENTILE DAILY P _{infection}	MEDIAN DAILY P _{infection}	90 th PERCENTILE DAILY P _{infection}
<i>Campylobacter</i>	2.11 x 10 ⁻¹²	4.26 x 10 ⁻⁵ 95% CI: 1.92 x 10 ⁻⁵ – 7.89 x 10 ⁻⁵	2.50 x 10 ⁻¹
<i>Cryptosporidium</i>	1.21 x 10 ⁻¹⁴	1.98 x 10 ⁻⁷ 95% CI: 8.31 x 10 ⁻⁸ – 3.71 x 10 ⁻⁷	3.43 x 10 ⁻³
rotavirus	5.62 x 10 ⁻¹³	8.47 x 10 ⁻⁶ 95% CI: 3.77 x 10 ⁻⁶ – 1.77 x 10 ⁻⁵	7.32 x 10 ⁻²

Model Sensitivity

For *Cryptosporidium* and rotavirus, most of the variation in the predicted risk of infection was explained by the *E. coli* count in IWS tap water (*Cryptosporidium*: 45.86%; rotavirus: 81.42%) followed by the pathogen to *E. coli* ratio (*Cryptosporidium*: 32.75%; rotavirus 9.79%). For *Campylobacter*, the opposite was observed with 85.44% of the variation explained by the *Campylobacter* to *E. coli* ratio followed by the *E. coli* count in IWS tap water with 8.52%. The dose response parameters for each pathogen explained most of the remaining uncertainty followed by the tap water consumption variable. The sensitivity analysis summarized in Tables B4, B5, and B6, highlights the importance of the *E. coli* counts in IWS tap water and the ratio of the reference pathogens to *E. coli* in estimating the risk of infection in the current assessment.

Global Population Served by IWS

Our preliminary estimate of the IWS population based on WHO reports and the 2015 JMP data and summarized in Table 1, found that approximately 1 billion people were likely exposed to IWS. The results of our more robust estimate made using IBNET and JMP data,

listed in Table B7, indicate that the global population served by IWS is 925 million (95% CI: 670 – 1,130 million) with almost half (44.2%) of those exposed living in South-east Asia and a significant number living in India (Figure B8).

Diarrheal Burden of Disease Calculations

Given the estimated population served by IWS and the median annual infection risk, the reference pathogens together account for 17.2 million (95% CI: 7.76 – 32.3) infections annually among IWS users. Of these infections, 83% are attributable to *Campylobacter*, 17% to rotavirus, and less than 1% to *Cryptosporidium*. These infections cause 4.52 million (95% CI: 2.04 – 8.36) cases of diarrhea annually with *Campylobacter* accounting for 95% of these cases while *Cryptosporidium* and rotavirus account for 1% and 4% each. These cases of diarrhea cause 109,000 DALYs (95% CI: 48,800 – 223,000) and 1,560 deaths (95% CI: 699 – 3,150) per year. Burden of disease estimates based on the median infection risks are summarized by WHO region in Table 4. Rotavirus accounts for 82.1% of annual diarrheal DALYs and deaths, while *Campylobacter* accounts for 18.1% of DALYs and deaths. In this exposure scenario, *Cryptosporidium* accounts for less than 1% of both annual DALYs and deaths among users of IWS. The burden of disease stratified by etiology is tabulated in Table B8. The predominance of rotavirus in the annual diarrheal disease burden is driven by its high DALY weighting in LIMCs (0.482 per case) along with its high LMIC case fatality rate (0.6%). *Campylobacter*'s burden of disease is driven by its high risk of infection, one order of magnitude greater than rotavirus, and population susceptibility of 100%. While it is also assumed that 100% of the population is susceptible to diarrheal disease from *Cryptosporidium* infection, the median infection risk for the organism is two orders of magnitude less than that of *Campylobacter*.

TABLE 4: Annual infections, diarrheal cases, DALYs, and deaths attributable to IWS as calculated using the median daily probability of infection and its associated 95% confidence interval for *Campylobacter*, *Cryptosporidium*, and rotavirus assuming consumption of fecally contaminated tap water from an IWS.

REGION	POPULATION SERVED BY IWS (MILLIONS)	ANNUAL INFECTIONS (MILLIONS)	ANNUAL DIARRHEAL CASES (MILLIONS)	ANNUAL DEATHS	ANNUAL DALYs (THOUSANDS)
Africa	116	2.16 95% CI: 0.973 – 4.06	0.566 95% CI: 0.256 – 1.05	196 95% CI: 88 – 395	13.7 95% CI: 6.12 – 28.0
Americas, LMI	47.0	0.874 95% CI: 0.394 – 1.64	0.229 95% CI: 0.104 – 0.424	79 95% CI: 36 – 160	5.55 95% CI: 2.48 – 11.3
Eastern Mediterranean, LMI	103	1.91 95% CI: 0.864 – 3.60	0.503 95% CI: 0.227 – 0.930	174 95% CI: 78 – 351	12.2 95% CI: 5.43 – 24.8
Europe, LMI	71.0	1.32 95% CI: 0.596 – 2.48	0.346 95% CI: 0.157 – 0.641	120 95% CI: 54 – 242	8.38 95% CI: 3.75 – 17.1
South-East Asia	409	7.60 95% CI: 3.43 – 14.3	2.00 95% CI: 0.902 – 3.69	691 95% CI: 309 – 1,390	48.3 95% CI: 21.6 – 98.6
Western Pacific, LMI	179	3.33 95% CI: 1.50 – 6.26	0.874 95% CI: 0.395 – 1.62	302 95% CI: 135 – 609	21.1 95% CI: 9.44 – 43.2
Global	925	17.2 95% CI: 7.76 – 32.3	4.52 95% CI: 2.04 – 8.36	1,560 95% CI: 699 – 3,150	109 95% CI: 48.8 – 223

The cumulative distributions of the annual burden of diarrheal disease for each reference pathogen, shown in Figures B5, B6, and B7, indicate that the annual burden for *Campylobacter* exceeds the WHO health threshold (10^{-6} DALYs/person-year) at the 39th percentile, *Cryptosporidium* at the 62nd percentile, and rotavirus at the 33rd percentile. The cumulative distributions of total diarrheal DALYs and deaths among the 925 million global users of IWS, shown in Figure B9, indicate that the upper bounds, as defined by the 90th percentile, are 30.9 million diarrheal DALYs and 394,000 deaths.

Uncertainties and Limitations

As with all QMRA approaches, there are uncertainties and limitations in the input variables that should be accounted for when interpreting the results. A significant source of uncertainty for our risk is the absence of direct measurements of pathogen concentrations in IWS distribution networks. Without these measurements, across settings and time, we relied on estimated concentrations of reference pathogens by proxy using fecal indicator bacteria measurements and ratios of pathogens to indicators in possible sources of contamination. Concerning fecal indicator bacteria, we were only able to pool data from three high-quality studies conducted in India, Cambodia, and Vietnam. These studies represent a small portion of the geographical range of IWS, globally, and include no data from South America and sub-Saharan Africa. The *E. coli* datasets used in this analysis also do not include first flush data when fecal indicator concentrations may be much higher (22,82). Further, the pooled dataset consists of *E. coli* measurements from both urban and rural supplies, which prevents stratifying infection risk by urban and rural location, a potential risk factor for contamination in piped water supplies (7). Together, these two uncertainties prevent us from examining the variation in risk across geographic and human settlement location and we are confined to providing an estimate of risk across all IWS users.

Concerning ratios of pathogens to indicators in potential sources of contamination, the correlation between pathogens and indicators in any medium have proven highly variable (83). In raw sewage, the concentration of indicator bacteria is fairly constant whereas the concentration of pathogens varies as a function of the infection prevalence in the

contributing population (84,85). Thus, it is important to characterize the ratio using a distribution to capture this variability. There are few published datasets of pathogen to *E. coli* ratios in sewage particularly in LMICs; in this study, we derived ratios using datasets from the Netherlands and Germany. These datasets likely underestimate the pathogen loadings in sewage in LMICs where higher prevalence of diarrheal infection could result in increased pathogen concentrations relative to indicators in sewage (86). For example, the mean ratio of norovirus GII to *E. coli* measured in wastewater drains and wastewater-impacted streams was around 6.3×10^{-4} in Accra, Ghana, which is several orders of magnitude higher than the ratio assumed for rotavirus in this study (87). The pathogen to *E. coli* ratios used in this study likely lead to risk estimates that are conservative.

Sources of uncertainty can also be found in the assumptions underlying exposure assessment. First, in the absence of untreated tap water consumption data from LMIC settings, we modeled daily tap water consumption as a uniform distribution from 1 to 2 liters based on exposure scenarios articulated in EPA and WHO estimates (36,73). This probability distribution is not likely to be representative of water consumption behavior in settings where supplies are deficient and consumer behaviors include a complex system of household water management (6). Second, the scenario being modeled is the consumption of drinking water as it is delivered to the tap. This behavior is unlikely in an IWS where users, who are accustomed to supply interruptions, may obtain water from multiple sources and often store water in tanks, cisterns, and other containers for hours to days before the water is used. Household water handling and storage involve several risk factors for contamination, such as unsafe storage and access; including these behaviors in the model would likely increase the estimated risks of infection (88,89). On the other hand, some

households with IWS may employ point-of-use water treatment systems, which mitigate the risks posed by contamination if operated correctly and consistently over time. High-quality datasets of *E. coli* measurements in household storage facilities and household water treatment behavior in an IWS remain limited and make accounting for such variables in a risk framework difficult (24). It should be noted that this risk assessment does not include scenarios beyond daily consumption of drinking water. Therefore, the estimated risks of infection and subsequent burden of disease calculations do not include infection and disease from water quantity related behaviors such as food and hand washing or the use of water for household hygiene, which are likely modulated by the water scarcity associated with IWS.

Further uncertainty is introduced to the risk assessment by the population-specific dose-response functions for the reference pathogens used in the model. The dose-response data for each of the reference pathogens were collected in human feeding studies conducted in high-income settings with healthy, and generally, for rotavirus, male, adults. These dose-response functions may underestimate the risk of infection for persons living in LMICs, including children under five who suffer disproportionately from enteric disease, and attendant risks associated with non-diarrheal effects of exposure including the range of effects potentially associated with environmental enteric dysfunction (EED) and its potential downstream impacts (66,90). For each reference pathogen, the only disease endpoint considered was diarrhea, which neglects other, potentially more severe health outcomes such as stunting and chronic undernutrition related to EED (91). These dose-response functions also do not consider the risk of infection among people living in LMICs who may be more susceptible to infections due to compromised immune status or who,

conversely, may benefit from acquired immunity due to endemic exposure. Additionally, dose-response models do not yet take into account the effects of co-infection, which is prevalent in LMIC settings and may lead to increased risks of infection and longer-term sequelae. The risks associated with unsafe water are co-distributed in populations that are also at risk of undernutrition, high prevalence of co-infections, and other risk factors that would tend to exacerbate the effects of waterborne pathogen exposure. Risk estimates do not consider the elevated risks likely for infants, children, the undernourished, the immunocompromised, and those who are unlikely to receive timely treatment for diarrheal disease (e.g. oral rehydration therapy), which can dramatically reduce the risk of mortality among children in particular (92).

Besides the previously mentioned limitations in estimating the risks of infection, further sources of uncertainty in the burden of disease calculations include the both the estimates of the IWS population and the diarrheal disease weighting metrics. In regard to the population exposed to IWS, the JMP piped-on-premise measures do not include those who receive water from standpipes served by distribution systems. Additionally, the IBNET database relies on self-reported data from utilities that are mostly located in urban areas. Taken together, our estimates using these assumptions likely underestimate the population exposed to IWS. For the diarrheal disease per-case burden, the use of rotavirus per-case DALY weighting for LMICs instead of that for high-income countries increases the overall burden of disease and means the rotavirus burden has an outsized effect on the overall burden estimates. For instance, in LMICs, the rotavirus DALY weighting is 0.482 per case with a case fatality rate of 0.6%; in high-income countries, the recommended DALY weighting is only 0.0142 per case and the case fatality rate is 0.015% (81). We have

presented the burden of disease based on the LMIC metrics, but we also provide alternative calculations with the high-income parameters in Table B9.

Data Gaps

A recent review proposed a comprehensive research agenda relating to IWS (16). Our study further supports this agenda by identifying key data gaps for estimating the health risks attributable to IWS at the population level. First, there is a clear need for direct pathogen measurements from IWS networks in a range of settings, as water quality impacts may vary widely depending on local conditions. Such measurements could be used as direct input for a refined IWS risk assessment and could also be used to develop more robust pathogen to indicator ratios that can be applied to specific settings *vis-a-vis* fecal indicator measurements. Additionally, for enumeration of fecal indicators, larger volumes of water should be assayed to lower the detection limit to levels more appropriate for risk assessment. Another research area concerns consumer behavior with regard to tap water consumption, household water management and treatment, and household water contamination. Our risk assessment utilized tap water consumption data from settings that are probably not representative of the complex water management behavior often observed among IWS users. A more accurate estimate of the health risks associated with IWS must include these household behaviors in the exposure assessment model. This study also underscores the need for dose-response models that are specific to LMIC settings where acquired immunity, co-infections, and host susceptibility could dramatically alter the infection probabilities associated with ingesting microbial pathogens. Lastly, there is a need for a more robust estimate of the global population served by IWS. The estimate used in this analysis was based on the projection of IBNET data onto the JMP estimates of the

global population served by piped-on-premise water supplies, and a simple dichotomy between “intermittent” and “continuous” without accounting for the degree of intermittency (15). It is likely that this underestimates the total number of people served by IWS.

Policy Implications

Piped water supplies rely on multiple barriers including pipeline integrity, positive pressure, and chlorine residual to maintain the safety of the drinking water they deliver (93). These barriers, traditionally considered redundant, are more likely to fail simultaneously in the resource-constrained settings where IWS is prevalent. Our risk assessment indicates that the 925 million users of IWS are likely exposed to DALY burdens that exceed the WHO health threshold for each of the three reference pathogens considered. The predominance of risk due to the bacterial and viral pathogens in our estimate underscore the importance of an adequate chlorine residual in IWS distribution networks as a potential strategy to mitigate health impacts in the absence of massive investments to upgrade piped networks. Similarly, proper and consistent household water treatment and storage could mitigate the microbial risks of piped water supplies that are operated intermittently (94).

The Millennium Development Goal era has seen rapid expansion in coverage of piped water supplies, delivering a wide range of health and non-health benefits to communities (2). Increasing urbanization and population growth are likely to continue this trend. As more households connect to water supply networks, however, greater attention is needed on microbial risks associated with distribution systems, including those associated with

intermittent function. Accounting for these risks highlights the need for continued investment in provision of microbiologically and chemically safe water globally.

CHAPTER 3: LEVERAGING DEUF AND DDPCR TO DETECT WATERBORNE PATHOGENS AT LEVELS RELEVANT TO RISK-BASED STANDARDS: A CASE STUDY IN JAIPUR, INDIA

ABSTRACT

Quantitative microbial risk assessment relies upon robust characterization of human exposure to infectious microbes to make reasonable estimates of health risks. For assessments of the health risks attributable to distributed drinking water, the exposure assessment often relies on measures of fecal indicators and pathogen to indicator ratios as proxies for actual measurements of pathogens. In the work described here, we used dead-end ultrafiltration (DEUF) and droplet digital PCR (ddPCR) to detect and quantify genes associated with waterborne pathogens in an intermittent water supply in Jaipur, India. We adapted previously described qPCR assays to droplet digital format and determined 95% limits of detection (LOD) for each assay. We interpreted the results of our ddPCR experiments using both manual and model-based thresholding. Our results highlight several important findings. First, statistical interpretations of ddPCR results require that extraction methods previously developed for qPCR be optimized for ddPCR which is sensitive to slight variations in fluorescence amplitude in negative clusters. Second, groundwater from urban Jaipur, which is not treated prior to distribution, is likely an important source of waterborne disease as we detected genes associated with *Cryptosporidium* spp., *Giardia duodenalis*, and ST-ETEC in samples from tube wells. And lastly, the 95% LODs observed for the ddPCR assays combined with the DEUF workflow indicate that without significant additional improvements in the amount of drinking water

filtered and downstream concentration, waterborne pathogens can only be detected and at levels well above those required by risk-based standards. Thus, exposure assessments of distributed drinking water will continue to rely on various forms of indicator and waterborne pathogen models to make robust estimates of risk.

INTRODUCTION

Quantitative microbial risk assessment (QMRA) has been found to be a useful tool for managing drinking water safety (38). A recent comparison of QMRA-estimated risks and epidemiologic observations found reasonable agreement between the two for outbreaks of *Cryptosporidium* and *Giardia* (95). However, this agreement is predicated on the availability of dose-response models that characterize the population of interest and accurate measurements of the relevant pathogens in the exposure medium. Such measures of waterborne pathogens are often unavailable, especially when considering risk assessments of distributed drinking water. In water distribution networks, there is great uncertainty and variation in pathogen concentration due to the complex physicochemical processes underway in piped networks (96). Consequently, risk assessments of distributed drinking water have relied upon measures of fecal indicator organisms and ratios of waterborne pathogens to these indicators in potential sources of contamination with sewage often assumed as the worst case scenario (42,51,97,98). The end result of this approach is great uncertainty at the moment of exposure via ingestion of drinking water. Sensitivity analyses associated with two QMRAs of distributed drinking water indicate that the

estimated risks at the time of exposure are most sensitive to the concentration of the microbial contaminant (97,99). Therefore, measures of waterborne pathogens in distributed drinking water at levels relevant to public health are vital to improving our understanding of health risks attributable to piped water supplies.

The concentration of waterborne pathogens acceptable in distributed drinking water is defined by the acceptable risk of infection (100). The United States Environmental Protection Agency (US EPA) defines acceptable risk as an annual infection risk not to exceed 1 in 10,000 persons (40). However, given actual rates of waterborne illness in the US, a risk of 1 infection in 1,000 persons per year has been suggested (32). If we assume a constant and independent daily probability of infection, we can estimate the daily probability of infection required to realize an annual risk of 1 in 10,000 persons or 1,000 persons by calculating one minus the 365th root of one minus the daily probability of infection (30). An annual risk of infection of 1 in 10,000 equates to a daily risk of infection of 2.74×10^{-7} (1 in 3.65 million), while 1 in 1,000 annual equates to 2.74×10^{-6} (1 in 365 thousand). The WHO defines acceptable risk via a burden of disease metric using the disability adjusted life year (DALY) as 10^{-6} DALYs per person per year, which, when subject to the same assumptions regarding constant and daily disease burden, equates to a daily burden 2.74×10^{-9} DALYs. By adopting the reference pathogens of the Guidelines for Drinking-Water Quality (GDWQ) and their associated dose-response functions and disease burden metrics, and assuming a daily tap water consumption of 1 liter, we can estimate the concentration of each reference pathogen required to exceed the acceptable risk levels as shown in Table 5 (36).

TABLE 5: Concentration of reference pathogens at which the annual infection risk is equal to the EPA level of acceptable risk for drinking water, an annual infection risk of 1 in 10 persons, and the WHO acceptable risk level for drinking water as calculated assuming a daily consumption of 1 liter of drinking water and the mean dose-response parameters as reported in the previous chapter.

REFERENCE PATHOGEN	ANNUAL INFECTION RISK: 1 IN 10,000	ANNUAL INFECTION RISK: 1 IN 10	ANNUAL BOD: 10 ⁻⁶ DALYs
<i>Campylobacter</i>	3.14 CFU in 100,000 liters	3.32 CFU in 100 liters	2.28 CFU in 10,000 liters
<i>Cryptosporidium</i>	4.79 oocysts in 1,000,000 liters	5.05 oocysts in 1,000 liters	4.66 oocysts in 100,000 liters
rotavirus	5.87 FFU in 10,000,000 liters	6.19 FFU in 10,000 liters	8.27 FFU in 10,000,000 liters

Based on these estimates, detecting and quantifying waterborne pathogens at concentrations equivalent to currently formulated acceptable risk limits seems unrealistic. These calculations also indicate that traditional methods of assessing water quality such as enumerating fecal indicator organisms via membrane filtration or spread plating of drinking water are inadequate for quantifying risks to the lowest thresholds, because test volumes are often only a few hundred milliliters (101–105). Additional uncertainty regarding the relationship between fecal indicators and waterborne pathogens also makes it difficult to quantify risks to public health using traditional fecal indicator measurements (83). For example, if *Campylobacter* were enumerated via membrane filtration and culture performed on ten 100 mL volumes of water, the limit of detection (LOD) would be 1 CFU per liter and the lowest risk quantifiable at this limit using the same dose-response parameters would be a daily probability of infection of 8.4 persons in 1,000 and an annual probability of infection of 95.5 persons in 100 (almost 10,000 times the EPA acceptable risk level).

These brief calculations demonstrate the need for microbiological methods that will allow for the assay of large volumes of water and the sensitive and specific detection of waterborne pathogens in these volumes. The pairing of two newly developed methods in environmental microbiology hold promise for improving the reliable detection and quantification of waterborne pathogens in drinking water at concentrations more informative for assessing and managing risk.

First is the use of ultrafiltration to concentrate microbes from large volumes of water (106,107). Tangential-flow ultrafiltration (TFUF) has proven to be an effective method for the simultaneous recovery of bacteria, protozoa, and viruses from large-volume drinking water samples (106–108). In their experiments, Hill, Polaczyk, and colleagues found that after ultrafiltration backflushing with solutions of sodium polyphosphate (NaPP), Tween 80, and Antifoam A or Y-30 Antifoam Emulsion yielded higher recoveries than elution (106,107). Sodium polyphosphate acts as a dispersant that increases repulsion between negatively charged microbes and ultrafilter surfaces (109). Tween 80 is a surfactant that minimizes hydrophobic interactions between microbes and the ultrafilter surface, while Antifoam A or Y-30 antifoam emulsion minimize the foaming associated with the addition of the surfactant (110).

Because TFUF requires a complex configuration, dead-end ultrafiltration (DEUF) was developed as a simpler configuration for the recovery of microbes from water in field settings (111). In tests of low- and mid-range turbidity (0.29 NTU and 1.5 NTU, respectively) drinking water samples seeded with microbes, DEUF followed by backflushing led to average recoveries of 93% for *E. faecalis*, 57% for MS2 phage, 94%

for *C. perfringens* spores, and 87% for *C. parvum* oocysts (111). While for high-range turbidity (4.3 NTU) drinking water samples, mean recoveries were 78% for *E. faecalis*, 73% for MS2 phage, 57% for *C. perfringens* spores, and 83% for *C. parvum* oocysts. DEUF has also performed well in recovering microbes from surface water samples with turbidities much higher than those typically observed in drinking water. For surface water samples with mean turbidities ranging from 16 to 92 NTU, recovery efficiencies of 85%, 81%, 66%, 63%, and 49% were observed for enterococci, *E. coli*, MS2 phage, *C. perfringens* spores, and *C. parvum* oocysts, respectively (112).

The second promising development is the use of droplet digital polymerase chain reaction (ddPCR) for the absolute quantification of nucleic acid templates (113,114). As Pinheiro and Hindson describe, in ddPCR a typical qPCR reaction is fractionalized into tens of thousands of nano-liter sized reaction volumes. Thermal cycling is performed and, when present, the target is amplified in an individual droplet with an associated increase in the fluorescence amplitude of that droplet. Following thermal cycling, the fluorescence amplitude of individual droplets is measured using a flow cytometer and droplets are classified as either positive for the target or negative based on a fluorescence threshold. Finally, the absolute copy number of the target is estimated using a most probable number technique with an assumed Poisson distribution, the observed proportion of droplets positive, and the volume of each droplet.

Droplet digital PCR has been used to quantify *Salmonella* in river sediments, Shiga-toxin-producing *E. coli* (STEC) in bovine feces, foodborne pathogens in soft cheeses, and fecal-indicators in environmental water (115–117). The method shows great promise for the

detection of waterborne pathogens in drinking water due to its reported sensitivity and resilience to inhibitory substances (118–120).

Importantly, previous application of molecular methods downstream of DEUF indicate that ddPCR is a viable means of assaying DEUF concentrates. When DEUF concentrates were assayed using real-time (qPCR) and reverse transcriptase real-time (RT-qPCR) polymerase chain reaction, no inhibition was associated with the addition of the backflushing reagents (107). The quality of the tap water, however, was found to affect the performance of molecular methods when 100 liter samples were assayed using qPCR and RT-qPCR, but no single water quality parameter was associated with increasing crossing point thresholds (108). During recovery experiments, MS2 and *Cryptosporidium* were consistently detected in concentrates via qPCR and RT-qPCR (112). These data indicate that DEUF and ddPCR could be paired for sensitive detection of microbes in drinking water.

Herein, we report on the application of DEUF and ddPCR to detect and quantify genetic elements associated with waterborne pathogens in drinking water samples collected from the municipal water supply in Jaipur, India. Jaipur is the capital of the northern semi-arid state of Rajasthan and receives an average of 600 millimeters of rainfall annually (121). Until the early 2000s, the residents of Jaipur (currently 3 million) received their drinking water exclusively from groundwater sources (122,123). But, overexploitation of aquifers for irrigation has caused drastic water table decline making dependence on groundwater for drinking increasingly untenable (124,125). In 2006, Jaipur was allocated surface water from Bisalpur Dam, 120 km southwest of the city, and by 2009 the city was augmenting its groundwater wells with surface water (122). Before being pumped from the dam to the city, raw surface water is subjected to filtration and chlorination at the Surajpura Water

Treatment Plant (122). Once it arrives in the city, the surface water (275 MLD) along with extracted groundwater (97 MLD) is distributed to 162 water supply zones via a network of pipelines, elevated storage reservoirs (ESR), and pump stations (122). The distribution network is operated intermittently with each supply zone receiving 1.5 to 3 hours of pressurized distribution per day (122,126).

In conjunction with a study of geogenic contaminants and with assistance from staff of the Ground Water Department and the Public Health Engineering Department of the Government of Rajasthan, we conducted a cross-sectional sampling of municipal drinking water in Jaipur in May of 2017 (127).

MATERIALS AND METHODS

We collected paired grab (n=23) and DEUF (n=22) samples at various points in the Jaipur water distribution network including groundwater tube wells, surface water distribution pipelines, and mixed source water from ESRs. Although our sampling locations were a convenience sample, we sought to collect water from locations in the distribution system where groundwater and surface water were hydraulically isolated, as well as mixed source locations that were spatially distributed throughout the municipal water supply as shown in Figure 4.

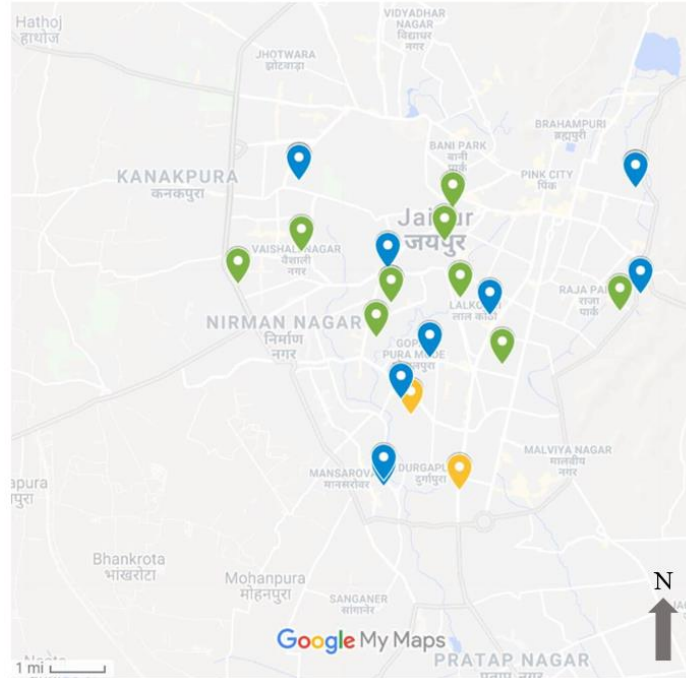


FIGURE 4: Sample collection locations from the cross-sectional sampling of the Jaipur municipal water supply with groundwater source samples shown in blue, surface source samples in green, and mixed source in yellow.

Physicochemical Measurements

At the time of sample collection, we measured physicochemical parameters including free chlorine, total chlorine, pH, conductivity, and total dissolved solids (TDS). We measured total and free chlorine using the US EPA DPD Colorimetric Method and a Hach Pocket Colorimeter II spectrophotometer (Hach, Loveland, CO) (128). We measured pH, conductivity, and TDS by electrode using a Hanna Low Range Combo Tester (Hanna Instruments, Woonsocket, RI) (129).

Fecal Indicator Bacteria by Culture

During each sampling event, we collected 500 mL grab samples in Whirl Pak bags pre-dosed with sodium thiosulfate to quench residual chlorine (Nasco, Fort Atkinson, WI). In

the time between collection and processing, we stored and transported all samples on ice. We enumerated total coliforms and *E. coli* in two 100 mL replicates from each grab sample by membrane filtration and selective media incubation per EPA Method 1604 (with modification) (104). We performed membrane filtration within 8 hours of sample collection using a Del Agua Filtration Set (Del Agua, Fyfield, UK). We incubated each replicate along with positive and negative controls at 35°C for 24 hours on Compact Dry-EC plates (Hardy Diagnostics, Santa Maria, CA). Following incubation, we counted both total coliforms and *E. coli* on each plate following the manufacturer's instructions for color-based identification and reported the counts in colony forming units (CFU) per 100 mL.

Dead-End Ultrafiltration

At each sampling location in Jaipur, we also concentrated suspended solids including microbes from large volumes of drinking water by DEUF following a CDC protocol. Because of the superior hydraulic performance noted by Smith and Hill, we used the Rexeed 25S dialyzer (Asahi Kasei Kuraray Medical Company, Tokyo, Japan), with a molecular weight cut off (MWCO) of 30 kilodaltons and an effective surface area of 2.5 m², for DEUF (111). Where possible, we connected to water sampling points via sterile Masterflex L/S 36 Platinum-Cured Silicone tubing (Cole Parmer, Vernon Hills, IL) and utilized the pressure in the piped network to force the drinking water through the dialyzer. When system pressure was inadequate, we collected drinking water in sterilized plastic buckets and filtered from the buckets using a Geotech Geopump (Geotech, Denver, CO). We recorded the volume of water we filtered using a totalizing flow meter (Clark Solutions, Hudson, MA). Immediately following filtration of each drinking water sample, we flushed

the dialyzer with a 500 mL 1% sodium thiosulfate solution to quench residual chlorine. In the laboratory, we backflushed each dialyzer into a sterile 500 mL bottle using a 500 mL 0.5% Tween 80, 0.01% NaPP, and, 0.001% Antifoam Y-30 emulsion solution as described by Smith and Hill (111). The volume of solution recovered during backflush ranged from 350 to 450 mL. We enumerated total coliforms and *E. coli* in the backflush solution using membrane filtration and culture as previously described and a series of 1:10 dilutions to achieve countable plates. Following culture-based enumeration in the backflush, we estimated the number of *E. coli* and total coliforms per 100 mL volume of drinking water using the backflush volume, bacterial recoveries as noted by Smith and Hill, and the total volume of drinking water filtered.

PEG Precipitation and Ultracentrifugation

After enumerating the *E. coli* and total coliforms in the backflush, we further concentrated it using polyethylene glycol (PEG) precipitation and ultracentrifugation (106,112). As described by Mull and Hill, we performed PEG precipitation using a 12% PEG 8000, 0.9 M NaCl, and 1% bovine serum albumin solution. After addition of the reagents and mixing thoroughly by slow hand shaking, we incubated the samples overnight at 4°C. Following incubation, we centrifuged 300 mL of the PEG-precipitated solution in six 50 mL centrifuge tubes at 10,000 times gravity at 4°C for 30 minutes. We slowly poured off the resulting supernatant and resuspended the pelleted material in each tube using 667 uL of a 1X phosphate-buffered saline (PBS), 0.01% Tween 80, and 0.001% Y-30 emulsion solution. We then recombined the resuspended pellets from each tube and the resulting volume of resuspended concentrate ranged from 3 to 4 mL. We stored the DEUF concentrate for each sample at -20°C until we pre-treated using UNEX lysis buffer and

bead beating as described in the next section. Following pre-treatment, we transported the samples at room temperature for approximately 36 hours to the laboratory in Atlanta, GA where we froze them at -80°C until extraction and further molecular analysis.

Nucleic Acid Extraction

We performed nucleic acid extraction using a universal extraction buffer, UNEX (Microbiologics, St. Cloud, MN), as developed by the CDC. In comparative testing of the UNEX buffer compared to other commercially available kits, samples extracted from large volume water samples with UNEX had comparable crossing threshold (CT) values during qPCR (130). UNEX buffer has also been previously used to extract concentrates from DEUF for downstream molecular analysis (108,112). We extracted the DEUF concentrate following the protocol for nucleic acid extraction from parasites in water samples as detailed in the manufacturer's instructions for use. Briefly, we added 500 uL of DEUF concentrate, 500 uL of UNEX buffer, and 5 uL of Inforce 3 Bovine Vaccine (Zoetis, Parsippany, NJ) (our process control containing bovine respiratory syncytial virus, BRSV, and bovine herpes virus, BoHV) to an SK-38 bead tube (Bertin Corp, Rockville, MD) and incubated the mixture for 15 minutes at room temperature and then bead beat the mixture for 2 minutes. Following bead beating, we centrifuged the bead beating tubes and transferred 500 uL of the supernatant into a HiBind mini column (Omega BioTek, Norcross, GA) and centrifuged at 10,000 times gravity for 1 minute. We then washed the column with 100% ethanol and 70% ethanol consecutively. Finally, we eluted the purified nucleic acids from the column using 100 uL of 10 mM Tris-1mM EDTA (pH 8.0) buffer and measured the resulting concentration of DNA or RNA using a Qubit Fluorometer (Invitrogen, Carlsbad, CA). With each extraction batch we included a negative extraction

control and an Inforce 3 only control (500 uL of molecular water and 5 uL of Inforce 3). We stored the purified nucleic acid from each sample and the controls at -80°C until molecular analysis or reverse transcription.

Reverse Transcription

Prior to ddPCR assay for RNA targets, we performed reverse transcription of RNA to cDNA using a High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (ThermoFisher Scientific, Waltham, MA). We performed reverse transcription (RT) per the manufacturer's instructions by adding 10 uL of nucleic acid extract to 10 uL of RT master mix. The kit utilizes a random hexamer primer method to initiate cDNA synthesis through the enzymatic activity of MultiScribe Reverse Transcriptase. Following completion of RT, we stored the resulting cDNA at -80°C until ddPCR interrogation with a maximum storage length of five days.

Microbes of Interest and Molecular Assays for ddPCR

We selected microbes of interest for molecular analysis on the basis of diarrheal disease etiologies from several sites in southeast Asia as observed in the Global Enteric Multicenter Study (GEMS) and the Malnutrition and Enteric Disease Study (MAL-ED) as summarized in Figures C1, C2, and C3 (67,131). Based on their associations with diarrheal disease in South Asian settings, we selected *Shigella*/enteroinvasive *E. coli* (EIEC) and enterotoxigenic *E. coli* with heat stable toxin (ST-EPEC) as bacterial microbes of interest, *Cryptosporidium* spp. and *Giardia* spp. as protozoan targets, and norovirus GI, norovirus GII, and adenovirus A-F as viral targets of interest. We did not select rotavirus as a viral

target due to the rollout of rotavirus vaccination programs in India which could confound our detection of wildtype rotavirus in drinking water via ddPCR (132).

To test the DEUF concentrate for the selected pathogens, we adopted previously published qPCR assays and adapted them to ddPCR. To maximize the possibility of detection of adenovirus, we selected an assay targeting adenovirus types A to F via the hexon gene, which encodes a major capsid protein (133). For norovirus GI and GII, we selected assays which target the ORF1/ORF2 junctions for each genotype (134,135). To detect *Shigella* spp. and enteroinvasive *E. coli* (EIEC), we selected an assay targeting the invasion plasmid antigen H gene (136). We tested DEUF concentrates for enterotoxigenic *E. coli* (ETEC) via an assay targeting the gene that encodes human heat stable toxin (STh) (137). We tested for *Cryptosporidium* using an assay targeting the genetic sequence encoding the 18S rRNA as developed for molecular testing in the GEMs and MAL-ED trials (138,139). Lastly, we tested for *Giardia duodenalis* using an assay that targets the beta giardin gene (140).

In addition to waterborne pathogens, we also performed molecular analysis for genes associated with microbial indicators *E. coli* and male-specific coliphage (MS2). For *E. coli*, we selected a qPCR assay targeting the ybbW gene, a putative allantoin transport protein, for adaptation to ddPCR because the assay has demonstrated 100% sensitivity and specificity for the bacteria as compared to 16S and 23S assays which demonstrate lower specificities for *E. coli* (141–144). For the detection and quantification of MS2, we selected an assay targeting the MS2g1 gene, which encodes a maturation protein, as the ddPCR assay based on published use of MS2 as an internal control in reverse transcriptase polymerase chain reaction (RT-PCR) assays for viral targets (67,145). We screened all primer, probe, and control material sequences via NCBI BLASTn to corroborate the

reported specificities and ordered all positive control genetic materials from Integrated DNA Technologies (Coralville, IA). All assays and relevant information including GenBank Accession numbers, sequences, and alignment positions are summarized in Table 6 for pathogens and Table C1 for control assays.

We performed all ddPCR workflow following the manufacturer's recommended protocol using a QX200 Droplet Generator, PX1 PCR Plate Sealer, C1000 Touch Thermal Cycler, and QX200 Droplet Reader (Bio-Rad, Hercules, CA) (146). For singleplex probe-based assays (MS2g1; beta giardin), we added each primer and the probe to achieve final concentrations of 900 nM and 250 nM, respectively. Whereas for multiplexed probe-based assays (norogI/norogII; hexon/ipaH; Crypto. 18S rRNA/STh), we added each primer and probe to achieve final concentrations of 650 nM and 250 nM, respectively. For the ybbW EvaGreen assay, we added each primer to achieve final concentrations of 250 nM. For all assay types and formats the final reaction volume was 20 uL including molecular-grade water, ddPCR Supermix for Probes or ddPCR EvaGreen Supermix (Bio-Rad, Hercules, CA), primers, probe, and 4 uL of extract or RT product from each sample. We included negative extraction controls, sample blanks, no-template controls, and an assay appropriate positive control with each ddPCR experiment. Each plate also included three randomly selected technical replicates and three randomly selected biological replicates for a total replication rate of 27%. We estimated the 95% limit of detection (LOD) for each ddPCR assay using positive control materials prepared in a serial dilution series and probit analysis as described by Stokdyk *et al.* (147).

Adapting qPCR Assays to ddPCR

We optimized the annealing temperature for each assay using temperature gradients following the manufacturer's recommendations (146). We used a two-step optimization by first running reactions with positive control materials and a 10°C annealing temperature gradient followed by a second narrower 4°C annealing temperature gradient. Following the second temperature gradient experiment, we selected the annealing temperature that achieved the greatest difference between the positive and negative droplet fluorescence amplitudes. To further minimize the quantity of droplets with ambiguous fluorescence amplitudes between the positive and negative clusters, known as rain, we set the temperature ramp rate to 2°C per second consistent with the findings of a systematic investigation of rain in a ddPCR assay for *Listeria monocytogenes* (148). Per the findings of Witte and colleagues, we also increased the denaturation temperature from 94 to 95°C for G/C rich templates and increased the annealing and extension time from 1 to 2 minutes for amplicons longer than 100 bp. The final thermal cycling conditions for each assay are summarized in Tables C2 and C3.

TABLE 6: Primer and probe sequence information for qPCR assays as adapted to ddPCR for interrogation of DEUF concentrate from drinking water samples collected in Jaipur, India.

MICROBE	GENE	GEN BANK ACCESSION		SEQUENCE	SEQUENCE POSITION	REFERENCE
adenovirus A-F	hexon	AC_000008	Forward Reverse Probe	GGACGCCTCGGAGTACCTGAG ACNGTGGGGTTTCTGAACTTGTT CTGGTGCAGTTCGCCCGTGCCA**	18895-18915 18990-18968 18923-18944	(133)
norovirus GI	ORF1- ORF2 junction	MG049693.1	Forward Reverse Probe	GCCATGTTCCGNTGGATG TCCTTAGACGCCATCATCAT TGTGGACAGGAGATCGCAATCTC**	5266-5283 5361-5342 5303-5325	(134)
norovirus GII	ORF1- ORF2 junction	AF145896.1	Forward Reverse Probe	CARGARBCNATGTTYAGRTGGATGAG TCGACGCCATCTTCATTCA TGGGAGGGCGATCGCAATCT**	5003-5028 5100-5080 5048-5067	(135)
<i>Shigella</i> /EIEC	ipaH	M76445.1	Forward Reverse Probe	ACCATGCTCGCAGAGAACT TACGCTTCAGTACAGCATGC TGGCGTGTCCGGAGTGACAGC**	1345-1364 1525-1506 1401-1421	(136)
ST-ETEC	STh	M29255.1	Forward Reverse Probe	TCCTGAAAGCATGAATAGTAGCAATTAC TTAATAGCACCCGGTACAAGCA ACAACACAATTCACAGCA*	171-198 243-222 199-216	(137)
<i>Cryptosporidium</i> spp.	18S rRNA	AF093491.1	Forward Reverse Probe	GGGTTGTATTTATTAGATAAAGAACCA AGGCCAATACCCTACCGTCT TGACATATCATTCAAGTTTCTGAC*	197-223 322-303 268-291	(138,139)
<i>Giardia duodenalis</i>	beta-giardin	AY072727	Forward Reverse Probe	GGCCCTCAAGAGCCTGAAC GGGCGATCGTCTCCTTCTC CTCGAGACAGGCATC*	402-420 544-526 424-438	(140)
MS2	MS2g1	NC_001417	Forward Reverse Probe	TGGCACTACCCCTCTCCGTATTCACG GTACGGGCGACCCACGATGAC CACATCGATAGATCAAGGTGCCTACAAGC**	160-185 258-237 201-229	(145)
<i>E. coli</i>	ybbW	NC_000913.3	Forward Reverse	TGATTGGCAAATCTGGCCG GAAATCGCCAAATCGCCAT	538033-538052 538224-538243	(141)

Interpreting ddPCR Results

In ddPCR, the concentration of a gene target in the reaction mix is estimated by means of a most probable number according to a Poisson distribution as shown in equation 1, where y is the number of droplets void of the target, m is the total number of droplets, and v is the volume of each droplet in μL (146).

$$\text{Conc} \left(\frac{gc}{\mu\text{L}} \right) = - \ln \left(\frac{y}{m} \right) * \frac{1}{v} \quad (1)$$

Droplets are classified as either positive or negative for the gene target based on their measured fluorescence amplitude following thermal cycling and a threshold; thus, thresholding is critical for accurate quantification in ddPCR assays (149). However, user-based thresholding creates opportunity for bias in the interpretation of ddPCR results (150). We estimated the concentration of target genes in our ddPCR assays using both a manual user-determined threshold in QuantaSoft (V1.7.4; Bio-Rad, Hercules, CA) and a model-based classification called Umbrella implemented in RStudio (Version 1.1.456; RStudio Team, 2015) (151). Briefly, the Umbrella package as devised by Jacobs *et al.* utilizes the fluorescence amplitude distributions of droplets from negative controls and a sample well to estimate the proportion of droplets negative in the sample and then uses Bayes' theorem to estimate the probability that each individual droplet is negative for the target based on the observed fluorescence amplitude for that droplet. The package then reports an estimate of the concentration based on the regression estimated proportion of droplets negative ("robust estimate") and an estimate based on the number of droplets with less than a 5% probability of being negative, which we refer to as the Umbrella threshold. For full details

see the description provided by Jacobs *et al.* in their manuscript and supporting information (151).

RESULTS AND DISCUSSION

During May 2017, we collected 24 paired grab and DEUF samples from throughout the Jaipur municipal water supply as shown in Figure 4. Due to a pump failure at the first sample tube well, DEUF samples 1 and 2 are a composite sample from two different tube wells located within a few hundred meters of one another, yielding a total of 23 unique grab samples and DEUF samples. All samples were collected from the piped water supply, but from points such that various source water types could be isolated. The sample set included 10 samples from groundwater sources (tube wells), 8 samples from surface water sources (ESRs and pipelines), 4 samples of mixed sources (1 household storage tank and distribution pipelines), and 1 sample from an unknown mixture.

Physicochemical Results

We measured total and free chlorine in 22 samples. Due to possible interference from oxidized iron and manganese in the samples, any chlorine value less than 0.1 mg/L was denoted as less than the method reporting limit (MRL). Of the nine groundwater samples where total and free chlorine were measured, all but one was less than the MRL for free chlorine and all but two for total chlorine. We detected free chlorine above the MRL in 88% (7/8) of surface source and 75% (3/4) of mixed source drinking water samples from the distribution network. We detected total chlorine above the MRL in all samples from the distribution network. The Indian Standard Drinking Water Specification requires that

treated water intended for drinking have a minimum free chlorine residual of 0.2 mg/L and suggests a minimum of 0.5 mg/L when disinfection of viruses is desired (152). We observed less than 0.2 mg/L of free chlorine in two of the 12 samples collected from the distribution network and less than 0.5 mg/L of free chlorine in five of the 12. No sample exceeded the WHO free chlorine guideline value of 5 mg/L (36). Box and whisker plots summarizing our findings are shown in Figure 5.

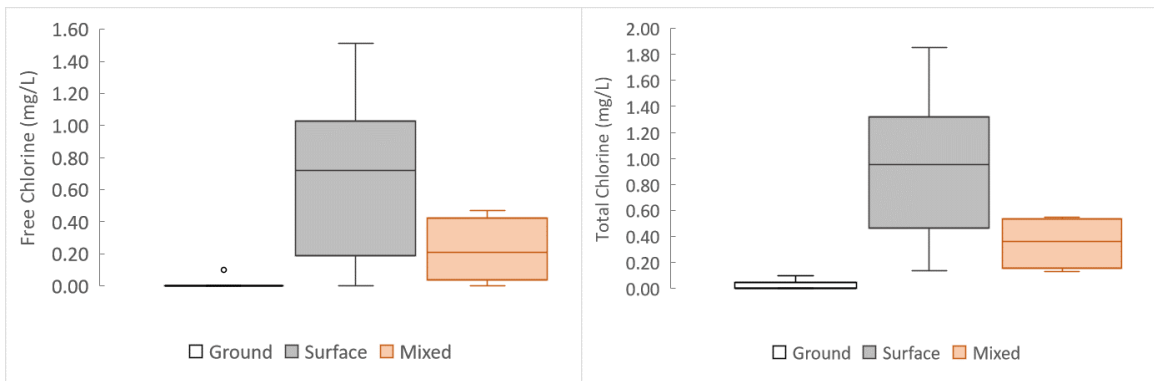


FIGURE 5: Box and whisker plots of free and total chlorine measures in groundwater samples from tubewells (n=9), surface source (n=8), and mixed source (n=4) drinking water samples from the municipal water supply in Jaipur.

We measured pH and conductivity in a total of 23 samples. The pH and conductivity measurements, summarized in Figure 6, reflect samples collected from two distinct sources groundwater and surface water, and samples from a mixture of the two. The pH in groundwater samples was consistently lower than the pH measured in surface water and the mixed water is between the two. The same pattern is observed in the conductivity measures, with measurements in surface water consistently lower than ground and mixed water.

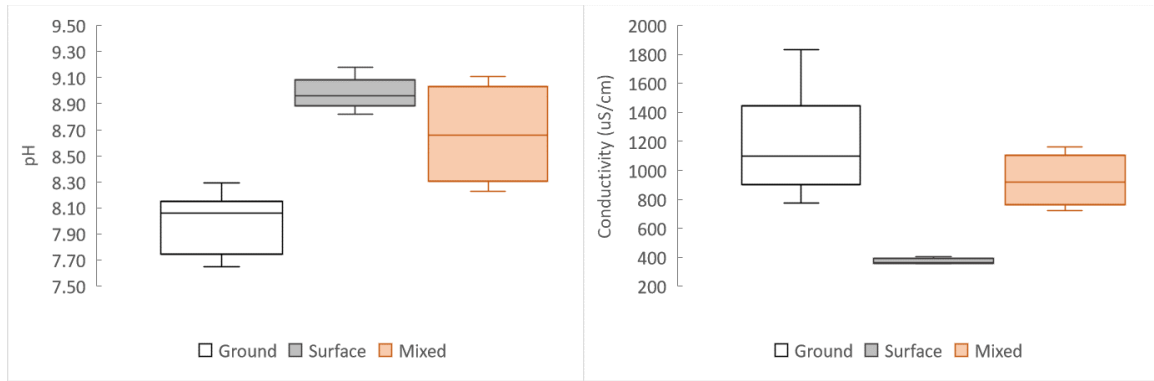


FIGURE 6: Box and whisker plots of pH and conductivity measures in groundwater samples from tubewells (n=10), surface source (n=9), and mixed source (n=4) drinking water samples from the distribution network in Jaipur.

Fecal indicators by Culture Results

We enumerated total coliforms and *E. coli* in 20 grab samples -- 9 in groundwater from tube wells, 7 in surface water from the distribution network, and 4 from mixed source water from the distribution network. The frequency distributions of the observed counts are shown in Figure 7. Total coliforms and *E. coli* were below the detection limit for all samples from surface water. For mixed source drinking water, *E. coli* was below the detection limit in each sample and total coliforms were below the detection limit in 75% of samples. In groundwater on the other hand, total coliforms were detected in 44% of grab samples and *E. coli* was detected in 22% of samples. The Indian Standard Drinking Water Specification requires that *E. coli* is not detected in 100 mL of any water intended for drinking. During our sampling event, 90% of the grab samples met this criteria (152). Per the WHO's Guidelines for Drinking-Water Quality (GDWQ), for a city with a population greater than 100,000 persons such as Jaipur, 90% of samples negative for *E. coli* would warrant a "Fair" rating (36).

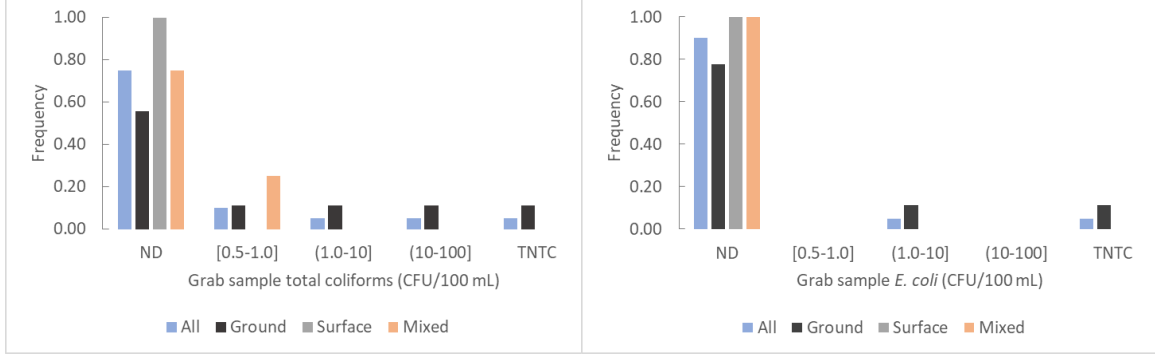


FIGURE 7: Frequency distribution of total coliform and *E. coli* counts as observed in grab samples from groundwater (n=9), surface source (n=7), and mixed source (n=4) drinking water samples from the distribution network in Jaipur.

We collected 23 DEUF samples and enumerated total coliforms and *E. coli* from the backflush. We filtered a total of 2,407.9 liters during DEUF sampling with an average of 104.7 liters, a minimum volume of 60.0 liters, and a maximum of 161.1 liters. To back calculate the number to bacteria in the original drinking water matrix in CFU per 100 mL (C_{DW}) we took the count of bacteria per mL in the backflush (C_{BF}) and multiplied by the total volume of the backflush in mL (V_{BF}). We then divided the result of this product by the product of the recovery efficiency (η , assumed to be 100% to estimate a conservative lower bound), the volume of drinking water filtered in liters (V_{DEUF}) and a conversion factor equivalent to 10 100-mL volumes per 1 liter as shown in equation 2.

$$C_{DW} = \frac{C_{BF} * V_{BF}}{\eta * V_{DEUF} * 10 \frac{100ml \text{ volumes}}{liter}} \quad (2)$$

The frequency distributions of total coliform and *E. coli* counts in DEUF samples are shown in Figure 8. We detected total coliforms in 83% of the DEUF samples including all four mixed source samples. We did not detect culturable *E. coli* in any surface water DEUF sample but we did detect it in 70% of groundwater DEUF samples and 25% of mixed

DEUF samples. *E. coli* were too numerous to count (TNTC) in 2 of 10 groundwater, and 1 of 4 mixed water DEUF samples. Only five DEUF samples, all from groundwater, yielded countable quantities of *E. coli* with a mean of 48.4 CFU per 100 mL and a standard deviation of 60.2 CFU per 100 mL.

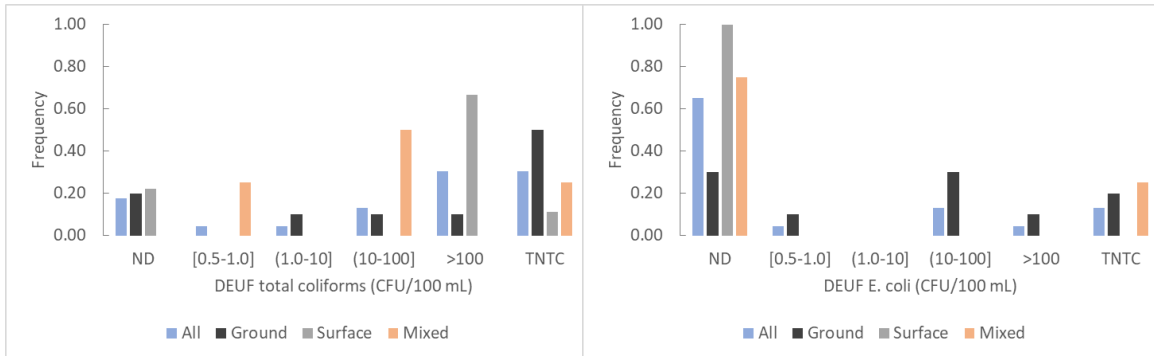


FIGURE 8: Frequency distribution of total coliform and *E. coli* counts as observed in DEUF samples from groundwater (n=10), surface source (n=9), and mixed source (n=4) drinking water samples from the distribution network in Jaipur.

ddPCR LODs

Based on serial dilution experiments with control materials spiked into sample blanks, we estimated the per reaction concentration at which 95% of ddPCR replicates are expected to be positive (95% LOD). The LODs for each assay are summarized in Table C2. The lowest LOD observed was 1.5 gene copies (gc) per reaction for hexon (adenovirus A-F); while the highest observed was 5.9 gc per reaction for the ORF1-2 target of norovirus GII. Based on the concentration factors associated with our previously described workflow and the range of drinking water volumes filtered, the ddPCR sample LODs can be propagated to equivalent LODs in the drinking water column. For example, the beta-giardin assay (LOD 1.7 gc/reaction) combined with the 161-liter filtration volume observed for one sample would yield an equivalent water column LOD of 6.3 gc per liter of drinking water,

assuming 100% recovery efficiencies for DEUF backflushing, PEG precipitation, and ultracentrifugation (detailed concentration factor calculations are shown in Table C4). For an RNA target from this same sample, such as norovirus GI (LOD 5.7 gc/reaction) the equivalent water column LOD would be 42.5 gc per liter of drinking water. As shown in Table C4, once the sample 95% LOD is established for each assay, the equivalent LOD in the drinking water column is a function of the volume of drinking water filtered by the dialyzer, the efficiency of DEUF recovery, the efficiency of the PEG precipitation and ultracentrifugation recovery, and the volumetric ratios of material flows into and out of each processing step.

Interpreting ddPCR Results

We began our analysis of the ddPCR results by inspecting the results of the qualitative process control spike of RNA virus BRSV via the Inforce 3 bovine vaccine. We interpreted successful detection of this control within a sample by ddPCR to indicate successful extraction, reverse transcription, and amplification. Following ddPCR, we detected BRSV in each sample by both manual and Umbrella thresholding with zero false positive droplets in the negative extraction controls (NECs) and no-template control (NTCs). Example Umbrella output along with detailed explanations of both manual thresholding and Umbrella are shown in Figures C4, C5, and C6. We observed perfect correlation between the number of droplets determined to be positive by manual thresholding and droplets called positive by Umbrella thresholding (Figure C7).

We analyzed the results of ddPCR for each molecular target in a systematic manner. First, we considered the results from the NECs, NTCs, and positive controls (PCs). Metrics of interest included the droplet false positive rate in the NECs and NTCs by both manual and

Umbrella thresholding and the successful detection of positive droplets in the PC. We compared manual and Umbrella thresholding results using five classifications for each reaction on the 96-well plate. First were wells we defined as “problematic” where the Umbrella threshold estimated concentration was one or more orders of magnitude greater than the manual thresholding concentration. Next were wells we termed “discordant detections” where the target was detected by one thresholding method but not by the other. Third were wells we termed “concordant detections” where the target was detected by both thresholding methods. Finally, for both the manual thresholding and Umbrella thresholding we also tallied the detections below and above the 95% LOD. We report the tabulated classifications stratified by assay in Table 7.

We observed zero false positive droplets in all negative controls across all assays for manual thresholding. Whereas for Umbrella thresholding we observed 1 false positive droplet in a negative extraction control for MS2g1, 5 false positive droplets (4 in an NEC, and 1 in an NTC) for ipaH, and 5 false positive droplets in an NEC for hexon. In each of these instances, the misclassification is caused by increased variance in the negative amplitude cluster of the NEC or NTC compared to another NTC on the plate. This increase causes a few of the outlying droplets associated with the negative cluster to be classified as positive in the Umbrella output. As an example, variances in the fluorescence amplitude in the negative clusters of the negative controls for the hexon experiment are shown in Figure C8.

TABLE 7: ddPCR analysis performance metrics tabulated for manual and Umbrella thresholding stratified by molecular assay as observed in ddPCR assays of DEUF concentrate collected from the municipal drinking water supply in Jaipur, India.

MOLECULAR ASSAY	MANUAL THRESHOLD	UMBRELLA THRESHOLD
MS2g1		
<i>false positive rate</i>	0 of 48,902	1 of 48,902
<i>problematic wells</i>		1
<i>discordant pos.</i>		1
<i>concordant pos.</i>		7
<i>pos. below 95% LOD</i>	7	8
<i>pos. above 95% LOD</i>	0	0
beta giardin		
<i>false positive rate</i>	0 of 45,494	0 of 45,494
<i>problematic wells</i>		2
<i>discordant pos.</i>		8
<i>concordant pos.</i>		1
<i>pos. below 95% LOD</i>	0	1
<i>pos. above 95% LOD</i>	1	7
NoV GII		
<i>false positive rate</i>	0 of 55,415	0 of 55,415
<i>problematic wells</i>		0
<i>discordant pos.</i>		6
<i>concordant pos.</i>		1
<i>pos. below 95% LOD</i>	7	1
<i>pos. above 95% LOD</i>	0	0
NoV GI		
<i>false positive rate</i>	0 of 55,415	0 of 55,415
<i>problematic wells</i>		0
<i>discordant pos.</i>		1
<i>concordant pos.</i>		1
<i>pos. below 95% LOD</i>	2	1
<i>pos. above 95% LOD</i>	0	0
ybbW		
<i>false positive rate</i>	0 of 48,374	0 of 48,374
<i>problematic wells</i>		1
<i>discordant pos.</i>		0
<i>concordant pos.</i>		21
<i>pos. below 95% LOD</i>	1	1
<i>pos. above 95% LOD</i>	21	20
ipaH		
<i>false positive rate</i>	0 of 51,710	5 of 51,710
<i>problematic wells</i>		0
<i>discordant pos.</i>		2
<i>concordant pos.</i>		4
<i>pos. below 95% LOD</i>	2	4
<i>pos. above 95% LOD</i>	2	2

TABLE 7: Continued

hexon		
<i>false positive rate</i>	0 of 51,710	5 of 51,710
<i>problematic wells</i>		1
<i>discordant pos.</i>		7
<i>concordant pos.</i>		0
<i>pos. below 95% LOD</i>	0	4
<i>pos. above 95% LOD</i>	0	3
STh		
<i>false positive rate</i>	0 of 52,408	0 of 52,408
<i>problematic wells</i>		3
<i>discordant pos.</i>		6
<i>concordant pos.</i>		7
<i>pos. below 95% LOD</i>	7	3
<i>pos. above 95% LOD</i>	1	3
Crypto. 18S rRNA		
<i>false positive rate</i>	0 of 52,408	0 of 52,408
<i>problematic wells</i>		2
<i>discordant pos.</i>		4
<i>concordant pos.</i>		5
<i>pos. below 95% LOD</i>	6	1
<i>pos. above 95% LOD</i>	1	6

During our analysis we observed 10 problematic wells out of 288 wells across all assays. In each instance, the large discrepancy between the Umbrella threshold quantification and the manual threshold quantification was attributable to significant differences between the fluorescence amplitude distribution of the negative cluster as observed in the sample compared to the negative controls. These differences were either greater variation of amplitudes in the negative cluster or a bimodal distribution in the negative cluster leading to a significant number of droplets in the negative cluster being misclassified as positive when compared to an NTC. The Umbrella results for each of these 10 wells was eliminated from further analysis. An additional 14 wells were eliminated from further analysis due to droplet counts below 10,000.

In our analysis of the remaining wells, we observed a total of 35 discordant detections in 288 wells. Of these Umbrella results from 7 were eliminated from further analysis due to misalignment between negative clusters confounding droplet classification as described for problematic wells. The difference in positive droplet quantities was only one in 22 of 35 wells and was greater than one (maximum of 10) in 6 wells. Excluding the wells eliminated, the discrepancy is driven by differences in the threshold between negative and positive droplets as determined by the user and as estimated by Umbrella. The small discrepancies in positive droplet numbers in discordant detections indicate that in general the manual thresholds were in good agreement with the statistically derived Umbrella threshold.

The total number of concordant wells between manual thresholding and Umbrella thresholding was 229 out of 288. Of these wells, 47 were concordant detections and 182 were concordant non-detects. This agreement rate of almost 80% again indicates that user-drawn manual thresholds and Umbrella thresholds were in good agreement overall. If we consider only detections, 63% (47/75) agreed between both thresholding approaches.

Jaipur Drinking Water Quality as Assessed by ddPCR

Based on our systematic review of the ddPCR data, we report our findings concerning the microbial water quality in Jaipur, as observed via molecular evidence, using a strength-of-evidence paradigm as shown in Figure C9. We consider the strongest possible evidence to be detection by both manual and Umbrella thresholding at a concentration above the 95% LOD for an assay with zero false positives in negative controls by either thresholding method. To fulfill this criterion the number of droplets classified as positive by both manual thresholding and Umbrella thresholding (<5% probability of being void of target) must be greater than the number of positive droplets required for successful detection in 95% of

replicates with a simultaneous observance of zero false positives in controls. Due to our observation of false positives in negative controls by the Umbrella thresholding method for the MS2g1, ipaH, and hexon assays, results for these targets are excluded from the strongest evidence category.

However, within the strongest evidence category, we observed one groundwater (GW) sample to be positive for beta giardin at a concentration of 9.8 to 228.5 gene copies (gc) per reaction. Given the concentration factor of the workflow and the 122.4 L DEUF volume, assuming 100% recovery efficiency during DEUF the drinking water column equivalent concentrations are 48 to 1,119 gc per liter. We also observed culturable *E. coli* counts that were TNTC in both the grab sample and the DEUF backflush from this location.

Our analysis also indicates that 19 of 22 DEUF samples across all three source water types are positive for the *E. coli* associated gene ybbW at the strongest level of evidence. This is a much higher proportion of samples positive than observed with culture-based evidence where 8 samples were positive for *E. coli* by culture and 12 samples were positive for total coliforms by culture. This finding likely reflects the presence of *E. coli* in a viable but not culturable (VBNC) state within the distribution network or the capture of inactivated but intact *E. coli* cells during sampling. With the development of the ybbW assay, Walker *et al.* fit a linear regression to the log of ybbW per PCR reaction versus the log₁₀ of CFU per reaction for samples with more than 10 CFU per 100 mL as observed in environmental water samples ($R^2 = 0.673$) (141). For our samples from drinking water in Jaipur we also attempted to fit such a model for *E. coli* by culture and ddPCR (n=5; CFU per 100 mL range 0.88 to 153.63), but the model fit was poor ($R^2 = 0.244$), as shown in Figure C10.

At the strongest level of evidence, we also found one groundwater sample positive for the gene encoding human heat-stable toxin (STh) associated with enterotoxigenic *E. coli* (ETEC) and one groundwater sample positive for the 18S rRNA gene of *Cryptosporidium*. Our findings for each molecular target by source water type and strength of evidence category are summarized in Table 8.

TABLE 8: Proportion of samples positive and concentration ranges for gene targets in groundwater (GW), surface water (SW) and mixed source (M) samples collected from the Jaipur municipal water supply and assayed by ddPCR. Reported results are stratified by strength of the ddPCR evidence.

		Number Positive			Concentration Range (gc per rxn)		
		GW (n=9)	SW (n=8)	M (n=4)	GW	SW	M
MS2g1 (MS2)	Strongest	-	-	-	-	-	-
	4	-	-	-	-	-	-
	3	-	-	-	-	-	-
	2	-	-	-	-	-	-
	Weakest	2	4	0	1.4 - 2.8	1.4 - 3.1	-
beta giardin (<i>Giardia</i>)	Strongest	1	0	0	9.8 - 228.5	-	-
	4	1	4	0	14.8	1.8 - 9.0	-
	3	0	0	0	-	-	-
	2	0	0	0	-	-	-
	Weakest	0	0	0	-	-	-
ORF1-2 (noro GI)	Strongest	0	0	0	-	-	-
	4	0	0	0	-	-	-
	3	0	1	0	-	1.2 - 2	-
	2	0	0	0	-	-	-
	Weakest	0	0	0	-	-	-
ORF1-2 (noro GI)	Strongest	0	0	0	-	-	-
	4	0	0	0	-	-	-
	3	1	0	0	1.2 - 1.3	-	-
	2	0	0	0	-	-	-
	Weakest	0	0	0	-	-	-

TABLE 8: Continued

ybbW (<i>E. coli</i>)	Strongest	8	7	4	4.9 - 7432.8	12.3 - 620	20.5 - 177.8
	4	0	0	0	-	-	-
	3	0	1	0	-	1.9 - 2	-
	2	0	0	0	-	-	-
	Weakest	0	0	0	-	-	-
ipah (<i>Shigella</i> /EIEC)	Strongest	-	-	-	-	-	-
	4	-	-	-	-	-	-
	3	-	-	-	-	-	-
	2	2	0	0	24 - 164.8	-	-
	Weakest	0	2	0	-	1.4 - 3.5	-
hexon (adeno A-F)	Strongest	-	-	-	-	-	-
	4	-	-	-	-	-	-
	3	-	-	-	-	-	-
	2	0	0	0	-	-	-
	Weakest	0	0	0	-	-	-
STh (ETEC)	Strongest	1	0	0	4.6 - 824.7	-	-
	4	1	1	1	613.3	146.5	337.9
	3	2	1	0	1.4 - 1.6	1.4	-
	2	0	0	0	-	-	-
	Weakest	0	0	0	-	-	-
18S rRNA (<i>Cryptosporidium</i>)	Strongest	1	0	0	4.2 - 5.7	-	-
	4	1	3	0	38.4	14.9 - 108.0	-
	3	0	1	0	-	1.4	-
	2	0	0	0	-	-	-
	Weakest	0	0	0	-	-	-

Implications for Drinking Water Safety in Jaipur

Although our study was only a small cross-sectional sample of drinking water from Jaipur, our findings highlight some important observations regarding drinking water safety in the city. We observed culturable *E. coli* in 22% of grab samples and 70% of DEUF backflush samples originating from groundwater being pumped directly into the water distribution network. Simultaneously, we did not detect *E. coli* by culture in grab samples or DEUF

backflush samples from distributed surface water. Free chlorine was detected in 7 of 8 distributed surface water samples and 3 of 4 mixed source water samples, but whether the contact time and concentration resulting from the pumping of groundwater directly into the network is sufficient for inactivation of pathogens is unclear. Further, we observed one groundwater sample to be positive for beta giardin gene, one to be positive for an 18S rRNA gene, and one to be positive for STh gene associated with *Giardia duodenalis*, *Cryptosporidium* spp., and ST-ETEC respectively at concentrations above the 95% LOD. For each of these samples the associated DEUF backflush was positive for culturable *E. coli* and total coliforms, with the sample positive for beta giardin also being TNTC for *E. coli* and total coliforms in both grab samples and DEUF samples.

These detections are consistent with molecular evidence from several other studies in India. In Lucknow, ETEC genes ST1 and LT1 were detected in potable water from the municipal supply and in Kolkata 9% of *E. coli* isolated from potable water sources were positive for virulence genes including heat-stable toxins associated with ETEC (153,154). In Chennai, 58% of samples from the piped water supply were reportedly positive for *Cryptosporidium* oocysts by immunofluorescent antibody screening (155). The assumption that groundwater is less likely to be contaminated is subject to increasing scrutiny as studies have documented the presence of viruses in groundwater both municipal water supplies and private wells in the US, and protozoan pathogens in tube wells in rural India (47,156,157). In the US acute gastrointestinal illness (AGI) has been associated with both septic tanks and with nondisinfected groundwater drinking supplies (43,158). It is therefore plausible that groundwater in the urban environment, particularly in environments with poor or failing sewer infrastructure could be a significant source of waterborne disease. Sewers are

known to be sources of groundwater recharge in urban areas and to degrade the quality of urban groundwater (159,160). In India, Somasundaram and colleagues noted microbial contamination in a urban aquifer in Madras as indicated by high nitrate levels and the detection of microbes in wells (161). A study of the groundwater recharge in Hyderabad found that the anthropogenic sources of groundwater recharge were ten times greater than the natural ones and that leakage from the water distribution system and leakage of sewage from the sewer system comprised a large proportion of the annual recharge volume (162). In rural Rajasthan, bacterial contamination of groundwater used for drinking has been observed (163). A recent study of inorganic contaminants in groundwater in Rajasthan found that elevated levels of dissolved organic carbon and nitrate suggested anthropogenic sources of groundwater pollution such as domestic sewage (164).

Our findings in Jaipur are consistent with these observations and indicate that groundwater in Jaipur should not be neglected as a source of waterborne disease. Currently, the municipal government reports that groundwater accounts for 26% of daily water supply; however, this may be an underestimate due to informal networks of wells that are operated by individuals and not the government of Rajasthan. To manage the risk of waterborne disease associated with this groundwater, it should be treated prior to distribution. Typically disinfection with chlorine might be considered the best option, but the high concentrations of halides and organic contaminants observed in the groundwater increase the likelihood of disinfection byproduct formation (164). This risk is corroborated by a recent study of tap water in Jaipur that found both regulated disinfection byproducts, such as trihalomethanes, and unregulated byproducts, such as haloacetonitriles, that could pose significant risks to public health (165). Additionally, disinfection using chlorine would be

ineffective against protozoan pathogens such as *Giardia* and *Cryptosporidium* which we detected via molecular assays. Given these constraints, a filtration-based treatment may be the best solution to reduce the risk of waterborne disease associated with groundwater in Jaipur. In the near term correct and consistent use of non-chlorine household water treatment technologies can provide a reasonable reduction of the risks of waterborne disease (94,166). A compelling alternative to treating the groundwater would be preventing its contamination with sewage in the first place. Repairs and improvements to reduce leakage from existing sewer lines and connecting sewage sources that are currently unserved could reap the dual benefit of reducing the microbes that cause waterborne disease and reducing the dissolved organic carbon and nitrates that are precursors for disinfection byproducts. A quantitative analysis of the tradeoffs between disinfection byproducts and waterborne disease specific to Jaipur would be useful for decision making.

Improvements in Interpreting ddPCR Results for DEUF Concentrate

Our results also emphasize critical points concerning the interpretation of ddPCR results to enumerate waterborne pathogens. Manual thresholding by the user is subject to confirmation bias and is detrimental to reproducibility. Many tools have been proposed to improve the rigor and reproducibility of droplet classification including k-means clustering, extreme value theory thresholding, and model-based clustering as we have used here (149–151). However, in our application of Umbrella we had to rely on user judgement downstream of the model to screen both problematic and discordant positive wells (20.4% of total wells) for further analysis. This screening is subject to the same bias as manual thresholding. Since many of the discordant and problematic wells were attributable to variations in the fluorescence amplitude distribution of the negative clusters, improvements

in the consistency of these distributions across samples and controls would greatly improve Umbrella thresholding. Our experience suggests two avenues of improvement. First, selecting sample-appropriate comparators for negative controls on each plate is critical. While NTCs are appropriate for assessing cross contamination, they may not be the best comparison for negative clusters in real samples. Instead multiple sample blanks consisting of all the background material expected in a real sample but negative for the target should be used for each assay on each plate. Second, optimizing extraction methods for ddPCR to produce consistent negative droplet clusters should be prioritized. In our case, we selected the UNEX buffer and protocol based on comparisons between CT values from qPCR assays. But it is possible that another extraction method may yield cleaner samples with less noise in the negative droplet cluster. Additionally, while various statistical approaches to interpreting ddPCR results improve the reproducibility of the process, they don't answer the fundamental question of whether droplets are truly positive or negative for the target.

Difficulty of Making Risk-Relevant Measurements of Waterborne Pathogens

Finally, our work highlights an important limitation of leveraging DEUF and ddPCR to make risk-relevant measurements of waterborne pathogens. We observed a 95% LOD of 3.9 gc per reaction for the *Cryptosporidium* 18S rRNA gene in ddPCR interrogation of DEUF concentrate. Based on this LOD, the concentration factors for our workflow, the largest volume of drinking water filtered in our study (161 liters), and the mean recovery of oocysts observed by Smith and Hill (87%), we calculate that the lowest drinking water equivalent LOD we achieved is 16.7 gc per liter of drinking water (111). Using a figure of 20 copies of 18S rRNA per oocyst we arrive at a concentration of 0.84 oocysts per liter. As with any molecular method we must make assumptions regarding the viability and

infectivity of these oocysts. But, if as an upper-bound, we assume these oocysts are 100% viable and infective and use the dose-response model and assumptions described previously we arrive at an LOD for the daily risk of infection of 25% and an equivalent annual risk of 100%. Subject to the same assumptions, if we assume filtration of 1,500 liters, as in EPA Method 1615, we arrive at an LOD of 0.045 oocysts per liter and the daily risk of infection of 1.5% and an equivalent annual risk of 99.6% (167). Even with DEUF and ddPCR, detecting *Cryptosporidium* oocysts at a concentration equivalent to an annual risk of infection of 1% would require filtering 209,588 liters of drinking water. This seems improbable without dramatic changes to the described workflow to increase the concentration factors, such as improving volumetric ratios of input material to output material in both DEUF backflush and secondary concentration. However, such improvements are also likely to further concentrate inhibitors present in the drinking water, so the net effect could be detrimental to downstream molecular assays. Additionally, translation from gene copies to the units necessary for input to dose-response models still requires tenuous assumptions relating genetic elements and infectious units. These results suggest that when annualized risk is the metric of performance, quantitative risk assessment for distributed drinking water will continue to rely on assumptions regarding exposure assessment often mediated by indicator organisms and pathogen to indicator ratios rather than direct measures of waterborne pathogens.

CHAPTER 4: MICROBIAL WATER QUALITY IN INTERMITTENT VERSUS CONTINUOUS WATER SUPPLY IN NAGPUR, INDIA

ABSTRACT

No city in India yet provides continuous water supply (CWS) to its inhabitants. As a result, some 460 million Indians are exposed to intermittent water supply (IWS), which has been associated with degraded water quality and increased risk of waterborne diseases such as typhoid and cholera. We leveraged the ongoing transition of the Nagpur municipal water supply from intermittent to continuous supply to conduct a natural experiment comparing microbial water quality in intermittent versus continuous supply zones. During sampling periods in 2015 and 2017, we collected 56 grab samples and 90 large volume dead-end ultrafiltration samples (total sampling volume of 6,925 liters). In addition to measuring traditional water quality parameters such as free and total chlorine, turbidity, heterotrophic plate count, thermotolerant coliforms, and *E. coli*, we also assayed DEUF concentrates for genes associated with seven waterborne pathogens by ddPCR. We detected genes associated with ETEC, *Shigella* spp., norovirus GI and GII, adenovirus, *Cryptosporidium* spp., and *Giardia duodenalis* in samples collected from household taps served by intermittent supply. Fecal indicator bacteria results suggest increased water quality degradation from elevated storage reservoir (ESR) to household tap in intermittent versus continuous water supply zones. But many of these observations are not statistically significant after accounting for random variation. We did observe statistically significant differences between the proportion of grab samples positive for *E. coli* (IWS: 43.8%; CWS: 3.57%) and DEUF concentrates positive for gene targets associated with waterborne

pathogens (IWS 61.1%; CWS: 11.1%) in samples collected from household taps served by intermittent versus continuous supply. Although our DEUF-based comparisons suffer from the limitations associated with a small sample size, the evidence should be considered in appreciation of the large volume of water sampled. For example, we did not detect any culturable *E. coli* in nine DEUF drinking water samples totaling 1,006 liters from household taps served by CWS. Despite the limitations of our study, this work contributes to an increasingly large body of literature that suggests IWS is associated with increased risk of water quality degradation and waterborne disease transmission in underserved settings.

INTRODUCTION

The intermittent delivery of drinking water in piped supplies is the result of a complex amalgamation of political, institutional, ecological, and infrastructural constraints whose end result is the inability or unwillingness to maintain positive pressures within a water distribution network 24 hours-per-day seven days-per-week 365 days-per-year (15,168). Lapses in positive pressure expose an IWS to numerous opportunities for contamination and degradation of drinking water quality including: intrusion of contaminants during low pressure periods, regrowth during stagnation, scouring during re-pressurization, and recontamination during household storage (23). Numerous studies have documented decreased residual chlorine and increased counts of various fecal indicators including total coliforms, thermotolerant coliforms, and *E. coli* in supplies operated intermittently in various contexts throughout the globe including rural settings, urban settings, and refugee

camp (26,27,74,75,169–174). However, such contamination does not appear to be ubiquitous as a study of several supply zones within an IWS in Panama documented only 0.9% of grab samples positive for total coliforms and one sample positive for *E. coli* (n=423). Even in a large study in India where 31.7% of grab samples from an IWS were positive for *E. coli*, the proportion of samples positive varied by hydraulic zone within the network (24). The spatial and temporal variation in water quality is likely due to the complex hydraulic conditions which preside within an IWS and the varying distribution of sources of fecal contamination within and between settings (175).

The heterogeneity observed in water quality data is also found in the association between IWS and diarrheal disease observed during epidemiological studies. Several studies have found significant associations between IWS and diarrhea among adults, children, and residents of refugee camps and slums (176–181). Another study documented a significant increase in the incidence of cholera cases following water supply interruption and another noted an increase in markers associated with environmental enteropathy (EE) for children exposed to IWS (28,182). However, a large longitudinal trial associated with a transition from IWS to continuous water supply (CWS) in Hubli-Dharwad, India found no association between water delivery mode and diarrhea (29).

IWS is especially relevant in India where an estimated 63% of the population served by a piped supply has water available for only 3 hours per day or less (183,184). Despite the success of several pilot programs of CWS, no major city in India yet offers its citizens access to CWS (185–187). Surveys of various India cities have noted an average duration of 4 hours per day and a range of 0.5 to 10 hours per day compared to an Asia Pacific average of 19 hours per day (184,188).

The city of Nagpur is now poised to become the first major city in India to transition its entire water supply from IWS to CWS. The 2.5 million people living in the city exert a daily water demand of 660 million liters which is withdrawn from three surface water sources (Kanhana River, Gorewada Lake, and Pench Dam) and treated at five water treatment plants (WTPs) prior to distribution via a network of master balancing reservoirs (MBR), elevated service reservoirs (ESR), and a 2,100 km pipeline network (89,189). In 2007 the Nagpur Municipal Corporation (NMC) reported an average supply time of 4.3 hours per day, unaccounted for water (UFW) of 51.9%, with 40% of household connections metered and 3.2 staff persons per 1,000 connections (188). From 2009 to 2011, a CWS demonstration completed in one area of the water supply showed remarkable improvements in service level benchmarks through a 270 million INR (USD \$4 million) improvement package that included 100% metering, and replacement of 30% of the pipes in the demonstration zone (190). Following the successful demonstration, NMC contracted with Orange City Water (OCW) to transition the entire water supply from IWS to CWS beginning in 2011 at an estimated cost of 3.87 billion INR (USD \$70.5 million) with contractual obligations to reduce NRW and improve tariff collection efficiency (186). In association with the transition, NMC and the National Environmental Engineering Research Institute (NEERI) also implemented a water safety plan (WSP) for the Nagpur municipal water supply following the WHO WSP Manual (191). In the period from 2011 to 2016, 496 km of pipeline was replaced and CWS was implemented in 10 of 64 hydraulically isolated areas -- called "command areas" by OCW (192).

The transition of the Nagpur water supply from IWS to CWS afforded us the opportunity to observe the water quality under the two modes of delivery with attention to measures

more directly relevant to waterborne pathogens than traditional fecal indicators. Two previous studies of drinking water quality in Nagpur have noted considerable contamination in the IWS water delivery system from treatment to the point-of-use (88,89). A study conducted during the CWS pilot in Hubli-Dharwad found a significant difference in drinking water quality between the IWS and CWS supply zones (24). However, each of these studies utilized measures of fecal indicator bacteria, which, while informative, are only proxies for fecal contamination and do not provide robust information on the presence of waterborne pathogens. A recent risk assessment found that the use of indicator bacteria and ratios of pathogens to indicators were two significant sources of uncertainty in quantifying the burden of diarrheal disease associated with IWS (97). While another review identified uncertainty about the risk posed by protozoan and viral pathogens introduced via intrusion as an important opportunity to increase our understanding of IWS and health (16). We leveraged the transition of the Nagpur water supply from IWS to CWS to conduct a natural experiment to document water quality differences between IWS and CWS water supply zones with special attention to molecular measures of genes associated with waterborne pathogens.

MATERIALS AND METHODS

Here we report on our use of traditional physicochemical and indicator-based water quality measures along with dead-end ultrafiltration (DEUF) and droplet digital PCR (ddPCR) to assess the water quality of drinking water delivered via intermittent versus continuous water supply in Nagpur, India. The transition from IWS in Nagpur officially begin in 2011 and called for the entire supply to be converted to CWS by 2016. However, project delays

associated with line rehabilitation and meter installation were such that, as of 2017, only 13 of 54 command areas had been converted to CWS. Our study design evolved as the project itself evolved.

Study Design

We originally conceived of the study as a controlled-before-and-after design with cross sectional measurements of water quality in one control IWS command area at baseline and endline and concurrent measurement of water quality in one intervention command area that was IWS at baseline and CWS at endline. We initiated our work in 2015 with cross sectional water quality sampling in two IWS command areas within the Nagpur water supply: Reshim Bagh and Karve Nagar. During each sampling event, we collected samples from various taps in the command area and the associated ESR such that we could isolate water quality changes occurring downstream of the ESR. However, when we returned in 2017 for the endline sampling we found that delays in the conversion project meant that neither of the command areas had yet converted to CWS. Therefore, we adapted our design to a cross-sectional study of water quality along the delivery system with the addition of data from two more command areas: Laxmi Nagar Old, which is CWS, and Khamla, which is IWS. As shown in Figure 9, both the Khamla and Laxmi Nagar Old command areas receive treated drinking water from the PENCH II WTP. From the WTP, finished drinking water is sent via transmission main to the ESRs that serve Laxmi Nagar Old and Khamla command areas. From these ESRs drinking water is distributed continually in Laxmi Nagar Old and intermittently in Khamla. As a result, our final study includes cross-sectional water quality data from ESRs and taps in three IWS zones (Khamla, Karve Nagar, and Reshim Bagh), water quality data from the ESR and taps from one CWS zone (Laxmi Nagar Old),

and water quality data through the treatment train at the Pench II WTP. Together this data allows us to examine how the water quality changes through treatment and distribution under both IWS and CWS modes. Each sample we collected consisted of a grab sample and a paired large-volume DEUF sample of approximately 20 liters in 2015 or 100 or more liters in 2017. We also measured physicochemical parameters at the time of sample collection.

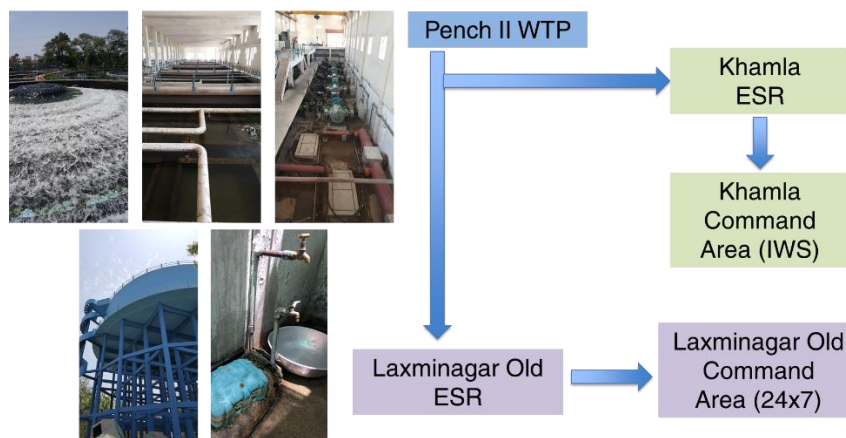


FIGURE 9: A schematic of the study design and water quality sampling as executed during the 2017 sampling period. We collected samples along the drinking water delivery system for an IWS command area (Khamla - green) and a CWS command area (Laxmi Nagar Old - purple). During each sampling event in the command areas we collected drinking water from the ESR (bottom left photo) and taps in the command area (lower right photo).

Physicochemical Measurements

At the time of sample collection, we measured physicochemical parameters including turbidity, free chlorine, total chlorine, pH, conductivity, and total dissolved solids (TDS). We did not measure turbidity during sample collection in 2015. But for samples collected in 2017, we measured turbidity by nephelometric turbidity unit (NTU) using a Hach 2100Q portable turbidimeter (Hach, Loveland, CO) (193). We measured total and free chlorine using the US EPA DPD Colorimetric Method implemented via a Hach Pocket Colorimeter

II spectrophotometer (Hach, Loveland, CO) (128). We measured temperature, pH, conductivity, and TDS by electrode using a Hanna Low Range Combo Tester (Hanna Instruments, Woonsocket, RI) (129). For a subset of samples collected from household taps in 2017, we measured water service pressure using a Dickson PR125 pressure logger (Dickson/Unigage Inc., Addison, IL).

Culture Methods

During each sampling event, we collected both grab samples and large-volume samples for DEUF. In 2015 sampling, we collected 300 mL grab samples in Whirl Pak bags pre-dosed with sodium thiosulfate to quench residual chlorine (Nasco, Fort Atkinson, WI); whereas in 2017, we collected 500 mL grab samples in sterile plastic bottles pre-dose with sodium thiosulfate. In the time between collection and further processing, grab samples were stored and transported on ice. In 2015, we enumerated *E. coli* in two 100 mL, 10 mL, and 1 mL replicates by membrane filtration and selective media incubation per EPA Method 1604 (with modification) (104). We incubated each replicate along with both positive and negative controls at 35°C for 24 hours on HiCrome M-TEC Agar (M1571; HiMedia, Mumbai, India). In 2017, we enumerated thermotolerant coliforms and *E. coli* in two 100 mL replicates from each grab sample by membrane filtration followed by incubation at 44.5°C for 24 hours on HiCrome Coliform HiVeg Agar with SLS (MV1300; HiMedia, Mumbai, India). Following incubation, we counted both thermotolerant coliforms and *E. coli* on each plate following the manufacturer's instructions for color-based identification, including the use of Kovac's Indole (R008; HiMedia, Mumbai, India) for confirmation of presumptive *E. coli* colonies on HiVeg agar. We reported the counts in colony forming units (CFU) per 100 mL. In a subset of grab samples collected in 2017, we also enumerated

heterotrophs by membrane filtration of 1 mL, 500 uL, and 250 uL volumes in duplicate followed by incubation at 35°C for 24 hours on R2A Agar (M1205; HiMedia, Mumbai, India) per the manufacturer's instructions.

Dead-End Ultrafiltration

We concentrated the microbes from large-volume drinking water samples by DEUF following a CDC protocol. We used Rexeed 25S dialyzers (Asahi Kasei Kuraray Medical Company, Tokyo, Japan) (111). In 2015, we collected 20 liter samples in sterile cubitainers pre-dosed with sodium thiosulfate and performed DEUF upon return to the lab. Whereas in 2017, where possible, we connected to sampling points via sterile Masterflex L/S 36 Platinum-Cured Silicone tubing (Cole Parmer, Vernon Hills, IL) and utilized service pressure in the piped network to force the drinking water through the dialyzer. When system pressure was inadequate or unavailable, such as at the Pench II WTP, we collected drinking water in sterilized plastic buckets and filtered from the buckets using a Geotech Geopump (Geotech, Denver, CO). In all instances, we recorded the volume of water filtered using a totalizing flow meter (Clark Solutions, Hudson, MA). For each sample, we quenched residual chlorine in the dialyzer by filtering 500 mL of 1% sodium thiosulfate solution immediately after filtering the sample. After DEUF, we backflushed each dialyzer into a sterile 500 mL bottle using a 500 mL 0.5% Tween 80, 0.01% NaPP, and, 0.001% Antifoam Y-30 emulsion solution with recovery volumes ranging from 350 to 450 mL (111). In 2017, we enumerated thermotolerant coliforms and *E. coli* in the backflush solution using 1 mL and 100 uL duplicates subjected to membrane filtration and culture as previously described. In addition to fecal indicators, we cultured 1 mL, 100 uL, and 10 uL duplicates of DEUF backflush for *Salmonella* and *Shigella* and pathogenic *E. coli* using

membrane filtration combined with incubation at 35°C for 24 hours on SS Agar (M108; HiMedia, Mumbai, India) and MUG Sorbitol Agar (M1205; HiMedia, Mumbai, India), respectively. Following incubation, we identified presumptive colonies of *Salmonella*, *Shigella*, and pathogenic *E. coli* per the manufacturer's instructions for each agar and subjected isolates to confirmatory biochemical testing via the HiSalmonella Identification Kit (KB011; HiMedia, Mumbai, India) or HiLMViC Biochemical Test Kit (KB001; HiMedia, Mumbai, India) as appropriate.

PEG Precipitation and Ultracentrifugation

After culture-based enumeration of both 2015 and 2017 samples, we further concentrated the DEUF backflush using polyethylene glycol (PEG) precipitation and ultracentrifugation (106,112). We performed PEG precipitation using a 12% PEG 8000, 0.9 M NaCl, and 1% bovine serum albumin solution. We mixed each reagent sequentially into the backflush volume slowly to ensure complete dissolution and incubated samples overnight at 4°C. Following overnight incubation, we aliquoted 300 mL of the PEG-precipitated solution into six 50 mL centrifuge tubes and spun them at 10,000 times gravity at 4°C for 30 minutes. We slowly poured off the resulting supernatant and resuspended the pelleted material using 677 uL of 1X phosphate-buffered saline (PBS), 0.01% Tween 80, and 0.001% Y-30 emulsion solution per centrifuge tube. We then recombined the resuspended pellets from each of the six centrifuge tubes. The volume of the resulting suspension, which we term the DEUF concentrate, ranged from 3 to 4 mL. We stored DEUF concentrate at -20°C until we pre-treated each sample with UNEX lysis buffer and bead beating as described in the next section. Immediately following pre-treatment, we transported the

samples at room temperature for approximately 36 hours to the laboratory in Atlanta, GA where they were frozen at -80°C until extracted for further molecular analysis.

Nucleic Acid Extraction

We performed nucleic acid extraction using UNEX buffer (Microbiologics, St. Cloud, MN), as developed by the CDC (130). We extracted the DEUF concentrate following the protocol for nucleic acid extraction from parasites in water samples as detailed in the manufacturer's instructions for use. We added 500 uL of DEUF concentrate and 500 uL of UNEX buffer, and 5 uL of Inforce 3 Bovine Vaccine (Zoetis, Parsippany, NJ), our process control containing bovine respiratory syncytial virus (BRSV) and bovine herpes virus (BoHV), to an SK-38 bead tube (Bertin Corp, Rockville, MD) and completed the lysis, purification, and elution of nucleic acids per the manufacturer's instructions. With each extraction batch, we included a negative extraction control and an Inforce 3 only control (500 uL of molecular water and 5 uL of Inforce 3). We stored the purified nucleic acid from each sample and the controls at -80°C until molecular analysis or reverse transcription.

Reverse Transcription

Prior to ddPCR assay for RNA targets, we performed reverse transcription of RNA to cDNA using a High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (ThermoFisher Scientific, Waltham, MA). We performed reverse transcription (RT) per the manufacturer's instructions by adding 10 uL of nucleic acid extract to 10 uL of RT master mix. Following RT, we stored the resulting cDNA at -80°C until ddPCR interrogation was completed within 5 days.

ddPCR

We selected microbes of interest for molecular analysis on the basis of diarrheal disease etiologies from several sites in southeast Asia as observed in the Global Enteric Multicenter Study (GEMS) and the Malnutrition and Enteric Disease Study (MAL-ED) as summarized in Figures C1, C2, and C3 (67,131) and previously described. We adapted existing qPCR assays to ddPCR and used them to interrogate nucleic acids from each sample for genetic targets associated with a set of waterborne pathogens and fecal indicators *E. coli* and MS2. The relevant genetic, thermal cycling, and performance information for all the assays we used, including for controls, are summarized in Tables 6, and C1 to C3.

We performed all ddPCR workflow following the manufacturer's recommended protocol using a QX200 Droplet Generator, PX1 PCR Plate Sealer, C1000 Touch Thermal Cycler, and QX200 Droplet Reader (Bio-Rad, Hercules, CA) (146). For singleplex probe-based assays (MS2g1; beta giardin), we added each primer and the probe to achieve final concentrations of 900 nM and 250 nM, respectively. While for multiplex probe-based assays (norovirus GI/norovirus GII; hexon/ipaH; Cryptosporidium 18S rRNA/STh), we added primers and probes to achieve final concentrations of 650 nM and 250 nM, respectively. For the ybbW EvaGreen assay, we added each primer to achieve final concentrations of 250 nM. For each assay, the final reaction volume was 20 μ L including molecular-grade water, ddPCR Supermix for Probes or ddPCR EvaGreen Supermix (Bio-Rad, Hercules, CA), primers, probe, and 4 μ L of extract or RT product from each sample. We included negative extraction controls, sample blanks, no-template controls, and an assay appropriate positive control on each ddPCR plate. We also included randomly selected technical replicates and biological replicates at rates of at least 10% each on each ddPCR plate.

In ddPCR, the concentration of a gene target in the reaction mix is estimated by means of a most probable number according to a Poisson distribution using the proportion of droplets classified as negative based on their observed fluorescence amplitude following PCR thermal cycling (146). Thus, determining the fluorescence threshold for classifying droplets as either positive or negative is critical for accurate quantification of the target gene (149). We estimated the concentration of target genes in our ddPCR assays using both a manual user-determined threshold in QuantaSoft (V1.7.4; Bio-Rad, Hercules, CA) and a model-based classification called Umbrella implemented in RStudio (Version 1.1.456; RStudio Team, 2015) (151). The Umbrella package reports an estimate of the concentration based on the number of droplets with less than a 5% probability of being negative, given their fluorescence amplitude and observed fluorescence amplitude distribution in negative controls. We refer to this estimate as the Umbrella threshold. For full details see the description provided by Jacobs *et al.* in their manuscript and supporting information (151).

RESULTS AND DISCUSSION

During water quality sampling in Nagpur, we collected a total of 90 DEUF and paired grab samples, plus an additional 56 grab samples. In 2015, we collected samples from June 23 to July 22 including 17 paired DEUF and grab samples from Karve Nagar command area (IWS) and 16 from the Reshim Bagh command area both of which are served by IWS. As previously mentioned, we performed DEUF in the lab during the 2015 sampling period with an average DEUF volume of 18.16 liters (max: 22.18, min: 14.38). In 2017, we collected samples from April 20 to October 10. We collected 21 paired DEUF and grab

samples from three points in the Pench 2 WTP, 19 paired samples from the Khamla command area (IWS), and 17 paired samples from the Laxmi Nagar Old command area (CWS). Because we performed DEUF in the field in 2017, the average volume was 110.97 liters (max: 121.47, min: 89.31). In addition to these paired samples, we collected 33 grab samples in Khamla and 23 grab samples in Laxmi Nagar Old. We selected sampling locations within each command area with the assistance of utility personnel and sought to distribute our samples throughout each area such that the data are spatially representative of the command area from which we collected them (Figures D1, D2, and D3).

Service Pressure

While collecting grab samples in the Khamla and Laxmi Nagard Old command areas, we measured the water pressure at household taps. For 30 observations in the Khamla area (IWS) the mean pressure during supply hours was 6.6 psi (4.65 m) (95% CI: 5.28 – 7.92 psi). Whereas for 21 observations in the Laxmi Nagar old area (CWS), the mean pressure was 4.5 psi (3.17 m) (95% CI: 3 – 6 psi). Boxplots of the observed service pressures are shown in Figure D4. While the service pressures were not different after accounting for random variation, the lower average service pressure in the CWS area is explained by the OCW's policy of lowering the service pressure in areas that are continuously pressurized in order to minimize water loss through leakage as explained to us by the command area supervisor.

Free and Total Chlorine

During our sampling, we detected free chlorine in 91.5% of samples from IWS areas (n=82) and 90.0% of samples in CWS areas (n=40). The average free chlorine concentration in IWS tap samples (n=65) was 0.20 mg/L (95% CI: 0.15 – 0.25) and in

CWS tap samples (n=28) was 0.22 mg/L (95% CI: 0.17 – 0.27) (Figure 10). Concurrently, we detected total chlorine in 98.8% of IWS samples and 100% of CWS samples. The Indian Standard for Drinking Water requires that all distributed water intended for human consumption have a free residual chlorine concentration of 0.2 mg/L (152). We observed that 59.8% of samples from the IWS area and 57.5% of samples from the CWS area had free chlorine residuals below this standard.

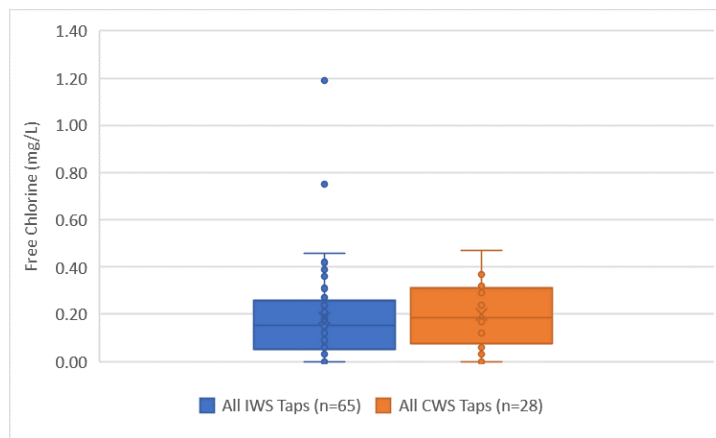


FIGURE 10: Boxplots of free chlorine concentration as observed in grab samples collected from taps served by IWS versus taps served by CWS in Nagpur.

Turbidity

We measured turbidity at the ESRs and household taps in both Khamla (IWS) and Laxmi Nagard Old (CWS) command areas. In the IWS area we observed a slight increase in the mean turbidity between the ESR and the household taps; while in the CWS area we observed a slight decrease (Figure D5). However, these differences are not statistically significant after accounting for random variation. The average turbidity at household taps in the IWS area was 1.56 NTU (95% CI: 1.22 – 1.90) and at household taps in the CWS area was 1.01 NTU (95% CI: 0.73 – 1.29) (Figure 11). While the difference in these means

is not statistically significant, both of these averages exceed the standard of less than 1 NTU for distributed drinking water establish by the India Standard specification (152).

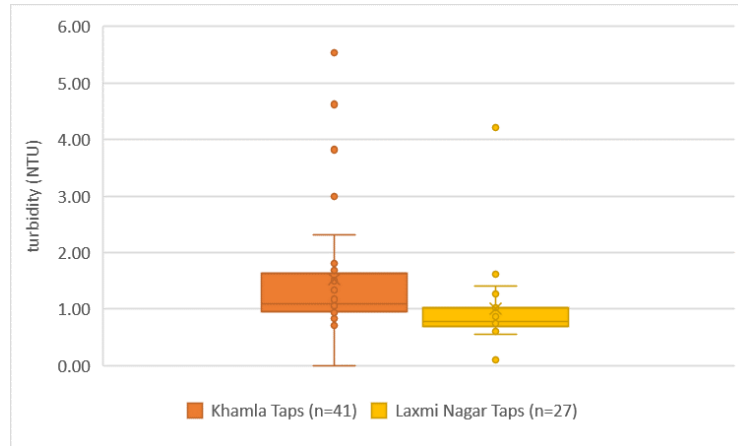


FIGURE 11: Turbidity measures as observed at household taps in IWS (Khamla) and CWS (Laxmi Nagar Old) command areas in Nagpur.

Fecal Indicators by Culture

During 2017 sampling, we enumerated heterotrophs in 23 grab samples from the IWS area and 13 from the CWS area. Each of these samples was positive for heterotrophic bacteria. The average heterotrophic plate count in the IWS command area was 120.7 per mL (95% CI: 82.9 – 158.5) and in the CWS command area was 121 per mL (95% CI: 65.7 – 176.3) (Figure D6). We discontinued our enumeration after 36 grab samples due to the labor intensity of preparing the dilutions required to achieve countable plates.

We enumerated thermotolerant coliforms (TTC) in both grab samples and DEUF backflush for samples collected in the Khamla (IWS) and Laxmi Nagar Old (CWS) areas. After enumerating TTCs in DEUF backflush, we back calculated the number of TTCs in the drinking water matrix using the equations shown in Table C4 assuming 100% recovery of coliforms during DEUF backflush. Of the 27 grab samples from household taps in the IWS

Khamla area, 48.2% (95% CI: 30.7 – 66.0) were positive for TTC with 5 samples too numerous to count (TNTC). For the 8 countable samples, the mean count was 24.8 CFU/100 mL. In the 24 grab samples from household taps in the CWS Laxmi Nagar Old area, only 3 were positive for TTCs, 12.5% (95% CI: 4.3 – 31.0) with a mean count of 0.6 CFU/100 mL. TTC counts as we observed in grab samples from IWS versus CWS taps are shown in Figure 12.

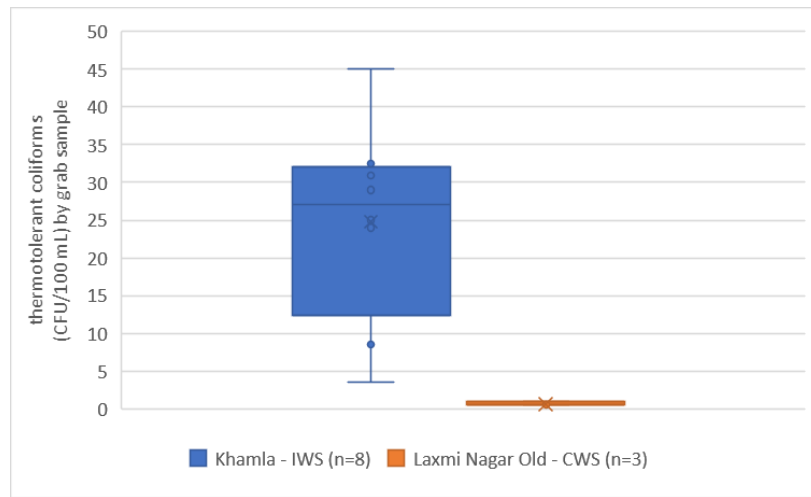


FIGURE 12: Counts of thermotolerant coliforms as observed in grab samples from household taps in an IWS versus a CWS service area in Nagpur.

In DEUF backflush from the IWS command area, 22% of samples were positive for TTC at the ESR (mean 20.1 CFU/100 mL) and 89% were positive at the household tap (mean 87.8 CFU/100 mL). While in DEUF backflush from the CWS command area, 56% of samples were positive for TTC at the ESR (mean 31.2 CFU/100 mL) and 44% were positive at household taps (mean = 10.9 CFU/100 mL). The proportion of DEUF samples positive for TTCs through the drinking water delivery chain is displayed in Figure D7. TTC

counts as enumerated via DEUF at ESRs and household taps in the IWS area (Khamla) and the CWS area (Laxmi Nagar Old) are shown in Figure 13. Although the means suggest an increase in TTC count between the ESR and household taps in IWS versus a decrease in CWS, as can be seen in Figure D8, these differences are not significant after accounting for the variation in the counts.

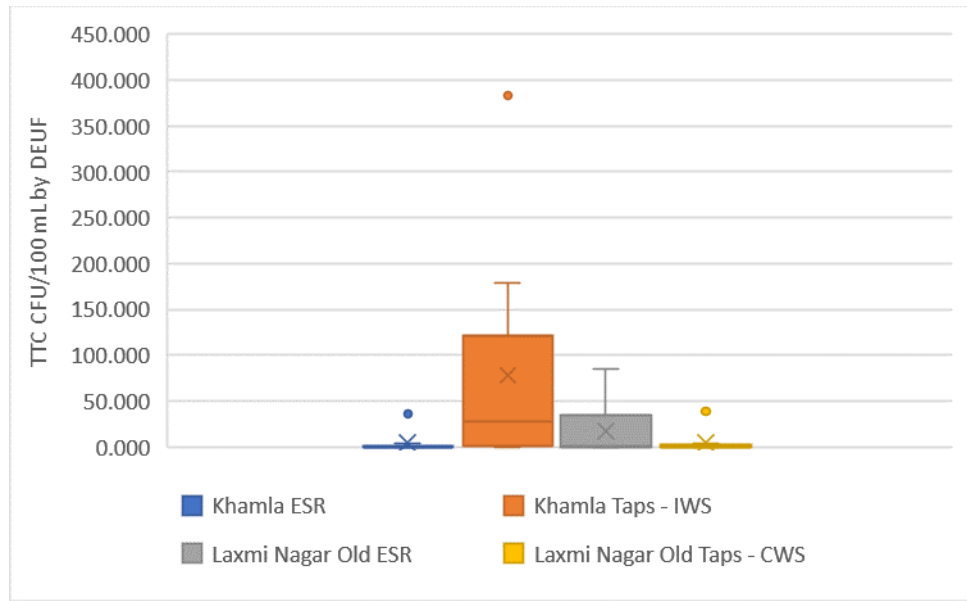


FIGURE 13: TTC counts as observed via DEUF at ESRs and household taps in IWS (Khamla) and CWS (Laxmi Nagar Old) service areas in Nagpur.

We enumerated *E. coli* in grab samples from household taps in IWS areas Karve Nagar (n=10) in 2015, Reshim Bagh (n=13) in 2015, and Khamla (n=41) in 2017, and CWS area Laxmi Nagar Old (n=28) in 2017. Across all grab samples from household taps served by IWS (n=64), 43.8% were positive for *E. coli* (95% CI: 32.3 – 55.9); whereas for CWS, (n=28) 3.6% were positive for *E. coli* (95% CI: 0.6 – 17.7). The proportion of grab samples positive for *E. coli* and *E. coli* counts stratified by command area are shown in Figures D9

and D10. Boxplots of *E. coli* counts as observed in grab samples across all IWS household taps and CWS household taps are displayed in Figure 14.

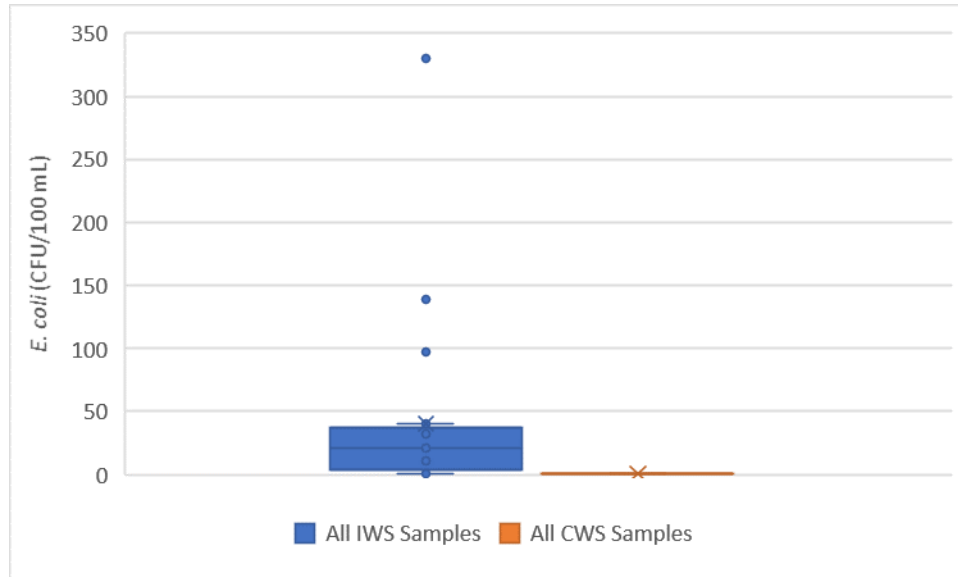


FIGURE 14: *E. coli* counts as observed in grab samples collected from household taps in IWS (n=28 countable) versus CWS service areas (n=1 countable) in Nagpur.

We enumerated *E. coli* in DEUF backflush from samples collected throughout the IWS and CWS drinking water supply chain during the 2017 sampling. In the IWS command area, 11% of DEUF backflush samples collected at the ESR (n=9) were positive for *E. coli* while 56% were positive at the household tap (n=9). In the CWS command area, 56% of DEUF backflush samples collected at the ESR (n=9) were positive and none were positive at the household tap (n=9) (Figure D11). *E. coli* counts as enumerated via DEUF at ESRs and household taps in the IWS area (Khamla) and the CWS area (Laxmi Nagar Old) are shown in Figure 15. Both the proportions of samples positive and the mean counts suggest an increase in *E. coli* between the ESR and household taps in IWS versus a decrease in

CWS; however, as can be seen in Figure D12, these differences are not significant after accounting for the variation in the counts.

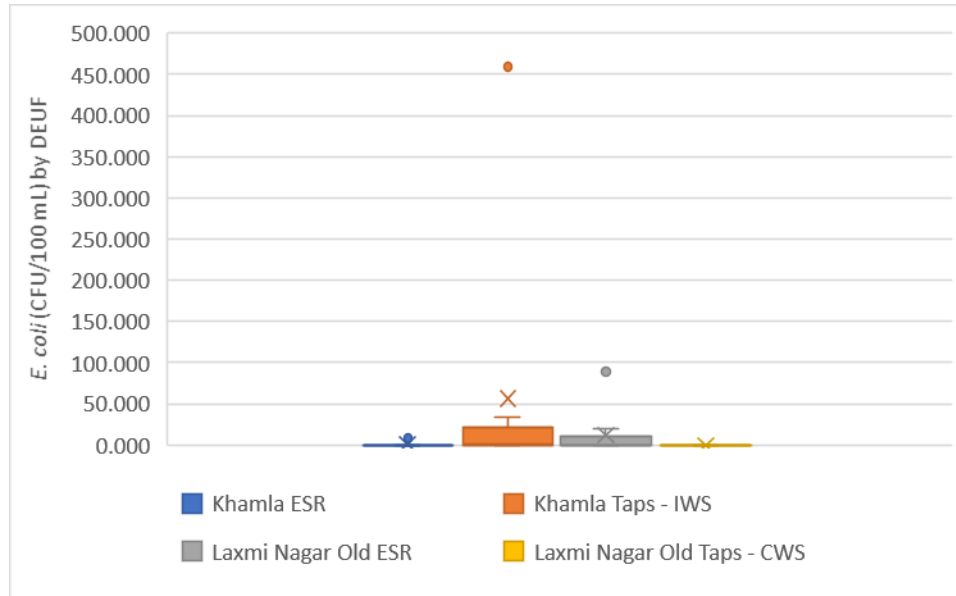


FIGURE 15: *E. coli* counts as observed via DEUF at ESRs and household taps in IWS (Khamla) and CWS (Laxmi Nagar Old) service areas in Nagpur.

Biochemical Testing of Isolates

Prior to testing bacterial isolates from DEUF backflush via the HiSalmonella and HilMViC kits, we tested archived isolates of *Salmonella typhi*, *Shigella boydii*, and *E. coli* O157:H7. The *S. typhi* isolate was identified as *S. typhimurium/S. enterica* serotype choleraesuis by the HiSalmonella Kit. The *S. boydii* isolate was identified as *S. boydii/S. flexneri/S. dysenteriae* by the HilMViC Kit, and the *E. coli* O157:H7 isolate was identified as *E. coli* by the same. We tested 28 isolates from SS Agar via the HiSalmonella kit; 25 were inconclusive, one was identified as *S. choleraesuis* subspecies Arizonae/Diarizonae, and two were identified as *S. choleraesuis* subspecies Indica. We tested 35 isolates from both SS Agar and MUG agar via the HilMViC Kit. Of these tests 15 were inconclusive, nine

were identified as *Proteus mirabilis*, four as *Citrobacter freundii*, three as *Citrobacter diversus*, three as *Klebsiella pneumoniae*, and one as *Cedecea lapagei*.

ddPCR Analysis of DEUF Samples

Prior to assaying the DEUF concentrate from each sample for gene targets, we analyzed cDNA from each sample for the BRSV process control. BRSV is a qualitative control where we interpret successful detection via ddPCR to indicate a successful extraction, reverse transcription, and amplification within the matrix of that individual sample. Our ddPCR assay for BRSV successfully detected the target spiked into each sample and all negative extraction controls were void of the target. For this assay, we found strong agreement between ddPCR quantification by manual thresholding and Umbrella thresholding as can be seen in Figure D13. For the MS2g1, beta giardin, norovirus GI ORF1-2, norovirus GII ORF1-2, ipaH, STh, and *Cryptosporidium* 18S rRNA assays we observed no false positive droplets in negative controls by either thresholding method. While in the hexon assay we observed one false positive droplet in 265,592 negative control droplets by manual thresholding and zero by Umbrella thresholding. We also only observed large discrepancies between manual and Umbrella quantification in 19 out of a total of 1,152 ddPCR wells (2 beta giardin, 8 STh, and 9 *Cryptosporidium* 18S rRNA). In each case inspection of the results from the problematic wells revealed that baseline shifts, bimodal negative distributions, or increased variation in sample negative distributions compared to negative controls caused misclassification of droplets. In each instance, the Umbrella results for that well were discarded and manual thresholding results were used for quantification of the target. The exception to this excellent performance is the ybbW assay where significant baselines shifts and dispersion within sample negative clusters

compared to negative controls cause significant discrepancy between the manual and Umbrella thresholding. In this instance, we discarded the Umbrella thresholding entirely and report only the manual thresholding results. Table D1 summarizes the analytical performance we observed for each assay as previously described.

We assayed DEUF concentrates for two genetic targets associated with fecal indicators. Gene target MS2g1, associated with male-specific coliphage, was detected above the 95% LOD in only 1 sample which was collected from the Laxmi Nagar ESR (172 gc/L). However, ybbW, a gene associated with *E. coli*, was detected above the 95% LOD in 45 DEUF concentrate samples. Boxplots of the observed ybbW concentration in the IWS and CWS drinking water supply trains suggest that the concentration of the gene decreased through the treatment train at Pench 2 WTP, but then increased en route to both the IWS and CWSs ESR (Figure 16). In the IWS command area, the mean ybbW concentration increased further from the ESR to household taps, while in the CWS area it decreased. These differences are not statistically significant after accounting for variation in the ybbW count at each sampling point. A comparison of the proportion of DEUF samples positive for culturable *E. coli* versus the ybbW gene, Figure 17, indicates that the gene was present even when culturable *E. coli* was not and that it was sometimes present in hundreds of copies per liter even without the detection culturable *E. coli* (CWS Taps). We observed a positive correlation (0.61) between gene copies of ybbW per ddPCR reaction and *E. coli* counts by culture (CFU/100mL), but unlike Walker *et al.* we found that a linear regression was a poor fit, $R^2 = 0.23$, of ybbW (log gc/reaction) versus culturable *E. coli* (log CFU/reaction) counts (Figure D14).

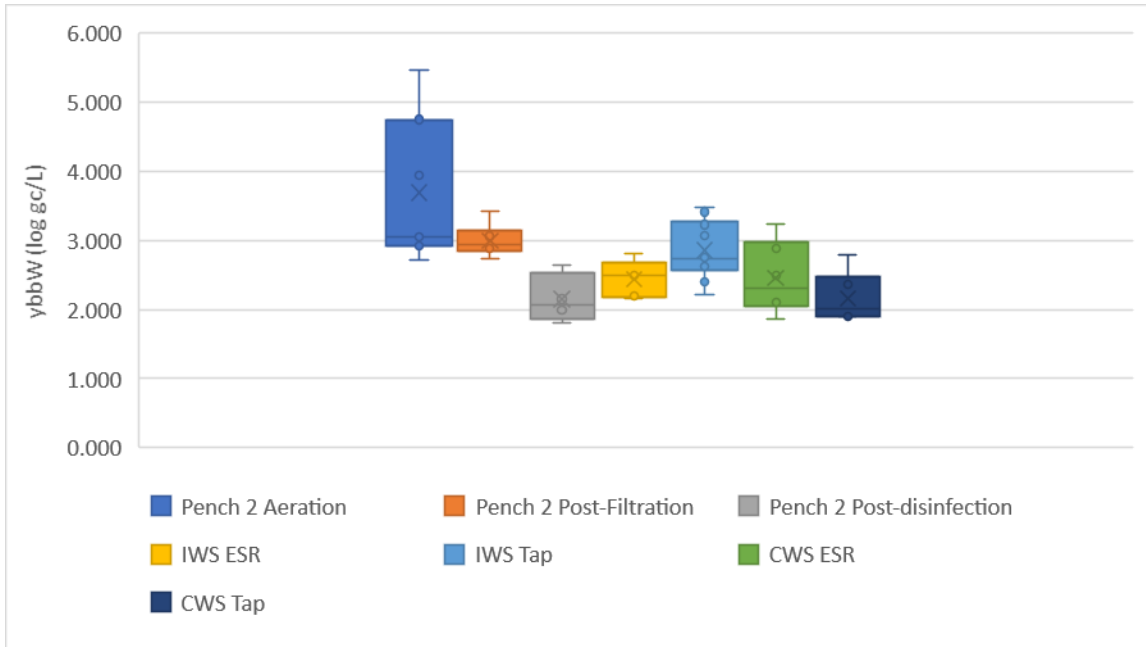


FIGURE 16: Box plots of observed gene copies of ybbW per liter of drinking water at sampling points along the IWS and CWS drinking water delivery chain in DEUF concentrate collected from the Nagpur water supply in 2017.

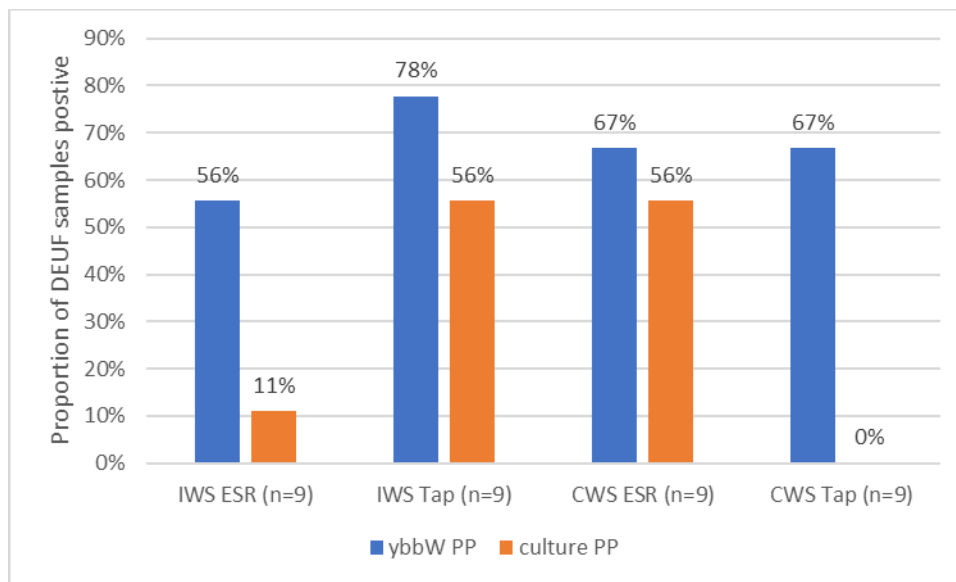


FIGURE 17: The proportion of DEUF samples positive for ybbW, a gene associated with *E. coli*, versus proportion of samples in which *E. coli* was detected by culture.

We detected genes associated with bacterial pathogens in 12 DEUF samples. We detected ipaH, a gene associated with *Shigella* and Enteroinvasive *E. coli*, above the 95% LOD in one sample from a tap in the Reshim Bagh (IWS) command area (821 gc/L). We detected, STh, associated with ETEC, above the 95% LOD in 11 DEUF samples: 6 from IWS taps (range: 167 – 1429 gc/L), 3 from ESRs (range 72 – 84 gc/L), and 2 from the Pench 2 WTP (range: 72 – 109 gc/L). For viral pathogens, we detected associated genes in 15 DEUF samples. We detected the ORF1-2 gene associated with norovirus GI in samples from 3 IWS taps (range: 2,927 – 19,144 gc/L) and the ORF1-2 gene associated with norovirus GII in samples from 4 IWS taps (range: 3,746 – 79,901 gc/L) and 1 ESR (2,643 gc/L). We detected the hexon gene associated with adenovirus A to F in a total of 7 samples: 3 samples from IWS taps (range: 246 – 689 gc/L), 1 sample from an ESR (24 gc/L), and 3 samples from the Pench 2 WTP (range: 26 – 44 gc/L). Lastly, we detected genes associated with protozoan pathogens in 10 DEUF concentrate samples. We detected beta giardin, associated with *Giardia duodenalis*, in samples from 5 IWS taps (range: 23 – 106 gc/L) and 1 ESR (79 gc/L). We detected the gene encoding 18S rRNA associated with *Cryptosporidium* in one IWS tap sample (429 gc/L), one CWS tap sample (67 gc/L), and one Pench 2 WTP sample (88 gc/L).

The total number of DEUF samples positive above the 95% LOD by ddPCR for each gene target stratified by sampling point within the IWS and CWS drinking water supply system of Nagpur are summarized in Table D2. While Table D3 summarizes the number of DEUF samples positive above the 95% LOD by ddPCR for each gene target and all gene targets associated with pathogens in IWS versus CWS command areas. The proportions of DEUF samples positive for each gene target and for an aggregate of all gene targets associated

with pathogens as observed at IWS taps versus CWS taps are summarized in Figure 18. Except for *Cryptosporidium* 18S rRNA, gene targets associated with pathogens were detected more frequently in IWS taps. Meanwhile, a gene associated with a pathogen, *Cryptosporidium*, was only detected in one sample from a CWS tap. Detection of this protozoan pathogen is not improbable given its resistance to residual chlorine. As shown in Figure 19, after aggregating all detections of genes associated with pathogens across both IWS and CWS taps and ESRs, we observe an increase in the proportion of samples positive from the IWS ESR (26.7% 95% CI: 10.9 – 52.0) to the associated household taps (61.1% 95%CI: 44.9 – 75.2) and a decrease from the CWS ESR (33.3% 95%CI: 12.1 – 64.6) to the associated household taps (11.1% 95%CI: 2.0 – 43.5). The Wilson Score intervals estimated for the proportion of samples positive at each of these sampling points indicates a significant difference between the number of IWS tap samples positive for any gene target associated with a pathogen and the number of CWS tap samples positive for the same.

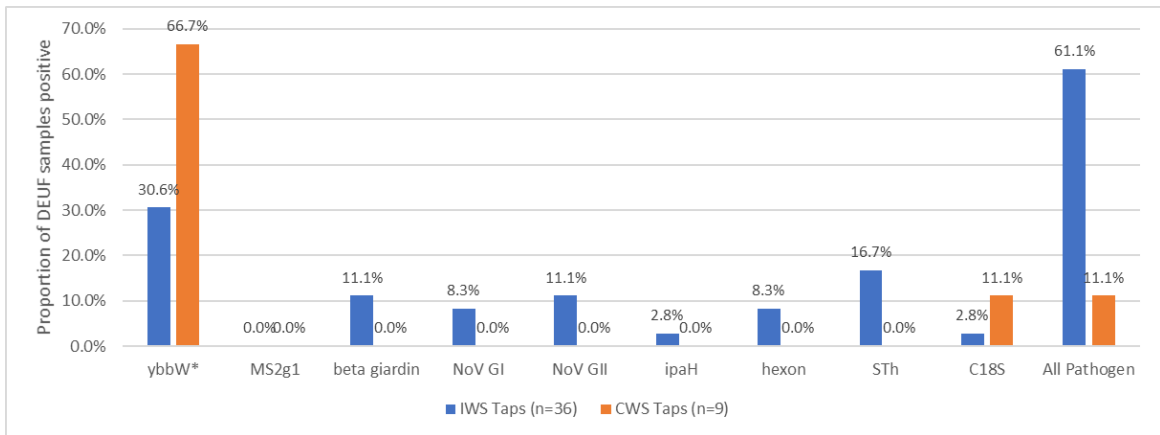


FIGURE 18: The proportion of DEUF samples positive for each genetic target as observed in taps served by IWS versus those served by CWS in Nagpur. The “All Pathogen” category is the aggregate of all the detections of a gene target associated with a pathogen across all IWS versus CWS taps.

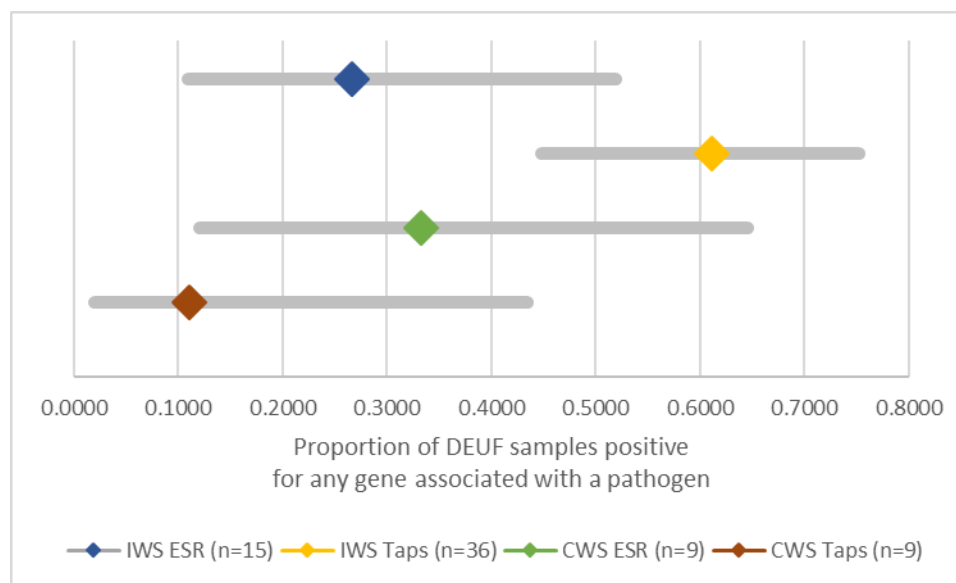


FIGURE 19: The proportion of DEUF samples positive for any gene associated with a pathogen and the associated Wilson Score Interval as observed in samples from an IWS ESR, IWS Taps, a CWS ESR, and CWS taps in Nagpur.

Drinking Water Quality in IWS vs CWS

Our observations highlight several important features of the quality of drinking water delivered by IWS versus CWS. In general, our data suggest that microbial water quality was degraded in IWS zones as mean counts of thermotolerant coliforms and *E. coli* increased from the ESR to the household taps, but these findings were not statistically significant after we accounted for the variation in the observed counts. Our data also suggest an increase in the variance of water quality parameters, including free chlorine, turbidity, and microbial counts at household taps in IWS zones compared to CWS zones. Interestingly, in our limited enumerations of heterotrophs we found almost identical counts per mL in IWS and CWS zones. We did observe a statistically significant increase in the proportion of grab samples positive for *E. coli* from taps served by IWS versus CWS. In the Indian context, our findings are consistent with a previous study of drinking water

quality in Hubli-Dharwad that documented significant degradation of drinking water quality in IWS zones compared to a CWS demonstration zone (24).

Waterborne Pathogens in IWS vs CWS

Our study is the first to provide observations concerning the presence of waterborne pathogens in IWS versus CWS zones in the same context. We detected gene targets associated with seven different fecal oral pathogens at taps in IWS zones compared to one pathogen detected in a sample from a tap in the CWS zone. We detected genes associated with bacterial, viral, and protozoan pathogens indicating that all microbial classes could be associated with waterborne disease transmitted via IWS. Each of the genes we detected is associated with a pathogen that contributes to the burden of diarrheal disease among children in South Asia as observed in the MAL-ED and GEMs trials (67,131). We observed that the proportion of samples positive for any gene associated with a pathogen in samples collected from IWS taps was greater than the proportion from CWS taps even after accounting for random variation. Our observation indicates that IWS users in the study context are subject to an increased likelihood of exposure to waterborne pathogens compared to CWS users. Of course, our findings, like all environmental surveillance using molecular methods, rely upon the implicit assumption that the observed gene counts are positively correlated with viable and infectious organisms within the sample.

Heterogeneity in Water Quality in IWS

Our findings also emphasize the intra-system and inter-system heterogeneity that has been observed in drinking water quality in IWSs. We observed variations in water quality between hydraulically isolated IWS zones within the municipal water supply. For example, 80% of grab samples were positive for *E. coli* in one IWS zone of the Nagpur water supply

and 28% were positive in another. And in Hubli-Dharwad the proportion of samples positive for *E. coli* ranged from 4% in one IWS ward to 80% in another (24). These intra-system water quality differences between hydraulically isolated areas are statistically significant. We also observe variation in inter-system water quality in IWSs. Erickson *et al.* detected culturable total coliforms in only 4 samples and *E. coli* in only 1 sample of 423 grab samples from an IWS in Panama (82). These findings indicate that the quality of water delivered by an IWS is likely dependent on context-specific variables. Given the complex hydraulic functioning of IWS distribution networks, it is likely that these context-specific variables must be defined at the level of municipal water supply hydraulic zones (175). Such variables are likely to include metrics regarding fecal pollution sources within the zone, operations and maintenance, and hydraulics.

Limitations

Several limitations of our current work require that we be circumspect in interpreting our findings. Firstly, is the small sample size. Due to constraints of both time and resources, we collected only 90 DEUF samples and 56 grab samples in the course of our study. However, these 90 DEUF samples represent a total volume of 6,925 liters collected from the Nagpur municipal water supply, which, assuming a 100 mL grab sample volume, is equivalent to 69,250 grab samples. While the volume filtered in a single DEUF sample is not independent in time and space, it does represent a significantly greater proportion of the delivered drinking water than a single grab sample. For example, we collected only 9 DEUF samples from CWS household taps, but these samples constitute a total volume of 1,006 liters and we did not detect a single culture *E. coli* in that volume. This is a meaningful observation despite the collection of only 9 samples.

The second limitation of our work is a result of the use of molecular assays to detect gene targets associated with waterborne pathogens. The use of such assays is predicated on the assumption that the detection of genes correlates with the presence of viable pathogens capable of causing an infection. The validity of this assumption likely varies based on microbial class and the environmental compartment. In our study, we observed that a gene associated with *E. coli* was positively correlated with the number of culturable *E. coli* in that sample; however, we also observed instances where we detected up to hundreds of copies of the gene in samples where we did not detect culturable *E. coli*. Given that viruses and protozoa are more resistant to disinfection than bacteria, we assume that the same positive correlation observed for *E. coli* would also apply for gene targets associated with pathogens in those classes (41,194).

Another limitation is that we conducted sampling in only three different IWS zones and one CWS zone and across two different time periods including two months in 2015 and six months in 2017. Although we worked with utility staff to identify zones that were representative of the water supply, it is possible that the limited number of zones might limit the generalizability of our findings. Also, during the time frame of sampling, the Nagpur drinking water supply was undergoing many changes associated with the transition from IWS to CWS. These changes included changes in personnel and operations and management protocols which could confound our findings. Additionally, the city of Nagpur is well supplied by surface water sources for its drinking water and is not subject to the scarcity that is normally cited as the reason for IWS. Collectively, these characteristics of our study design and site may limit the generalizability of our findings.

Lastly, a fundamental limitation of both our study and the study in Hubli-Dharwad is that these interventions did not consist of only changes to the water delivery mode. These conversions to CWS were in the context of large multi-million-dollar upgrades of the existing pipe networks in the zones receiving CWS. These upgrades are necessary to prevent the tremendous increases in non-revenue water via leakage that would accompany pressurization for 24 hours instead of only a few hours each day (195). These conversions also included overhauls in the operations and maintenance practices of the entities responsible for the drinking water supply. We know of only one study where IWS zones were temporarily converted to CWS without accompanying distribution line replacements. In this study, Andey and Kelkar found that the proportion of samples positive for thermotolerant coliforms decreased from a range of 17% to 74% to a range of 0% to 10% following the switch to CWS (196). However, this improvement in water quality was accompanied by an increase in non-revenue water from 19.5% to 47.8% due to increased leakage during the continuous supply period. Therefore, the findings of our study must be interpreted as water quality improvements associated with a conversion from IWS to CWS accompanied by the replacements of distribution lines necessary to limit non-revenue-water.

Policy Implications for IWS in India

Our results strongly suggest that in Nagpur the conversion from IWS to CWS effectively reduced microbial contamination in drinking water as indicated by reductions in the number of samples positive for *E. coli* and genes associated with waterborne pathogens. While these results are directly attributable to physical improvements in the water infrastructure, it is extremely important to note that these physical improvements flow from

improvements in governance and management. Indian utilities are characterized by inefficiency, increasing costs, and flat revenues with only 20% of connections metered and non-revenue water rates of 40% or more (186,197). During the conversion in Nagpur, increased metering and decreased non-revenue water has been critical to the financial solvency of the project and improving compliance with service level benchmarks (190). Our findings indicate that in Nagpur increased revenue through improved metering and billing is allowing the utility to invest in infrastructure improvements that are subsequently delivering water of improved microbiological quality. This experience in Nagpur could provide a road map for other Indian utilities to follow in transitioning from IWS to CWS.

In the near term, the health impacts of IWS could be reduced through two different approaches one at the municipal level and one at the household level. First, at the municipal level it may be possible to identify hotspots of fecal contamination with water supply zones and implement specific interventions to reduce them. For example, targeted repairs of failing sewer lines proximate to water lines or extending sewer connections to unsewered sources of fecal contamination could reduce the flow of fecal contaminants into drinking water lines. Replacing segments of sewer lines or failing latrine pits is likely to be more cost effective than the replacement of entire water distribution networks required to reduce non-revenue water prior to implementing continuous supply.

A second approach is improvement of water quality at the household level immediately prior to the point-of-use. Previous studies of microbial water quality in Nagpur have documented degradation of water quality from the moment of delivery at the tap to the moment of consumption at the point of use, especially in low-income households (88,89). Risk assessments of household water treatment indicate that health gains for such treatment

technologies are driven by consistent and effective use (94,166). However in India, treatment of drinking water at the household level is low and variable with 25% of households in Mumbai, 44% in Bengaluru, and 33% in a Dehli slum indicating they treat their drinking water at the household (198,199). While household water treatment could be an important temporary solution to decrease waterborne disease associated with IWS, it is clear that designs must be implemented that balance ease of use and efficacy to maximize benefits (200).

Despite the limitations of the current study, our findings provide additional evidence that in certain settings IWS is associated with degraded quality of drinking water compared to CWS. Further, our findings indicate that this degradation of water quality includes an increased prevalence of waterborne pathogens at the household taps served by IWS. Our findings are most relevant to India, where some 460 million urban dwellers are served by IWS. These IWS users are likely exposed to an increased number of waterborne pathogens via drinking water and are therefore subject to an increased risk of diarrheal disease as a result. However, the success of transition from IWS to CWS in Nagpur in reducing microbial contamination of drinking water at the household indicates that the project provides a template that could be replicated elsewhere.

CHAPTER 5: RISKS OF INFECTION WITH WATERBORNE PATHOGENS ATTRIBUTABLE TO INTERMITTENT WATER SUPPLY IN INDIA

ABSTRACT

Quantitative microbial risk assessments (QMRA) of distributed drinking water frequently rely on measures of fecal indicator bacteria and pathogen to indicator ratios in sources of contamination, such as sewage, to estimate risks of infection associated with water distribution systems. In this QMRA we used droplet digital PCR measures of gene targets associated with six different fecal-oral pathogens, experimental details regarding dead-end ultrafiltration (DEUF) workflows, and previously published dose-harmonization models to estimate the risk of infection associated with tap water delivered to households via IWS. Our Monte Carlo simulations estimated daily and annual probabilities of infections associated with pathogens known to be associated with mild to severe diarrhea in South Asia. Our model indicates substantial risks of infection associated with even one day of ingesting tap water from an IWS. For all but enterotoxigenic *E. coli* (ETEC), the 10th percentile daily probability of infection exceeded the US EPA annual risk threshold of 1 infection in 10,000. Sensitivity analyses indicate that risk estimates associated with protozoan and bacterial pathogens were sensitive to dose-harmonization parameters including measures of viability. All estimated risks were sensitive to gene target measures by ddPCR, and the volume of drinking water sampled by DEUF. Given the substantial daily risk of infection at the 10th percentile, we estimated annual risks of infection using a daily risk fixed at this level for each pathogen. Our calculations indicate that IWS could account for 11 million *Giardia* infections, 60 million *Cryptosporidium* infections, and 2.17

million *Shigella* infections annually among the 460 million Indians served by IWS. These risks are several orders of magnitude higher than those estimated by a previous QMRA using reference pathogens and *E. coli* counts because pathogen counts as estimated by DEUF and ddPCR are greater than those estimated by pathogen to *E. coli* ratios in sewage. Despite limitations the current risk assessment adds further evidence that IWS is very likely associated with infection risks that exceed the 1 in 10,000 threshold established by the US EPA.

INTRODUCTION

Piped water distribution networks rely on positive pressure within them to prevent the intrusion of contaminants from the environment into the drinking water they distribute. Among these potential contaminants are microbes capable of causing diarrheal disease if ingested by end users. Even brief interruptions in pressure have been linked to increased risk of gastrointestinal illness (GII) among end users (RR=3.26 95% CI: 1.48 – 7.19) (14). It is not surprising, then, that chronic sustained interruptions in pressure like those experienced by users of intermittent water supply (IWS) have been associated with increased risks of GII (OR = 1.61 95%CI: 1.26 – 2.07) (14). More recently IWS has been linked with increased risk of typhoid fever in Hubli-Dharward, India, increased incidence of cholera in the Democratic Republic of the Congo, increased diarrhea among children in Guatemala and Addis Ababa, and increased gut inflammation among children in Peru (28,29,180–182).

In Hubli-Dharwad the association of IWS with increased risk of typhoid fever was corroborated by an increase in the prevalence of *E. coli* in water samples collected from IWS wards compared to CWS wards (24). Numerous studies of water quality have documented fecal contamination in IWSs as indicated by the presence of various fecal indicators including *E. coli*, thermotolerant coliforms, and total coliforms (23,75,169,170). However, this contamination appears contextually dependent, as a study of water quality in an IWS in Panama detected culturable *E. coli* in only one sample of 423 collected (82). But, in the Indian context, the presence of fecal contamination in IWSs and increasing levels of such contamination as drinking water moves through the delivery chain to the point-of-use is documented consistently (88,89,196,199). For the 460 million urban dwellers exposed to IWS in India, both the epidemiological and microbiological data indicate that the drinking water they consume could constitute a non-negligible risk to their health in the form of diarrheal disease (201). The impact of such waterborne disease on the Indian economy has been estimated to be the loss of 90 million workdays per year at a cost of 6 billion INR in production loss and treatment costs (202).

Quantitative microbial risk assessment (QMRA) provides a useful tool for quantifying the public health impacts of IWS in India. QMRA has been applied to quantify and manage the health risks associated with drinking water (38,203). A QMRA found that the annual infection risks associated with pathogenic *E. coli* in distributed drinking water exceeds the WHO threshold of 10^{-6} DALYs per year among users of the municipal drinking water supply in Mysore, India (98). While another assessment estimated that IWS could be responsible for 17.2 million annual infections and 4.52 million cases of diarrhea among the 950 million people exposed to such supplies (97).

These QMRAs, like many others of drinking water distribution, relied upon the use of observed fecal indicator counts in distribution networks along with assumed ratios of these indicators to waterborne pathogens to estimate the number of pathogens ingested at the moment of exposure in a process known as exposure assessment (42,51,204). Sensitivity analyses in two risk assessments of distributed drinking water found that estimated risks were most sensitive to the concentration of the pathogen and subsequently *E. coli* counts and ratios of pathogens to *E. coli* (97,99).

In the current study we make use of data from a characterization of microbial water quality in an IWS in Nagpur, India to populate a stochastic model and estimate the risk of infection among those exposed to IWS in India. Importantly, our QMRA makes use of quantitative measures of gene targets associated with waterborne pathogens as observed in samples collected from household taps served by IWS and enumerated via dead-end ultrafiltration and droplet digital PCR. Like fecal indicator organisms, these genes are still proxies for infectious waterborne pathogens; however, our use of them to perform hazard identification and exposure assessment represents another option to estimate the public health risks associated with IWS.

MATERIALS AND METHODS

We developed our risk assessment model using the hazard identification, exposure assessment, dose-response, and risk characterization framework (30). As shown in Figure 20, we also performed dose harmonization as an additional module within this framework to convert from the experimental unit of measure, gene copies, to infectious units

appropriate for dose-response. In total our model includes 22 stochastic variables, and 21 equations to estimate the probability of infection and/or diarrhea associated with six different fecal-oral pathogens as the result of consuming tap water from an IWS. We provide a descriptive summary of each stochastic variable in Table 9 and the equations are summarized in Tables E1, E2, and E3. We executed the Monte Carlo simulations of our model in Oracle Crystal Ball (Release 11.1.2.4.600, Redwood City, CA) within Microsoft Excel 2016 (Version 16.0.11929.20300, Redmond, WA).

Hazard Identification

During our sampling from household taps served by IWS in Nagpur, we detected genes associated with *Giardia*, *Cryptosporidium*, norovirus GI, norovirus GII, adenovirus A-F, ETEC, and *Shigella*/EIEC. With the exception of *Giardia*, each of these fecal-oral pathogens was associated with diarrheal disease in under-5 children as observed at South Asian study sites during both GEMS and MAL-ED (Figures C1, C2, and C3) (66,68,131). Our detection of these genes is consistent with previous surveillance of drinking water sources in South Asia including the detection of *E. coli* virulence genes in piped water in Lucknow and Kolkata, isolation of *Shigella flexneri* from a piped water supply in Gayeshpur, and detection of *Cryptosporidium* oocysts and *Giardia* cysts in potable water supplies in Chennai and Odisha (153–155,157,205,206). Based on these observations, each of these pathogens is a plausible agent for waterborne disease associated with IWS in South Asia in general and India in particular.

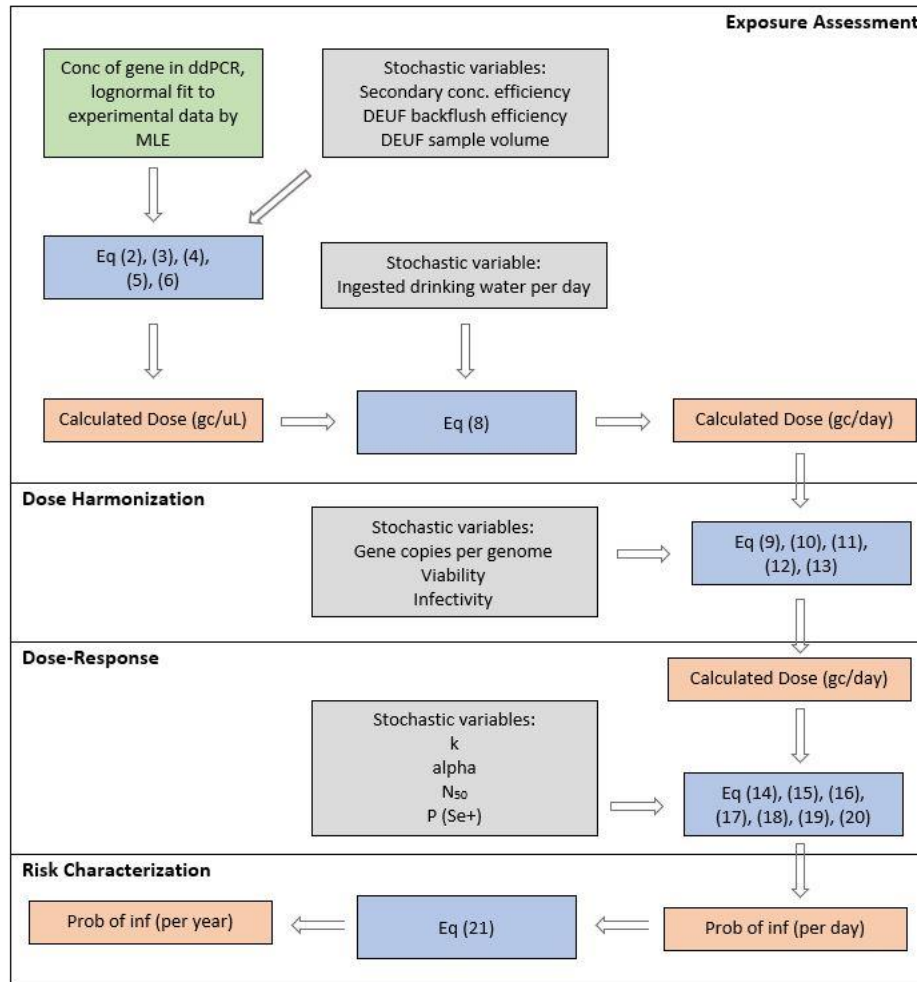


FIGURE 20: Schematic of the mathematical framework developed to estimate the probability of infection for each of six waterborne pathogens detected at household taps served by IWS in India.

Protozoan Pathogens

Giardia duodenalis is a protozoan pathogen that causes self-limiting watery diarrhea following ingestion and infection (207). Evidence indicates that assemblages A and B infect humans at prevalence rates of 8% to 30% in low- and middle-income countries and 1 to 8% in high-income countries (208,209). *Giardia* is excreted in the feces of infected individuals in the form of cysts which are resistant to environmental degradation and disinfection via chlorine (41,210). These cysts have been detected in tap water samples,

but only rarely (211,212). Despite rare detection in finished drinking water, there were 199 documented outbreaks of waterborne giardiasis between 2004 and 2010 (213).

Cryptosporidium is also a protozoan parasite that causes self-limiting diarrhea in healthy adults following the ingestion of oocysts shed in the feces of infected mammals (214). Unlike *Giardia*, it can also cause serious long term sequelae in children, the elderly, and the immunocompromised (215–218). During the GEMS study, *Cryptosporidium* was among the leading causes of diarrhea in children under 5 (66). Based on the attributable fraction of diarrhea observed in GEMS, *Cryptosporidium* could plausibly account for 4.7 million cases of diarrheal disease in South Asia (219). Although 37 species have been identified, *C. parvum* and *C. hominis* are the most prominent pathogenic species in humans (220). Oocysts shed in the feces of infected individuals are resistant to degradation in the environment and are resistant to drinking water treatment (194,221). Baldursson *et al.* found that *Cryptosporidium* accounted for 63% of waterborne outbreaks caused by protozoa (213).

Viral Pathogens

Noroviruses are non-enveloped single-stranded RNA viruses in the family *Caliciviridae* (222). In healthy adults, infection with norovirus via direct person-to-person, foodborne, or waterborne transmission causes the acute onset of diarrhea and projectile vomiting. However, among children norovirus is the third most common cause of diarrheal mortality (223). Although there are seven genogroups, infections in humans are only caused by genogroups GI, GII, and GIV (224). Norovirus particles are more resistant to chlorine than other viruses with 2 to 4 log reductions observed for chlorine at a concentration of 1 mg/L

(225). There have been many documented outbreaks of waterborne norovirus associated with drinking water (226,227).

Adenoviruses are non-enveloped double-stranded DNA viruses that have been classified into six subgenera with subgenus F, including types 40 and 41, associated with gastroenteritis transmitted via fecal-oral routes (228). Types 40 and 41 are second among the primary causes of gastroenteritis in children (229). Human adenoviruses have been detected in both groundwater and piped water supplies (156,230). While adenoviruses appear quite stable under environmental conditions types, 40 and 41 are readily inactivated by free chlorine with 3 log inactivation observed following 5 seconds at 0.2 mg/L (231). The role of adenovirus in waterborne disease remains unclear with no reported waterborne outbreaks (232).

Bacterial Pathogens

Shigella spp. are gram negative, facultative anaerobes in the *Enterobacteriaceae* family (233). The four species within the genus include *S. dysenteriae*, *S. flexneri*, *S. boydii*, and *S. sonnei* with *S. flexneri* being more prevalent in LMICs and *S. sonnei* more prevalent in high-income countries (234). *Shigella* infections cause diarrhea ranging from mild and watery to severe dysentery (*S. dysenteriae*) (235). There is evidence of an environmentally persistent *Shigella* spp. isolated from surface water in Bangladesh (236,237). Under environmental stress, *Shigella* have been shown to enter a viable but not culturable (VBNC) state (238). Although *Shigella* are inactivated by free chlorine, various species have been associated with outbreaks of disease via drinking water (239,240).

Pathogenic *E. coli* are a group of serotypes within the *Enterobacteriaceae* family (241). They include several types such as enteropathogenic (EPEC), enterotoxigenic (ETEC), enteroaggregative (EAaggEC), enteroinvasive (EIEC), enterohemorrhagic (EHEC), and diffusely adherent (DAEC). Pathogenic *E. coli* are transmitted via contaminated water and food. Infection with most serotypes results in diarrhea; however, infection with EHEC, which includes *E. coli* O157:H7, can cause serious disease including hemolytic uremic syndrome. Although they are inactivated by free chlorine, they are also able to enter a VBNC state with observed recovery in cultivability (238,242). Waterborne outbreaks of pathogenic *E. coli* have been associated with extreme rainfall events and drinking water supplies that are not chlorinated (243–245).

Exposure Assessment

Exposure assessment describes the process by which the ingested dose of each organism at the moment of exposure is computed. In our case, we begin with the concentration of the gene target associated with each pathogen as measured via ddPCR. For each pathogen we used maximum likelihood estimation (MLE) to parameterize a log normal distribution that maximizes the likelihood of observing the measured concentrations (gene copies per uL of reaction mix). For three genetic targets we performed MLE using probability density function values for concentrations above the 95% limit of detection (LOD) and cumulative density values for censored observations – beta giardin (5 above LOD and 37 censored); hexon (3 above LOD and 33 censored); STh (5 above LOD and 30 censored). For *Cryptosporidium* (Crypto.) 18S rRNA and ipaH, we observed only one sample positive above the 95% LOD from IWS taps, so we included additional detections with at least two positive droplets observed in ddPCR – Crypto. 18S rRNA (7 above threshold and 27

censored); ipaH (6 above threshold and 30 censored). For norovirus GI and GII, we pooled the observed concentrations above the 95% LOD for each genogroup into a single model (4 above LOD and 32 censored). Plots of the MLE lognormal model fit to the observed concentrations for each gene target are shown in Figures E1 through E6. The mean and standard deviation of each lognormal distribution are summarized in Table 9.

The gene target concentration in the ddPCR reaction mix is then propagated through equation (1) (DNA target) or equation (2) (RNA target to account for reverse transcription) followed by equations (3), and (4). These equations are formulated on the basis of mass balance through ddPCR reaction, elution of the nucleic acid, and lysis of the sample. The output from equation (4) is then used in equation (5) assuming mass balance during PEG precipitation and ultracentrifugation. We estimated the efficiency of this process using normal distributions fit to mean observed recoveries for MS2 (surrogate for viruses) and *Cryptosporidium* (surrogate for protozoa) during experiments by Mull and Hill (112). In the absence of *E. coli* data, we assumed the recovery for bacteria was equivalent of that of protozoa. The volume of drinking water filtered by DEUF is estimated using equation (6) with the volumes filtered modeled using uniform distributions based on values measured in the field and the dummy variable A is a Bernoulli trial with the probability of success equal to the proportion of samples that were collected in 2015. To estimate the concentration of each gene target in the drinking water, the output from equation (5) and equation (6) are used in equation (7). The efficiency of backflushing for viruses, protozoa, and bacteria was modeled as a normal distribution fit to observations by Smith and Hill (111). The output from equation (7) is the gene copies of each target in one uL of drinking water. This output is multiplied by a conversion factor and the volume of drinking water

ingested daily as observed in a study of the population of Mysore, India to calculate the gene copies of each target ingested daily (246). The detailed equations we used to model the exposure that results from the consumption of IWS tap water are summarized in Table E1.

Dose Harmonization

The dose calculated during the exposure assessment is in units of gene copies per day. The gene copies per day for each target must be converted to the appropriate infectious unit for the dose-response function associated with that pathogen. In the case of *Giardia* each cyst is equivalent to 16 gene copies of beta giardin (equation 9) (247). However, not all of these cysts are viable infectious units. We calculated the expected number of viable cysts based on the beta-Binomial (binomial distribution where the probability of success is beta distributed) relationship observed by Teunis using the stochastic parameters described in Table 9 (248). For *Cryptosporidium* each oocyst is equivalent to 20 gene copies of the small subunit rRNA gene (equation 10) (249–251). Again, not every oocyst is a viable infectious unit and we estimate the number of viable oocysts using the beta-Binomial distribution as parameterized by Teunis (Table 9) (248).

For norovirus dose harmonization, we assumed each gene copy was equivalent to one genome and thereby one infectious viral particle. The infectious unit for adenovirus is the median tissue culture infective dose (TCID₅₀) (252). We converted hexon gene copies to TCID₅₀ using conversion factor of 700 gene copies per TCID₅₀ as show in equation (11) (253,254).

Because *Shigella* and pathogenic *E. coli* are susceptible to inactivation by chlorine, we used paired enumerations of culturable *E. coli* in DEUF backflush and the *E. coli* associated gene *ybbW* to estimate the numbers of viable *Shigella* and *ETEC* in the drinking water as shown in equations (12) and (13) (141). These equations are formulated around the assumption that the ratio of viable *E. coli* to *E. coli* genomes would be equivalent to the ratio of viable pathogen to pathogen genomes. We used MLE techniques to fit lognormal distributions to the observed count of *ybbW* per uL in ddPCR wells (10 above LOD, 25 censored) and the observed count of culturable *E. coli* at IWS taps (5 above 95% LOD, 4 censored) as described previously and shown in Figures E7 and E8. The *STh* and *ipaH* gene copies per genome were modeled as uniform distributions between 1 and 16 and 5 and 14, respectively (136,255,256). Equations (12) and (13) also include a variable *B*, a binomial distribution with the probability of success equal to the proportion of IWS tap samples that were positive for culturable *E. coli* given the presence of *ybbW*. The effect of this variable is to convert the absolute viable counts to an expected value based on the likelihood of observing viable bacteria in samples positive for gene targets. The detailed equations each dose harmonization as implemented in our model are described in Table E3 and the stochastic variables and their associated parameters are listed in Table 9.

Dose-Response

We estimated the probability of infection given a dose of infectious *Giardia* cysts using an exponential model (equation 14) as parameterized by Rose *et al.* with parameter *k* lognormally distributed (31). We also calculated the probability of infection resulting from ingestion of *Cryptosporidium* oocysts using an exponential model with lognormally distributed parameter *k* (equation 15) (70).

Norovirus dose-response models are contentious owing to large uncertainties about model parameters and mechanistic versus empirical interpretation (257,258). On the basis of the review undertaken by Van Abel, we adopted two dose response models for drinking water – one model as a “high risk” representation and one model as a “low risk” representation (258). For our high-risk model, we adopted the fractional Poisson as described by Messner (equation 16) assuming a mean aggregate size, μ_a , of 1 – no aggregation (259). This assumption is reasonable based on the observed pH values of the drinking water during sampling which were well above the isoelectric point of norovirus. We estimated the fraction of the Indian population that is susceptible to norovirus infection (Se^+) as a normal distribution with mean and standard deviation as observed in a cohort of children in India (260). For our low-risk model we adopted the approximate beta-Poisson as parameterized by MLE in Van Abel’s analysis (equation 17) with point estimates for parameters alpha and the median infectious dose (N_{50}) (258). We elected to model norovirus dose-response parameters as fixed values to avoid injecting further variation into dose-response models that we selected to provide an envelope of possible risks.

To quantify the risk of infection with adenovirus we adopted the exact beta-Poisson model (equation 18) proposed for oral ingestion as described by Teunis with fixed values for parameters alpha and beta (261). This dose response model is based on infection with adenovirus 4 and 7, which are not typically associated with gastroenteritis (262). Estimating the probability of infection requires evaluating the confluent hypergeometric function (Kummer’s function) of the first kind. We evaluated this function using the XNumbers (XN.xlam v.6.0, Leonardo Volpi) visual basic application in Excel 2010. We embedded this VBA function within our Monte Carlo simulation.

We calculated the probability of diarrhea following ingestion of ETEC using an approximate beta-Poisson model (equation 19) fit to pooled data from several dose-response trials using MLE (71,263–266). To incorporate variability in the parameters we model the log of alpha and the log of the median infectious dose as normally distributed. It should be noted that the estimated probability for this model is that of diarrhea and not infection. Lastly, we estimated the probability of infection with *Shigella* using the approximate beta-Poisson model (equation 20) with the log of parameters alpha and the median infectious dose normally distributed (267,268). Detailed equations for each dose-response model we employed are summarized in Table E3 with the associated stochastic variables summarized in Table 9.

Risk Characterization

The output from the dose-response equations is the daily probability of infection, or diarrhea in the case of ETEC, that results from the ingestion of a dose attributable to drinking water from an IWS. As we have described our estimated dose represents the confluence of both experimental workflows, human behavior via drinking water consumption, and biological attributes of the pathogens of interest via dose harmonization and dose response. Each of these elements is characterized by variability and uncertainty. To propagate this uncertainty and variability into our estimates of risk, we executed our model as a Monte Carlo simulation in Crystal Ball where each stochastic variable is drawn in independent trials and the estimating equations are calculated for each draw. To estimate summary statistics of the daily probability of infection, including deciles, the mean and associated confidence interval, and the median and associated confidence interval, we bootstrapped the model by executing 200 samples of 1000 trials each and then calculating

the resulting summary statistics and their associated hyper-statistics. To convert the daily risk estimate into an annual risk estimate, we executed the Monte Carlo simulation and generated 10,000 trial values of the probability of infection. We then assumed independence of each daily risk and estimated the annual probability of infection by subsampling 365 daily probabilities from the 10,000 generated by the model and calculated the annual probability of infection as shown in equation (21). We repeated this process 200 times and then calculated the resulting summary statistics for the annual probability of infection. We executed all experiments using Crystal Ball's Monte Carlo sampling method with an initial seed value of 999.

$$P_{inf,annual} = 1 - \prod_1^n (1 - P_{inf,daily,i}), n = 365 \quad (21)$$

TABLE 9: A summary of the stochastic variables as implemented in our risk assessment model to estimate the probability of infection associated with IWS tap water in India.

Model Variable	Probability Distribution	Reference
Exposure Assessment		
Concentration in ddPCR reaction mix by ddPCR (gc/uL)	<i>Giardia</i> (beta giardin), LN (0.0451, 0.1140) Crypto. (18S rRNA), LN (0.0513, 0.0532) norovirus GI/GII (ORF1-2), LN (0.4309, 15.40) adeno A-F (hexon), LN (0.2002, 51.29) ETEC (STh), LN (0.1914, 1.474) <i>Shigella</i> (ipaH), LN (0.0380,0.0690)	MLE fit to experimental data
PEG efficiency (%),	viral, N (0.82,0.20) protozoan, N (0.89,0.20) bacterial, N (0.89,0.20)	(112)
Blackfush efficiency (%)	viral, N (0.57,0.077) protozoan, N (0.87,0.18) bacterial, N (0.85,0.07)	(111)
Volume of backflush total (L)	U (0.350,0.450)	experimental data
2015 DEUF or 2017 DEUF, A (1 or 0)	BT (0.37)	experimental data
2015 DEUF Volume (L)	U (13.42,22.56)	experimental data
2017 DEUF Volume (L)	U (83.35, 122.28)	experimental data
Ingestion drinking water (L/day)	N (1.125,0.4)	(246)
Dose Harmonization and Infectious Unit		
Probability of culturable <i>E. coli</i> present given ybbW detected, B	Binom (0.7142)	experimental data
Culturable <i>E. coli</i> at IWS taps (CFU/100mL)	LN (145.0, 6863)	MLE fit to experimental data
ybbW at IWS taps (gc/uL)	LN (2.880, 244.4)	MLE fit to experimental data
STh gene copy per genome	U (1, 16)	(255,256)
ipaH gene copy per genome	U (5, 14)	(136)
Prob. of viable Crypto. oocyst, PVO	Beta (0, 1, 1.65, 2.46)	(248)
Prob. of viable <i>Giardia</i> cyst, PVC	Beta (0, 1, 3.00, 17.4)	(248)
Proportion of viable oocyst in dose (%)	Binom (PVO)	(248)
Proportion of viable cyst in dose (%)	Binom (PVC)	(248)
Dose-Response		
<i>Cryptosporidium parvum</i> d-r parameter, k	LN (0.3973, 2.6836)	(70)
<i>Giardia duodeanlis</i> d-r parameter, k	LN (0.0208, 0.0064)	(31)
norovirus, high-risk model, susceptible proportion, P(Se+)	N (0.6098, 0.0280)	(260)
ETEC dose-response parameters, alpha, N ₅₀	log alpha N (-1.123, 0.1653) log N ₅₀ N (6.230, 0.643)	(263–266)
<i>Shigella spp.</i> dose-response parameters, alpha, N ₅₀	log alpha N (-0.5768, 0.0961) log N ₅₀ N (3.170, 0.1397)	(267,268)

LN = lognormal (mean,sd); N = normal (mean,sd); U = uniform (max,min); BT = Bernoulli trial (Prob of success); Beta = beta distribution (min, max, alpha, beta); Binom = binomial distribution (Prob of success)

RESULTS AND DISCUSSION

Daily Probability of Infection

Using our Monte Carlo simulation as described, we estimated the daily and annual risks of infection associated with the consumption of tap water from an IWS for six different fecal-oral pathogens. For each one we estimated the mean, median, 10th percentile, and 90th percentile of the daily probability of infection via bootstrapping. Our model estimated substantial risks associated with IWS for each pathogen. For both viral pathogens and *Cryptosporidium*, the estimated daily risks at the 10th percentile exceed the annual risk of 1 in 10,000 specified by the US EPA. For *Giardia* the 10th percentile daily risk of infection was 6.62×10^{-5} or 1 in 15,106. While for ETEC and *Shigella* the 10th percentile risks were 2.52×10^{-9} (1 diarrheal case in 396 million) and 1.26×10^{-5} (1 in 79,365). Both the low-risk and high-risk dose-response models for norovirus predicted significant risk of infection even at the 10th percentile (high-risk model: 1 in 7; low-risk model: 1 in 1,800). The 10th percentile daily risk of infection for adenovirus was also significant at 1 in 585. At the median percentile of risk only the risk of diarrhea associated with ETEC remains below 1 in 10,000. The cumulative distributions of the daily risk of infection (diarrhea for ETEC) associated with exposure to each pathogen via IWS tap water are shown in Figures E9 through E15. The observed means for each pathogen except the norovirus high-risk model are in the 68th to 93rd percentile reflecting the right-skewed nature of the lognormal models used to estimate microbial counts. For the norovirus high-risk cumulative distribution, the mean is in the 27th percentile owing to the fractional-Poisson dose-response model and the susceptible population (mean = 61%). Based on these results,

summarized in Table 10, even a single day of ingesting tap water as delivered by an IWS would represent a significant risk of infection.

TABLE 10: Measures of central tendency of the daily risk of infection (diarrhea for ETEC) and the 10th and 90th percentiles as estimated by bootstrapping of the Monte Carlo model of risk associated with ingesting IWS tap water in India.

Pathogen	10th Percentile	Median (95% CI)	Mean (95% CI)	90th Percentile
<i>Giardia</i>	6.62×10^{-5}	7.24×10^{-4} ($6.33 \times 10^{-4} - 8.19 \times 10^{-4}$)	3.90×10^{-3} ($3.07\text{E-}3 - 4.91\text{E-}3$)	8.06×10^{-3}
<i>Cryptosporidium</i>	3.83×10^{-4}	9.48×10^{-3} ($7.88 \times 10^{-3} - 1.17 \times 10^{-2}$)	7.32×10^{-2} ($6.72 \times 10^{-2} - 8.50 \times 10^{-2}$)	2.04×10^{-1}
norovirus High Risk	1.48×10^{-1}	5.91×10^{-1} ($5.87 \times 10^{-1} - 5.95 \times 10^{-1}$)	5.05×10^{-1} ($4.92 \times 10^{-1} - 5.18 \times 10^{-1}$)	6.38×10^{-1}
norovirus Low Risk	5.43×10^{-4}	1.83×10^{-2} ($1.45 \times 10^{-2} - 2.17 \times 10^{-2}$)	4.89×10^{-2} ($4.49 \times 10^{-2} - 5.30 \times 10^{-2}$)	1.48×10^{-1}
adenovirus	1.71×10^{-3}	8.71×10^{-3} ($7.76 \times 10^{-3} - 9.56 \times 10^{-3}$)	2.12×10^{-2} ($1.91 \times 10^{-2} - 2.34 \times 10^{-2}$)	5.25×10^{-2}
ETEC (diarrhea)	2.52×10^{-9}	1.34×10^{-6} ($9.00 \times 10^{-7} - 1.84 \times 10^{-6}$)	4.59×10^{-3} ($2.27 \times 10^{-3} - 7.55 \times 10^{-3}$)	6.62×10^{-4}
<i>Shigella</i>	1.26×10^{-5}	3.20×10^{-3} ($2.23 \times 10^{-3} - 4.46 \times 10^{-3}$)	7.97×10^{-2} ($6.61 \times 10^{-2} - 9.14 \times 10^{-2}$)	3.01×10^{-1}

Sensitivity Analysis

For each model we assessed the sensitivity of the estimated risk of infection to each stochastic input parameter by means of the rank order correlation. The rank order correlations we observed during the Monte Carlo simulation for each pathogen and dose-response model are displayed in Figures E16 to E22. For the protozoan pathogens *Giardia* and *Cryptosporidium* the estimated risks were strongly correlated with the dose-response parameter k (0.78 for Crypto; 0.16 for *Giardia*), the gene target concentration as measured by ddPCR (0.74 for *Giardia*; 0.33 for Crypto), DEUF sample volume (0.46 for *Giardia*; 0.34 for Crypto), and the viability models (0.30 for *Giardia*; 0.26 for Crypto). In the high-risk model (Figure E18), the estimated risk of infection with norovirus was most sensitive to the gene target ddPCR concentration (0.66) followed by the proportion of the population that is susceptible (0.47), the DEUF sample volume (0.21) and the daily volume of

drinking water ingested (0.12). As shown in Figure E19, in the low-risk norovirus model, the estimated risk remained strongly correlated with the gene target ddPCR concentration (0.92), DEUF sample volume (0.29) and the daily drinking water consumption (0.15). The estimated risk of infection for adenovirus was most sensitive to DEUF volume (0.63), gene target ddPCR concentration (0.62), and drinking water ingestion (0.29). For bacterial pathogens ETEC and *Shigella*, both estimated risks were sensitive to the ybbW concentration by ddPCR (-0.59 ETEC; -0.67 *Shigella*), the *E. coli* count by membrane filtration (0.56 ETEC; 0.62 *Shigella*), and the measured concentration of virulence genes by ddPCR (0.40 ETEC; 0.26 *Shigella*). However, they differed in their sensitivity to the approximate beta-Poisson dose response parameters with ETEC most sensitive to the median infectious dose (-0.28) and *Shigella* most sensitive to the parameter alpha (-0.08).

Annual Probability of Infection

Given the substantial daily risks of infection we observed for each pathogen, it is not surprising that the annualized risks of infection, as calculated by equation 21 and shown in Table 11, approach certainty at even the lowest percentiles. Based on this estimate everyone exposed to IWS should experience an infection with *Giardia*, *Cryptosporidium*, norovirus, adenovirus, ETEC, and *Shigella* in a year. Given the extremity of this estimation, we also calculated the annual probability of infection given a constant daily probability of infection fixed at the 10th percentile, median, mean, and 90th percentile as shown in Table 12. Even with the daily probability of infection fixed at the 10th percentile, the annual risk of infection for every pathogen except ETEC exceeds the US EPA acceptable risk threshold of 1 in 10,000 by at least an order of magnitude.

TABLE 11: The annual risk of infection for each pathogen attributable to IWS as estimated by subsampling of the daily probability of infection.

Pathogen	10th Percentile (%)	Median (%) (95% CI)	Mean (%) (95% CI)	90th (%) Percentile
<i>Giardia</i>	68.3	77.2 (76.2 – 78.1)	76.9 (75.9 – 77.8)	86.0
<i>Cryptosporidium</i>	100	100	100	100
norovirus (high-risk)	100	100	100	100
norovirus (low-risk)	100	100	100	100
adenovirus	99.9	100	100	100
ETEC (Prob diarrhea)	46	81.9 (77.1 – 87.0)	76.2 (73.3 – 79.1)	99.4
<i>Shigella</i>	100	100	100	100

TABLE 12: The annual probability of infection for each pathogen attributable to IWS in India as calculated with the daily probability of infection fixed at the denoted value.

Pathogen	Daily P fixed at 10th Percentile (%)	Daily P fixed at Median (%)	Daily P fixed at Mean (%)	Daily P fixed at 90th Percentile
<i>Giardia</i>	2.39	23.2	76.0	94.8
<i>Cryptosporidium</i>	13.0	96.9	100	100
norovirus High Risk	100	100	100	100
norovirus Low Risk	18.0	99.9	100	100
adenovirus	46.5	95.9	100	100
ETEC (Prob diarrhea)	9.2×10^{-5}	0.05	8.13	21.5
<i>Shigella</i>	0.46	69.0	100	100

Comparison with Previous Risk Assessment of IWS

We previously estimated the risk of infection and waterborne disease via IWS using *E. coli* counts pooled from several different IWSs and the observed pathogen to *E. coli* ratios in sewage (97). The median daily probabilities of infection based on this estimate were approximately 10^{-5} for *Campylobacter*, 10^{-7} for *Cryptosporidium*, and 10^{-6} for rotavirus.

Within microbial classes we compare the current median risk of infection for bacteria (ETEC and *Shigella*) at 10^{-6} to 10^{-3} , for protozoa (Crypto. and *Giardia*) at 10^{-3} to 10^{-4} , and viruses (adenovirus and norovirus) at 10^{-3} to 10^{-1} . Our current risk assessment consistently places the median daily risk of infection several orders of magnitude higher than the previous across each microbial class. For protozoa, the 10th percentiles of the daily risk of infection in the current estimate are equivalent to the 65th to 85th percentiles of the previously estimated risk of infection for *Cryptosporidium*. The 10th percentiles of risk for viruses in the current estimate (norovirus low-risk and adenovirus) are equivalent to the 60th and 70th percentiles estimated for rotavirus in the previous assessment. Lastly, the 10th percentiles of risk associated with bacteria in the current estimate are equivalent to the 20th and 45th percentiles of risk estimated for *Campylobacter*.

These large differences in percentiles could be the result of the water quality data used for the current assessment truly being worse than the water quality observed in the previous assessment; however, the median *E. coli* count in the current model 3.05 CFU/100 mL is equivalent to the 58th percentile of *E. coli* counts in the previous model and compares reasonably well with its median value of 1.58 CFU/100 mL. If we populate point estimates from the previous model using the median *E. coli* count for the current assessment, the estimated daily risks of infection are the same order of magnitude for each reference pathogen except rotavirus, where the estimated risk is one order of magnitude higher. Therefore, the discrepancy between the two models is not explained by microbial water quality as measured by culturable *E. coli*.

Infections Attributable to IWS in India

Based on the equivalence of the 10th percentile of risk for bacterial pathogens in the current assessment and the 20th to 45th percentiles of risk in the previous assessment, we adopt the annual infection of risks based on the daily infection risk fixed at the 10th percentile for estimating a plausible number of infections among the population of India. We consider this a reasonable approach because both the previous and current estimate make use of *E. coli* by culture method to estimate the number of viable bacterial pathogens. Additionally, the median infectious doses of *Shigella* and *Campylobacter* are roughly equivalent (*Campylobacter*: 1690; *Shigella*: 1479) making the equivalence of the 10th percentile for *Shigella* in the current estimate and the 45th percentile for *Campylobacter* in the previous comparable. Based on the annual risk of infection from the current model, IWS could account for up to 11 million *Giardia* infections, 60 million *Cryptosporidium* infections, 83 million norovirus infections, 214 million adenovirus infections, and 2.17 million *Shigella* infections among the 460 million urban dwellers in India. However, there are important uncertainties and limitations to consider with these estimates.

Limitations and Uncertainties

Despite our best efforts to implement a sophisticated risk assessment model populated with data obtained using advanced environmental microbiology sampling techniques, the improbability of the resulting risk estimates emphasizes the limitations of such techniques.

First, even when using advanced techniques such as DEUF to concentrate microbes from drinking water, most measures are censored at concentrations that represent substantial levels of risk. For example, we estimate the 95%LOD for the ipaH gene assay to be 0.205 gc/uL. In the current model framework this would equate to a 3.1% chance of infection per

day, which is the 69th percentile of the daily risk of infection. Across all gene target 95% LODs the equivalent percentiles of daily risk range from the 55th (adenovirus) to the 75th (*Giardia*).

During the sensitivity analysis we found the estimated risk for each class of pathogen was sensitive to the volume of water sampled via DEUF. This is because greater volumes of water filtered equates to lower limit of detection in the drinking water matrix. While it may be possible to sample several thousand liters of water with this method, the limitations are fundamentally found in the assay of only microliters via ddPCR while exposure is considered in liters per day of ingestion (111,246). Because there are 1 million microliters in a liter, any concentration method must achieve at least a 1 to 500,000 concentration to assay the equivalent of one exposure in a single PCR well (assuming a 2uL sample addition). If we factor in limits of detection and even quantification the necessary concentration is even greater (147). The end result of these constraints are data sets such as the one we used to parameterize our model where more than 75% of measures are censored and measures above the LOD only characterize the highest levels of risk.

Next are the limitations of using molecular measures of genetic targets as proxies for viable and infectious microorganisms. While it may seem that such genes are less distal proxies than indicator organisms, the tradeoffs between the molecular versus indicator techniques are uncertain. The previous IWS risk assessment made use of *E. coli* counts and pathogen to indicator ratios observed in sewage where viability is explicitly considered in the measures. Whereas, in our current assessment we relied upon genetic measures and mathematical models or assumptions regarding viability.

Our current sensitivity analysis indicated that the estimated risk levels for every pathogen were sensitive to the concentration of target genes as measured by ddPCR. In the case of protozoan and bacterial pathogens, we attempted to convert genetic targets to infectious units using previously described quantitative relationships between genes and infectious units. In both cases the sensitivity analysis corroborated that estimated risks were sensitive to such parameters. In the model the mean *Cryptosporidium* oocyst viability is 40.4% and the mean *Giardia* cyst viability is 14.7%. These estimates are based on observations in surface water and may greatly overestimate the viability of oocysts and cysts following water treatment distribution (248,269).

In the case of viral pathogens, we assumed that each genetic copy detected represented an infectious viral particle. Given the low concentration of free chlorine observed at IWS taps (59.8% of samples less than 0.2 mg/L) this assumption may be plausible for norovirus, which is more resistant to chlorine (225). But for adenovirus which is readily inactivated by free chlorine even at concentrations as low as 0.2 mg/L this assumption is more suspect. Given the certainty of infection with both viruses predicted by our model, it seems unlikely that every genetic element detected in distributed drinking water samples represents an infectious viral unit.

For bacterial pathogens, we incorporated estimates of viability using culture-based *E. coli* measures in combination with ddPCR measures of genetic targets. Both *Shigella* and pathogenic *E. coli* are readily inactivated by free chlorine, but they are also capable of entering a VBNC state due to environmental stress. We attempted to reflect this in our model through culture-based counts of *E. coli*. Although the annual risks of infection still

approach certainty, the use of both genetic and culture methods in tandem provides more plausibility to this estimate.

In addition to the limitations of environmental microbiology, there are the fundamental limitations of QMRA. At the center of every QMRA model are dose-response functions. Such functions are usually fit to dose-response trial data with limited dose levels and small sample sizes at each dose level (270). Further, these models are most uncertain at low doses which are most relevant to exposures via drinking water. The end result are dose-response models that are contentious such as norovirus (257,258). In our model we incorporated a high-risk model assuming no aggregation of viral particles (1 viral particle per dose unit), which is expected to accurately reflect norovirus behavior in environmental compartments. On the other hand, the aggregation parameter value suggested by MLE on the dose-response trial data is 1,106 viral particles per dose unit. In the case of the adenovirus model we employed, it was parameterized based on pooled data from various experiments and included data from an oral ingestion experiment of adenovirus 4 and 7 where all exposures resulted in asymptomatic infection (261). These difficulties aside, dose-response models are also usually fit to trial data associated with health adults in high-income countries and likely greatly underestimate the risk for vulnerable populations including children under 5, the elderly, and the immunocompromised. Additionally, infection in these models is considered an independent event, but, for children and adults living in low-income settings, coinfection is likely and the effects of such on dose-response remain largely unknown.

Additional limitations in QMRA are associated with the exposure assessment. Some limitations are fundamental such as our current approach of modeling daily exposure in a single consumption of the entire day's volume of water rather than many small-volume

exposures in a single day. It is possible that ingesting water in more frequent smaller volumes could decrease the consumed dose for each event and thereby decrease the overall risk. In our model we adopted the single daily consumption as a risk-conservative approach.

The exposure assessment portion of our model features additional limitations specific to IWS. In our model we assessed the water quality at the water meter and therefore excluded the household water management activities necessitated by intermittent supply (271). These household level behaviors could increase the risk of infection via contamination at the household level or decrease it through the efficacious use of point-of-use treatment technologies (88,272). The net detriment or benefit of such behavior remains unknown.

Despite the limitations and uncertainties, our assessment highlights several important findings regarding the risk of infection associated with IWS in India. First, our model indicates that IWS could plausibly account for 11 million infections of *Giardia*, 60 million *Cryptosporidium* infections, and 2.17 million *Shigella* infections among those living with IWS in the cities of India. Further, the risk of infection associated with five out of six pathogens in the model exceeds the US EPA acceptable risk threshold of 1 in 10,000 by orders of magnitude. While the estimated risks of infection for the viral pathogens included in our model seem improbable, it is not unlikely that the risk associated with these microbial classes also exceeds the 1 in 10,000 threshold since they are each more resistant to chlorine disinfectant and environmental degradation than bacteria such as *Shigella*. Thus, endemic waterborne disease associated with IWS is likely a significant contributor to diarrheal disease in India.

CONCLUSION

During the Millennium Development Goal era from 2000 to 2015, 1.2 billion people were newly connected to a piped-on-premise water supply. Despite the improvements in accessibility and water quality associated with piped supplies, they can also become efficient transmitters of waterborne disease as during a *Cryptosporidium* outbreak in Milwaukee that sickened up to 403,000 (5). Piped water supplies are still subject to frequent fecal contamination as evidenced by the presence of indicator bacteria (273). The public health impacts of such contamination are corroborated by epidemiology trials that have estimated a 1.34-fold increase in the risk of gastrointestinal illness among people consuming tap water from deficient water distribution systems (14). Such trials have also found that interruptions in water supplies that are normally operated continuously are associated with a 3.26-fold increase in GII (14).

In many settings piped water supplies are intentionally interrupted to manage scarcity of water or other resources in a delivery mode known as intermittent water supply (IWS) (15). IWS is reported in 43% of the countries represented in the International Benchmarking Network (17). Studies of water quality in supplies operated intermittently indicate that such supplies are often contaminated with various fecal indicators, but that such contamination is not homogenous, and contamination levels can vary both between and within IWSs. Epidemiology trials indicate that IWS is associated with the transmission of typhoid fever, cholera, and diarrhea and gut inflammation among children in certain contexts (28,29,180–182). But the largest longitudinal trial comparing diarrheal disease rates among users of

IWS versus CWS in India found no association between exposure to IWS and diarrhea (29).

Given this heterogeneity, quantitative microbial risk assessment offers another strategy for quantifying the public health risks of IWS (30). Such risk assessments rely on robust environmental microbiology data to characterize human exposures to pathogens and estimate the subsequent risk of infection and/or diarrheal disease.

We first used QMRA to estimate the global diarrheal burden of disease associated with IWS. We used a large dataset of *E. coli* measures in samples collected from IWS networks and pathogen to *E. coli* ratios in sewage to estimate daily and annual probabilities of infection associated with three reference pathogens – *Campylobacter*, *Cryptosporidium*, and rotavirus. The median daily risks of infection for each was estimated to be 4.25×10^{-5} , 1.98×10^{-7} , and 8.47×10^{-6} , respectively. The annual risks of infection for each pathogen exceeds the US EPA standard of 1 in 10,000. Based on the annual risks of infection with the three reference pathogens considered and WHO DALY weightings, IWS could account for 17.2 million infections globally which cause 4.52 million cases of diarrhea, 109,000 DALYs, and 1,560 deaths.

During our first IWS QMRA, we found that our estimated risk was sensitive to the assumed ratios of pathogens to *E. coli* and the *E. coli* counts from IWS networks. We used this approach to formulate an initial estimate because robust datasets pertinent to waterborne pathogens in IWSs were not available. In response to this lack of data, we developed and field tested a method to quantify gene targets associated with waterborne pathogens using

dead-end ultrafiltration (DEUF), PEG precipitation and ultracentrifugation, and droplet digital PCR (111,112,118).

During a cross-sectional sampling event in May of 2017 we collected 23 DEUF samples and paired grab samples from an IWS in Jaipur, India. We adapted qPCR assays targeting genes associated with enterotoxigenic *E. coli*, *Shigella* spp., *Cryptosporidium*, *Giardia*, norovirus GI and GII, and adenovirus A-F to ddPCR format and determined the 95% LOD for each one (147). We assayed DEUF concentrates from samples collected in Jaipur and estimated the concentration of genetic targets using both manual and Umbrella thresholding. Our results indicate that in Jaipur, groundwater pumped directly into the water distribution network without disinfection likely poses a significant risk of infection as we detected genes associated with *Cryptosporidium*, *Giardia*, and ETEC in samples from tube wells that were also positive for culturable *E. coli*. During our study we also found that DEUF and ddPCR combined could detect gene targets associated with waterborne pathogens with increased sensitivity compared to culture based methods; however, the limits of detection were still such that detecting pathogens at concentrations relevant to risk-based standards requires filtering thousands to hundreds of thousands of liters of drinking water. These findings indicate that risk assessments of distributed drinking water will likely continue to rely on culture-based and molecular-based methods to make robust estimates of risk.

We also leveraged our DEUF and ddPCR workflow in a natural experiment to compare microbial water quality between IWS and CWS as the city of Nagpur, India works to transition its entire water supply to CWS. During sampling in 2015 and 2017 we collected grab samples and DEUF samples totaling a volume of 6,925 liters from household taps

served by IWS and CWS. We measured physicochemical water quality parameters, fecal indicator bacteria by culture-based methods, and gene targets associated with six fecal-oral pathogens by ddPCR. The fecal indicator and physicochemical data suggest increased degradation of water quality between the storage reservoir and households taps in IWS versus CWS zones; however, these differences are not statistically significant. We did observe a statistically significant greater proportion of grab samples positive for *E. coli* from household taps in IWS zones compared to CWS zones. We also detected genes associated with ETEC, *Shigella*, norovirus GI and GII, adenovirus, *Cryptosporidium*, and *Giardia* at increase prevalence in samples from IWS taps compared to CWS taps. Despite the sample size limitations of our study, the results suggest that in the Indian context IWS is associated with increased likelihood of exposure to waterborne pathogens compared to CWS.

Using the microbial data we collected in Nagpur, we formulated a QMRA to estimate the health risks associated with IWS using molecular biology data as the basis for exposure assessment. We used Monte Carlo simulation with 22 stochastic variables and 21 equations to quantify human exposure to infectious pathogens via IWS tap water. Our model estimates substantial risks of infection from even a single exposure to drinking water delivered by an IWS in India. The daily 10th percentile of infection risk for *Giardia*, *Cryptosporidium*, norovirus, adenovirus, and *Shigella* exceeds the US EPA annual risk threshold of 1 in 10,000. Given this tremendous risk level, we estimated the annual probability of infection assuming the daily infection risk was fixed at the 10th percentile for each pathogen. Based on this risk level and dose-harmonization approaches to estimate viability, IWS could account for up to 11 million *Giardia* infections, 60 million

Cryptosporidium infections, and 2.17 million *Shigella* infections annually among the 460 million Indians consuming water from an IWS. These risks are several orders of magnitude greater than those estimated via fecal indicator counts and pathogen to indicator ratios. Despite the uncertainties of translating gene copy numbers to viable infectious units, our results add further evidence that the risk of infection associated with IWS exceeds the 1 in 10,000 US EPA standard.

The results of our work indicate that operating piped water supplies intermittently causes an increase in the risk of exposure to fecally contaminated water and fecal-oral pathogens. QMRA indicates that this increased exposure leads to public health risks that exceed both US EPA and WHO normative thresholds. While it would be interesting to conduct a QMRA comparing IWS and CWS directly, neither our current study in Nagpur nor the previous study in Hubli-Dharwad detected *E. coli* frequently enough to allow for parameterization of a risk assessment model. At taps served by CWS in Nagpur, we failed to detect a single culturable *E. coli* in over 1,000 liters of filtered drinking water.

Importantly, our work also highlights a path forward for reducing waterborne disease associated with drinking water in low- and middle-income countries where the number of people exposed to IWS is likely to increase due to urbanization, climate change driven water scarcity, and deferred investment in infrastructure (198). IWS in these settings is driven by causes that are rooted in technical engineering, institutional administration, and human behavior (15). In Nagpur, systematic improvements in each of these areas allowed for infrastructure improvements made transitioning to CWS financially and technically feasible which subsequently improved microbial water quality. First metering and billing compliance were improved to increase revenues. Next, these revenues were used to repair

and replace the piped distribution networks. Then and only then could CWS be implemented and water quality improved. This chain of events outlines a path forward for other utilities attempting to transition from IWS to CWS and our microbiological findings indicate that the improvements in water quality are significant with direct relevance to waterborne pathogens. Since these institutional and infrastructural improvements require long timelines, in the near term two strategies can be implemented to improve microbiological water quality. First, sources of fecal pollution can be reduced through targeted improvements of sanitation infrastructure such as failing sewer lines or pit latrines in close proximity to waterlines. Second, water quality can be improved via household water treatment at the point-of-use through the design and implementation of technologies that offer efficacious treatment through convenient to use devices.

Finally, our work surrounding IWS indicates that infrastructure-centric metrics such as “improved” versus “unimproved” will not suffice to increase access to safe drinking water. The Sustainable Development Goals recognize this by defining the highest level of access to drinking water as “safely managed drinking water services” which is defined as “located on premises, available when needed, and free from contamination” (274,275). Our results indicate that IWSs are unlikely to meet these standards. Further, our findings indicate that measuring microbes at concentrations relevant to risk-based standards will remain difficult if not impossible in most settings. In the near term, measuring safety in drinking water supplies will continue to rely on inference from culture-based evidence with large uncertainties and unclear relevance, especially given the current definitions of acceptable risk.

APPENDIX A

TABLE A1: A summary of observed proportions of samples positive and summary statistics for measures of various fecal indicators in IWS networks throughout the globe.

Study	Location	FIB	Sample Size	PP	Reported Statistics*
(27)	Trujillo, Peru	FC TC	21 21	NR	Geometric Mean: 4.0 Min: 0 Max: 75 (CFU/100mL) Geometric Mean: 9.0 Min: 0 Max: 1700 (CFU/100mL)
(170)	Karachi, Pakistan	FC TC	40 276**	15% 81%	NR (n=20) Min: 3 Max: 1100 (MPN/100 mL)
(26)	Dushanbe, Tajikistan	FC	NR	97%	Mean: 175 Min: 0 Max: 400 (CFU/100mL)
(173)	Beruit, Lebanon	HPC	12	NR	Mean: 124 Min: 17 Max: 513 (CFU/100mL)
(23)	Amman, Jordan Nablus, Palestine	HPC HPC	NR NR	NR NR	Mean: 40 (CFU/100mL) Mean: 2300 (CFU/100mL)
(276)	Beruit, Lebanon Beruit, Lebanon	FC TC	91 91	4% 10%	NR NR
(179)	Khan Yunis Governorate, Gaza Strip	FC TC	NR NR	12% 25%	NR NR
(196)	Ghaziabad, India Jaipur, India Nagpur, India Panaji, India	FC FC FC FC	NR NR NR NR	76% 27% 63% 40%	NR NR NR NR
(277)	Hyderabad, India	TTC EC	51 51	4% 2%	Mean: 880 (MPN/100mL) NR
(88)	Nagpur, India	TTC	NR	NR	Histogram by risk categories
(24)	Hubli-Dharwad, India	EC TC	589 589	32% 65%	Median: 0.5 Min: 0.5 Max: >2420 (MPN/100mL) Median: 18.5 Min: 0.5 Max: >2420 (MPN/100mL)
(74)	Da Nang Province, Vietnam	EC	553	NR	Geometric mean: 16
(89)	Nagpur, India	FS TTC	NR NR	NR NR	Mean: 20 Max: 290 (CFU/100mL) Mean: 220 Max: 610 (CFU/100mL)
(278)	Ghana, Tamale	EC TC	NR NR	NR NR	Histogram by risk categories Histogram by risk categories
(75)	Kandal Province, Cambodia	EC	142	NR	Arithmetic Mean: 520 Geometric Mean: 120 Median: 10 (CFU/100mL) COV: 5.7 Histogram by risk categories

TABLE A1: Continued

(169)	Maputo, Mozambique	TC	60	22%	NR
(82)	Arraijan, Panama	TC	423	0.9%	NR
		EC	423	0.2%	NR

FC – fecal coliform; TC – total coliform; HPC – heterotrophic plate count; EC – *E. coli*; FS – fecal streptococci; TTC – thermotolerant coliform; NR – not reported; PP – proportion positive

*As reported in the literature

**Multiple bodies performed sampling during the same study with one reporting statistics for TC (n=20) and another reporting proportion positive (n=276)

Bold datasets were used to develop the probability density function for *E. coli* in the QMRA Monte Carlo simulation in Chapter 2.

TABLE A2: A summary of epidemiological evidence linking IWS and diarrheal disease.

Study	Location	Design	Population	Exposure/Control	Outcome	Measure of effect
(176)	Mexico	CS	Children under-5 (n=732)	Consumption of water from IWS/Consumption of water from CWS	Caretaker reported diarrhea	aROR = 2.00 (95% CI: 1.16 – 3.70) (logistic regression)
(177)	Palestine	CS	Households (n=1625)	Consumption of water from IWS/Consumption of water from CWS	Self-reported diarrhea	ROR = 1.53 (95% CI: 1.15 – 2.03)
(178)	Palestine	CS	Gaza residents (n=141)	Consumption of water from greater than 1-day intermittency/Consumption of water from 1 day or less intermittency	Self-reported diarrhea	RR = 1.33 (95% CI: 0.92 – 1.91)
(179)	Palestine	CS	Khan Yunis Governorate residents (n=200)	Consumption of water from greater than 1-day intermittency/Consumption of water from 1 day or less intermittency	Self-reported diarrhea	RR = 1.49 (95% CI: 1.06 – 2.09)
(28)	Democratic Republic of the Congo	E	Patients admitted to Cholera Treatment Center (n=5,745)	Incidence of suspected cholera cases following a day without tap water supply/ Incidence of suspected cholera cases following a day with no supply interruption	Suspected case of cholera as indicated by acute watery diarrhea	IRR = 2.55 (95% CI: 1.54 – 4.24)
(29)	India	LC	Households with children under 5 (n=3,922)	Consumption of water from IWS/ Consumption of water from CWS	(1) diarrhea (2) typhoid fever	(1) PR = 1.08** (95% CI: 0.96 – 1.20) (2) CIR = 1.72** (95% CI: 1.28 – 2.44)

*CS – Cross-section; LC – longitudinal cohort; E – ecological

**Calculated from inverse of reported PR and CIR for comparison to other studies.

APPENDIX B

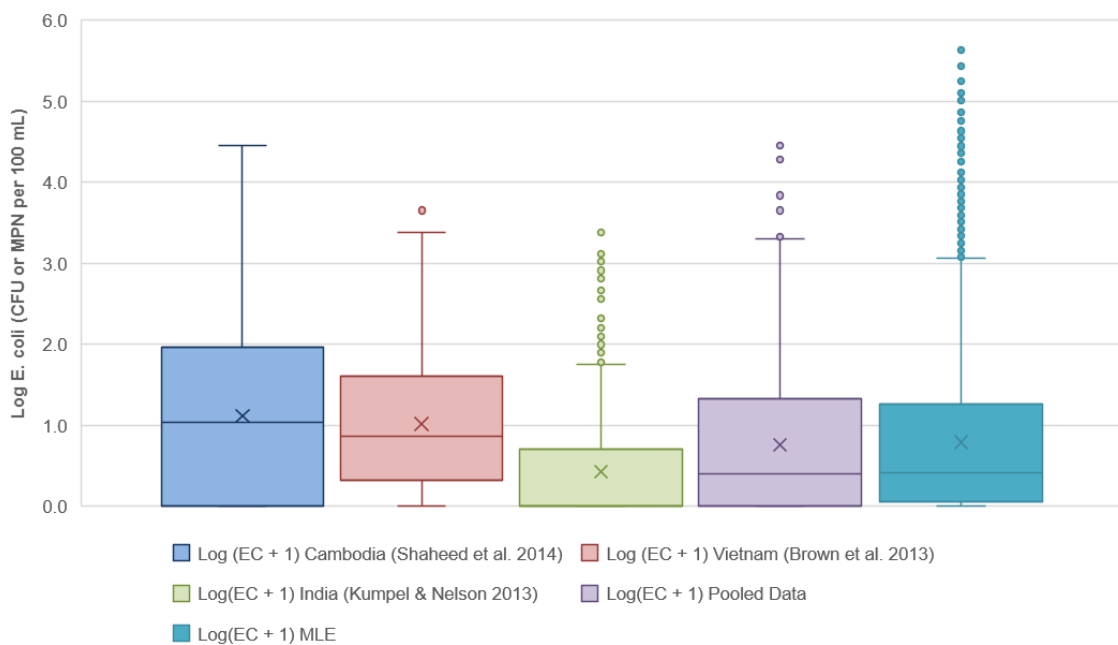


FIGURE B1: Boxplots of the field observed log *E. coli* counts and the modeled log *E. coli* counts from the normal distribution as parameterized using maximum likelihood estimation (MLE).

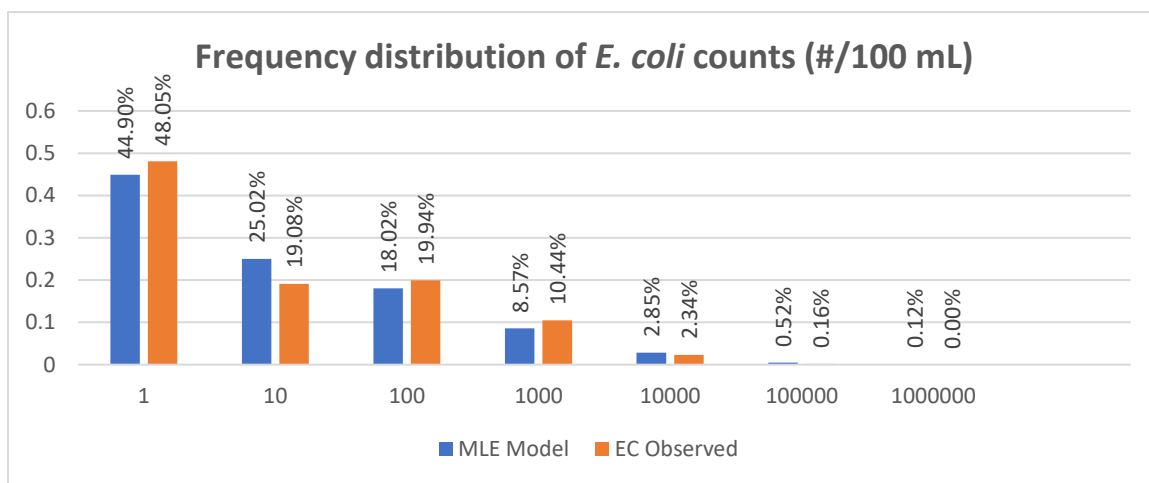


FIGURE B2: Frequency distributions of field-observed and maximum likelihood estimated *E. coli* counts at taps in an IWS.

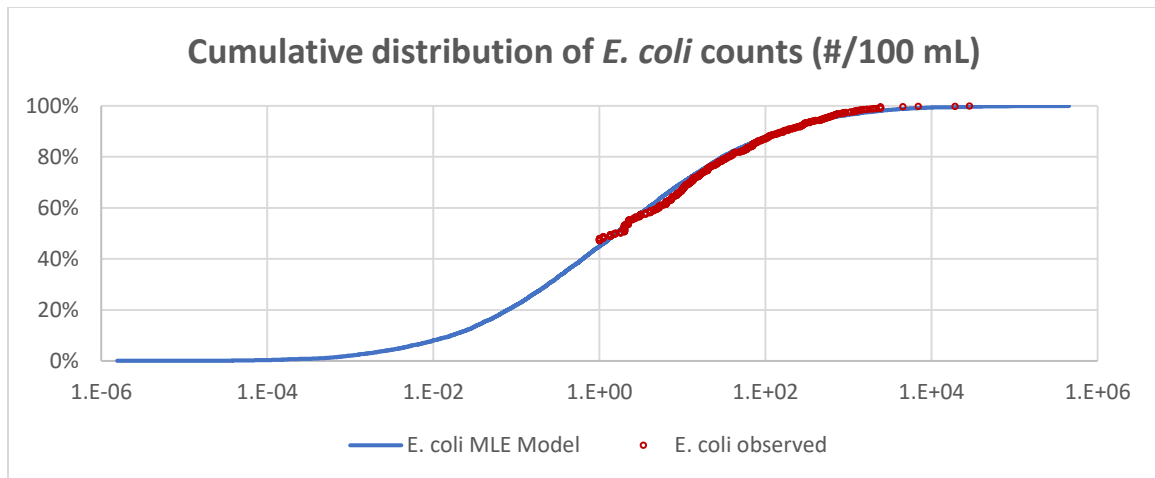


FIGURE B3: Cumulative distribution of *E. coli* counts in an IWS as observed in the field and modeled using a lognormal distribution parameterized using maximum likelihood estimation (MLE).

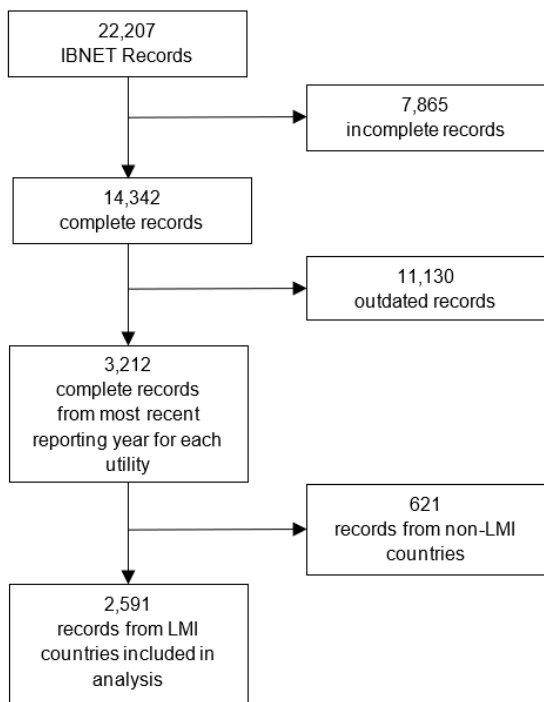


FIGURE B4: Record screening flow chart for IBNET records used to estimate the global population served by IWS.

TABLE B1: Point estimates of the daily and annual probability of infection, and annual disease burden assuming the median *E. coli* concentration and raw sewage as the source of fecal contamination in an IWS. These calculations demonstrate the mathematical framing and calculations associated with the QMRA model. Annual disease burdens highlighted in bold exceed the normative burden of disease threshold of 10^{-6} DALYs per person per year.

STATE VARIABLES	<i>Cryptosporidium</i>	<i>Campylobacter</i>	Rotavirus
(A) Median <i>E. coli</i> concentration (CFU/100 mL) in IWS tap water	1.3	1.3	1.3
(B) Median pathogen to EC ratio in sewage	2.11E-07	5.79E-05	5.61E-07
(C) Pathogen concentration (N/mL) in IWS tap water (A*B)	2.80E-09	7.70E-07	7.46E-09
(D) Median tap water consumption (mL/day)	1500	1500	1500
(E) Pathogen dose (N) (C*D)	4.20E-06	1.16E-03	1.12E-05
(F) Probability of infection (daily)	2.40E-07	2.28E-05	6.65E-06
(G) Probability of infection (annual) (1-[1-F] ³⁶⁵)	8.77E-05	8.28E-03	2.42E-03
(H) Annual disease burden (DALY per person per year)	9.21E-08	1.14E-05	7.60E-05

TABLE B2: Point estimates of the daily and annual probability of infection, and annual disease burden assuming the mean *E. coli* concentration and raw sewage as the source of fecal contamination in an IWS. Annual disease burdens highlighted in bold exceed the normative burden of disease threshold of 10^{-6} DALYs per person per year.

STATE VARIABLES	<i>Cryptosporidium</i>	<i>Campylobacter</i>	Rotavirus
(A) Mean <i>E. coli</i> concentration (CFU/100 mL) in IWS tap water	404.0	404.0	404.0
(B) Median pathogen to EC ratio in sewage	2.11E-07	5.79E-05	5.61E-07
(C) Pathogen concentration (N/mL) in IWS tap water (A*B)	8.51E-07	2.34E-04	2.27E-06
(D) Median tap water consumption (mL/day)	1500	1500	1500
(E) Pathogen dose (N) (C*D)	1.28E-03	3.51E-01	3.40E-03
(F) Probability of infection (daily)	7.30E-05	6.74E-03	2.01E-03
(G) Probability of infection (annual) (1-[1-F] ³⁶⁵)	2.63E-02	9.15E-01	5.20E-01
(H) Annual disease burden (DALY per person per year)	2.76E-05	1.26E-03	1.63E-02

TABLE B3: Point estimates of the daily and annual probability of infection, and annual disease burden assuming the median *E. coli* concentration and raw sewage as the source of fecal contamination in an IWS, as characterized in the Guidelines for Drinking Water Quality 4th Edition Table 7.6. Annual disease burdens highlighted in bold exceed the normative burden of disease threshold of 10^{-6} DALYs per person per year.

STATE VARIABLES	<i>Cryptosporidium</i>	<i>Campylobacter</i>	Rotavirus
(A) Median <i>E. coli</i> concentration (CFU/100 mL) in IWS tap water	1.3	1.3	1.3
(B) Mean pathogen to EC ratio in untreated wastewater	1.00E-06	1.00E-04	5.05E-07
(C) Pathogen concentration (N/mL) in IWS tap water (A*B)	1.33E-08	1.33E-06	6.72E-09
(D) Median tap water consumption (mL/day)	1500	1500	1500
(E) Pathogen dose (N) (C*D)	2.00E-05	2.00E-03	1.01E-05
(F) Probability of infection (daily)	1.14E-06	3.94E-05	5.99E-06
(G) Probability of infection (annual) (1-[1-F] ³⁶⁵)	4.16E-04	1.43E-02	2.18E-03
(H) Annual disease burden (DALY per person per year)	4.37E-07	1.97E-05	6.84E-05

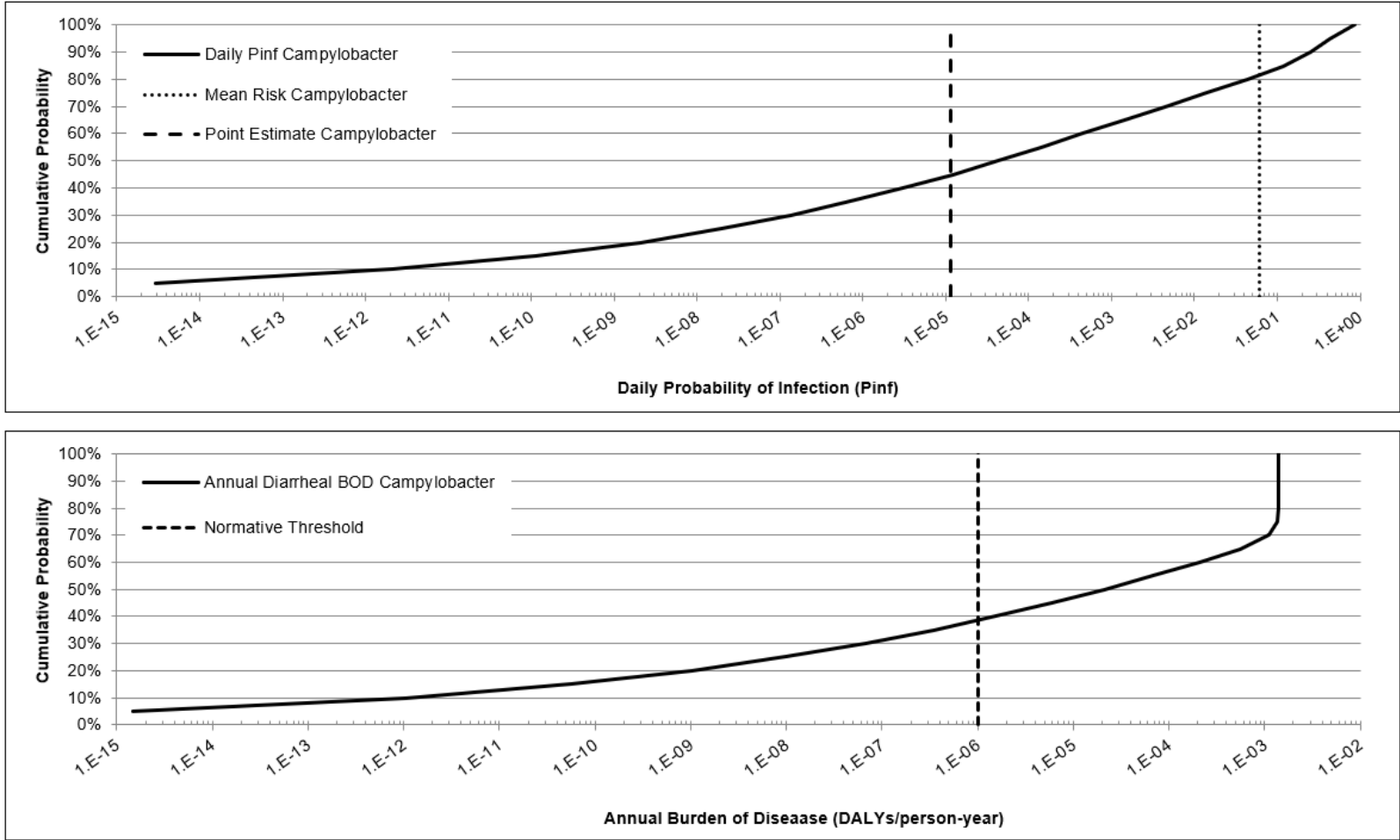


FIGURE B5: Cumulative distributions of the daily probability of infection with *Campylobacter* and the associated annual burden of disease assuming consumption of fecally contaminated tap water from an IWS.

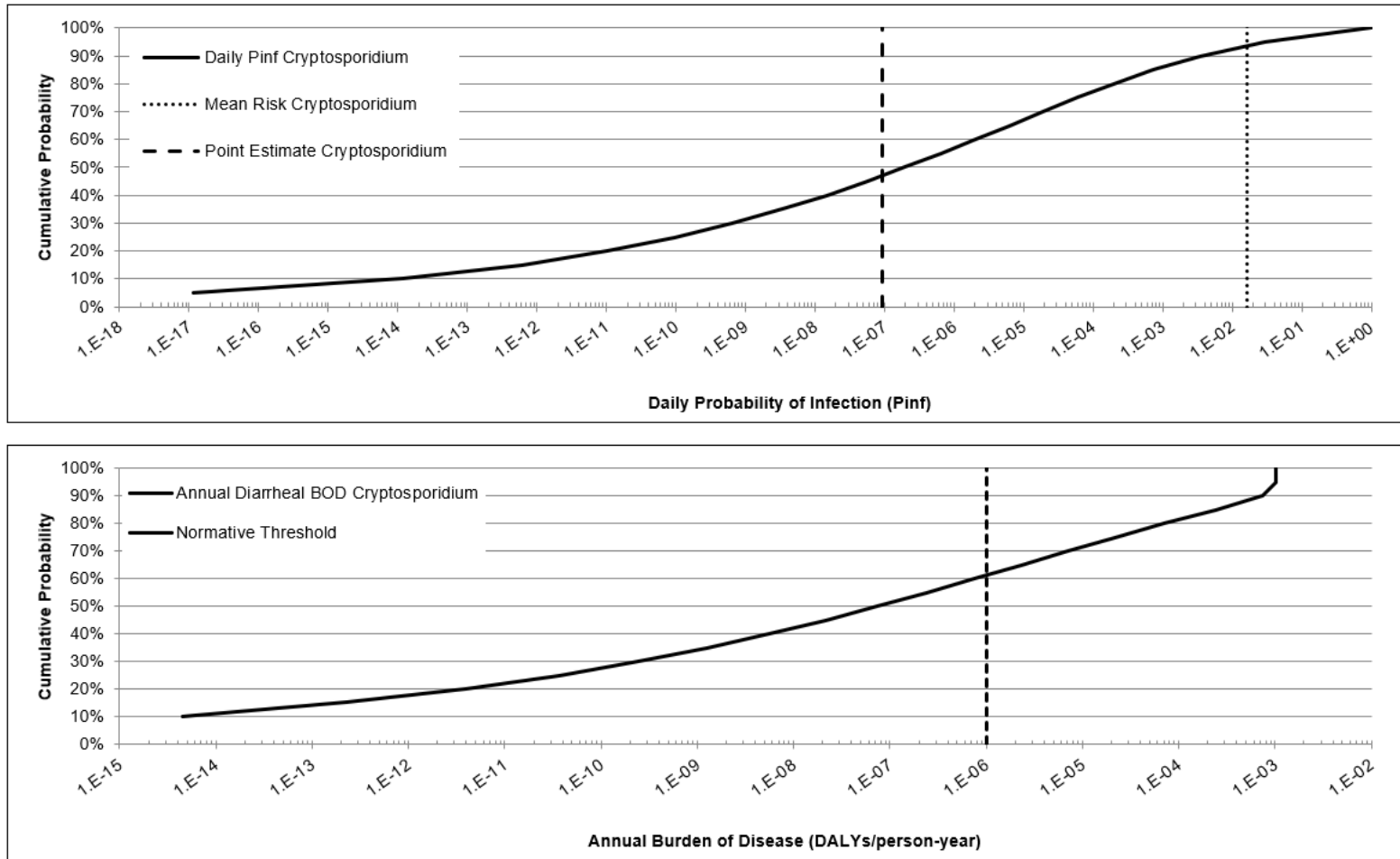


FIGURE B6: Cumulative distributions of the daily probability of infection with *Cryptosporidium* and the associated annual burden of disease assuming consumption of fecally contaminated tap water from an IWS.

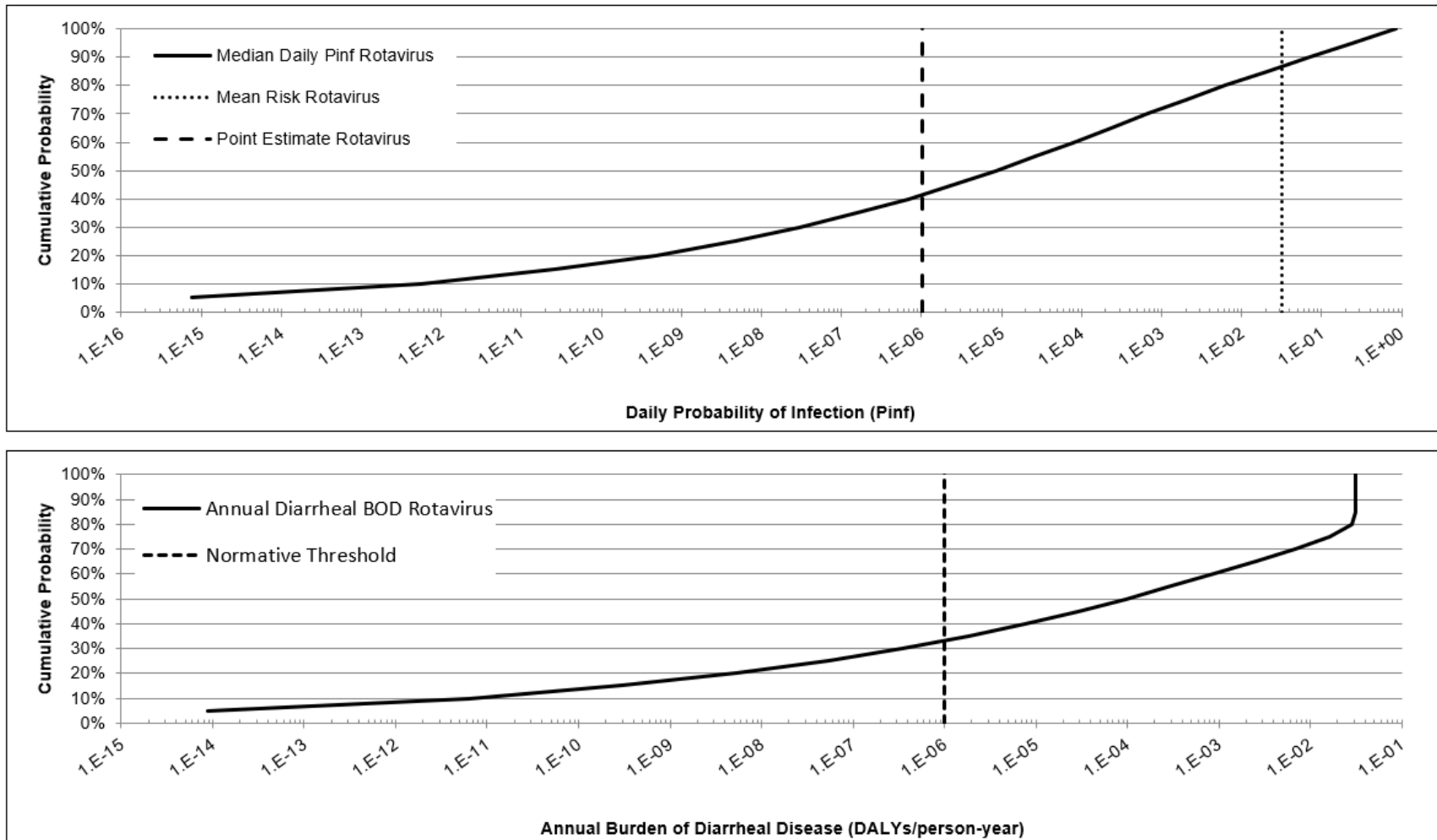


FIGURE B7: Cumulative distributions of the daily probability of infection with rotavirus and the associated annual burden of disease assuming consumption of fecally contaminated tap water from an IWS.

TABLE B4: Results of the tornado analysis and rank correlation to assess the sensitivity of the estimated risk of infection for *Campylobacter* to input parameters when considering consumption of fecally contaminated tap water by an IWS user.

INPUT VARIABLE	EXPLAINED VARIATION (%)	EXPLAINED VARIATION (Cumulative %)	RANK CORRELATION
<i>Campylobacter</i> to <i>E. coli</i> ratio	85.49	85.49	0.71
<i>E. coli</i> count in IWS tap water	8.52	94.01	0.50
<i>Campylobacter</i> alpha	5.40	99.41	-0.35
<i>Campylobacter</i> N50	0.58	99.99	-0.26
tap water consumption	0.01	100.00	0.06

TABLE B5: Results of the tornado analysis testing the sensitivity of the estimated risk of infection for *Cryptosporidium* to input parameters when considering consumption of fecally contaminated tap water by an IWS user.

INPUT VARIABLE	EXPLAINED VARIATION (%)	EXPLAINED VARIATION (Cumulative %)	RANK CORRELATION
<i>E. coli</i> in IWS tap water	45.86	45.86	0.59
<i>Cryptosporidium</i> to <i>E. coli</i> ratio	32.75	78.61	0.56
<i>Cryptosporidium</i> k	21.34	99.95	0.51
tap water consumption	0.05	100.00	0.04

TABLE B6: Results of the tornado analysis testing the sensitivity of the estimated risk of infection for rotavirus to input parameters when considering consumption of fecally contaminated tap water by an IWS user.

INPUT VARIABLE	EXPLAINED VARIATION (%)	EXPLAINED VARIATION (Cumulative %)	RANK CORRELATION
<i>E. coli</i> count in IWS tap water	81.42	81.42	0.75
rotavirus to <i>E. coli</i> ratio	9.79	91.21	0.43
rotavirus alpha	7.35	98.56	-0.37
rotavirus N50	1.35	99.91	-0.24
tap water consumption	0.09	100.00	0.07

TABLE B7: Estimate of the global population served by IWS by projecting IBNET reported percentages of IWS population onto JMP 2015 Update “piped on premise” water supply populations.

REGION	POPULATION SERVED BY IWS	PERCENTAGE OF TOTAL POPULATION SERVED BY IWS
Africa	116,000,000 95% CI: 106 – 126 million	12.5%
Americas, LMI	47,000,000 95% CI: 40 – 54 million	5.1%
Eastern Mediterranean, LMI	103,000,000 95% CI: 66 – 143 million	11.1%
Europe, LMI	71,000,000 95% CI: 63 – 80 million	7.7%
South-east Asia	409,000,000 95% CI: 369 – 447 million	44.2%
Western Pacific, LMI	179,000,000 95% CI: 130 – 235 million	19.4%
GLOBAL	925,000,000 95% CI: 670 – 1,130 million	

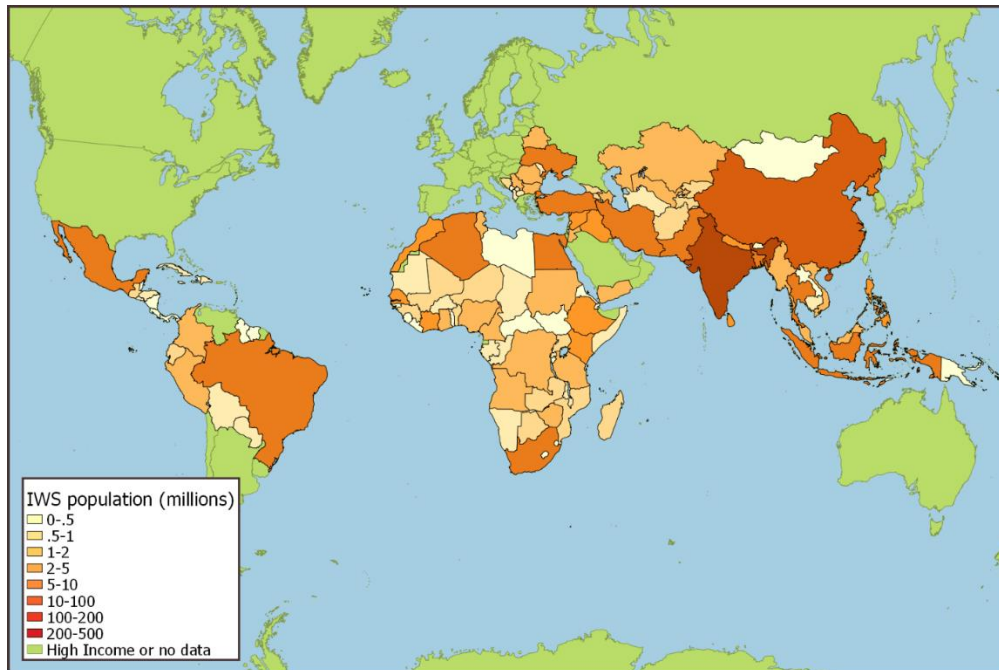


FIGURE B8: Map of the geographic range and estimated magnitude of the population served by IWS by country.

TABLE B8: The annual burden of diarrheal disease attributable to IWS based on the median daily probabilities of infection for *Campylobacter*, *Cryptosporidium* and rotavirus assuming consumption of fecally contaminated tap water from an IWS. The rotavirus burden is calculated using metrics for LMICs.

REGION	IWS POPULATION	ANNUAL INFECTIONS (Median P _{inf})				ANNUAL CASES OF DIARRHEA (Median P _{inf})				ANNUAL DEATHS (Median P _{inf})				ANNUAL DIARRHEAL DALYs (Median P _{inf})			
		CAMPY	CRYPTO	ROTA	ALL	CAMPY	CRYPTO	ROTA	ALL	CAMPY	CRYPTO	ROTA	ALL	CAMPY	CRYPTO	ROTA	ALL
Africa	116,000,000	1,789,771	8,383	358,068	2,160,000	536,931	5,868	23,274	566,000	56	0	140	196	2,470	9	11,218	13,700
Americas, LMI	47,000,000	725,166	3,397	145,079	874,000	217,550	2,378	9,430	229,000	23	0	57	79	1,001	3	4,545	5,550
Eastern Mediterranean, LMI	103,000,000	1,589,194	7,444	317,939	1,910,000	476,758	5,210	20,666	503,000	50	0	124	174	2,193	8	9,961	12,200
Europe, LMI	71,000,000	1,095,464	5,131	219,162	1,320,000	328,639	3,592	14,246	346,000	34	0	85	120	1,512	5	6,866	8,380
South East Asia	409,000,000	6,310,487	29,557	1,262,497	7,600,000	1,893,146	20,690	82,062	2,000,000	198	0	492	691	8,708	30	39,554	48,300
Western Pacific, LMI	179,000,000	2,761,803	12,936	552,535	3,330,000	828,541	9,055	35,915	874,000	87	0	215	302	3,811	13	17,311	21,100
Global	925,000,000	14,271,885	66,847	2,855,280	17,200,000	4,281,565	46,793	185,593	4,520,000	448	0	1,114	1,560	19,695	69	89,456	109,000

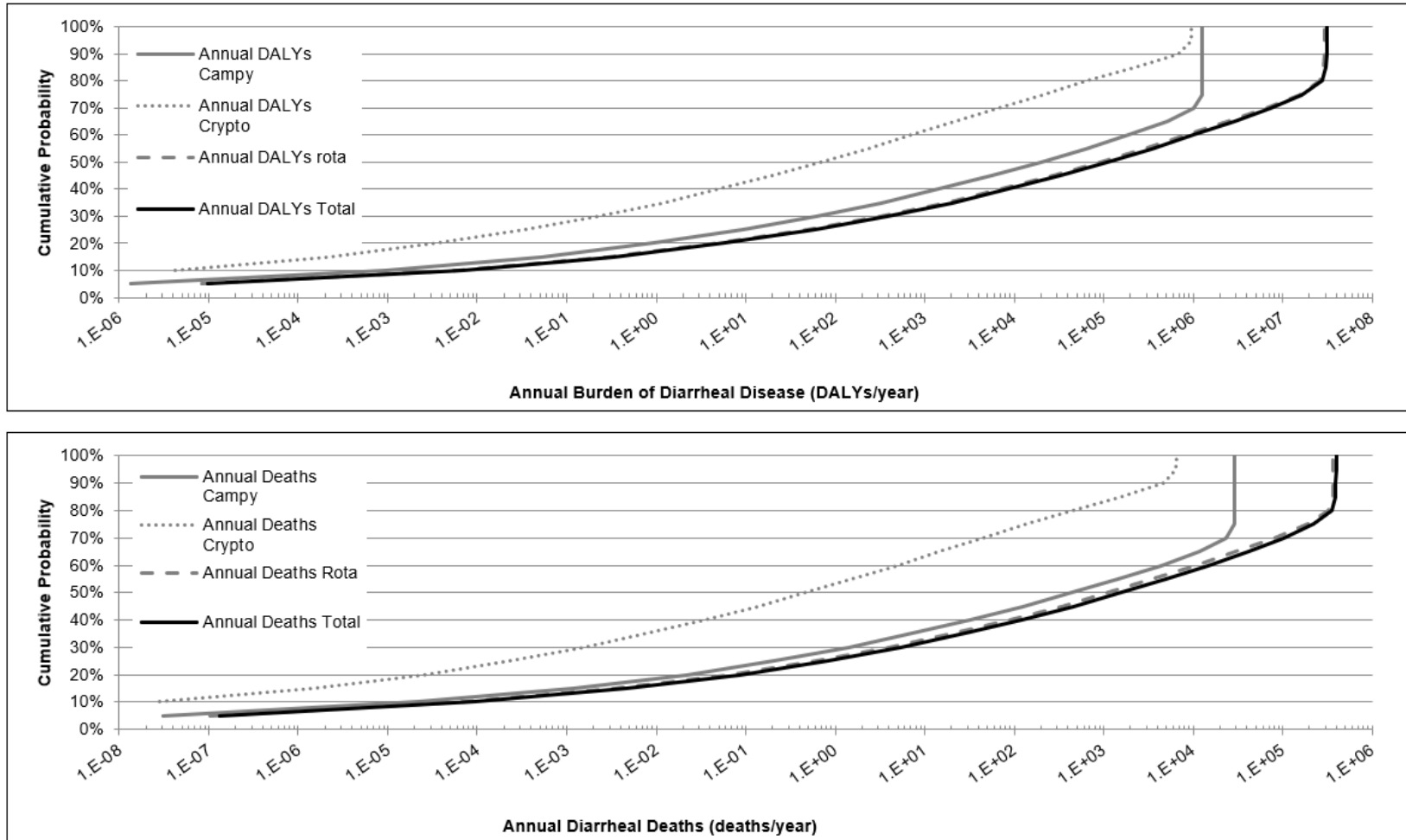


FIGURE B9: Cumulative distributions, by etiology and in total, of the annual diarrheal burden of disease and deaths attributable to consumption of fecally contaminated tap water among the 925 million global users of IWS.

TABLE B9: The annual burden of diarrheal disease attributable to IWS based on the median daily probabilities of infection for *Campylobacter*, *Cryptosporidium* and rotavirus assuming consumption of fecally contaminated tap water from an IWS. The rotavirus burden is calculated using metrics for high-income countries.

REGION	IWS POPULATION	ANNUAL INFECTIONS (Median P_{inf})				ANNUAL CASES OF DIARRHEA (Median P_{inf})				ANNUAL DEATHS (Median P_{inf})				ANNUAL DIARRHEAL DALYs (Median P_{inf})			
		CAMPY	CRYPTO	ROTA	ALL	CAMPY	CRYPTO	ROTA	ALL	CAMPY	CRYPTO	ROTA	ALL	CAMPY	CRYPTO	ROTA	ALL
Africa	116,000,000	1,789,771	8,383	358,068	2,160,000	536,931	5,868	10,742	554,000	56	0	2	58	2,470	9	153	2,630
Americas, LMI	47,000,000	725,166	3,397	145,079	874,000	217,550	2,378	4,352	224,000	23	0	1	23	1,001	3	62	1,070
Eastern Mediterranean, LMI	103,000,000	1,589,194	7,444	317,939	1,910,000	476,758	5,210	9,538	492,000	50	0	1	51	2,193	8	135	2,340
Europe, LMI	71,000,000	1,095,464	5,131	219,162	1,320,000	328,639	3,592	6,575	339,000	34	0	1	35	1,512	5	93	1,610
South East Asia	409,000,000	6,310,487	29,557	1,262,497	7,600,000	1,893,146	20,690	37,875	1,950,000	198	0	6	204	8,708	30	538	9,280
Western Pacific, LMI	179,000,000	2,761,803	12,936	552,535	3,330,000	828,541	9,055	16,576	854,000	87	0	2	89	3,811	13	235	4,060
Global	925,000,000	14,271,885	66,847	2,855,280	17,200,000	4,281,565	46,793	85,658	4,410,000	448	0	13	461	19,695	69	1,216	21,000

APPENDIX C

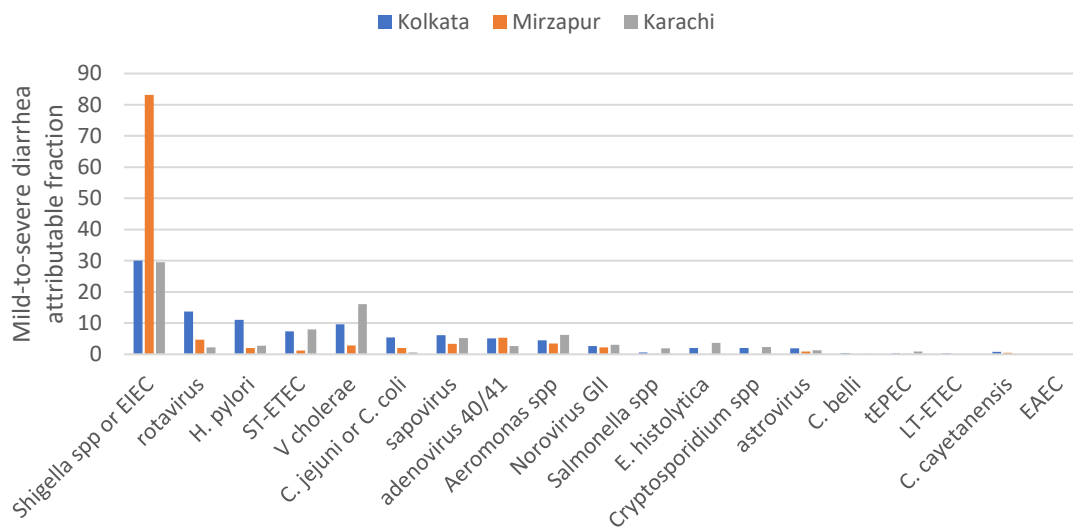


FIGURE C1: Attributable fraction of mild-to-severe diarrhea in 24- to 59-month olds as measured by molecular methods at three southeast Asian study sites during GEMS. Adapted from (131).

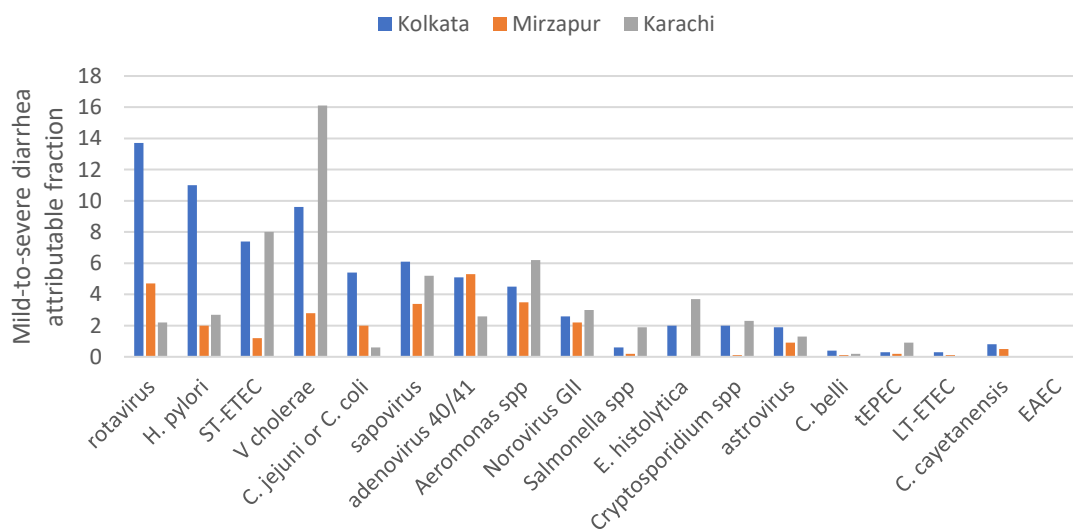


FIGURE C2: Attributable fraction of mild-to-severe diarrhea in 24- to 59-month olds as measured by molecular methods are three southeast Asian study sites during GEMS with *Shigella* omitted to allow closer examination of other etiological agents. Adapted from (131).

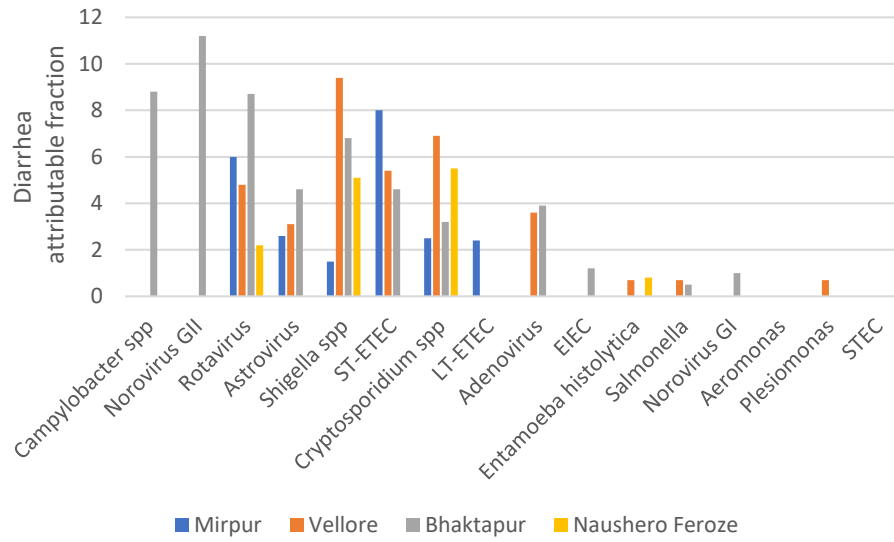


FIGURE C3: Attributable fraction of diarrhea in 12- to 24-month olds as measured by molecular methods are four southeast Asian study sites during MAL-ED. Adapted from (67).

TABLE C1. Primer and probe sequences for probe-based qPCR assays adapted to ddPCR and used as controls in interrogation of DEUF concentrate samples from drinking water.

MICROBE	GENE		SEQUENCE	REFERENCE
Hepatitis G (RT inhibition)	5' non-coding region	Forward Reverse Probe	CGGCCAAAAGGTGGTGGATG CGACGAGCCTGACGTCGGG AGGTCCCTCTGGCGCTTGTGGCGAG	(279)
Lambda gDNA (PCR inhibition control)	CAL1	Forward Reverse Probe	AGACGAATGCCAGGTCATCTGAAACAG CTTTTGCTCTGCGATGCTGATACCG CGTCAACGGCATCCACGAAGGCGACAGA	(280)
Bovine respiratory syncytial virus (Positive RNA extraction control)	beta-actin	Forward Reverse Probe	GCAATGCTGCAGGACTAGGTATAAT ACACTGTAATTGATGACCCATTCT ACCAAGACTTGTATGATGCTGCCAAAGCA**	(281)
Bovine herpes virus (Positive DNA extraction control)	glycoprotein B	Forward Reverse Probe	TGTGGACCTAAACCTCACGGT GTAGTCGAGCAGACCCGTGTC AGGACCGCGAGTTCTTGCCAC**	(282)

TABLE C2: Optimized thermal cycling conditions and 95% LODs for qPCR assays as adapted to ddPCR.

MICROBE	GENE	CONTROL MATERIAL	REPORTER	AMPLICON LENGTH	THERMAL CYCLING	95% LOD (gc/reaction)
adenovirus A-F	hexon	ultramer DNA	FAM	96	95 C 10 min; 95 C 30s, 58.7 C 1 min (40x); 98 C 10 min; 4 C hold	1.5
norovirus GI	NoV GI ORF1-2	ultramer DNA	HEX	96	95 C 10 min; 95 C 30s; 54.6 C 1 min (40x); 98 C 10 min; 4 C hold	5.7
norovirus GII	NoV GII ORF1-2	ultramer DNA	FAM	98	95 C 10 min; 95 C 30s, 54.6 C 1 min (40x); 98 C 10 min; 4 C hold	5.9
<i>Shigella</i> /EIEC	ipaH	gBlock DNA	HEX	181	95 C 10 min; 95 C 30s, 58.7 C 2 min (40x); 98 C 10 min; 4 C hold	4.1
ST-ETEC	STh	gBlock DNA	FAM	73	95 C 10 min; 94 C 30s, 54.6 C 1 min (40x); 98 C 10 min; 4 C hold	3.2
<i>Cryptosporidium</i> spp.	Crypto 18S	gBlock DNA	HEX	80	95 C 10 min; 94 C 30s, 54.6 C 1 min (40x); 98 C 10 min; 4 C hold	3.9
<i>Giardia duodenalis</i>	beta giardin	gBlock DNA	FAM	143	95 C 10 min; 95 C 30s, 58.7 C 1 min (40x); 98 C 10 min; 4 C hold	1.7
MS2	MS2g1	gBlock DNA	HEX	99	95C 10 min; 95 C 30s, 59.9 C 1 min (40x); 98 C 10 min; 4 C hold	5.6
<i>E. coli</i>	ybbW	gBlock DNA	EvaGreen	211	95 C 10 min; 95 C 30s, 59 C 2 min (40x); 4 C 5 min; 90 C 5 min; 4 C hold	3.3

TABLE C3: Optimized thermal cycling conditions for qPCR control assays as adapted to ddPCR.

MICROBE	GENE TARGET	CONTROL MATERIAL	REPORTER	AMPLICON LENGTH	THERMAL CYCLING
Bovine respiratory syncytial virus	BRSV	Inforce 3	FAM	124	95 C 10 min; 95 C 30s, 56.6 C 2 min (40x); 98 C 10 min; 4 C hold
Bovine herpes virus	BoHV	Inforce 3	HEX	97	95 C 10 min; 95 C 30s, 56.6 C 1 min (40x); 98 C 10 min; 4 C hold
Lambda phage	Cal 1	lambda gDNA	FAM	151	95 C 10 min; 95 C 30s, 56.6 C 2 min (40x); 98 C 10 min; 4 C hold
Hep G	HepG 5'UTR	ultramer RNA	HEX	205	95 C 10 min; 95 C 30 s, 56.6 C 2 min (40x); 98 C 10 min; 4 C hold

TABLE C4: Concentration factor calculations for the DEUF and ddPCR workflow as used to analyze concentrated water samples from Jaipur, India.

<p>Concentration Factor 1 (CF1), DEUF and Backflush:</p> $CF1 = \frac{Conc_{BF}}{Conc_{DW}} = \frac{V_{DW} * \eta_{BF}}{V_{BF}}$
<p>CF2, PEG Precipitation and Ultracentrifugation:</p> $CF2 = \frac{V_{BF} * \eta_{PEG}}{V_{Concentrate}} = \frac{300 \text{ mL} * 1.00}{4.5 \text{ mL}} = 66.7$
<p>CF3, Lysis:</p> $CF3 = \frac{V_{Conc \text{ into lysis}}}{V_{conc \text{ into lysis}} + V_{UNEX}} = \frac{500 \text{ uL}}{500 \text{ uL} + 500 \text{ uL}} = 0.5$
<p>CF4, Elution:</p> $CF4 = \frac{V_{lysate}}{V_{eluent}} = \frac{500 \text{ uL}}{100 \text{ uL}} = 5$
<p>CF5, Reverse Transcription:</p> $CF5 = \frac{V_{extract}}{V_{extract} + V_{RT-mix}} = \frac{10 \text{ uL}}{10 \text{ uL} + 10 \text{ uL}} = 0.5$
<p>TCF_{DNA}, Total Concentration Factor for DNA target:</p> $TCF_{DNA} = \frac{V_{DW} * \eta_{BF}}{V_{BF}} * 66.7 * 0.5 * 5 = 166.75 * \frac{V_{DW} * \eta_{BF}}{V_{BF}}$
<p>TCF_{RNA}, Total Concentration Factor for RNA target:</p> $TCF_{RNA} = \frac{V_{DW} * \eta_{BF}}{V_{BF}} * 66.7 * 0.5 * 5 * 0.5 = 83.38 * \frac{V_{DW} * \eta_{BF}}{V_{BF}}$

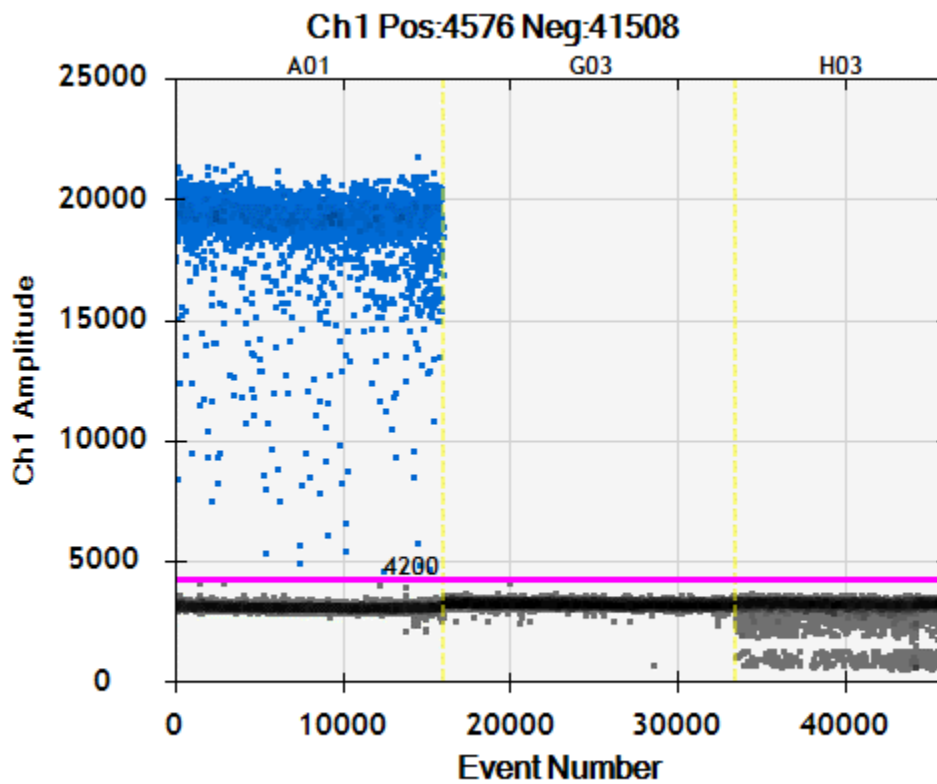


FIGURE C4: Example output from manual thresholding in QuantaSoft (Version 1.7.4.0917; Bio-Rad, Hercules, CA) during quantification of process control BRSV. Well A01 is a sample spiked with 5 uL of Inforce 3 bovine vaccine. Well G03 is a negative extraction control and Well H03 is a no-template control. In this case, the user drew the threshold for the plate at a Ch1 amplitude of 4200 on the y axis. Each droplet in a well is denoted as an “Event” on the x axis. Blue droplets are classified as positive while grey and black droplets are negative. The vertical dashed yellow lines denote different wells.

BRSV_Jaipur Well_A01 results

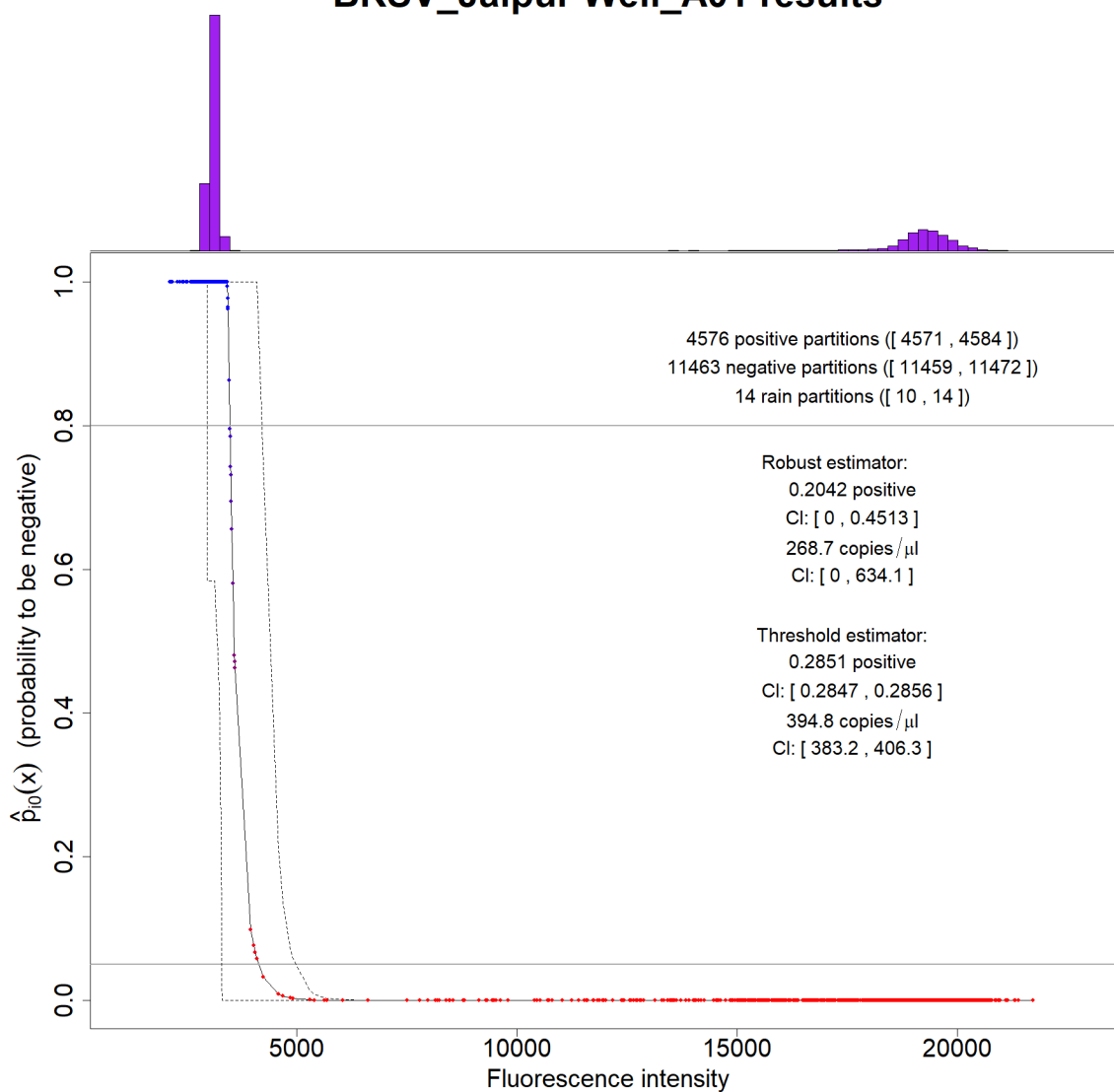


FIGURE C5: Umbrella output from Well A01, a sample spiked with 5 μ L of BRSV. The y-axis reports the probability of a partition being negative for the target based on the fluorescence amplitude distributions observed in the NTC and NEC given the observed fluorescence value noted on the x-axis. Each dot on the graph represents a droplet observed in the well. The blue droplets have less than a 20% probability to be false negatives and are classified as “negative partitions”. The red droplets have less than a 5% probability to be false positives and are classified as “positive partitions”. Droplets between these two thresholds are displayed as a color gradient between blue and red and are interpreted as ambiguous “rain partitions”. The fluorescence histogram at the top demonstrates the presence of a negative cluster and a positive cluster in the well.

BRSV_Jaipur Well_H03 results

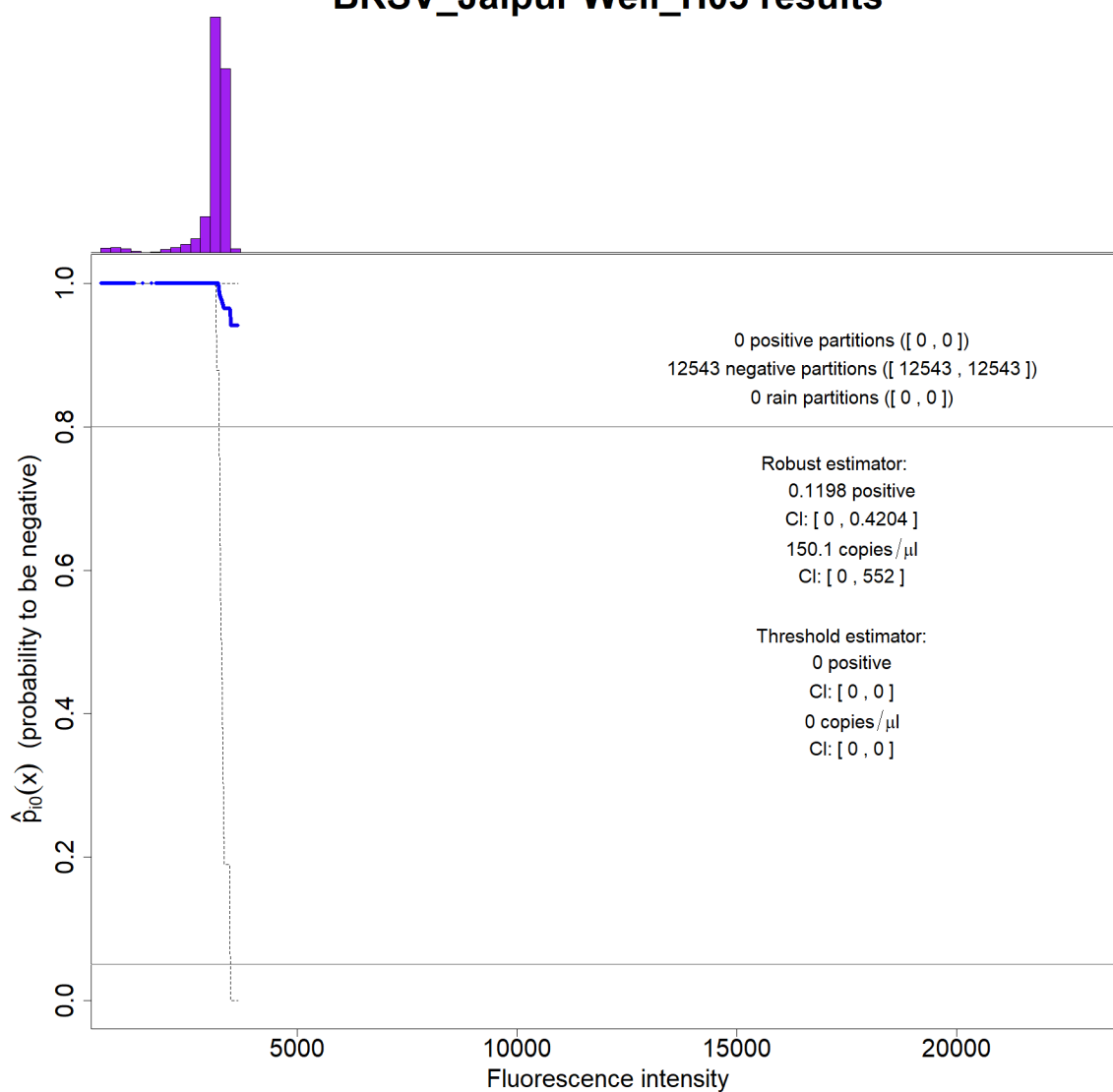


FIGURE C6: Umbrella results for well H03 an NTC from the BRSV ddPCR experiment. The histogram of the fluorescence values reflects the dual banded negative cluster observed from the QuantaSoft output in Figure C4. For this well all 12,543 droplets have been estimated to have less than a 20% probability of being false negatives and are therefore as negative and are shown in blue.

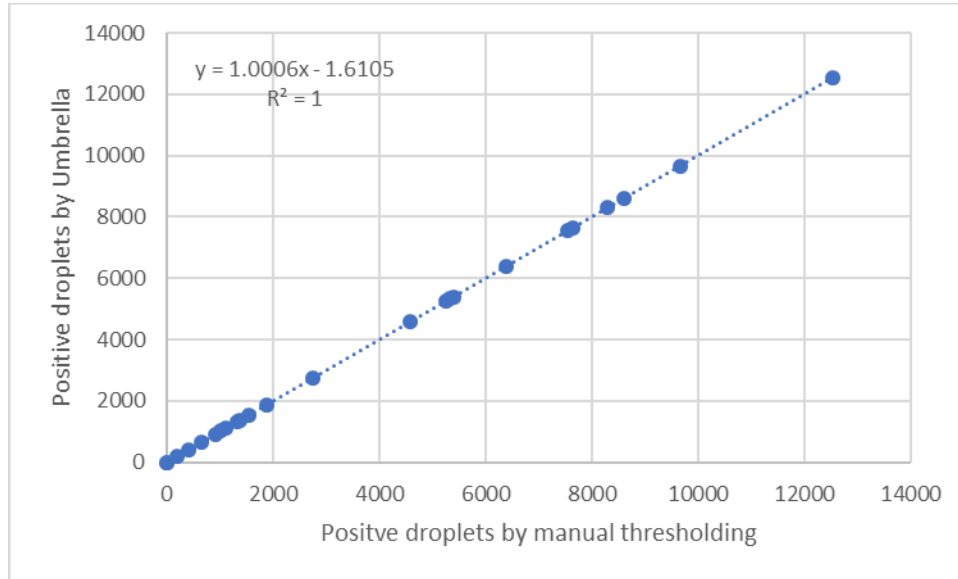


FIGURE C7: Correlation between the number of positive droplets by manual thresholding and Umbrella thresholding as observed for BRSV process control spiked into extracted DEUF concentrates from drinking water samples collected in Jaipur, negative extraction control, and no-template-control (n=24).

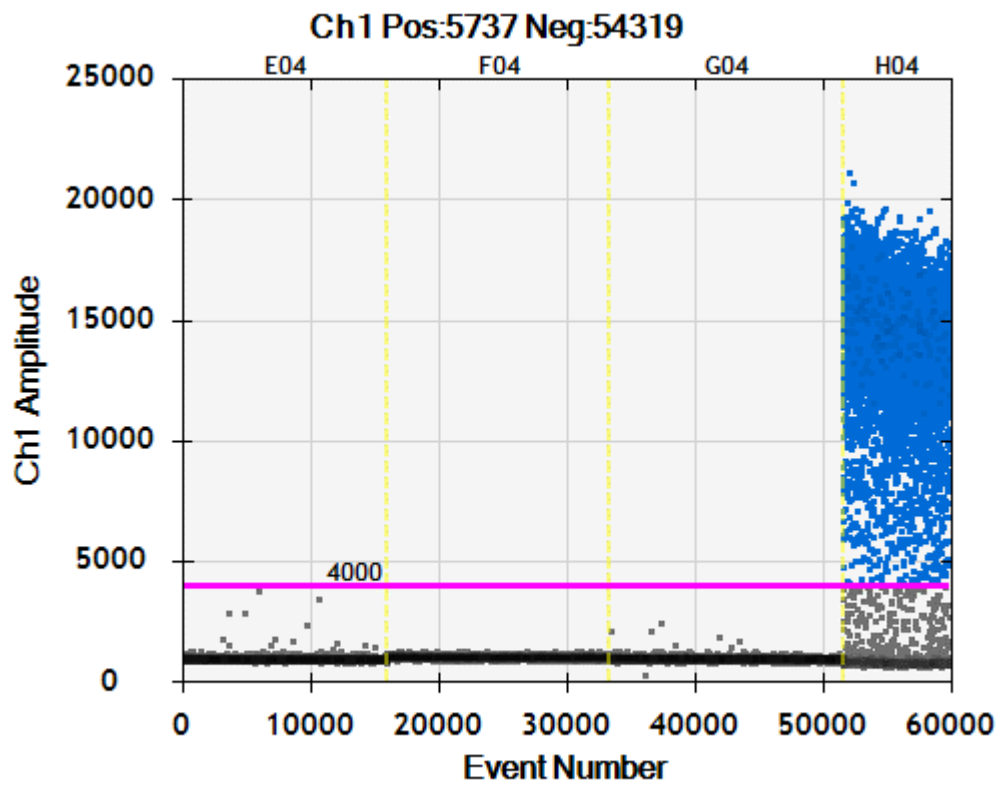


FIGURE C8: Variance in the fluorescence amplitude in the negative clusters as observed in the NEC (Well E04), two NTCs (Well F04 and G04), and a positive control (Well H04) for the hexon assay experiment.

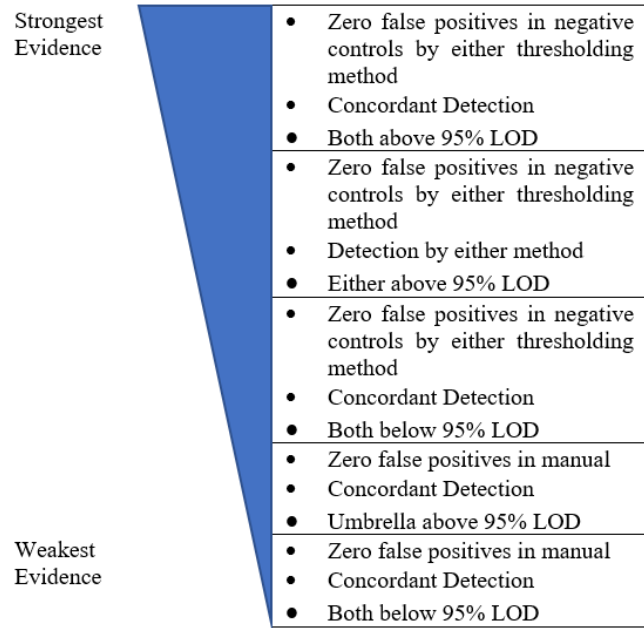


FIGURE C9: Proposed strength-of-evidence paradigm for reporting molecular evidence of waterborne pathogens in Jaipur drinking water as determined by ddPCR assay.

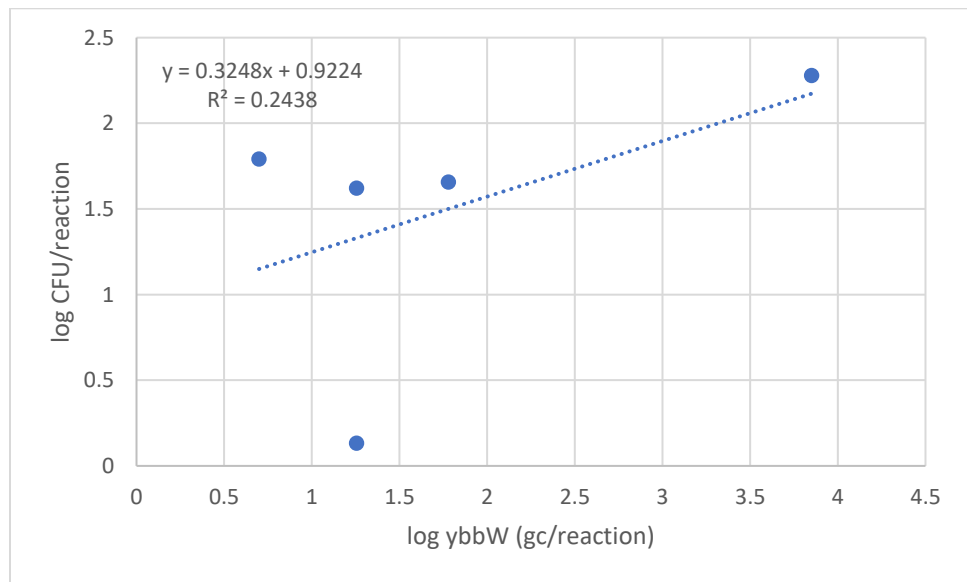


FIGURE C10: Linear model fit to the log of ybbW gc per PCR reaction versus the log of CFU per reaction as observed during drinking water sampling in Jaipur.

APPENDIX D

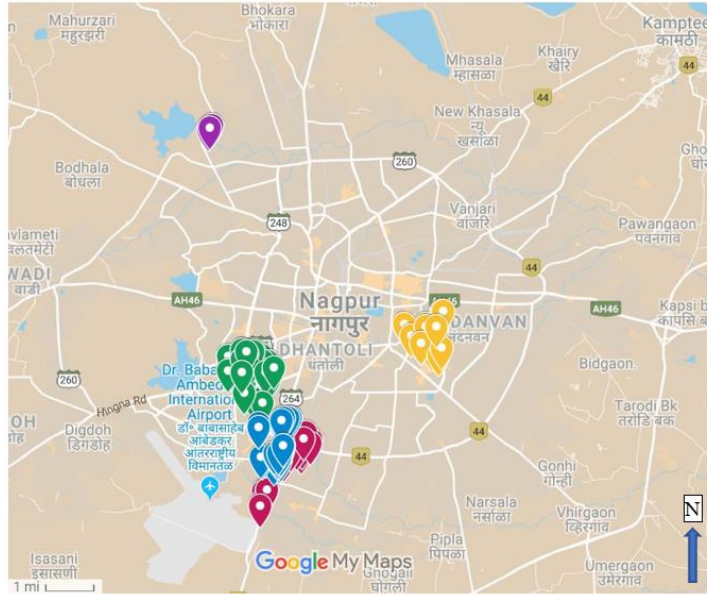


FIGURE D1: Overview of sample collection locations in Nagpur during 2015 and 2017 sampling. Yellow and red markers denote sample collections in Reshim Bagh (IWS) and Karve Nagar (IWS), respectively, during 2015. Purple markers in the upper left are samples collected from the Pench 2 WTP and green and blue markers are samples collected in the Laxmi Nagar Old (CWS) and Khamla (IWS) command areas, respectively, in 2017.

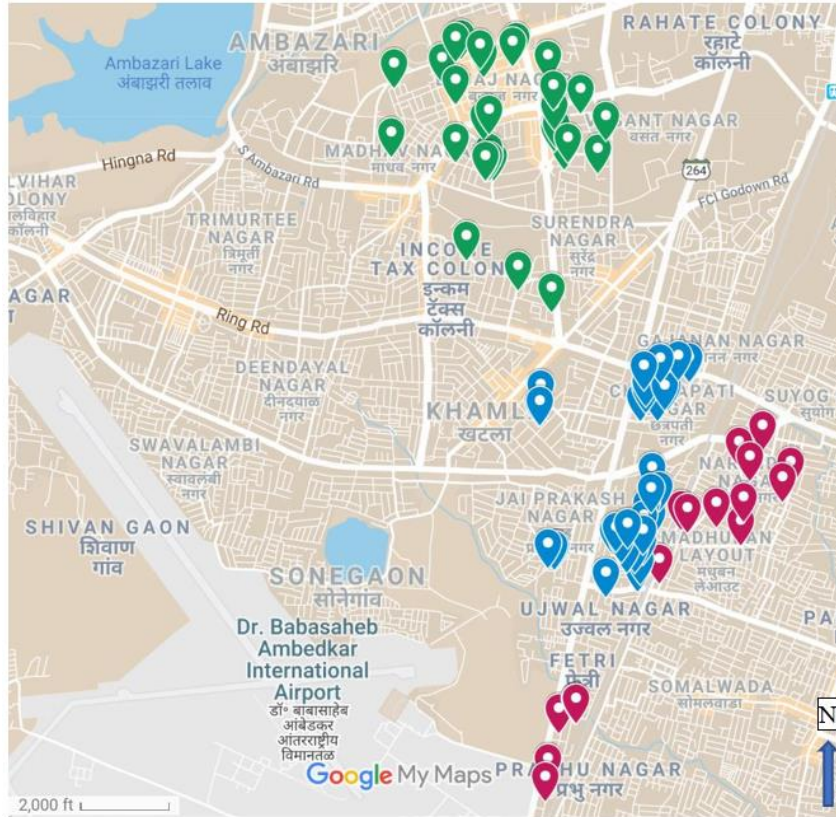


FIGURE D2: Sample collection locations in Karve Nagar (IWS) in 2015 (red markers), Khamla (IWS) in 2017 (blue markers), and Laxmi Nagar Old (CWS) in 2017 (green markers).

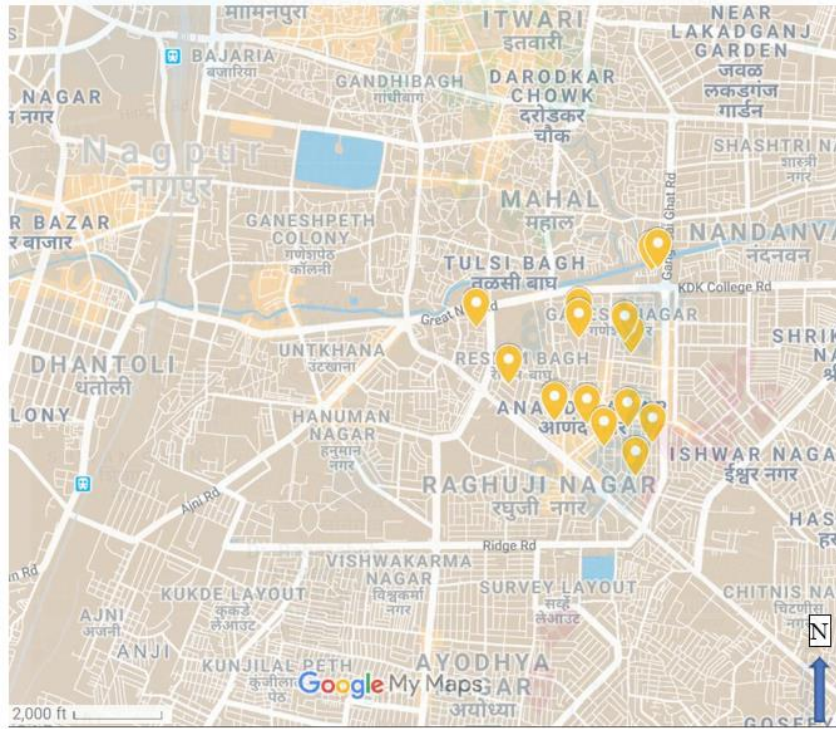


FIGURE D3: Sample collection locations in Reshim Bagh (IWS) in 2015 (yellow markers).

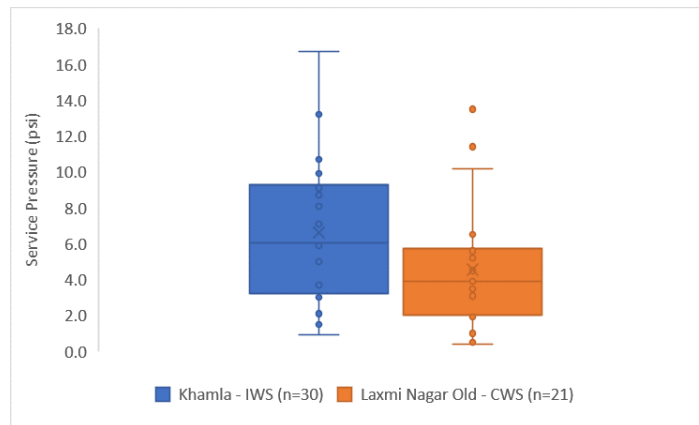


FIGURE D4: Boxplots of observed water pressure at household service taps in an IWS area (Khamla) and a CWS area (Laxmi Nagar Old) in Nagpur.

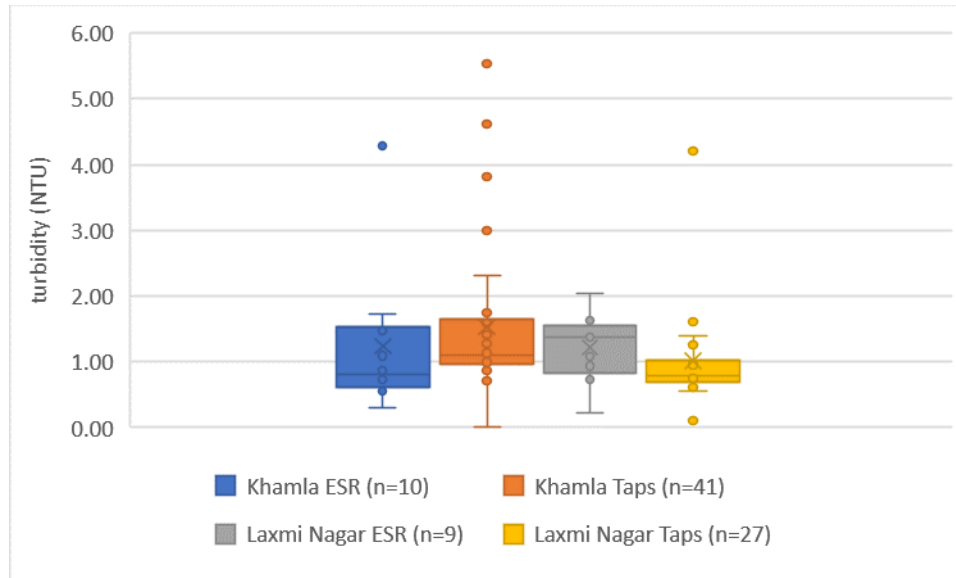


FIGURE D5: Boxplots of observed turbidity at Khamla ESR and household taps (IWS) and Laxmi Nagar Old ESR and household taps (CWS) in Nagpur.

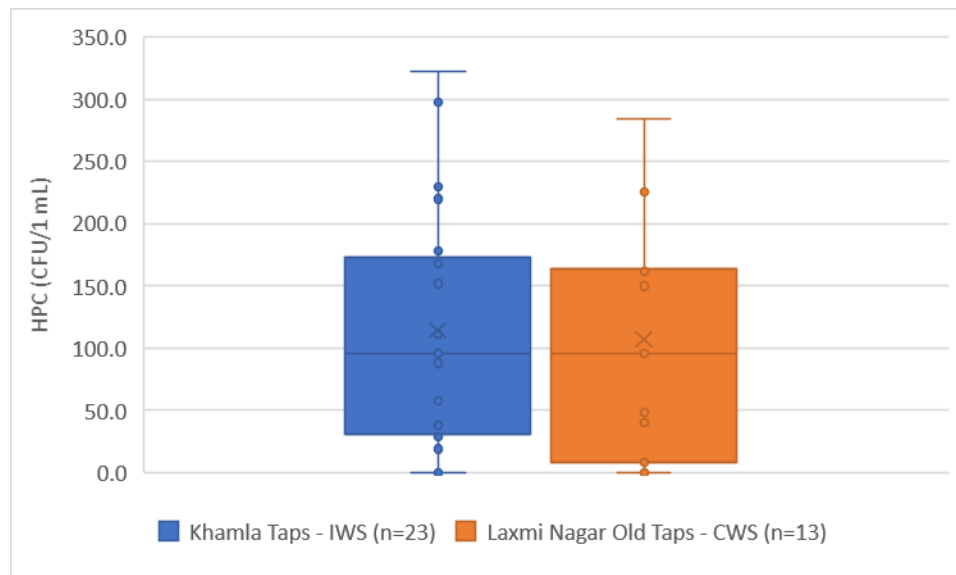


FIGURE D6: Heterotrophic plate counts (HPC) as observed in grab samples from an IWS and CWS service area during sampling in Nagpur in 2017.

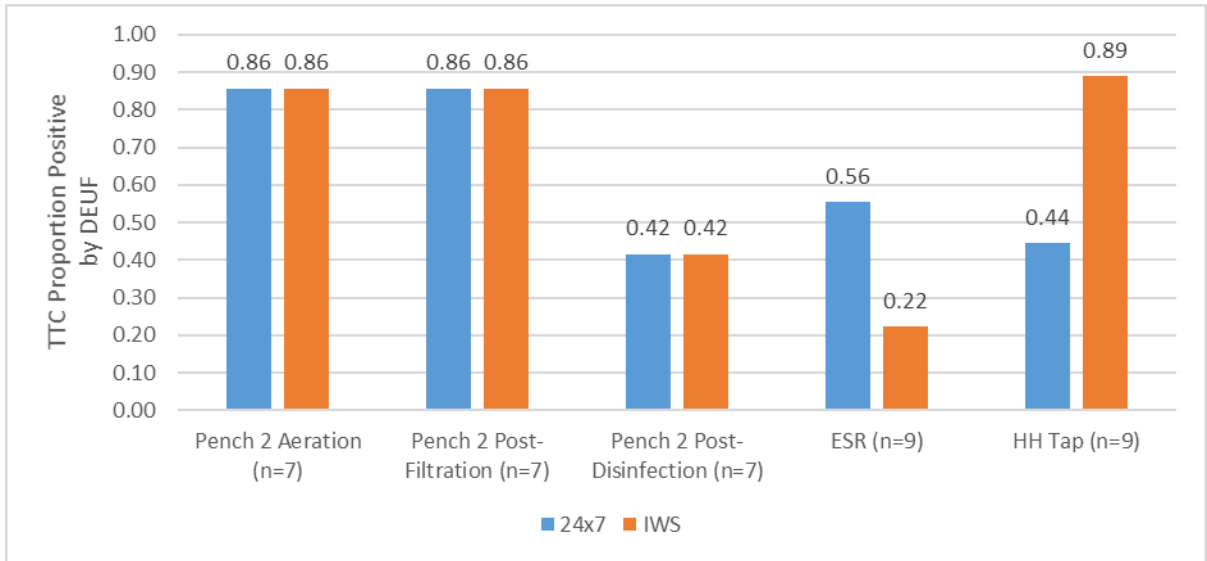


FIGURE D7: Proportion of DEUF backflush samples positive for thermotolerant coliforms along the CWS versus the IWS drinking water supply chain as observed during sampling in 2017.

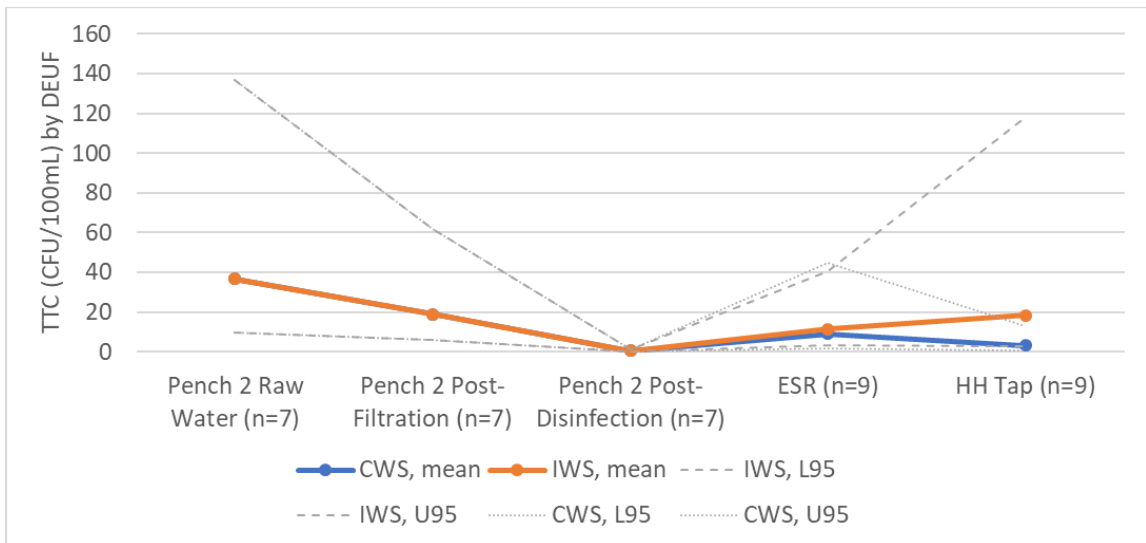


FIGURE D8: Mean TTC counts and associated 95% confidence intervals as observed in the IWS and CWS drinking water supply chains via DEUF in Nagpur.

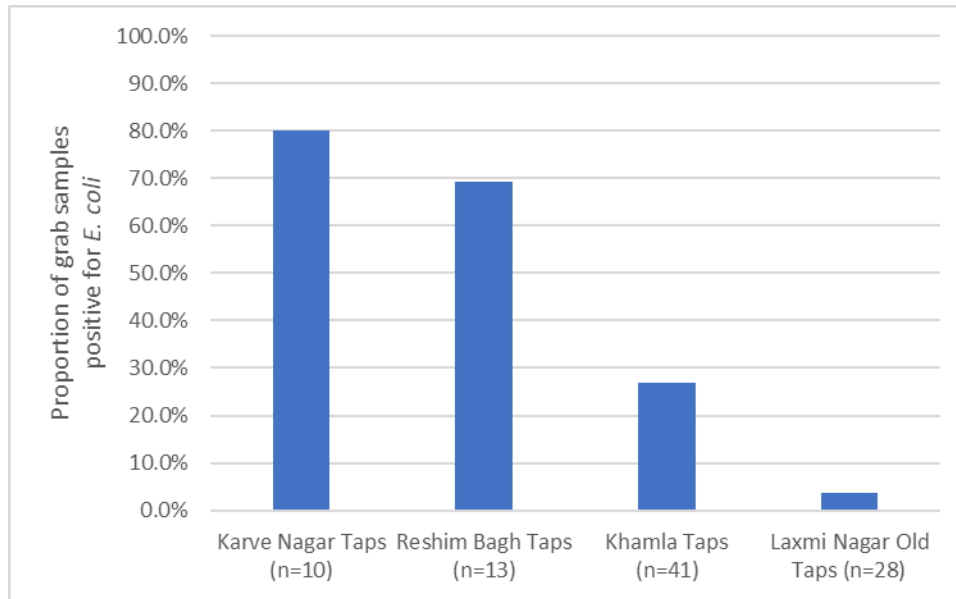


FIGURE D9: The proportion of grab samples collected from household taps positive for *E. coli* in three IWS service areas (Karve Nagar, Reshim Bagh, Khamla) versus a CWS service area (Laxmi Nagar Old).

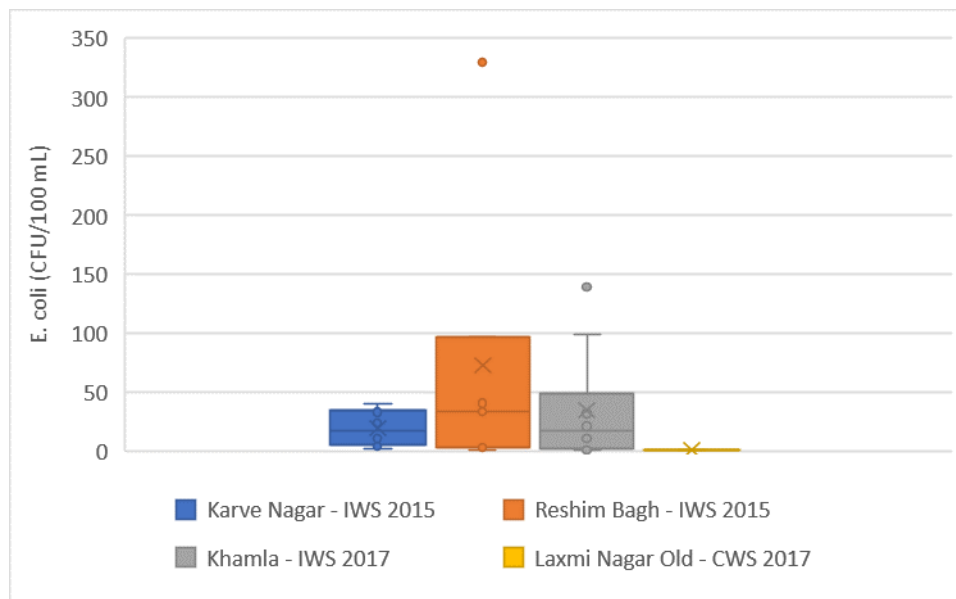


FIGURE D10: *E. coli* counts at household taps as observed in grab samples collected from Karve Nagar (IWS, n=8 countable), Reshim Bagh (IWS, n=9 countable), Khamla (IWS, n=11 countable), and Laxmi Nagar Old (CWS, n=1 countable).

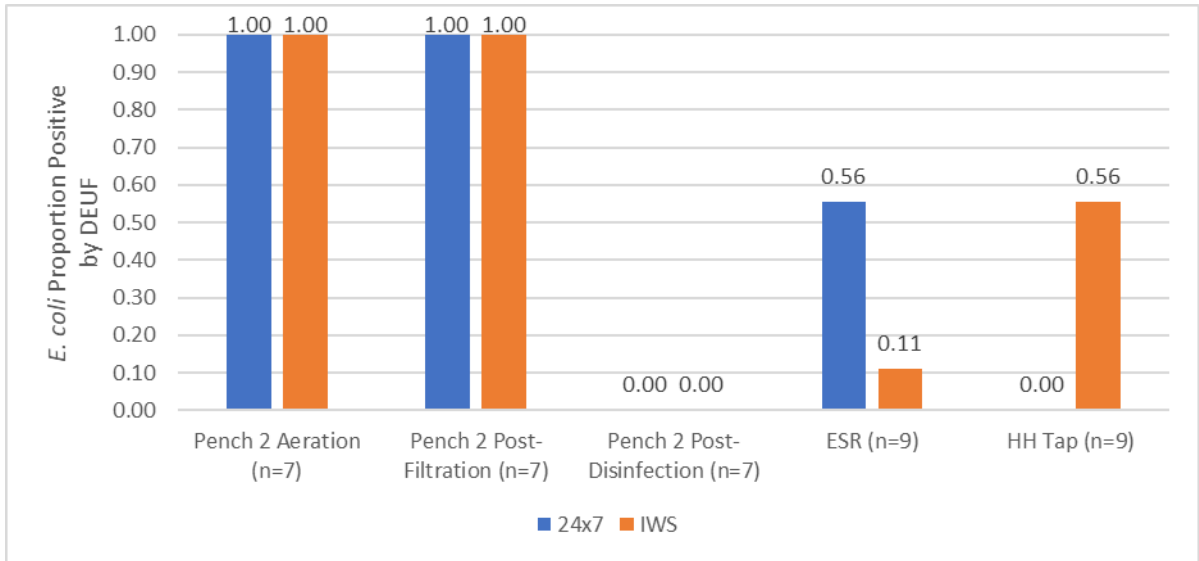


FIGURE D11: Proportion of DEUF backflush samples positive for *E. coli* along the CWS versus the IWS drinking water supply chain as observed during sampling in 2017.

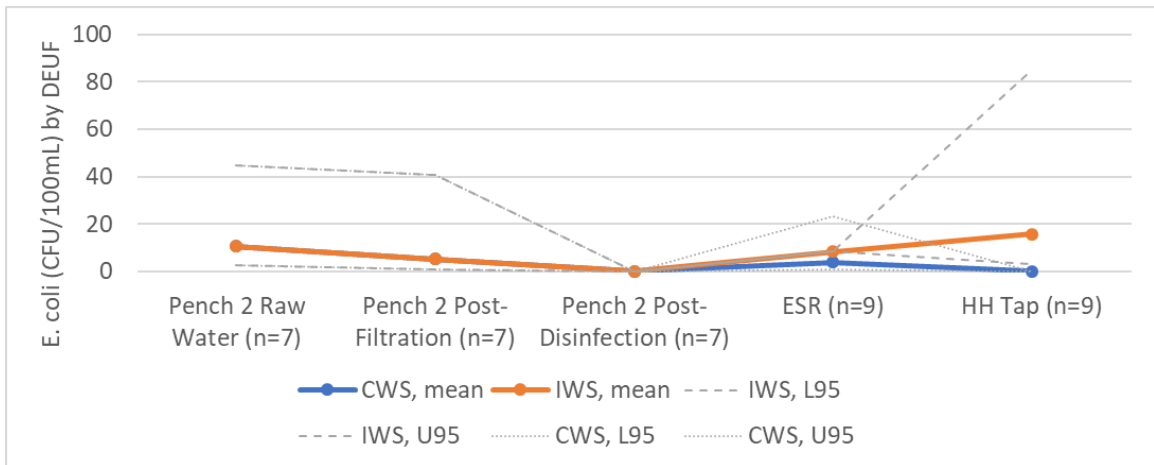


FIGURE D12: Mean *E. coli* counts and associated 95% confidence intervals as observed in the IWS and CWS drinking water supply chains via DEUF in Nagpur.

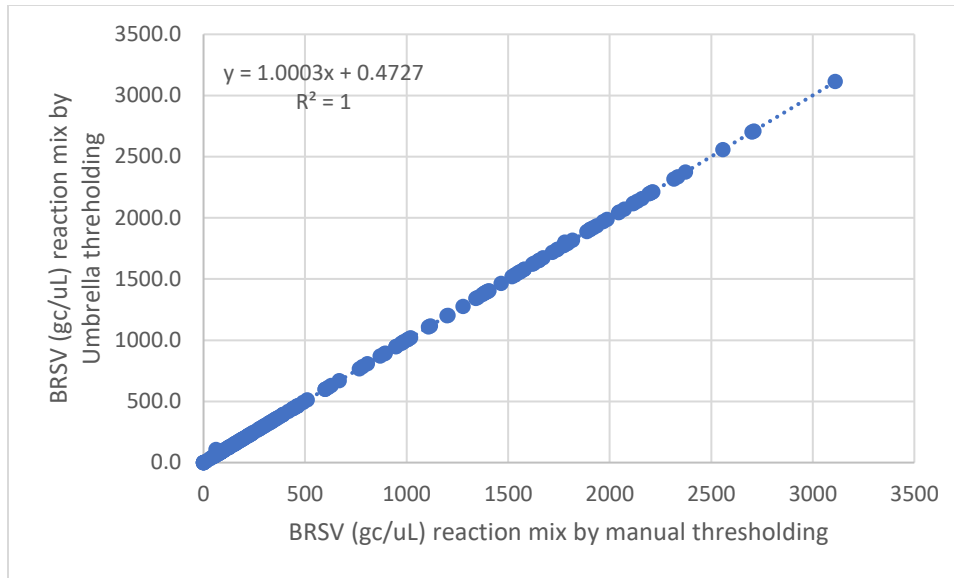


FIGURE D13: Quantification of process control BRSV in ddPCR reaction mix by manual thresholding and Umbrella thresholding demonstrated perfect linear correlation between the two thresholding methods.

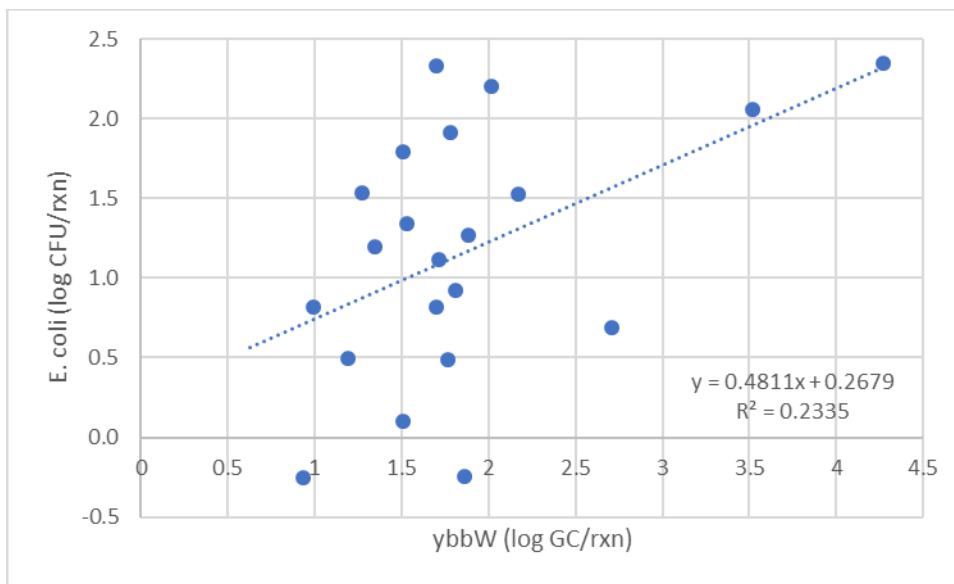


FIGURE D14: Linear model fit to the log of $ybbW$ gc per PCR reaction versus the log of CFU per reaction as observed in DEUF samples from the municipal water supply in Nagpur.

TABLE D1: Performance of ddPCR assays as observed during the interrogation of DEUF concentrate from samples collected from the municipal water supply in Nagpur, India.

MOLECULAR ASSAY	MANUAL THRESHOLD	UMBRELLA THRESHOLD
BRSV (Control)		
<i>false positive rate</i>	0	0
<i>problematic wells</i>		0
<i>discordant pos.</i>		0
<i>concordant pos.</i>		192
<i>pos. below 95% LOD</i>	n/a	n/a
<i>pos. above 95% LOD</i>	n/a	n/a
MS2g1		
<i>false positive rate</i>	0 of 271,027	0 of 271,027
<i>problematic wells</i>		0
<i>discordant pos.</i>		9
<i>concordant pos.</i>		6
<i>pos. below 95% LOD</i>	13	7
<i>pos. above 95% LOD</i>	0	1
beta giardin		
<i>false positive rate</i>	0 of 237,077	0 of 237,077
<i>problematic wells</i>		2
<i>discordant pos.</i>		7
<i>concordant pos.</i>		9
<i>pos. below 95% LOD</i>	16	11
<i>pos. above 95% LOD</i>	2	6
NoV GII		
<i>false positive rate</i>	0 of 253,343	0 of 253,343
<i>problematic wells</i>		0
<i>discordant pos.</i>		18
<i>concordant pos.</i>		9
<i>pos. below 95% LOD</i>	16	12
<i>pos. above 95% LOD</i>	3	5
NoV GI		
<i>false positive rate</i>	0 of 253,343	0 of 253,343
<i>problematic wells</i>		0
<i>discordant pos.</i>		22
<i>concordant pos.</i>		6
<i>pos. below 95% LOD</i>	17	12
<i>pos. above 95% LOD</i>	2	3
ybbW		
<i>false positive rate</i>	3 of 241,070	N/A
<i>problematic wells</i>		N/A
<i>discordant pos.</i>		N/A
<i>concordant pos.</i>		N/A
<i>pos. below 95% LOD</i>	39	1
<i>pos. above 95% LOD</i>	51	20
ipaH		
<i>false positive rate</i>	0 of 265,592	0 of 265,592
<i>problematic wells</i>		0
<i>discordant pos.</i>		9

TABLE D1: Continued

<i>concordant pos.</i>	6	
<i>pos. below 95% LOD</i>	9	10
<i>pos. above 95% LOD</i>	1	1
hexon		
<i>false positive rate</i>	1 of 265,592	0 of 265,592
<i>problematic wells</i>	0	
<i>discordant pos.</i>	9	
<i>concordant pos.</i>	17	
<i>pos. below 95% LOD</i>	20	10
<i>pos. above 95% LOD</i>	8	7
STh		
<i>false positive rate</i>	0 of 215,395	0 of 215,395
<i>problematic wells</i>	8	
<i>discordant pos.</i>	20	
<i>concordant pos.</i>	35	
<i>pos. below 95% LOD</i>	29	43
<i>pos. above 95% LOD</i>	8	11
Crypto. 18S rRNA		
<i>false positive rate</i>	0 of 250,899	0 of 250,899
<i>problematic wells</i>	9	
<i>discordant pos.</i>	11	
<i>concordant pos.</i>	11	
<i>pos. below 95% LOD</i>	15	4
<i>pos. above 95% LOD</i>	0	13

TABLE D2: Number of DEUF samples positive by ddPCR for each genetic target and an aggregate of all targets associated with pathogens (All Path) above the 95% LOD observed at each sampling point within IWS and CWS areas of Nagpur.

Command Area	ybbW*	MS2g1	beta giardin	NoV GI	NoV GII	ipaH	hexon	STh	C18S	All Path
Karve Nagar ESR (n=3)	0	0	0	0	0	0	0	0	0	0
Karve Nagar Taps (n=14)	3	0	1	1	1	0	1	0	1	5
Khamla ESR (n=9)	6	0	0	0	0	0	1	1	1	3
Khamla Taps (n=9)	7	0	1	0	1	0	1	3	0	6
Laxmi Nagar Old ESR (n=9)	6	1	0	0	1	0	0	2	0	3
Laxmi Nagar Old Taps (n=9)	6	0	0	0	0	0	0	0	1	1
Pench 2 Post Disinfection (n=7)	4	0	0	0	0	0	1	0	0	1
Pench 2 Post Filtration (n=7)	6	0	0	0	0	0	0	2	0	2
Pench 2 Aeration (n=7)	7	0	0	0	0	0	2	0	1	3
Reshim Bagh ESR (n=3)	0	0	1	0	0	0	0	0	0	1
Reshim Bagh Taps (n=13)	1	0	2	2	2	1	1	3	0	11

TABLE D3: Number of DEUF samples positive by ddPCR for each genetic target and an aggregate of all targets associated with pathogens (All Path) above the 95% LOD as observed in IWS versus CWS areas of Nagpur.

Command Area	ybbW*	MS2g1	beta giardin	NoV GI	NoV GII	ipaH	hexon	STh	C18S	All Pathogen
IWS ESR (n=15)	6	0	1	0	0	0	1	1	1	4
IWS Taps (n=36)	11	0	4	3	4	1	3	6	1	22
CWS ESR (n=9)	6	1	0	0	1	0	0	2	0	3
CWS Taps (n=9)	6	0	0	0	0	0	0	0	1	1

APPENDIX E

TABLE E1: Exposure assessment equations as implemented in the IWS in India QMRA.

<p>Concentration of DNA gene targets in the eluate (gc/uL):</p> $C_{eluate} = \frac{C_{ddPCR} * V_{ddPCR}}{\eta_{ddPCR} * V_{template}} \quad (1)$ <p>where, C_{ddPCR} is the concentration of the target per uL of reaction mix (gc/uL), stochastic variable V_{ddPCR} is the volume of the entire ddPCR reaction, 20 uL η_{ddPCR} is the efficiency of the PCR reaction, assumed 100% $V_{template}$ is the volume to the sample template added to the reaction, 4 uL</p>
<p>Concentration of cDNA gene targets in the eluate (gc/uL):</p> $C_{eluate} = \frac{C_{ddPCR} * V_{ddPCR} * V_{RT}}{\eta_{ddPCR} * \eta_{RT} * V_{template, ddPCR} * V_{template, RT}} \quad (2)$ <p>where, C_{ddPCR} is the concentration of the target per uL of reaction mix (gc/uL), stochastic variable V_{ddPCR} is the volume of the entire ddPCR reaction, 20 uL V_{RT} is the volume of the entire RT reaction, 20 uL η_{ddPCR} is the efficiency of the PCR reaction, assumed 100% η_{RT} is the efficiency of the reverse transcription reaction, assumed 100% $V_{template, ddPCR}$ is the volume to the RT product added to the ddPCR reaction, 4 uL $V_{template, RT}$ is the volume to the sample template added to the RT reaction, 10 uL</p>
<p>Concentration of gene target in the lysate (gc/uL):</p> $C_{lysate} = \frac{C_{eluate} * V_{eluate}}{\eta_{elution} * V_{eluent}} \quad (3)$ <p>where, C_{eluate} is the concentration of the target per uL of eluate (gc/uL), from eq. (1) or (2) V_{eluate} is the volume of the eluate added to the extraction column, 100 uL $\eta_{elution}$ is the efficiency of the elution reaction, assumed 100% V_{eluent} is the volume of the eluent added to the extraction tube, 500 uL</p>
<p>Concentration of the gene target in the DEUF concentrate (gc/uL):</p> $C_{DEUFconcentrate} = \frac{C_{lysate} * V_{lysate}}{\eta_{lysis} * V_{conc\ into\ lysis}} \quad (4)$ <p>where, C_{lysate} is the concentration of the target per uL of lysate (gc/uL), from eq. (3) V_{lysate} is the volume of entire lysis reaction, 1000 uL η_{lysis} is the efficiency of the lysis reaction, assumed 100% $V_{conc\ into\ lysis}$ is the volume of DEUF concentrate added to the lysis reaction, 500 uL</p>

TABLE E1: Continued

<p>Concentration of the gene target in the DEUF backflush (gc/uL):</p> $C_{DEUF\ Backflush} = \frac{C_{DEUF\ conc} * V_{resuspended\ conc}}{\eta_{PEG} * V_{backflush\ into\ PEG}} \quad (5)$ <p>where, $C_{DEUF\ conc}$ is the concentration of the target per uL in the DEUF concentrate, from eq. (4) $V_{resuspended\ conc}$ is the volume of the resuspended pellet following PEG precip, 4 mL η_{PEG} is the recovery efficiency of the PEG + ultracentrifugation process, %, stochastic variable $V_{backflush\ into\ PEG}$ is the volume of DEUF backflush subject to PEG + ultracentrifugation, 300 mL</p>
<p>Volume of drinking water sampled by DEUF (L):</p> $V_{DEUF} = A * V_{DEUF_{15}} + (1 - A) * V_{DEUF_{17}} \quad (6)$ <p>where, A is a dummy variable for sample collection in 2015 or 2017, stochastic variable $V_{DEUF_{15}}$ is the volume of water sampled by DEUF during 2015, stochastic variable $V_{DEUF_{17}}$ is the volume of water sampled by DEUF during 2017, stochastic variable</p>
<p>Concentration of the gene target in the drinking water (gc/uL):</p> $C_{DW} = \frac{C_{DEUF\ Backflush} * V_{Backflush}}{\eta_{Backflush} * V_{DW}} \quad (7)$ <p>where, $C_{DEUF\ Backflush}$ is the concentration of the target per uL in the DEUF backflush, from eq. (5) $V_{Backflush}$ is the volume of liquid recovered during backflushing, mL, stochastic variable $\eta_{Backflush}$ is the recovery efficiency of the DEUF backflush process, %, stochastic variable V_{DW} is the volume of drinking water filtered by DEUF, L, from eq. (6)</p>
<p>Dose of the gene target per day (gc/day):</p> $Dose \left(\frac{gc}{day} \right) = C_{DW} * 10^6 * V_{DW} \quad (8)$ <p>where, C_{DW} is the concentration of the target per uL of drinking water, from eq. (7) V_{DW} is the volume of drinking water ingested daily, stochastic variable</p>

TABLE E2: Dose harmonization equations as implemented in the IWS in India QMRA.

<p>Dose response harmonization for <i>Giardia duodenalis</i>:</p> $D_{Giardia}(\text{cysts per day}) = \text{beta giardin} * \frac{1 \text{ cyst}}{16 \text{ gc}} \quad (9)$ <p>Where, beta giardin is the dose of beta giardin gene (gc/day), from Eq. (8)</p>
<p>Dose response harmonization for <i>Cryptosporidium spp.</i>:</p> $D_{Crypto.}(\text{oocysts per day}) = \text{C18S rRNA} * \frac{\text{oocyst}}{20 \text{ gc}} \quad (10)$ <p>Where, C18S rRNA is the dose of 18s rRNA gene (gc/day), from Eq. (8)</p>
<p>Dose response harmonization for adenovirus 40/41:</p> $D_{Adeno}(\text{TCID}_{50} \text{ per day}) = \text{hexon} * \frac{\text{PFU}}{1000 \text{ genome}} * \frac{\text{TCID}_{50}}{0.7 \text{ PFU}} \quad (11)$ <p>where, TCID₅₀ is the median tissue culture infectious dose, the dose response unit for adenovirus hexon is the dose of hexon gene in (gc/day), from Eq. (8)</p>
<p>Dose response harmonization and viability for ST-EPEC (STh): (12)</p> $D_{ST-EPEC}(\text{CFU per day}) = B * \frac{\text{Viable}_{EC} * \text{STh}_{\text{dose}} * \text{ybbW}_{\text{gc/genome}} * \text{STh}_{\text{ddPCR}}}{\text{ybbW} * \text{STh}_{\text{gc/genome}} * \text{STh}_{\text{DW}}}$ <p>Where, B is a dummy variable indicating the presence of culturable <i>E. coli</i> at an IWS tap given detection of ybbW, stochastic Viable_{EC} is the culturable <i>E. coli</i> count as observed at IWS taps, CFU/L, stochastic variable STh_{dose}, dose of STh gene in drinking water (gc/day), from eq. (8) ybbW_{gc per genome}, ybbW gene per <i>E. coli</i> genome, 1 STh_{ddPCR} is the STh gene copies per uL of ddPCR reaction mix, stochastic ybbW, concentration of ybbW gene in ddPCR reaction mix (gc/uL), stochastic STh_{gc per genome}, STh gene copies per ST-EPEC genome, stochastic STh_{DW} is the STh gene concentration per liter of drinking water, from eq. (7)</p>

TABLE E2: Continued

Dose response harmonization and viability for *Shigella* (ipaH): (13)

$$D_{Shig}(CFU \text{ per day}) = B * \frac{Viable_{EC} * ipaH_{dose} * ybbW_{gc/genome} * ipaH_{ddPCR}}{ybbW * ipaH_{gc/genome} * ipaH_{DW}}$$

Where,

B is a dummy variable indicating the presence of culturable *E. coli* at an IWS tap given detection of ybbW, stochastic

Viable_{EC} is the culturable *E. coli* count as observed at IWS taps, CFU/L, stochastic variable

ipaH_{dose}, dose of ipaH gene in drinking water (gc/day), from eq. (8)

ybbW_{gc per genome}, ybbW gene per *E. coli* genome, 1

ipaH_{ddPCR} is the ipaH gene copies per uL of ddPCR reaction mix, stochastic

ybbW, concentration of ybbW gene in ddPCR reaction mix (gc/uL), stochastic

ipaH_{gc per genome}, ipaH gene copies per ST-EPEC genome, stochastic

ipaH_{DW} is the ipaH gene concentration per liter of drinking water, from eq. (7)

TABLE E3: Dose-response equations as implemented in the IWS in India QMRA.

<p>Probability of infection with <i>Giardia duodenalis</i> given dose of cysts:</p> $P_{inf} = 1 - \exp(-k * dose) \quad (14)$ <p>where, k is a parameter of the exponential model fit to dose-response data, stochastic dose is the ingested dose of viable cysts, stochastic variable, stochastic</p>
<p>Probability of infection with <i>Cryptosporidium</i> given ingestion of dose of oocysts:</p> $P_{inf} = 1 - \exp(-k * dose) \quad (15)$ <p>where, k is a parameter of the exponential model fit to dose-response data, stochastic dose is the ingested dose of viable oocysts, stochastic variable, stochastic</p>
<p>Probability of infection with norovirus given dose of infectious particles, high-risk model:</p> $P_{inf} = P * \left(1 - e^{-\frac{dose}{\mu_a}}\right) \quad (16)$ <p>where, μ_a is the mean aggregate size, 1 P is the fraction of population susceptible to infection (Se+), stochastic dose is the ingested dose, from eq. (8)</p>
<p>Probability of infection with norovirus given ingestion of infectious particles, low-risk model:</p> $P_{inf} = 1 - \left[1 + dose * \frac{\left(\frac{1}{2\alpha-1}\right)^{-\alpha}}{N_{50}}\right]^{-\alpha} \quad (17)$ <p>where, alpha is a parameter of the approximate beta Poisson model, 0.104 N_{50} is the median infectious dose, 25 302.7 dose is the ingested dose from eq. (8)</p>
<p>Probability of infection with adenovirus given ingestion of infectious particles:</p> $P_{inf}(D_{Adeno} \alpha, \beta) = 1 - {}_1F_1(\alpha, \alpha + \beta; -D) \quad (18)$ <p>where, cV is the dose, from eq. (11) α is the exact beta Poisson parameter, 5.11 β is the exact beta Poisson parameter, 2.80 ${}_1F_1$ is the confluent hypergeometric function of the first kind</p>

TABLE E3: Continued

<p>Probability of infection with ETEC given ingestion of viable CFU:</p> $P_{diarrhea} = 1 - \left[1 + dose * \frac{\left(\frac{1}{2\alpha-1}\right)}{N_{50}} \right]^{-\alpha} \quad (19)$ <p>where, alpha is a parameter of the approximate beta Poisson model, stochastic N₅₀ is the median infectious dose, stochastic dose is the dose of viable ETEC (CFU), from eq. (12)</p>
<p>Probability of infection with <i>Shigella spp.</i> given ingestion of viable CFU:</p> $P_{inf} = 1 - \left[1 + dose * \frac{\left(\frac{1}{2\alpha-1}\right)}{N_{50}} \right]^{-\alpha} \quad (20)$ <p>where, alpha is a parameter of the approximate beta Poisson model, stochastic (see table) N₅₀ is the median infectious dose, stochastic (see table) dose is the ingested dose of viable <i>Shigella</i> (CFU), from eq. (9)</p>

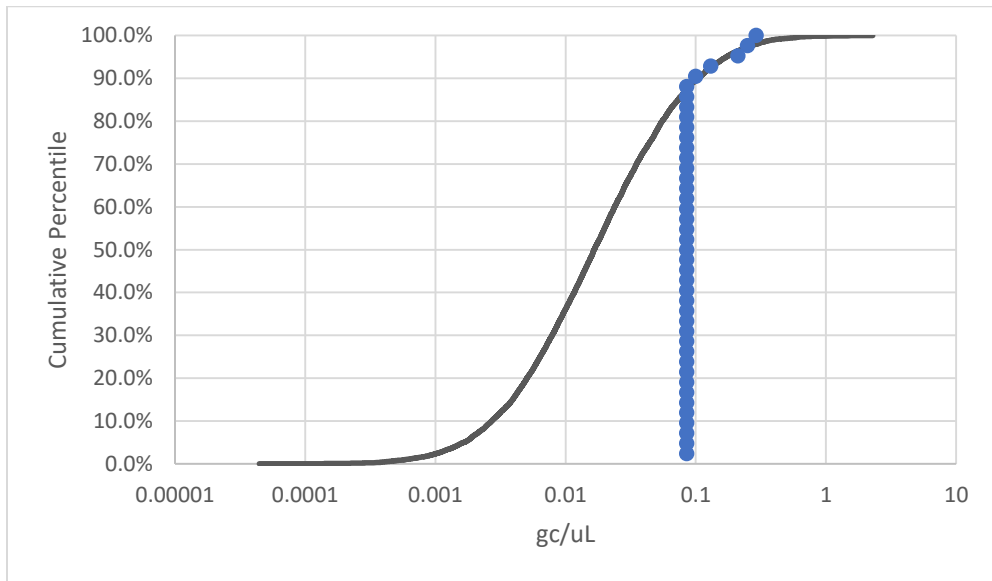


FIGURE E1: Lognormal model fit to observed concentrations of beta giardin (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.

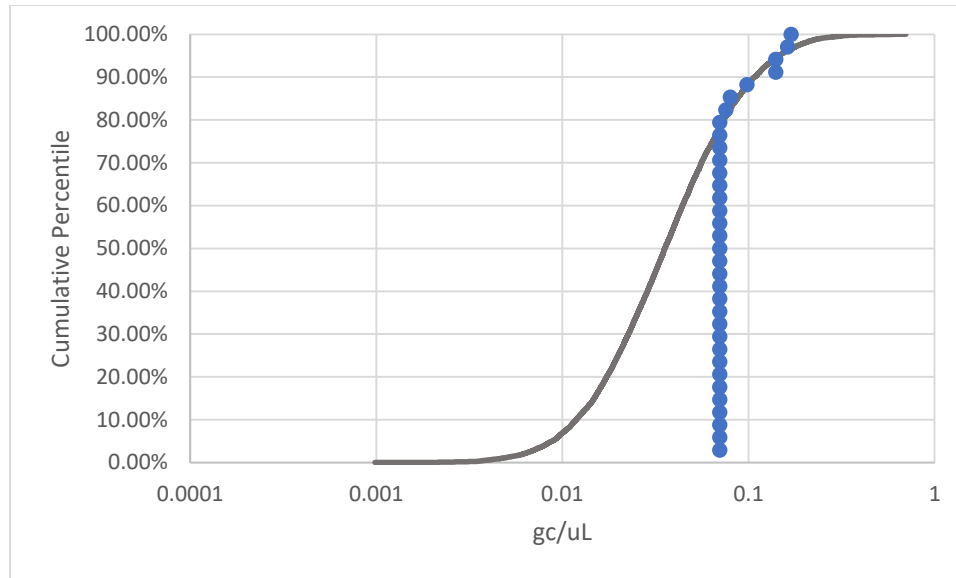


FIGURE E2: Lognormal model fit to observed concentrations of Crypto. 18S rRNA (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.

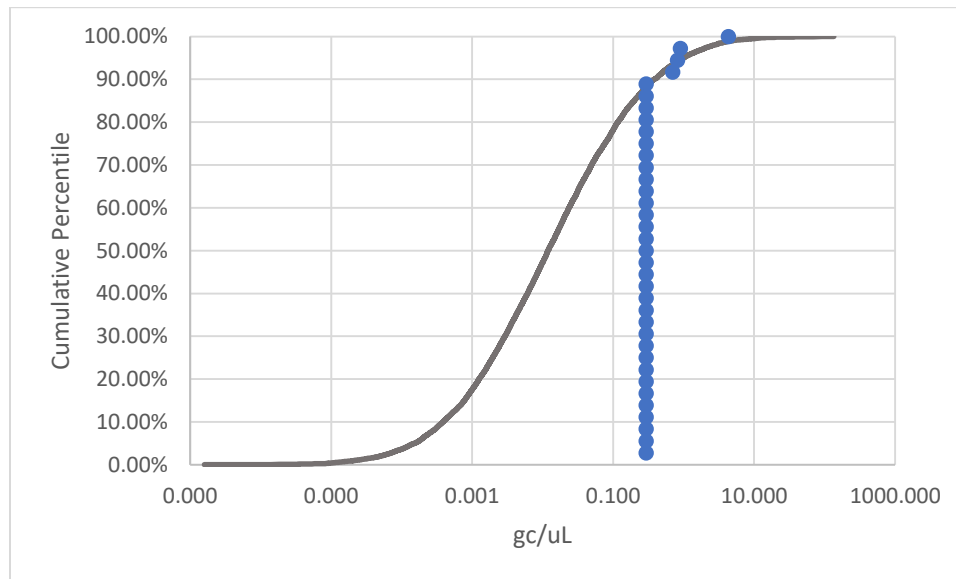


FIGURE E3: Lognormal model fit to observed concentrations of norovirus GI and GII ORF1-2 (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.

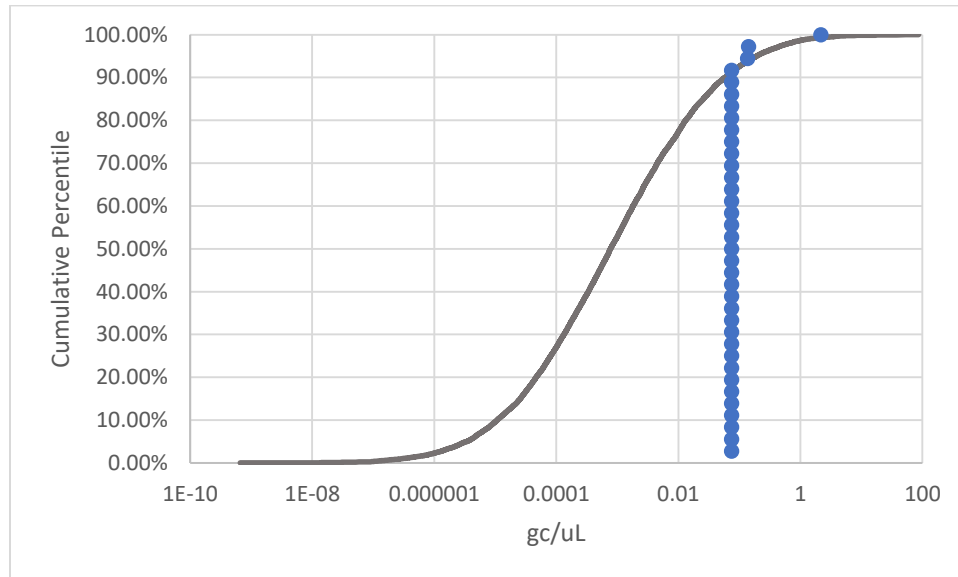


FIGURE E4: Lognormal model fit to observed concentrations of hexon (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.

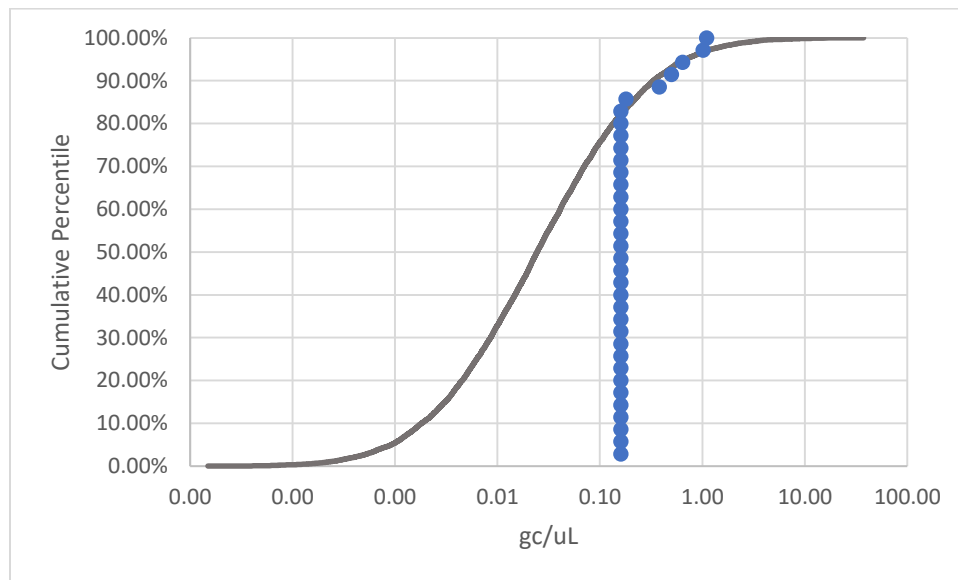


FIGURE E5: Lognormal model fit to observed concentrations of STh (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.

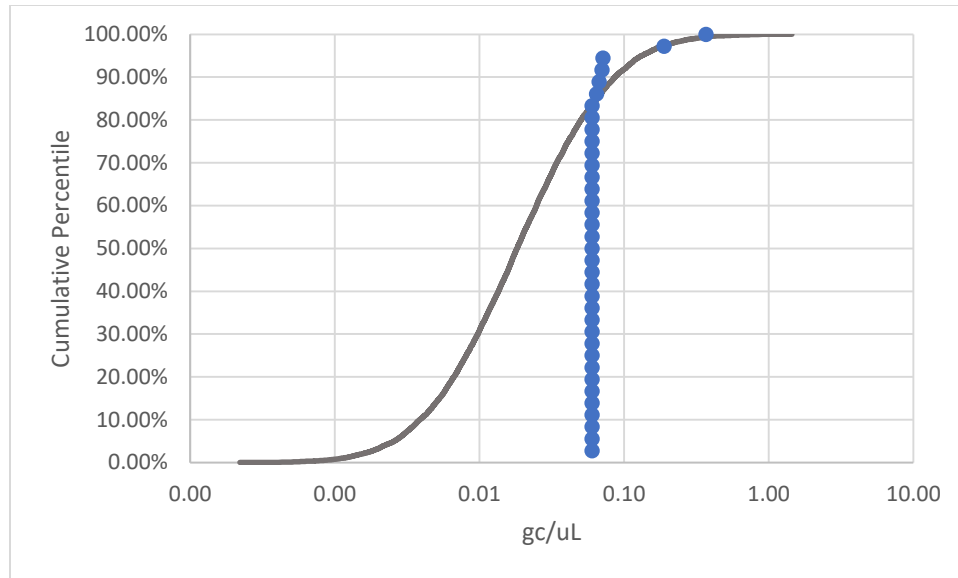


FIGURE E6: Lognormal model fit to observed concentrations of ipaH (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.

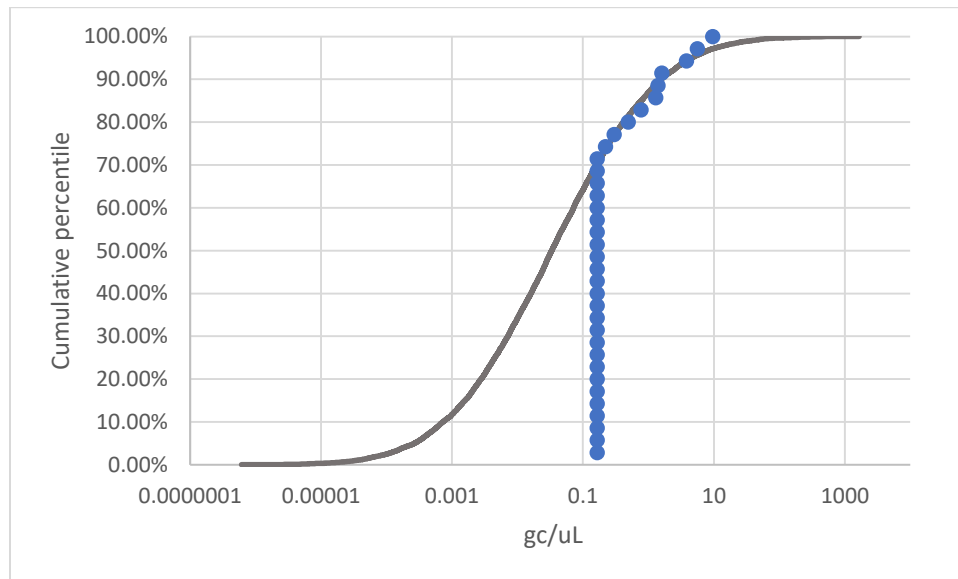


FIGURE E7: Lognormal model fit to observed concentrations of ybbW (gc/uL) in ddPCR wells associated with samples collected from IWS taps in Nagpur.

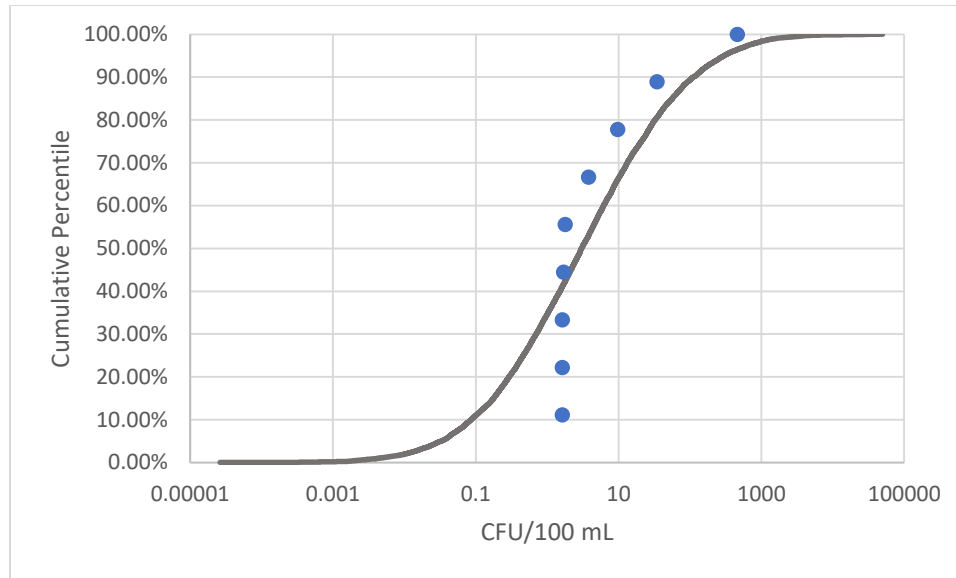


FIGURE E8: Lognormal model fit to observed counts of *E. coli* (CFU/100 mL) at IWS taps as enumerated via DEUF and membrane filtration.

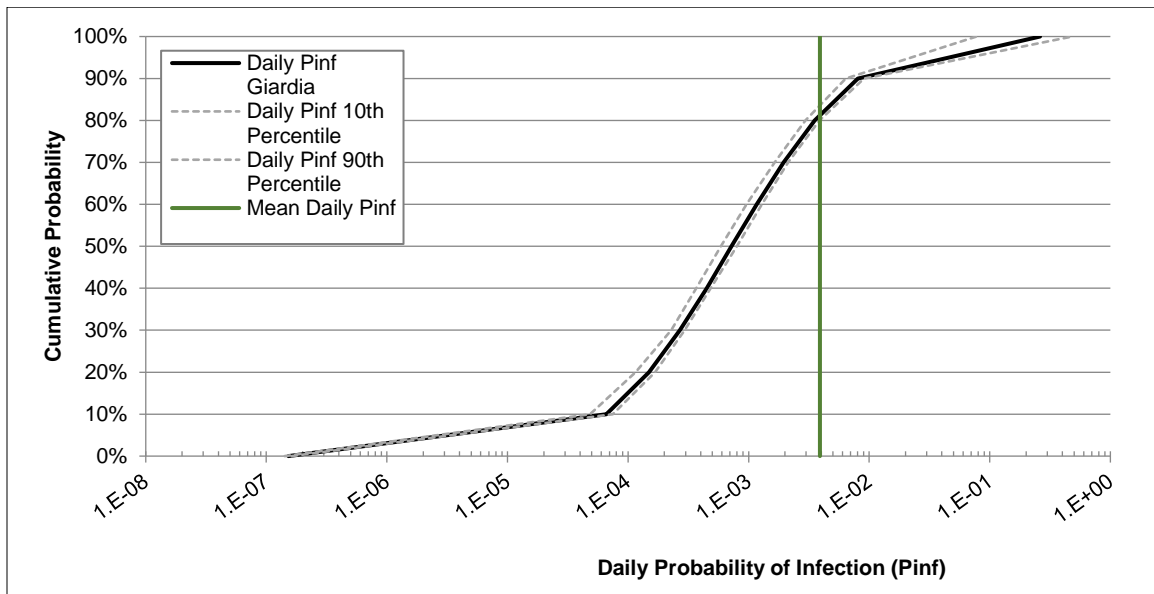


FIGURE E9: The cumulative distribution of the daily risk of infection for *Giardia* associated with ingestion of tap water supplied by an IWS in India.

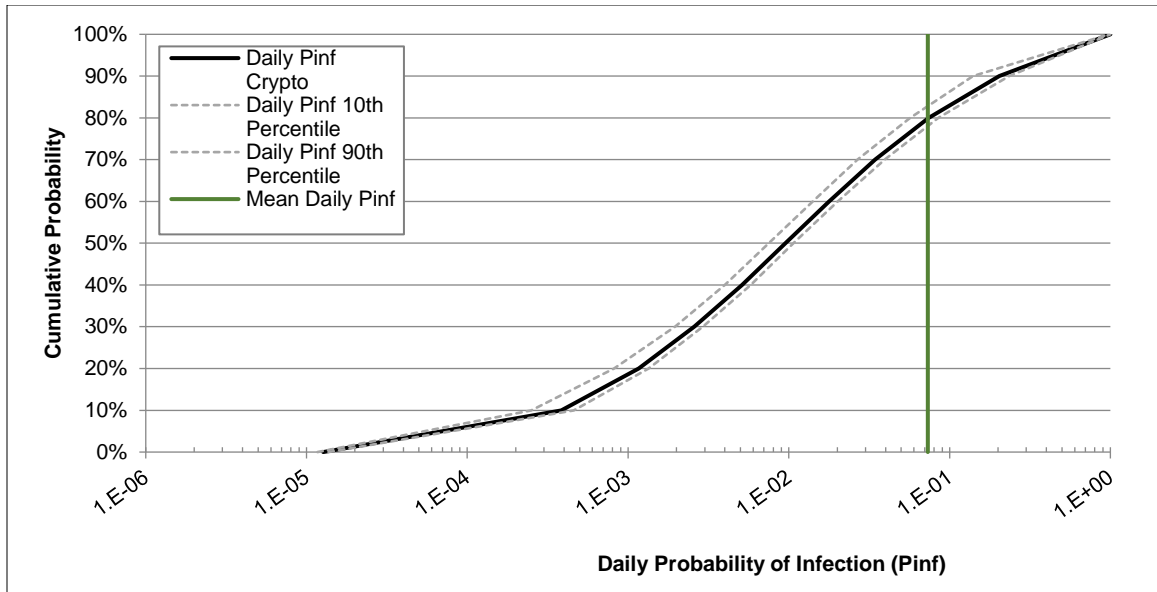


FIGURE E10: The cumulative distribution of the daily risk of infection for *Cryptosporidium* associated with ingestion of tap water supplied by an IWS in India.

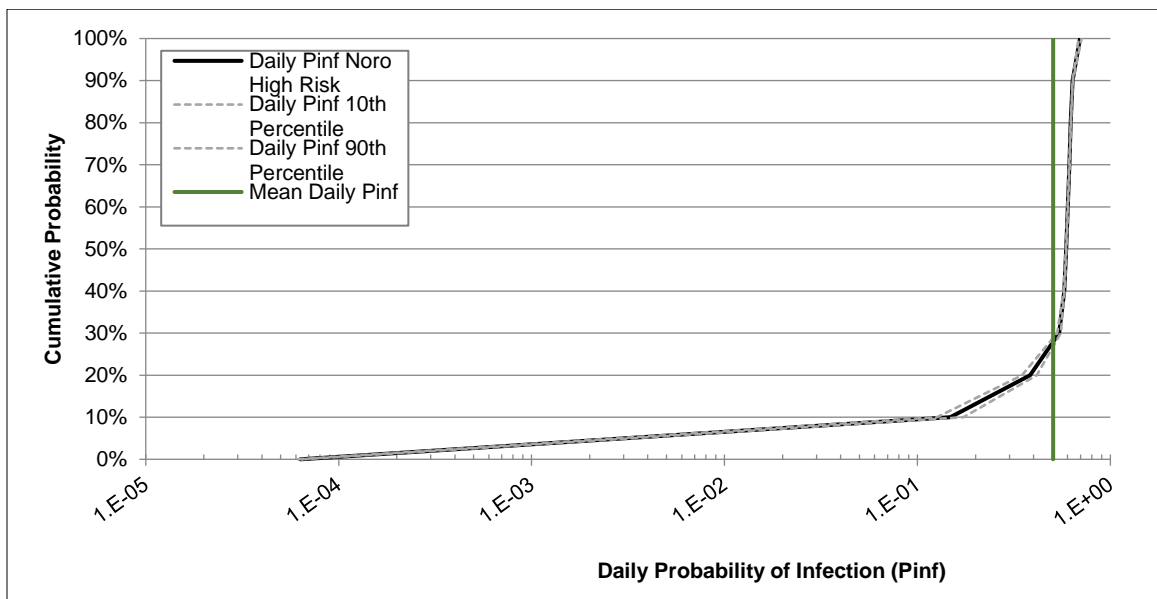


FIGURE E11: The cumulative distribution of the daily risk of infection for norovirus associated with ingestion of tap water supplied by an IWS in India as estimated by the high-risk dose-response model.

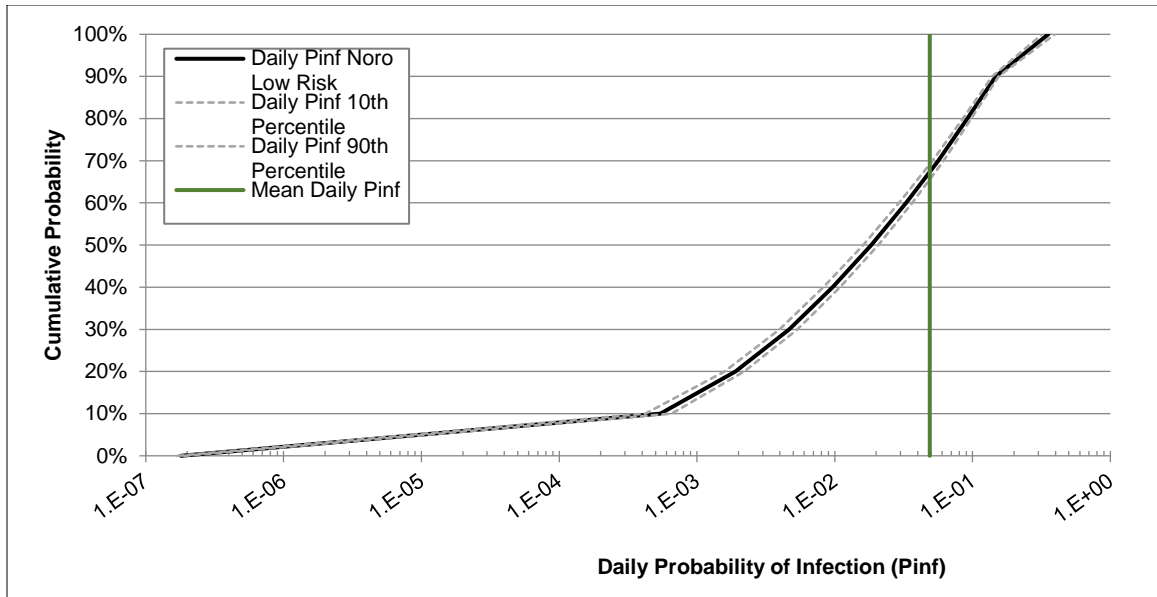


FIGURE E12: The cumulative distribution of the daily risk of infection for norovirus associated with ingestion of tap water supplied by an IWS in India as estimated by the low-risk dose-response model.

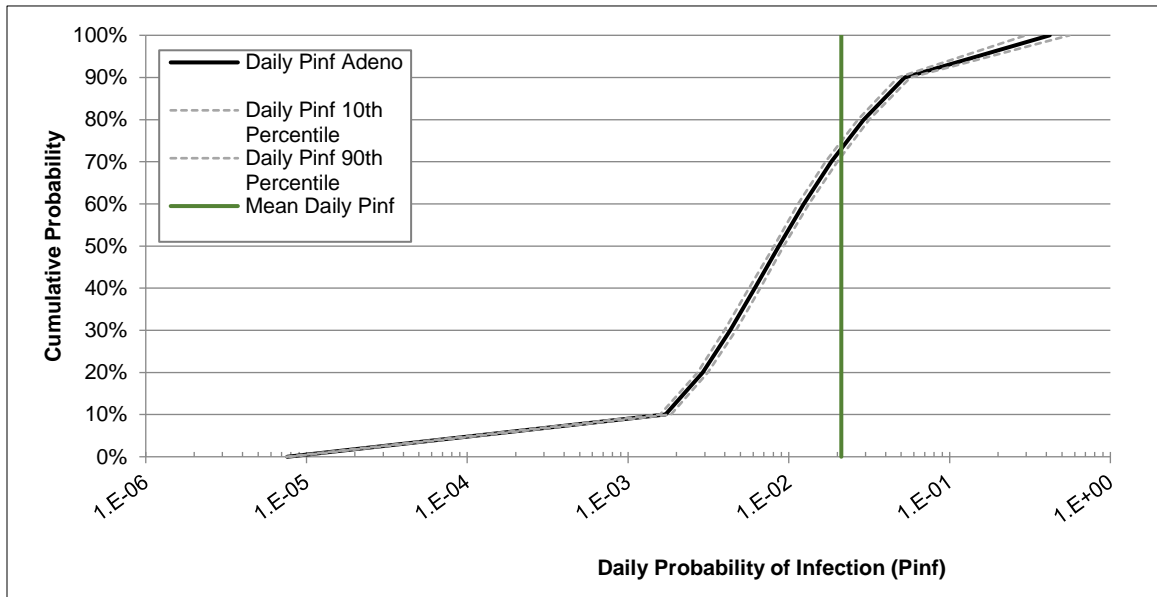


FIGURE E13: The cumulative distribution of the daily risk of infection for adenovirus associated with ingestion of tap water supplied by an IWS in India.

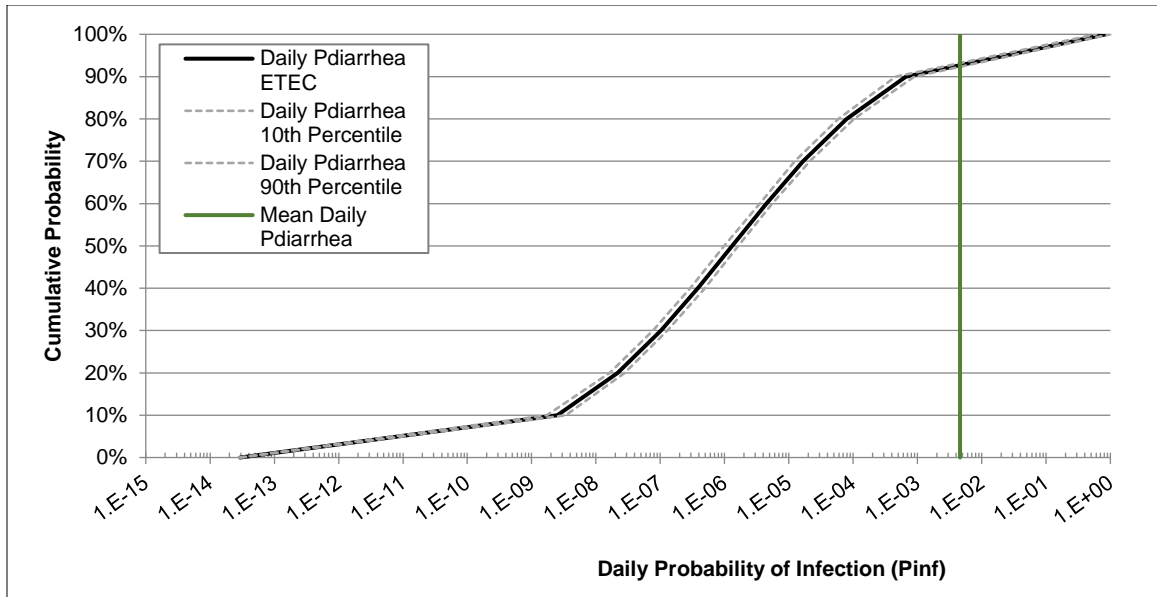


FIGURE E14: The cumulative distribution of the daily risk of diarrhea from infection with ETEC associated with ingestion of tap water supplied by an IWS in India.

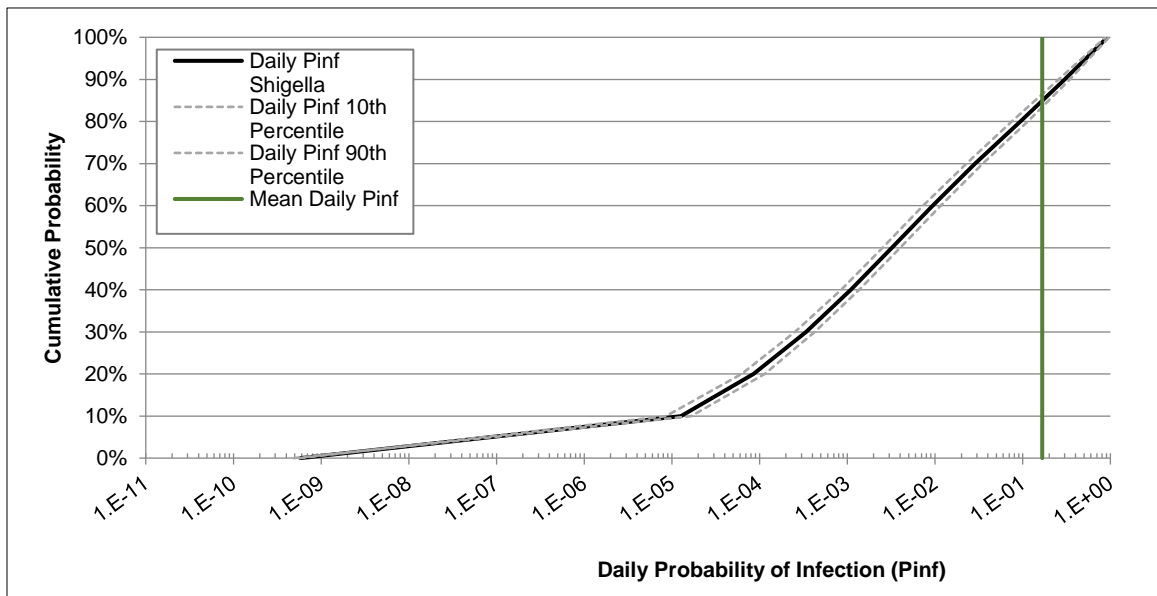


FIGURE E15: The cumulative distribution of the daily risk of infection with *Shigella* associated with ingestion of tap water supplied by an IWS in India.

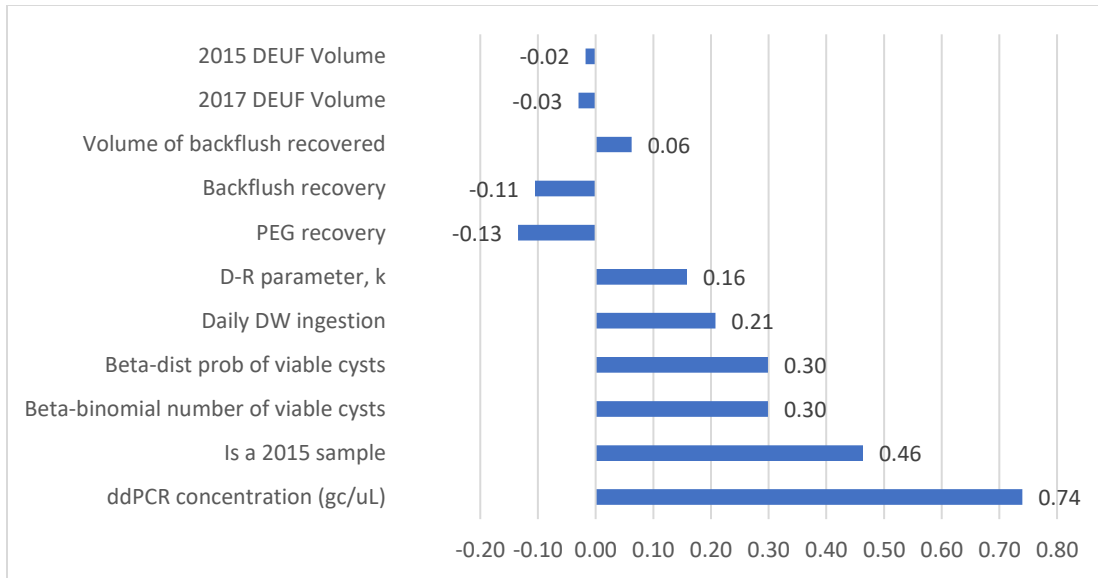


FIGURE E16: The sensitivity of the estimated risk of infection with *Giardia* to model input parameters as assessed by rank order correlation.

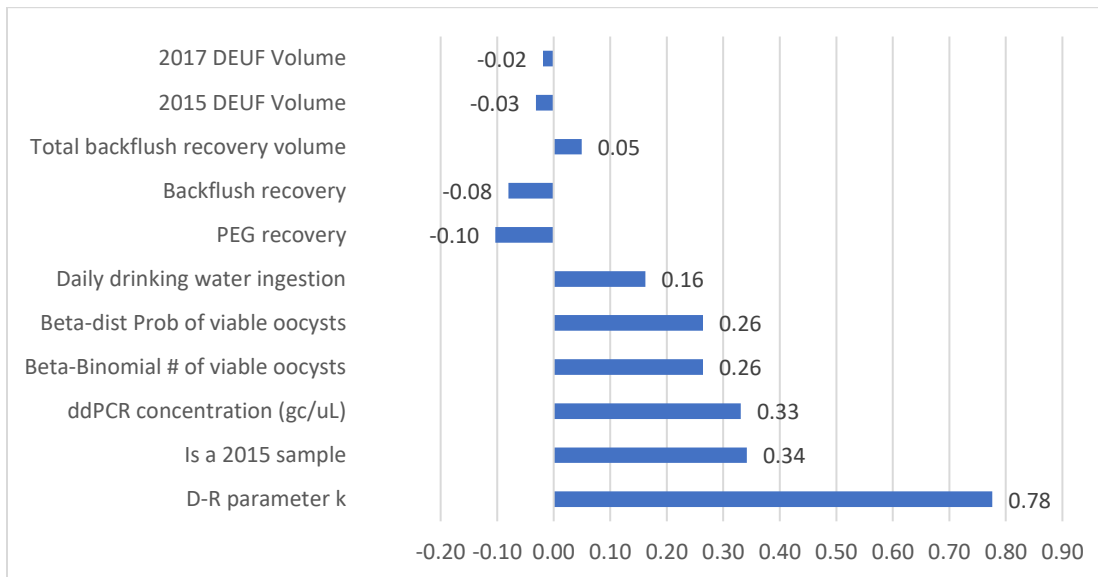


FIGURE E17: The sensitivity of the estimated risk of infection with *Cryptosporidium* to model input parameters as assessed by rank order correlation.

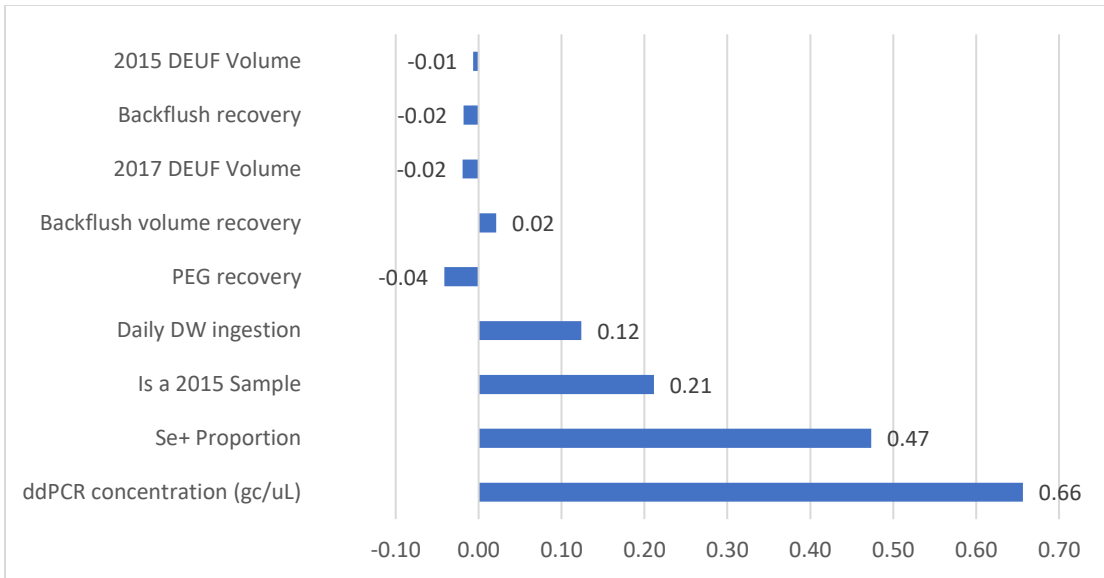


FIGURE E18: The sensitivity of the estimated risk of infection with norovirus (high-risk model) to input parameters as assessed by rank order correlation.

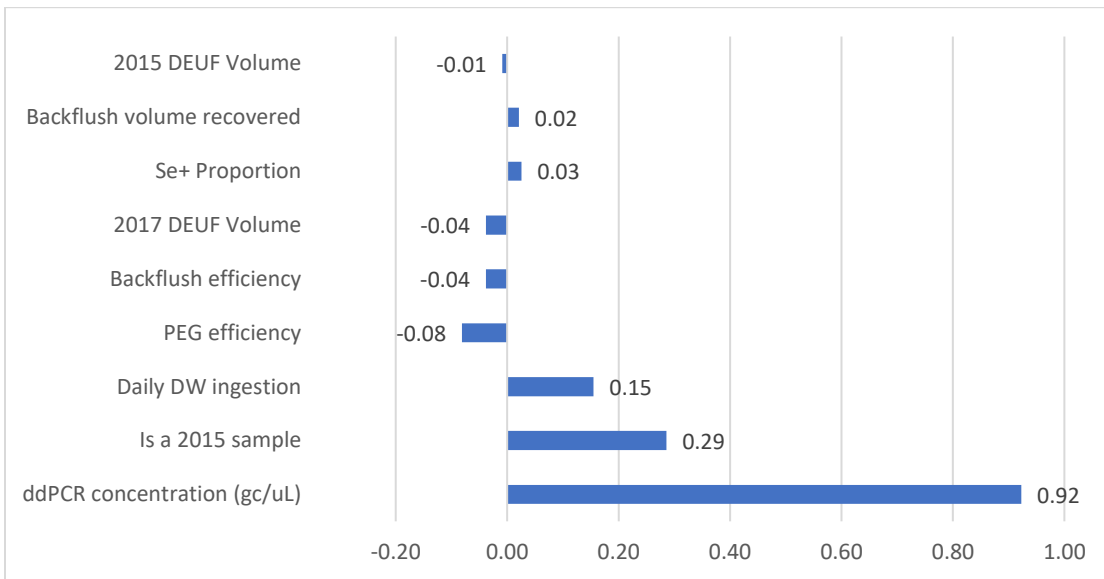


FIGURE E19: The sensitivity of the estimated risk of infection with norovirus (low-risk model) to input parameters as assessed by rank order correlation.

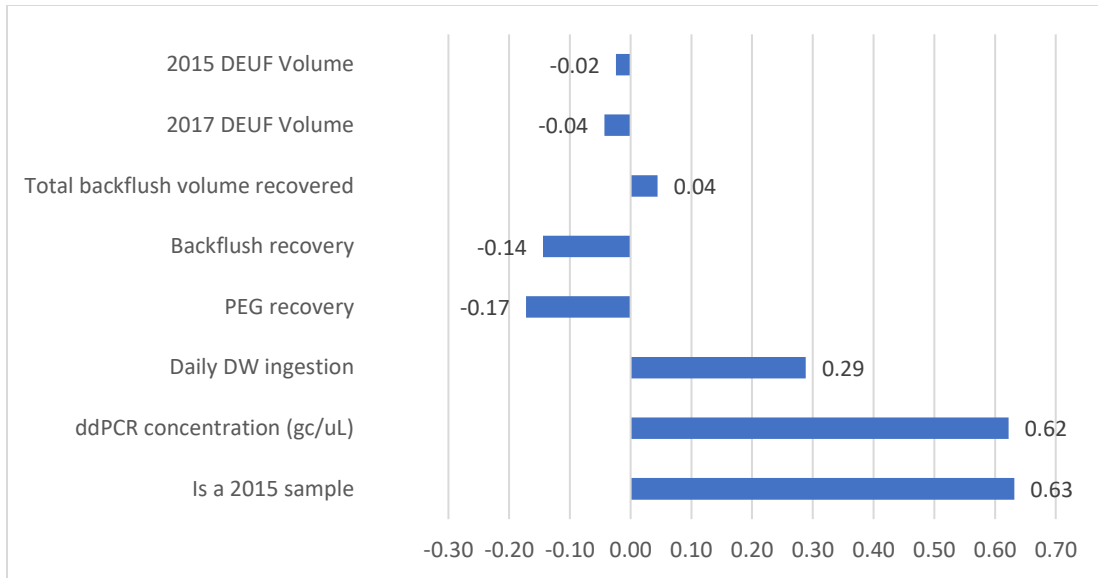


FIGURE E20: The sensitivity of the estimated risk of infection with adenovirus to input parameters as assessed by rank order correlation.

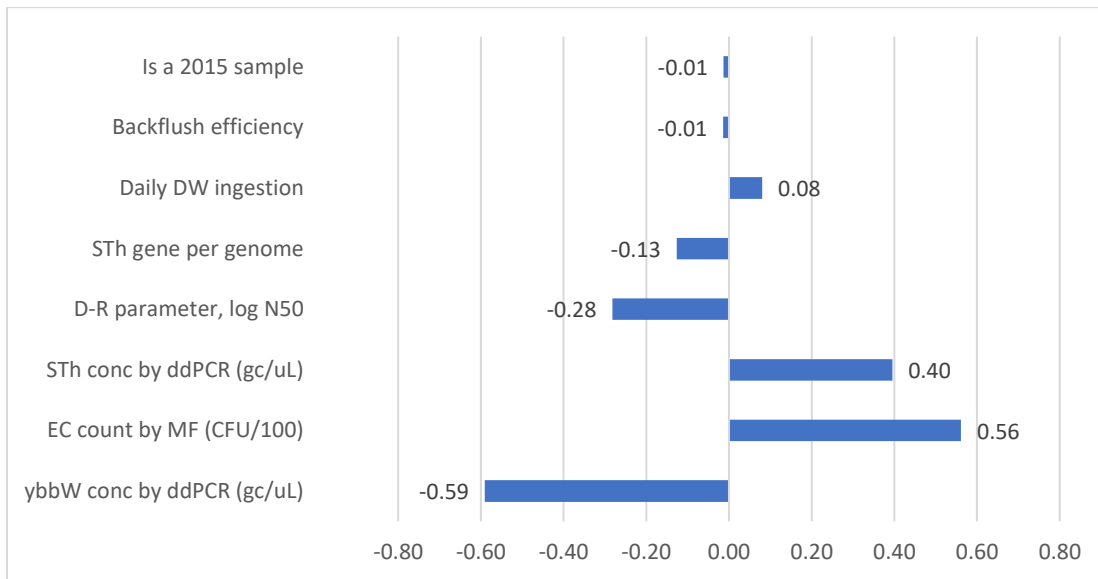


FIGURE E21: The sensitivity of the estimated risk of infection with ETEC to input parameters as assessed by rank order correlation.

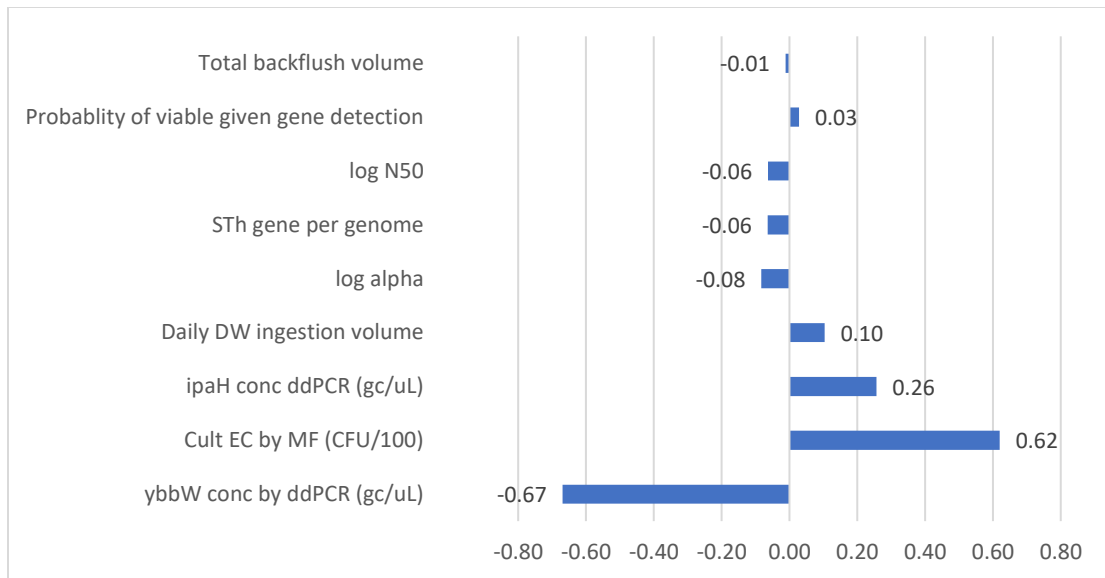


FIGURE E22: The sensitivity of the estimated risk of infection with *Shigella* to input parameters as assessed by rank order correlation.

REFERENCES

1. Hansen RD. Water and Waste Water in Imperial Rome. *JAWRA J Am Water Resour Assoc.* 1983;19(2):263–9.
2. WHO/UNICEF. 2015 Update and MDG Assessment. *World Heal Organ.* 2015;90.
3. WHO, UNICEF. Progress on Drinking Water, Sanitation and Hygiene: 2017 Update and SDG Baseline [Internet]. Geneva, Switzerland; 2017. Available from: http://apps.who.int/bookorders.%0Ahttp://apps.who.int/bookorders.%0Ahttps://www.unicef.org/publications/files/Progress_on_Drinking_Water_Sanitation_and_Hygiene_2017.pdf
4. National Research Council. *Drinking Water Distribution Systems: Assessing and Reducing Risks.* National Academies Press. 2006. 404 p.
5. MacKenzie WR, Hoxie NJ, Proctor ME, Gradus MS, Blair KA, Peterson DE, et al. A massive outbreak in Milwaukee of *Cryptosporidium* infection transmitted through the public water supply. *N Engl J Med.* 1994;331(3):161–7.
6. Shaheed A, Orgill J, Montgomery MA, Jeuland MA, Brown J. Why “improved” water sources are not always safe. *Bull World Health Organ.* 2014;92(4):283–9.
7. Bain R, Cronk R, Hossain R, Bonjour S, Onda K, Wright J, et al. Global assessment of exposure to faecal contamination through drinking water based on a systematic review. *Trop Med Int Heal [Internet].* 2014/05/09. 2014 Aug;19(8):917–27. Available from: <http://doi.wiley.com/10.1111/tmi.12334>
8. Payment P, Richardson L, Siemiatycki J, Dewar R, Edwardes M, Franco E. A randomized trial to evaluate the risk of gastrointestinal disease due to consumption of drinking water meeting current microbiological standards. *Am J Public Heal [Internet].* 1991/06/01. 1991;81(6):703–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2029037>
9. Payment P, Siemiatycki J, Renaud G, Franco E, Prevost M. A prospective epidemiological study of gastrointestinal health effects due to the consumption of drinking water. *Int J Environ Health Res.* 1997;7:5–31.
10. Hellard ME, Sinclair MI, Forbes AB, Fairley CK. A randomized, blinded, controlled trial investigating the gastrointestinal health effects of drinking water quality. *Env Heal Perspect [Internet].* 2001/09/21. 2001;109(8):773–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11564611>
11. Colford JM, Rees JR, Wade TJ, Khalakdina A, Hilton JF, Ergas IJ, et al. Participant blinding and gastrointestinal illness in a randomized, controlled trial of an in-home drinking water intervention. *Emerg Infect Dis.* 2002;8(1):29–36.

12. Colford JM, Saha SR, Wade TJ, Wright CC, Vu M, Charles S, et al. A pilot randomized, controlled trial of an in-home drinking water intervention among HIV + persons. *J Water Health*. 2005;3(2):173–84.
13. Colford JM, Roy S, Beach MJ, Hightower A, Shaw SE, Wade TJ. A review of household drinking water intervention trials and an approach to the estimation of endemic waterborne gastroenteritis in the United States. *J Water Health* [Internet]. 2006 Jul;04(Suppl 2):71. Available from: <http://jwh.iwaponline.com/cgi/doi/10.2166/wh.2006.018>
14. Ercumen A, Gruber JS, Colford JM. Water Distribution System Deficiencies and Gastrointestinal Illness: A Systematic Review and Meta-Analysis. *Environ Health Perspect* [Internet]. 2014/03/25. 2014 Jul;122(7):651–60. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24659576>
15. Galaitsi SE, Russell R, Bishara A, Durant JL, Bogle J, Huber-Lee A. Intermittent domestic water supply: A critical review and analysis of causal-consequential pathways. *Water (Switzerland)*. 2016;8(7).
16. Kumpel E, Nelson KL. Intermittent Water Supply: Prevalence, Practice, and Microbial Water Quality. *Env Sci Technol*. 2015/12/17. 2016;50(2):542–53.
17. Danilenko A, van den Berg C, Macheve B, Moffitt LJ. *The IBNET Water Supply and Sanitation Blue Book 2014*. Washington, D.C.; 2014.
18. PAHO, WHO. *Regional Report on the Evaluation 2000 in the Region of the Americas Water Supply and Sanitation Current Status and Prospects* [Internet]. 2001. Available from: <http://www.paho.org/English/HEP/HES/Reg-rep.pdf>
19. WHO, UNICEF. *Global Water Supply and Sanitation Assessment 2000 Report*. Geneva, Switzerland; 2000.
20. WHO. *Progress on Sanitation and Drinking Water 2015 Update and MDG Assessment*. 2015.
21. Mellor JE, Levy K, Zimmerman J, Elliott M, Bartram J, Carlton E, et al. Planning for climate change: The need for mechanistic systems-based approaches to study climate change impacts on diarrheal diseases. *Sci Total Environ* [Internet]. 2016;548–549(January):82–90. Available from: <http://dx.doi.org/10.1016/j.scitotenv.2015.12.087>
22. Kumpel E, Nelson KL. *Mechanisms Affecting Water Quality in an Intermittent Piped Water Supply*. 2014;
23. Coelho ST, James S, Sunna N, Abu Jaish A, Chatia J. Controlling water quality in intermittent supply systems. *Water Sci Technol Water Supply*. 2003;3(1–2):119–25.
24. Kumpel E, Nelson KL. *Comparing microbial water quality in an intermittent and*

- continuous piped water supply. *Water Res* [Internet]. 2013;47(14):5176–88. Available from: <http://dx.doi.org/10.1016/j.watres.2013.05.058>
25. Lee EJ, Schwab KJ. Deficiencies in drinking water distribution systems in developing countries. *J Water Heal*. 2005/08/04. 2005;3(2):109–27.
 26. Mermin JH, Villar R, Carpenter J, Roberts L, Samariddin a, Gasanova L, et al. A massive epidemic of multidrug-resistant typhoid fever in Tajikistan associated with consumption of municipal water. *J Infect Dis*. 1999;179(6):1416–22.
 27. Swerdlow DL, Mintz ED, Rodriguez M, Tejada E, Ocampo C, Espejo L, et al. Waterborne Transmission of Epidemic Cholera in Trujillo, Peru: Lessons for a Continent at Risk. *Lancet*. 1992;340:28–32.
 28. Jeandron A, Saidi JM, Kapama A, Burhole M, Birembano F, Vandeveld T, et al. Water supply interruptions and suspected cholera incidence: a time-series regression in the Democratic Republic of the Congo. *PLoS Med*. 2015/10/28. 2015;12(10):e1001893.
 29. Ercumen A, Arnold BF, Kumpel E, Burt Z, Ray I, Nelson K, et al. Upgrading a piped water supply from intermittent to continuous delivery and association with waterborne illness: a matched cohort study in urban India. *PLoS Med*. 2015/10/28. 2015;12(10):e1001892.
 30. Haas CN, Rose JB, Gerba CP. *Quantitative Microbial Risk Assessment* [Internet]. Hoboken, New Jersey: John Wiley & Sons, Inc; 2014 [cited 2018 Mar 5]. Available from: <http://doi.wiley.com/10.1002/9781118910030>
 31. Rose JB, Haas CN, Regli S. Risk Assessment and Control of Waterborne Giardiasis. *Am J Public Health*. 1991;81(6):709–13.
 32. Haas CN. Acceptable Microbial Risk. *J Am Water Works Assoc*. 1996;88:8.
 33. Haas CN, Rose JB, Gerba C, Regli S. Risk Assessment of Virus in Drinking Water. *Risk Anal* [Internet]. 1993;13(5):545–52. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8259444>
 34. Mena KD, Gerba CP, Haas CN, Rose JB. Risk Assessment Of Waterborne Coxsackievirus. *Am Water Work Assoc*. 2003;95(7):122–31.
 35. Storey M V, Ashbolt NJ, Stenström TA. Biofilms, thermophilic amoebae and *Legionella pneumophila* - a quantitative risk assesment for distributed water. *Water Sci Technol*. 2004;50(1):77–82.
 36. World Health Organization. WHO guidelines for drinking water quality. 2011;4th Editio(July).
 37. WHO. *Quantitative Microbial Risk Assessment: Application for Water Safety*

Management. Geneva, Switzerland: WHO; 2016. 204 p.

38. Petterson SR, Ashbolt NJ. QMRA and water safety management: Review of application in drinking water systems. *J Water Health*. 2016;14(4):571–89.
39. Rose JB, Gerba CP. Use of Risk Assessment for Development of Microbial Standards. 1991;24(2):29–34. Available from: <http://wst.iwaponline.com/content/24/2/29>
40. Macler BA, Regli S. Use of microbial risk assessment in setting US drinking water standards. *Int J Food Microbiol*. 1993;18:245–56.
41. United States Environmental Protection Agency. Surface Water Treatment Rule. *Fed Regist*. 1989;54(124):27485–541.
42. van Lieverloo JHM, Blokker EJM, Medema G. Quantitative microbial risk assessment of distributed drinking water using faecal indicator incidence and concentrations. *J Water Health [Internet]*. 2007 Sep;5(Suppl 1):S131. Available from: <http://jwh.iwaponline.com/cgi/doi/10.2166/wh.2007.134>
43. Lambertini E, Borchardt MA, Kieke BA, Spencer SK, Loge FJ. Risk of viral acute gastrointestinal illness from nondisinfected drinking water distribution systems. *Environ Sci Technol*. 2012;46(17):9299–307.
44. Teunis PFM, Medema GJ, Kruidenier L, Havelaar AH. Assessment of the risk of infection by *Cryptosporidium* or *Giardia* in drinking water from a surface water source. BA Bilthoven, The Netherlands; 1998.
45. Astrom J. Evaluation of the microbial risk reduction due to selective closure of water water intake before drinking water treatment. *J Water Heal*. 2016;5(1):S81–97.
46. Simmons G, Heyworth J, Rimajova M. Assessing the Microbial Health Risks of Tank Rainwater Used for Drinking Water. *Environ Heal*. 2001;1(3):57–64.
47. Borchardt MA, Bertz PD, Spencer SK, Battigelli DA. Incidence of enteric viruses in groundwater from household wells in Wisconsin. *Appl Env Microbiol [Internet]*. 2003;69(2):1172–80. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12571044>
48. Daniels ME, Smith WA, Jenkins MW. Estimating *Cryptosporidium* and *Giardia* disease burdens for children drinking untreated groundwater in a rural population in India. 2018;1–23.
49. Hunter PR, Zmirou-Navier D, Hartemann P. Estimating the impact on health of poor reliability of drinking water interventions in developing countries. *Sci Total Environ [Internet]*. 2009;407(8):2621–4. Available from: <http://dx.doi.org/10.1016/j.scitotenv.2009.01.018>

50. Gerba CP, Rose JB, Haas CN, Crabtree KD. Waterborne rotavirus: A risk assessment. *Water Res.* 1996;30(12):2929–40.
51. Westrell T, Bergstedt O, Stenström T, Ashbolt N. A theoretical approach to assess microbial risks due to failures in drinking water systems. *Int J Environ Health Res* [Internet]. 2003 Jun;13(2):181–97. Available from: <http://www.tandfonline.com/doi/abs/10.1080/0960312031000098080>
52. Howard G, Pedley S, Tibatemwa S. Quantitative microbial risk assessment to estimate health risks attributable to water supply: Can the technique be applied in developing countries with limited data? *J Water Health.* 2006;4(1):49–65.
53. Machdar E, van der Steen NP, Raschid-Sally L, Lens PNL. Application of Quantitative Microbial Risk Assessment to analyze the public health risk from poor drinking water quality in a low income area in Accra, Ghana. *Sci Total Environ* [Internet]. 2013 Apr;449:134–42. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0048969713001034>
54. National Research Council. *Science and Decisions: Advancing Risk Assessment* [Internet]. Washington, D. C.: The National Academies Press; 2009. 423 p. Available from: <http://www.nap.edu/catalog/12209>
55. Medema G, Teunis P, Blokker M, Deere D, Davison A, Charles P, et al. Risk Assessment of Cryptosporidium in Drinking Water. 2009.
56. Organization WH. Evaluating household water treatment options: Health-based targets and microbiological performance specifications. *World Heal Organ Publ.* 2011;68.
57. Jones IG, Roworth M. An outbreak of *Escherichia coli* 0157 and *Campylobacteriosis* associated with contamination of a drinking water supply. *Public Health.* 1996;110(5):277–82.
58. Holme R. Drinking water contamination in Walkerton, Ontario: Positive resolutions from a tragic event. *Water Sci Technol.* 2003;47(3):1–6.
59. Black RE, Levine MM, Clements M Lou, Hughes TP, Blaser MJ. Experimental *Campylobacter jejuni* Infection in Humans. *J Infect Dis.* 1988;157(3):472–9.
60. Chin J, editor. *Control of Communicable Diseases Manual*. 17th ed. Washington D. C.: American Public Health Association; 2000.
61. Leclerc H, Schwartzbrod L, Dei-Cas E. Microbial agents associated with waterborne diseases. *Crit Rev Microbiol* [Internet]. 2002;28(4):371–409. Available from: <https://www.scopus.com/record/display.uri?eid=2-s2.0-0036964601&origin=resultslist&sort=plf-f&src=s&st1=rotavirus&st2=drinking+water&nlo=&nlr=&nls=&sid=4611C10EA636ACA92D986B92B3067FEC.WIW7NKKKC52nnQNxjqAQrlA:570&sot=b&sdt>

=cl&cluster=scosubtype,%22re%22,t&

62. Sutmoller F, Azeredo RS, Lacerda MD, Barth OM, Pereira HG, Hoffer E, et al. An outbreak of gastroenteritis caused by both rotavirus and *Shigella sonnei* in a private school in Rio de Janeiro. *J Hyg (Lond)*. 1982;88(2):285–93.
63. Hopkins RS, Gaspard GB, Williams FP, Karlin RJ, Cukor G, Blacklow NR. A community waterborne gastroenteritis outbreak: Evidence for rotavirus as the agent. *Am J Public Health*. 1984;74(3):263–5.
64. Tao H, Changan W, Zhaoying F, Zinyi C, Xuejian C, Xiaoquang L, et al. WATERBORNE OUTBREAK OF ROTAVIRUS DIARRHOEA IN ADULTS IN CHINA CAUSED BY A NOVEL ROTAVIRUS. *Lancet [Internet]*. 1984 May;323(8387):1139–42. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0140673684913916>
65. Ward RL, Bernstein DI, Young EC, Sherwood JR, Knowlton DR, Schiff GM. Human rotavirus studies in volunteers: determination of infectious dose and serological response to infection. *J Infect Dis [Internet]*. 1986;154(5):871–80. Available from: <http://jid.oxfordjournals.org/content/154/5/871.full.pdf>
66. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. *Lancet*. 2013;382(9888):209–22.
67. Liu J, Platts-Mills JA, Juma J, Kabir F, Nkeze J, Okoi C, et al. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. *Lancet [Internet]*. 2016;388(10051):1291–301. Available from: [http://dx.doi.org/10.1016/S0140-6736\(16\)31529-X](http://dx.doi.org/10.1016/S0140-6736(16)31529-X)
68. Platts-Mills JA, Babji S, Bodhidatta L, Gratz J, Haque R, Havt A, et al. Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob Heal [Internet]*. 2015 Sep;3(9):e564–75. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S2214109X15001515>
69. Medema GJ, Teunis PFM, Havelaar AH, Haas CN. Assessment of the dose-response relationship of *Campylobacter jejuni*. *Int J Food Microbiol*. 1996;30(1–2):101–11.
70. Messner MJ, Chappell CL, Okhuysen PC. Risk assessment for *Cryptosporidium*: A hierarchical Bayesian analysis of human dose response data. *Water Res*. 2001;35(16):3934–40.
71. Rose JB, Wendt C, Weir MH. Microbial dose response models monograph. Vols. I & II. East Lansing, MI, USA; 2014.
72. Vairavamorthy K, Yan J, Gorantiwar SD. Modelling the risk of contaminant

- intrusion in water mains. *Proc Inst Civ Eng - Water Manag* [Internet]. 2007 Jun;160(2):123–32. Available from: <http://www.icevirtuallibrary.com/doi/10.1680/wama.2007.160.2.123>
73. US Environmental Protection Agency. *Exposure Factors Handbook: 2011 Edition*. US Environ Prot Agency [Internet]. 2011;EPA/600/R-(September):1–1466. Available from: c:%5CDocuments and Settings%5Cturner_j%5CDesktop%5CJT_Biblioscope_8_111003%5CJT_Bib8_111003%5Cattachments%5Cefh-complete.pdf
 74. Brown J, Hien VT, McMahan L, Jenkins MW, Thie L, Liang K, et al. Relative benefits of on-plot water supply over other “improved” sources in rural Vietnam. *Trop Med Int Heal*. 2013;18(1):65–74.
 75. Shaheed A, Orgill J, Ratana C, Montgomery MA, Jeuland MA, Brown J. Water quality risks of ‘improved’ water sources: evidence from Cambodia. *Trop Med Int Heal* [Internet]. 2014 Feb;19(2):186–94. Available from: <http://doi.wiley.com/10.1111/tmi.12229>
 76. Helsel DR. Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere* [Internet]. 2006 Dec;65(11):2434–9. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0045653506005157>
 77. Hoogenboezem W, Ketelaars H, Medema G, Rijs G, Schijven J. *Cryptosporidium en Giardina: voorkomen in rioolwater, mest en oppervlaktewater met zwem-en drinkwaterfunctie*. Amsterdam RIWA 2001. 2001;
 78. Holler C. Quantitative und qualitative Untersuchungen an *Campylobacter* in der Kanalisation einer Grobstadt. *Zentrabl Bakteriol Hyg B*. 1988;185((4-5)):307–25.
 79. WHO; UNICEF. *Data and estimates. Joint Monitoring Programme (JMP) for Water Supply and Sanitation* [Online] [Internet]. [cited 2015 Oct 15]. Available from: <https://www.wssinfo.org/data-estimates/>
 80. World Bank. *Benchmarking Database*. IBNET: International Benchmarking Network [Online] [Internet]. [cited 2015 May 14]. Available from: <https://database.ib-net.org/DefaultNew.aspx>
 81. A.H. Havelaar, J.M.Melse. *Quantifying Public Health Risk in the WHO Guidelines for Drinking Water Quality*. WHO RIVM Rep. 2003;1–49.
 82. Erickson JJ, Smith CD, Goodridge A, Nelson KL. Water quality effects of intermittent water supply in Arraiján, Panama. *Water Res* [Internet]. 2017;114:338–50. Available from: <http://dx.doi.org/10.1016/j.watres.2017.02.009>
 83. Wu J, Long SC, Das D, Dorner SM. Are microbial indicators and pathogens correlated? A statistical analysis of 40 years of research. *J Water Health* [Internet]. 2011 Jun;9(2):265. Available from:

<http://jwh.iwaponline.com/cgi/doi/10.2166/wh.2011.117>

84. Sedmak G, Bina D, MacDonald J, Couillard L. The occurrence of culturable viruses in source water for two drinking water treatment plants and the influent and effluent of a wastewater treatment plant in Milwaukee,. *Appl Environ Microbiol* [Internet]. 2005;71(2):1042–50. Available from: <http://aem.asm.org/content/71/2/1042.short>
85. Hellmér M, Paxéus N, Magnius L, Enache L, Arnholm B, Johansson A, et al. Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and norovirus outbreaks. *Appl Environ Microbiol*. 2014;80(21):6771–81.
86. Feachem R, Mara DD, Bradley DJ. *Sanitation and Disease*. Washington D. C.: John Wiley & Sons, Inc; 1983.
87. Silverman AI, Akrong MO, Amoah P, Drechsel P, Nelson KL. Quantification of human norovirus GII, human adenovirus, and fecal indicator organisms in wastewater used for irrigation in Accra, Ghana. *J Water Health* [Internet]. 2013 Sep;11(3):473–88. Available from: <https://iwaponline.com/jwh/article/11/3/473/1689/Quantification-of-human-norovirus-GII-human>
88. Elala D, Labhasetwar P, Tyrrel SF. Deterioration in water quality from supply chain to household and appropriate storage in the context of intermittent water supplies. *Water Sci Technol Water Supply*. 2011;11(4):400–8.
89. John V, Jain P, Rahate M, Labhasetwar P. Assessment of deterioration in water quality from source to household storage in semi-urban settings of developing countries. *Environ Monit Assess* [Internet]. 2014 Feb 19;186(2):725–34. Available from: <http://link.springer.com/10.1007/s10661-013-3412-z>
90. Ngure FM, Reid BM, Humphrey JH, Mbuya MN, Pelto G, Stoltzfus RJ. Water, sanitation, and hygiene (WASH), environmental enteropathy, nutrition, and early child development: making the links. *Ann N Y Acad Sci* [Internet]. 2014 Jan;1308(1):118–28. Available from: <http://doi.wiley.com/10.1111/nyas.12330>
91. Korpe PS, Petri WA. Environmental enteropathy: Critical implications of a poorly understood condition. *Trends Mol Med* [Internet]. 2012;18(6):328–36. Available from: <http://dx.doi.org/10.1016/j.molmed.2012.04.007>
92. Munos MK, Fischer Walker CL, Black RE. The effect of oral rehydration solution and recommended home fluids on diarrhoea mortality. *Int J Epidemiol*. 2010;39(SUPPL. 1).
93. WHO. *Safe Piped Water: Managing Microbial Water Quality in Piped Distribution Systems*. Ainsworth R, editor. London, UK: IWS Publishing; 2004. 168 p.
94. Brown J, Clasen T. High adherence is necessary to realize health gains from water quality interventions. *PLoS One*. 2012;7(5):e36735.

95. Burch T. Validation of Quantitative Microbial Risk Assessment Using Epidemiological Data from Outbreaks of Waterborne Gastrointestinal Disease. *Risk Anal* [Internet]. 2018; Available from: <http://doi.wiley.com/10.1111/risa.13189>
96. Besner MC, Prévost M, Regli S. Assessing the public health risk of microbial intrusion events in distribution systems: Conceptual model, available data, and challenges. *Water Res.* 2011;45(3):961–79.
97. Bivins AW, Sumner T, Kumpel E, Howard G, Cumming O, Ross I, et al. Estimating Infection Risks and the Global Burden of Diarrheal Disease Attributable to Intermittent Water Supply Using QMRA. *Environ Sci Technol* [Internet]. 2017 Jul 5;51(13):7542–51. Available from: <http://pubs.acs.org/doi/abs/10.1021/acs.est.7b01014>
98. George J, An W, Joshi D, Zhang D, Yang M, Suriyanarayanan S. Quantitative Microbial Risk Assessment to Estimate the Health Risk in Urban Drinking Water Systems of Mysore, Karnataka, India. *Water Qual Expo Heal* [Internet]. 2015;7(3):331–8. Available from: <http://dx.doi.org/10.1007/s12403-014-0152-4>
<http://link.springer.com/10.1007/s12403-014-0152-4>
99. Blokker M, Smeets P, Medema G. QMRA in the drinking water distribution system. *Procedia Eng* [Internet]. 2014;89:151–9. Available from: <http://dx.doi.org/10.1016/j.proeng.2014.11.171>
100. World Health Organization. Acceptable risk. Hunter PR, Fewtrell L, editors. *Water Quality: Guidelines, Standards and Health*. London, UK: IWA Publishing; 2001.
101. US EPA. Method 1602 : Male-specific (F +) and Somatic Coliphage in Water by Single Agar Layer (SAL) Procedure April 2001. EPA Document 821-R-01-029. Washington D. C.; 2001.
102. US EPA. Method 1601: Male-specific (F+) Coliphage in Water by Two-Step Enrichment Procedure. Washington D. C.; 2001. 1–30 p.
103. US EPA. Method 1603: Escherichia coli (E. coli) in Water by Membrane Filtration Using Modified membrane-Thermotolerant Escherichia coli Agar (Modified mTEC). Washington D. C.; 2014.
104. US EPA. Method 1604: Total Coliforms and Escherichia coli in Water by Membrane Filtration Using a Simultaneous Detection Technique (MI Medium) [Internet]. 2002. Available from: <http://www.epa.gov/nerlcwww/1604sp02.pdf>
105. US EPA. Method 1600 : Enterococci in Water by Membrane Filtration Using membrane-Enterococcus Indoxyl-\$-D-Glucoside Agar (mEI). 2002;(September).
106. Polaczyk AL, Narayanan J, Cromeans TL, Hahn D, Roberts JM, Amburgey JE, et al. Ultrafiltration-based techniques for rapid and simultaneous concentration of

- multiple microbe classes from 100-L tap water samples. *J Microbiol Methods* [Internet]. 2008 May 1 [cited 2018 Feb 9];73(2):92–9. Available from: <https://www.sciencedirect.com/science/article/pii/S0167701208000456?via%3Dihub>
107. Hill VR, Polaczyk AL, Hahn D, Narayanan J, Cromeans TL, Roberts JM, et al. Development of a rapid method for simultaneous recovery of diverse microbes in drinking water by ultrafiltration with sodium polyphosphate and surfactants. *Appl Env Microbiol* [Internet]. 2005;71(11):6878–84. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16269722>
 108. Hill VR, Kahler AM, Jothikumar N, Johnson TB, Hahn D, Cromeans TL. Multistate evaluation of an ultrafiltration-based procedure for simultaneous recovery of enteric microbes in 100-liter tap water samples. *Appl Env Microbiol* [Internet]. 2007;73(13):4218–25. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17483281>
 109. Sharma MM, Chang YI, Yen TF. Reversible and irreversible surface charge modification of bacteria for facilitating transport through porous media. *Colloids and Surfaces*. 1985;16:193–206.
 110. Morales-Morales HA, Vidal G, Olszewski J, Rock CM, Dasgupta D, Oshima KH, et al. Optimization of a reusable hollow-fiber ultrafilter for simultaneous concentration of enteric bacteria, protozoa, and viruses from water. *Appl Env Microbiol* [Internet]. 2003;69(7):4098–102. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12839786>
 111. Smith CM, Hill VR. Dead-end hollow-fiber ultrafiltration for recovery of diverse microbes from water. *Appl Env Microbiol* [Internet]. 2009/06/30. 2009;75(16):5284–9. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2725477/pdf/0456-09.pdf>
 112. Mull B, Hill VR. Recovery of diverse microbes in high turbidity surface water samples using dead-end ultrafiltration. *J Microbiol Methods* [Internet]. 2012 Dec;91(3):429–33. Available from: <http://dx.doi.org/10.1016/j.mimet.2012.10.001>
 113. Pinheiro LB, Coleman VA, Hindson CM, Herrmann J, Hindson BJ, Bhat S, et al. Evaluation of a Droplet Digital Polymerase Chain Reaction Format for DNA Copy Number Quantification. *Anal Chem*. 2011;84(2):1003–11.
 114. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem*. 2011;83(22):8604–10.
 115. Singh G, Sithebe A, Enitan AM, Kumari S, Bux F, Stenström TA. Comparison of droplet digital PCR and quantitative PCR for the detection of *Salmonella* and its application for river sediments. *J Water Health* [Internet]. 2017;15(4):505–8. Available from: <http://jwh.iwaponline.com/lookup/doi/10.2166/wh.2017.259>

116. Verhaegen B, De Reu K, De Zutter L, Verstraete K, Heyndrickx M, Van Coillie E. Comparison of droplet digital PCR and qPCR for the quantification of shiga toxin-producing *Escherichia coli* in bovine feces. *Toxins (Basel)*. 2016;8(5):1–11.
117. Cremonesi P, Cortimiglia C, Picozzi C, Minozzi G, Malvisi M, Luini M, et al. Development of a droplet digital polymerase chain reaction for rapid and simultaneous identification of common foodborne pathogens in soft cheese. *Front Microbiol*. 2016;7(October):1–10.
118. Hindson CM, Chevillet JR, Briggs HA, Gallichotte EN, Ruf IK, Hindson BJ, et al. Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* [Internet]. 2013 Oct 1;10(10):1003–5. Available from: <http://www.nature.com/articles/nmeth.2633>
119. Dingle TC, Sedlak RH, Cook L, Jerome KR. Tolerance of Droplet-Digital PCR vs Real-Time Quantitative PCR to Inhibitory Substances. *Clin Chem* [Internet]. 2013 Nov 1;59(11):1670–2. Available from: <http://www.clinchem.org/cgi/doi/10.1373/clinchem.2013.211045>
120. Rački N, Dreo T, Gutierrez-Aguirre I, Blejec A, Ravnikar M. Reverse transcriptase droplet digital PCR shows high resilience to PCR inhibitors from plant, soil and water samples. *Plant Methods* [Internet]. 2014;10(1):42. Available from: <http://plantmethods.biomedcentral.com/articles/10.1186/s13007-014-0042-6>
121. Roberts K, Reiner M, Gray K. Water Scarcity in Jaipur, Rajasthan, India. Jaipur, India; 2013.
122. Jain S. Drinking Water Management of Jaipur City: Issues & Challenges [Internet]. Vol. 26. Jaipur, India; 2011. Available from: http://icrier.org/pdf/Jaipur_service.pdf
123. India CO of. Census of India 2011 [Internet]. 2011. Available from: <https://www.census2011.co.in/>
124. Rodell M, Velicogna I, Famiglietti JS. Satellite-based estimates of groundwater depletion in India. *Nature* [Internet]. 2009;460(7258):999–1002. Available from: <http://dx.doi.org/10.1038/nature08238>
125. Dalin C, Wada Y, Kastner T, Puma MJ. Groundwater depletion embedded in international food trade. *Nature* [Internet]. 2017;543(7647):700–4. Available from: <http://dx.doi.org/10.1038/nature21403>
126. Andey SP, Kelkar PS. Influence of Intermittent and Continuous Modes of Water Supply on Domestic Water Consumption. *Water Resour Manag* [Internet]. 2009 Sep 23;23(12):2555–66. Available from: <http://link.springer.com/10.1007/s11269-008-9396-8>
127. Coyte RM, Jain RC, Srivastava SK, Sharma KC, Khalil A, Ma L, et al. Large-Scale Uranium Contamination of Groundwater Resources in India. *Environ Sci Technol*

Lett. 2018;5(6):341–7.

128. APHA, AWWA, WEF. 4500-Cl CHLORINE (RESIDUAL) G. DPD Colorimetric Method. In: Standard Methods for the Examination of Water and Wastewater. 2017. p. 4-67-4–68.
129. APHA, AWWA, WEF. 2510 Conductivity. In: Baird RB, Eaton AD, Rice EW, editors. Standard Methods for the Examination of Water and Wastewater. 23rd ed. Washington D. C.: APHA; 2000. p. 2-44-2–48.
130. Hill V, Narayanan J, Gallen R, Ferdinand K, Cromeans T, Vinjé J. Development of a Nucleic Acid Extraction Procedure for Simultaneous Recovery of DNA and RNA from Diverse Microbes in Water. *Pathogens* [Internet]. 2015;4(2):335–54. Available from: <http://www.mdpi.com/2076-0817/4/2/335/>
131. Platts-Mills JA, Liu J, Rogawski ET, Kabir F, Lertsethtakarn P, Sigua M, et al. Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of diarrhoea in children in low-resource settings: a reanalysis of the MAL-ED cohort study. *Lancet Glob Heal* [Internet]. 2018; Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2214109X18303498>
132. Nair NP, N SR, Giri S, Mohan VR, Parashar U, Tate J, et al. Rotavirus vaccine impact assessment surveillance in India : protocol and methods. 2019;1–8.
133. Jothikumar N, Cromeans TL, Hill VR, Lu X, Sobsey MD, Erdman DD. Quantitative real-time PCR assays for detection of human adenoviruses and identification of serotypes 40 and 41. *Appl Env Microbiol*. 2005/06/04. 2005;71(6):3131–6.
134. Jothikumar N, Lowther JA, Henshilwood K, Lees DN, Hill VR, Vinjé J. Rapid and sensitive detection of noroviruses by using TaqMan-based one-step reverse transcription-PCR assays and application to naturally contaminated shellfish samples. *Appl Environ Microbiol*. 2005;71(4):1870–5.
135. Kageyama T, Kojima S, Shinohara M, Uchida K, Fukushi S, Hoshino FB, et al. Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. *J Clin Microbiol* [Internet]. 2003;41(4):1548–57. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12682144%5Cn0299899393 on real time PCR.pdf
136. Lin WS, Cheng C-M, Van KT. A Quantitative PCR Assay for Rapid Detection of *Shigella* Species in Fresh Produce. *J Food Prot* [Internet]. 2010;73(2):221–33. Available from: <http://jfoodprotection.org/doi/abs/10.4315/0362-028X-73.2.221>
137. Grant MA, Hu J, Jinneman KC. Multiplex real-time PCR detection of heat-labile and heat-stable toxin genes in enterotoxigenic *Escherichia coli*. *J Food Prot* [Internet]. 2006;69(2):412–6. Available from: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt>

=Citation&list_uids=16496584

138. Liu J, Gratz J, Amour C, Kibiki G, Becker S, Janaki L, et al. A Laboratory-Developed TaqMan Array Card for Simultaneous Detection of 19 Enteropathogens. *J Clin Microbiol* [Internet]. 2013 Feb 1;51(2):472–80. Available from: <http://jcm.asm.org/cgi/doi/10.1128/JCM.02658-12>
139. Liu J, Kabir F, Manneh J, Lertsethtakarn P, Begum S, Gratz J, et al. Development and assessment of molecular diagnostic tests for 15 enteropathogens causing childhood diarrhoea: A multicentre study. *Lancet Infect Dis* [Internet]. 2014;14(8):716–24. Available from: [http://dx.doi.org/10.1016/S1473-3099\(14\)70808-4](http://dx.doi.org/10.1016/S1473-3099(14)70808-4)
140. Baque RH, Gilliam AO, Robles LD, Jakubowski W, Slifko TR. A real-time RT-PCR method to detect viable *Giardia lamblia* cysts in environmental waters. *Water Res* [Internet]. 2011;45(10):3175–84. Available from: <http://dx.doi.org/10.1016/j.watres.2011.03.032>
141. Walker DI, McQuillan J, Taiwo M, Parks R, Stenton CA, Morgan H, et al. A highly specific *Escherichia coli* qPCR and its comparison with existing methods for environmental waters. *Water Res* [Internet]. 2017;126:101–10. Available from: <https://doi.org/10.1016/j.watres.2017.08.032>
142. Silkie SS, Nelson KL. Concentrations of host-specific and generic fecal markers measured by quantitative PCR in raw sewage and fresh animal feces. *Water Res* [Internet]. 2009;43(19):4860–71. Available from: <http://dx.doi.org/10.1016/j.watres.2009.08.017>
143. Smith GJ, Helf M, Nesbet C, Betita HA, Meek J, Ferre F. Fast and Accurate Method for Quantitating *E. Coli* Host-Cell DNA Contamination in Plasmid DNA Preparations. *Biotechniques*. 1999;26(3):518–26.
144. Chern EC, Siefring S, Paar J, Doolittle M, Haugland RA. Comparison of quantitative PCR assays for *Escherichia coli* targeting ribosomal RNA and single copy genes. 2011;298–306.
145. Rolfe KJ, Parmar S, Mururi D, Wreghitt TG, Jalal H, Zhang H, et al. An internally controlled, one-step, real-time RT-PCR assay for norovirus detection and genogrouping. *J Clin Virol* [Internet]. 2007 Aug;39(4):318–21. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S1386653207001795>
146. Bio-Rad. Droplet Digital PCR Applications Guide. 2015.
147. Stokdyk JP, Firnstahl AD, Spencer SK, Burch TR, Borchardt MA. Determining the 95 % limit of detection for waterborne pathogen analyses from primary concentration to qPCR. *Water Res* [Internet]. 2016;96:105–13. Available from: <http://dx.doi.org/10.1016/j.watres.2016.03.026>

148. Witte AK, Mester P, Fister S, Witte M, Schoder D, Rossmanith P. A systematic investigation of parameters influencing droplet rain in the *Listeria monocytogenes* prfA assay -reduction of ambiguous results in ddPCR. PLoS One. 2016;11(12):1–13.
149. Jones M, Williams J, Gärtner K, Phillips R, Hurst J, Frater J. Low copy target detection by Droplet Digital PCR through application of a novel open access bioinformatic pipeline, “definetherain.” J Virol Methods [Internet]. 2014;202:46–53. Available from: <http://dx.doi.org/10.1016/j.jviromet.2014.02.020>
150. Trypsteen W, Vynck M, de Neve J, Bonczkowski P, Kiselinova M, Malatinkova E, et al. Ddpcrquant: Threshold determination for single channel droplet digital PCR experiments. Anal Bioanal Chem. 2015;407(19):5827–34.
151. Jacobs BKM, Goetghebeur E, Vandesompele J, De Ganck A, Nijs N, Beckers A, et al. Model-Based Classification for Digital PCR: Your Umbrella for Rain. Anal Chem. 2017;89(8):4461–7.
152. Bureau of Indian Standards. IS 10500:2012 Indian Standard Drinking Water-Specification [Internet]. New Delhi, India; 2012. Available from: <https://www.google.com/patents/US5773004>
153. Patel CB, Vajpayee P, Singh G, Upadhyay RS, Shanker R. Contamination of potable water by enterotoxigenic *Escherichia coli*: qPCR based culture-free detection and quantification. Ecotoxicol Environ Saf [Internet]. 2011;74(8):2292–8. Available from: <http://dx.doi.org/10.1016/j.ecoenv.2011.07.022>
154. Batabyal P, Mookerjee S, Sur D, Palit A. Diarrheogenic *Escherichia coli* in potable water sources of West Bengal, India. Acta Trop [Internet]. 2013;127(3):153–7. Available from: <http://dx.doi.org/10.1016/j.actatropica.2013.04.015>
155. Anbazhagi M, Loganathan D, Tamilselvan S, Jayabalou R, Kamatchiammal S, Kumar R. *Cryptosporidium* oocysts in drinking water supply of Chennai City, southern India. Clean - Soil, Air, Water. 2007;35(2):167–71.
156. Borchardt MA, Spencer SK, Kieke BA, Lambertini E, Loge FJ. Viruses in nondisinfected drinking water from municipal wells and community incidence of acute gastrointestinal illness. Env Heal Perspect [Internet]. 2012;120(9):1272–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22659405>
157. Daniels ME, Shrivastava A, Smith WA, Sahu P, Odagiri M, Misra PR, et al. *Cryptosporidium* and *Giardia* in Humans, Domestic Animals, and Village Water Sources in Rural India. Am J Trop Med Hyg [Internet]. 2015 Sep 2;93(3):596–600. Available from: <http://www.ajtmh.org/cgi/doi/10.4269/ajtmh.15-0111>
158. Borchardt MA, Chyou PH, DeVries EO, Belongia EA. Septic system density and infectious diarrhea in a defined population of children. Environ Health Perspect. 2003;111(5):742–8.

159. Lerner DN. Groundwater recharge in urban areas. Hydrol Process water Manag urban areas Lect Pap UNESCO/IHP Symp Duisburg, Lelystad, Amsterdam, Rotterdam, 1988. 1990;24(1):59–65.
160. Bishop PK, Misstear BD, White M, Harding NJ. Impacts of sewers on groundwater quality. *Water Environ J.* 1998;12(3):216–23.
161. Somasundaram MRaG and TJ. Somasundaram_et_al-1993-Groundwater.pdf.
162. Wakode HB, Baier K, Jha R, Azzam R. Impact of urbanization on groundwater recharge and urban water balance for the city of Hyderabad, India. *Int Soil Water Conserv Res.* 2018;6(1):51–62.
163. Suthar S, Chhimpa V, Singh S. Bacterial contamination in drinking water: a case study in rural areas of northern Rajasthan, India. *Environ Monit Assess* [Internet]. 2008;1–6. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19023670
164. Coyte RM, Singh A, Furst KE, Mitch WA, Vengosh A. Co-occurrence of geogenic and anthropogenic contaminants in groundwater from Rajasthan, India. *Sci Total Environ* [Internet]. 2019;688:1216–27. Available from: <https://doi.org/10.1016/j.scitotenv.2019.06.334>
165. Furst KE, Coyte RM, Wood M, Vengosh A, Mitch WA. Disinfection Byproducts in Rajasthan, India: Are Trihalomethanes a Sufficient Indicator of Disinfection Byproduct Exposure in Low- Income Countries? 2019;
166. Bivins A, Beetsch N, Majuru B, Montgomery M, Sumner T, Brown J. Selecting Household Water Treatment Options on the Basis of World Health Organization Performance Testing Protocols. *Environ Sci Technol.* 2019;53(9):5043–51.
167. Fout GS, Cashdollar JL, Griffin SM, Brinkman NE, Varughese EA, Parshionikar SU. EPA Method 1615. measurement of enterovirus and norovirus occurrence in water by culture and RT-qPCR. Part III. Virus detection by RT-qPCR. *J Vis Exp.* 2016;2016(107):1–13.
168. Kaminsky J, Kumpel E. Dry pipes: Associations between utility performance and intermittent piped water supply in low and middle income countries. *Water (Switzerland).* 2018;10(8).
169. Matsinhe NP, Juízo DL, Persson KM. THE EFFECTS OF INTERMITTENT SUPPLY AND HOUSEHOLD STORAGE IN THE QUALITY OF DRINKING WATER IN MAPUTO Förändringar i dricksvattenkvalitet i Maputo på grund av intermittent försörjning och tanklagring i hushållen. 2014;51–60.
170. Rahman A, Lee HKK, Khan MAA. Domestic Water Contamination in Rapidly Growing Megacities of Asia: Case of Karachi, Pakistan. *Environ Monit Assess.*

1997;44:339–60.

171. Alazzeh S, Galaitsi S, Bishara A, Al-Azraq N, Durant JL. Impacts of Intermittent Water Supply on Water Quality in Two Palestinian Refugee Camps. *Water* [Internet]. 2019;11(4):670. Available from: <https://www.mdpi.com/2073-4441/11/4/670>
172. Rubino F, Corona Y, Jimenez Perez JG, Smith C. Bacterial Contamination of Drinking Water in Guadalajara, Mexico. 2018;(December).
173. Tokajian S, Hashwa F. Water quality problems associated with intermittent water supply. *Water Sci Technol*. 2000;47(3):229–34.
174. Eshcol J, Mahapatra P, Keshapagu S. Is fecal contamination of drinking water after collection associated with household water handling and hygiene practices? A study of urban slum households in Hyderabad, India. *J Water Health* [Internet]. 2008/10/30. 2009;7(1):145–54. Available from: <http://jwh.iwaponline.com/content/ppiwajwh/7/1/145.full.pdf>
175. Taylor DDJ, Slocum AH, Whittle AJ. Analytical scaling relations to evaluate leakage and intrusion in intermittent water supply systems. *PLoS One* [Internet]. 2018;13(5):e0196887. Available from: <http://dx.plos.org/10.1371/journal.pone.0196887>
176. Cifuentes E, Suarez L, Solano M, Santos R. Diarrheal diseases in children from a water reclamation site in Mexico city. *Env Heal Perspect* [Internet]. 2002/10/04. 2002;110(10):A619-24. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1241048/pdf/ehp0110-a00619.pdf>
177. Abu Mourad T. Palestinian refugee conditions associated with intestinal parasites and diarrhoea: Nuseirat refugee camp as a case study. *Public Health* [Internet]. 2004 Mar;118(2):131–42. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0033350603002385>
178. Yassin MM, Amr SSA, Al-Najar HM. Assessment of microbiological water quality and its relation to human health in Gaza Governorate, Gaza Strip. *Public Health* [Internet]. 2006 Dec;120(12):1177–87. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0033350606002149>
179. Abu Amr SS, Yassin MM. Microbial contamination of the drinking water distribution system and its impact on human health in Khan Yunis Governorate, Gaza Strip: Seven years of monitoring (2000-2006). *Public Health* [Internet]. 2008;122(11):1275–83. Available from: <http://dx.doi.org/10.1016/j.puhe.2008.02.009>
180. Trudeau J, Aksan A-M, Vásquez WF. Water system unreliability and diarrhea incidence among children in Guatemala. *Int J Public Health* [Internet]. 2018 Mar 16;63(2):241–50. Available from: <http://link.springer.com/10.1007/s00038-017->

1054-6

181. Adane M, Mengistie B, Medhin G, Kloos H, Mulat W. Piped water supply interruptions and acute diarrhea among under-five children in Addis Ababa slums, Ethiopia: A matched case-control study. Hill PC, editor. PLoS One [Internet]. 2017 Jul 19;12(7):e0181516. Available from: <http://dx.plos.org/10.1371/journal.pone.0181516>
182. Exum NG, Lee GO, Olórtegui MP, Yori PP, Salas MS, Trigoso DR, et al. A Longitudinal Study of Household Water, Sanitation, and Hygiene Characteristics and Environmental Enteropathy Markers in Children Less than 24 Months in Iquitos, Peru. Am J Trop Med Hyg [Internet]. 2018 Apr 4;98(4):995–1004. Available from: <http://www.ajtmh.org/content/journals/10.4269/ajtmh.17-0464>
183. Desai S, Vanneman R. India Human Development Survey (IHDS). New Delhi; 2005.
184. McKenzie D, Ray I. Urban water supply in India: status, reform options and possible lessons. Water Policy [Internet]. 2009 Aug;11(4):442–60. Available from: <http://wp.iwaponline.com/cgi/doi/10.2166/wp.2009.056>
185. Water and Sanitation Program. The Karnataka Urban Water Sector Improvement Project: 24x7 Water Supply is Achievable. New Delhi; 2010.
186. WSP. Running Water in India's Cities : A Review of Five Recent Public-Private Partnership Initiatives. Washington D. C.; 2014.
187. Sangameswaran P, Madhav R, D'Rozario C. 24/7, 'Privatisation' and Water Reform Insights from Hubli-Dharwad. Economic and Political Weekly. 2008;(April):60–7.
188. Asian Development Bank. Benchmarking and Data Book of Water Utilities in India [Internet]. 2007. Available from: <http://www.adb.org/Documents/Reports/Benchmarking-Data-Book-Utilities-in-India/2007-Indian-Water-Utilities-Data-Book.pdf>
189. Nijhawan A, Jain P, Sargaonkar A, Labhasetwar PK. Implementation of water safety plan for a large-piped water supply system. Environ Monit Assess. 2014;186(9):5547–60.
190. Hastak S, Labhasetwar P, Kundley P, Gupta R. Changing from intermittent to continuous water supply and its influence on service level benchmarks: a case study in the demonstration zone of Nagpur, India. Urban Water J [Internet]. 2017 Aug 9;14(7):768–72. Available from: <http://dx.doi.org/10.1080/1573062X.2016.1240808>
191. Lockhart G, Oswald WE, Hubbard B, Medlin E, Gelting RJ. Development of indicators for measuring outcomes of water safety plans. J Water, Sanit Hyg Dev [Internet]. 2014;4(1):171. Available from:

<http://washdev.iwaponline.com/cgi/doi/10.2166/washdev.2013.159>

192. WHO. Strengthening Operations & Maintenance through Water Safety Planning: A Collection of Case Studies. 2018.
193. USEPA. Method 180.1: Determination of Turbidity by Nephelometry [Internet]. 1993. Available from: http://linkinghub.elsevier.com/retrieve/pii/0019057893900758%0Ahttps://www.epa.gov/sites/production/files/2015-08/documents/method_180-1_1993.pdf
194. Korich DG, Mead JR, Madore MS, Sinclair NA, Sterling CR. Effects of ozone, chlorine dioxide, chlorine, and monochloramine on *Cryptosporidium parvum* oocyst viability. *Appl Environ Microbiol*. 1990;56(5):1423–8.
195. Taylor DDJ, Al TET. Demand Satisfaction as a Framework for Understanding Intermittent Water Supply Systems. 2017;1–21.
196. Andey SP, Kelkar PS. Performance of water distribution systems during intermittent versus continuous water supply. *Am Water Work Assoc*. 2007;99(8):99–106.
197. Nyathikala SA, Kulshrestha M. Performance and productivity measurement of urban water supply services in India. *Water Sci Technol Water Supply* [Internet]. 2017;17(2):407–21. Available from: <http://ws.iwaponline.com/lookup/doi/10.2166/ws.2016.106>
198. Mitlin D, Beard VA, Satterthwaite D, Du J. Unaffordable and Undrinkable: Rethinking Urban Water Access in the Global South [Internet]. Washington D. C.; 2019. Available from: www.citiesforall.org
199. Sangma RVN, Rasania S, Prasuna J, Das R, Jais M. Bacteriological Quality of Drinking Water and Diarrhoeal Outcome among Under Five Children in Resettlement Colony, Delhi. *Proc 5th World Congr New Technol*. 2019;(August).
200. Hayashi MAL, Eisenberg MC, Eisenberg JNS. Linking Decision Theory and Quantitative Microbial Risk Assessment: Tradeoffs Between Compliance and Efficacy for Waterborne Disease Interventions. 2019;
201. The World Bank. World Bank Open Data [Internet]. 2018 [cited 2019 Sep 25]. Available from: <https://data.worldbank.org/>
202. Chaudhuri R. *Water: What are our rights to it?* Calcutta, India; 1998.
203. Medema G, Smeets P. Quantitative risk assessment in the Water Safety Plan: Case studies from drinking water practice. *Water Sci Technol Water Supply*. 2009;9(2):127–32.
204. Blokker M, Smeets P, Medema G. Quantitative microbial risk assessment of repairs of the drinking water distribution system. *Microb Risk Anal* [Internet].

2018;8(September 2017):22–31. Available from:
<https://doi.org/10.1016/j.mran.2017.12.002>

205. Bhunia R, Hutin Y, Ramakrishnan R, Pal N, Sen T, Murhekar M. A typhoid fever outbreak in a slum of South Dumdum municipality, West Bengal, India, 2007: Evidence for foodborne and waterborne transmission. *BMC Public Health*. 2009;9:1–8.
206. Ram S, Vajpayee P, Dwivedi PD, Shanker R. Culture-free detection and enumeration of STEC in water. *Ecotoxicol Environ Saf* [Internet]. 2011;74(4):551–7. Available from: <http://dx.doi.org/10.1016/j.ecoenv.2011.01.019>
207. Boarato-David É, Guimarães S, Cacciò S. *Giardia duodenalis*. In: Fayer R, Jakubowski W, editors. *Global Water Pathogen Project* [Internet]. Michigan State University; 2019. Available from: <https://www.waterpathogens.org/book/giardia-duodenalis>
208. Cacciò SM, Sprong H. Epidemiology of Giardiasis in Humans. *Giardia*. 2011;17–28.
209. Feng Y, Zhao X, Chen J, Jin W, Zhou X, Li N, et al. Occurrence, source, and human infection potential of *Cryptosporidium* and *giardia* spp. in source and tap water in Shanghai, China. *Appl Environ Microbiol*. 2011;77(11):3609–16.
210. U.S. EPA. *Guidance Manual for Compliance with the Filtration and Disinfection Requirements for Public Water Systems Using Surface Water Sources*. Am Water Work Assoc Denver, CO [Internet]. 1991;3–10. Available from: https://www.epa.gov/sites/production/files/2015-10/documents/guidance_manual_for_compliance_with_the_filtration_and_disinfection_requirements.pdf
211. Karanis P, Sotiriadou I, Kartashev V, Kourenti C, Tsvetkova N, Stojanova K. Occurrence of *Giardia* and *Cryptosporidium* in water supplies of Russia and Bulgaria. *Environ Res*. 2006;102(3):260–71.
212. Bajer A, Toczyłowska B, Bednarska M, Sinski E. Effectiveness of water treatment for the removal of *Cryptosporidium* and *Giardia* spp. *Epidemiol Infect*. 2012;140(11):2014–22.
213. Baldursson S, Karanis P. Waterborne transmission of protozoan parasites: review of worldwide outbreaks - an update 2004-2010. *Water Res*. 2011;45:6603–6614.
214. Betancourt W. *Cryptosporidium* spp. In: Fayer R, Jakubowski W, editors. *Global Water Pathogen Project* [Internet]. Michigan State University; 2019. Available from: <https://www.waterpathogens.org/book/cryptosporidium>
215. Checkley W, Epstein LD, Gilman RH, Black RE, Cabrera L, Sterling CR. Effects

- of *Cryptosporidium parvum* infection in Peruvian children: Growth faltering and subsequent catch-up growth. *Am J Epidemiol.* 1998;148(5):497–506.
216. Guerrant DI, Moore SR, Lima AAM, Patrick PD, Schorling JB, Guerrant RL. Association of early childhood diarrhea and cryptosporidiosis with impaired physical fitness and cognitive function four-seven years later in a poor urban community in northeast Brazil. *Am J Trop Med Hyg.* 1999;61(5):707–13.
 217. Hunter PR, Hughes S, Woodhouse S, Syed Q, Verlander NQ, Chalmers RM, et al. Sporadic cryptosporidiosis case-control study with genotyping. *Emerg Infect Dis.* 2004;10(7):1241–9.
 218. Gerba CP, Rose JB, Haas CN. Sensitive populations: Who is at the greatest risk? *Int J Food Microbiol.* 1996;30(1–2):113–23.
 219. Sow SO, Muhsen K, Nasrin D, Blackwelder WC, Wu Y, Farag TH, et al. The Burden of *Cryptosporidium* Diarrheal Disease among Children < 24 Months of Age in Moderate/High Mortality Regions of Sub-Saharan Africa and South Asia, Utilizing Data from the Global Enteric Multicenter Study (GEMS). Kosek M, editor. *PLoS Negl Trop Dis* [Internet]. 2016 May 24;10(5):e0004729. Available from: <https://dx.plos.org/10.1371/journal.pntd.0004729>
 220. Xiao L. Molecular epidemiology of cryptosporidiosis: An update. *Exp Parasitol* [Internet]. 2010;124(1):80–9. Available from: <http://dx.doi.org/10.1016/j.exppara.2009.03.018>
 221. Betancourt WQ, Rose JB. Drinking water treatment processes for removal of *Cryptosporidium* and *Giardia*. *Vet Parasitol.* 2004;126(1-2 SPEC.ISS.):219–34.
 222. Katayama H, Vinjé J. Norovirus and other Caliciviruses. In: Meschke JS, Girones R, editors. *Global Water Pathogen Project* [Internet]. Michigan State University; 2019. Available from: <https://www.waterpathogens.org/book/norovirus-and-other-caliciviruses>
 223. Lanata CF, Fischer-Walker CL, Olascoaga AC, Torres CX, Aryee MJ, Black RE. Global Causes of Diarrheal Disease Mortality in Children <5 Years of Age: A Systematic Review. Sestak K, editor. *PLoS One* [Internet]. 2013 Sep 4;8(9):e72788. Available from: <https://dx.plos.org/10.1371/journal.pone.0072788>
 224. Siebenga JJ, Vennema H, Zheng D, Vinjé J, Lee BE, Pang X, et al. Norovirus Illness Is a Global Problem: Emergence and Spread of Norovirus GII.4 Variants, 2001–2007. *J Infect Dis.* 2009;200(5):802–12.
 225. Shin GA, Sobsey MD. Inactivation of norovirus by chlorine disinfection of water. *Water Res* [Internet]. 2008;42(17):4562–8. Available from: <http://dx.doi.org/10.1016/j.watres.2008.08.001>
 226. Kaplan JE, Goodman RA, Schonberger LB, Lippy EC, Gary GW. Gastroenteritis

- due to Norwalk virus: An outbreak associated with a municipal water system. *J Infect Dis.* 1982;146(2):190–7.
227. Kukkula M, Maunula L, Silvennoinen E, von Bonsdorff C. Outbreak of Viral Gastroenteritis Due to Drinking Water Contaminated by Norwalk-like Viruses. *J Infect Dis.* 1999;180(6):1771–6.
228. Allard A, Vantarakis A. Adenoviruses. In: Meschke JS, Girones R, editors. *Global Water Pathogen Project* [Internet]. Michigan State University; 2019. Available from: <https://www.waterpathogens.org/book/adenoviruses>
229. Uhnoo I, Wadell G, Svensson L, Johansson ME. Importance of enteric adenoviruses 40 and 41 in acute gastroenteritis in infants and young children. *J Clin Microbiol.* 1984;20(3):365–72.
230. Kishida N, Morita H, Haramoto E, Asami M, Akiba M. One-year weekly survey of noroviruses and enteric adenoviruses in the Tone River water in Tokyo metropolitan area, Japan. *Water Res* [Internet]. 2012;46(9):2905–10. Available from: <http://dx.doi.org/10.1016/j.watres.2012.03.010>
231. Cromeans TL, Kahler AM, Hill VR. Inactivation of adenoviruses, enteroviruses, and murine norovirus in water by free chlorine and monochloramine. *Appl Environ Microbiol.* 2010;76(4):1028–33.
232. Guillot E, Loret J-F. *Waterborne Pathogens: Review for the Drinking Water Industry.* London, UK: IWA Publishing; 2010. 193 p.
233. The HC, Thanh DP, Holt KE, Thomson NR, Baker S. The genomic signatures of *Shigella* evolution, adaptation and geographical spread. *Nat Rev Microbiol* [Internet]. 2016;14(4):235–50. Available from: <http://dx.doi.org/10.1038/nrmicro.2016.10>
234. Kotloff KL, Winickoff JP, Ivanoff B, Clemens JD, Swerdlow DL, Sansonetti PJ, et al. Global burden of *Shigella* infections: Implications for vaccine development and implementation of control strategies. *Bull World Health Organ.* 1999;77(8):651–66.
235. Schroeder GN, Hilbi H. Molecular pathogenesis of *Shigella* spp.: Controlling host cell signaling, invasion, and death by type III secretion. *Clin Microbiol Rev.* 2008;21(1):134–56.
236. Faruque SM, Khan R, Kamruzzaman M, Yamasaki S, Ahmad QS, Azim T, et al. Isolation of *Shigella dysenteriae* type 1 and *s. flexneri* strains from surface waters in Bangladesh: Comparative molecular analysis of environmental *Shigella* isolates versus clinical strains. *Appl Environ Microbiol.* 2002;68(8):3908–13.
237. Rahman MZ, Azmuda N, Hossain MJ, Sultana M, Khan SI, Birkeland NK. Recovery and characterization of environmental variants of *Shigella flexneri* from surface water in Bangladesh. *Curr Microbiol.* 2011;63(4):372–6.

238. Oliver JD. The viable but nonculturable state in bacteria. *J Microbiol.* 2005;43(February):93–100.
239. King CH, Shotts EB, Wooley RE, Porter KG. Survival of coliforms and bacterial pathogens within protozoa during chlorination. *Appl Environ Microbiol.* 1988;54(12):3023–33.
240. Green DM, Scott SS, Mowat DAE, Shearer EJM, Thomson JM. Water-borne outbreak of viral gastroenteritis and Sonne dysentery. *J Hyg (Lond).* 1968;66(3):383–92.
241. Garcia-Aljaro C, Momba M, Muniesa M. Pathogenic members of *Escherichia coli* & *Shigella* spp. Shigellosis. In: Pruden A, Ashbolt N, Miller J, editors. *Global Water Pathogen Project* [Internet]. Michigan State University; 2019. Available from: <https://www.waterpathogens.org/book/ecoli>
242. Rahman I, Shahamat M, Chowdhury MAR, Colwell RR. Potential Virulence of Viable but Nonculturable. *Microbiology.* 1996;62(1):115–20.
243. Rangel JM, Sparling PH, Crowe C, Griffin PM, Swerdlow DL. Epidemiology of *Escherichia coli* O157:H7 Outbreaks, United States, 1982-2002 [Internet]. Vol. 11. 2005. Available from: <http://digitalcommons.unl.edu/publichealthresources/73>
244. Auld H, MacIver D, Klaassen J. Heavy rainfall and waterborne disease outbreaks: The Walkerton example. *J Toxicol Environ Heal - Part A.* 2004;67(20–22):1879–87.
245. LICENCE K, OATES KR, SYNGE BA, REID TMS. An outbreak of *E. coli* O157 infection with evidence of spread from animals to man through contamination of a private water supply . *Epidemiol Infect.* 2001;126(1):135–8.
246. George J, Suriyanarayanan S. Seasonal patterns of drinking water consumption pattern in Mysore City, Karnataka, South India. *J Water, Sanit Hyg Dev* [Internet]. 2016;6(4):569–75. Available from: <http://washdev.iwaponline.com/cgi/doi/10.2166/washdev.2016.082>
247. Bernander R, Palm JED, Svärd SG. Genome ploidy in different stages of the *Giardia lamblia* life cycle. *Cell Microbiol.* 2001;3(1):55–62.
248. Teunis PFM, Medema GJ, Kruidenier L, Havelaar AH. Assessment of the risk of infection by *Cryptosporidium* or *Giardia* in drinking water from a surface water source. *Water Res.* 1997;31(6):1333–46.
249. Hadfield SJ, Pachebat JA, Swain MT, Robinson G, Cameron SJS, Alexander J, et al. Generation of whole genome sequences of new *Cryptosporidium hominis* and *Cryptosporidium parvum* isolates directly from stool samples. *BMC Genomics* [Internet]. 2015;16(1). Available from: <http://dx.doi.org/10.1186/s12864-015-1805-9>

250. Le Blancq SM, Khramtsov N V., Zamani F, Upton SJ, Wu TW. Ribosomal RNA gene organization in *Cryptosporidium parvum*. *Mol Biochem Parasitol*. 1997;90(2):463–78.
251. Jothikumar N, Da Silva AJ, Moura I, Qvarnstrom Y, Hill VR. Detection and differentiation of *Cryptosporidium hominis* and *Cryptosporidium parvum* by dual TaqMan assays. *J Med Microbiol*. 2008;57(9):1099–105.
252. Couch RB, Knight V, Douglas RG, Black SH, HAMORY BH. The minimal infectious dose of adenovirus type 4; the case for natural transmission by viral aerosol. *Trans Am Clin Climatol Assoc*. 1969;80:205–11.
253. Dulbecco R, Ginsberg HS. The nature of viruses. In: *Virology*. Philadelphia: Harper and Row; 1980.
254. He JW, Jiang S. Quantification of enterococci and human adenoviruses in environmental samples by real-time PCR (*Applied and Environmental Microbiology* (2005) 71:5 (2250-2255)). *Appl Environ Microbiol*. 2009;75(2):557.
255. Youmans BP, Ajami NJ, Jiang ZD, Petrosino JF, DuPont HL, Highlander SK. Development and accuracy of quantitative real-time polymerase chain reaction assays for detection and quantification of enterotoxigenic *Escherichia coli* (ETEC) heat labile and heat stable toxin genes in travelers' diarrhea samples. *Am J Trop Med Hyg*. 2014;90(1):124–32.
256. Lothigius Å, Janzon A, Begum Y, Sjöling Å, Qadri F, Svennerholm AM, et al. Enterotoxigenic *Escherichia coli* is detectable in water samples from an endemic area by real-time PCR. *J Appl Microbiol*. 2008;104(4):1128–36.
257. Schmidt PJ. Norovirus Dose-Response: Are Currently Available Data Informative Enough to Determine How Susceptible Humans Are to Infection from a Single Virus? *Risk Anal* [Internet]. 2015 Jul;35(7):1364–83. Available from: <http://doi.wiley.com/10.1111/risa.12323>
258. Van Abel N, Schoen ME, Kissel JC, Meschke JS. Comparison of Risk Predicted by Multiple Norovirus Dose-Response Models and Implications for Quantitative Microbial Risk Assessment. *Risk Anal*. 2017;37(2):245–64.
259. Messner MJ, Berger P, Nappier SP. Fractional poisson - A simple dose-response model for human norovirus. *Risk Anal*. 2014;34(10):1820–9.
260. Menon VK, George S, Sarkar R, Giri S, Samuel P, Vivek R, et al. Norovirus gastroenteritis in a birth cohort in southern India. *PLoS One*. 2016;11(6):1–18.
261. Teunis P, Schijven J, Rutjes S. A generalized dose-response relationship for adenovirus infection and illness by exposure pathway. *Epidemiol Infect*. 2016;144(16):3461–73.

262. Couch RB, Chanock RM, Cate TR, Lang DJ, Knight V, Huebner RJ. INFECTION OF THE INTESTINAL TRACTI.
263. Coster TS, Wolf MK, Hall ER, Cassels FJ, Taylor DN, Liu CT, et al. Immune response, ciprofloxacin activity, and gender differences after human experimental challenge by two strains of enterotoxigenic *Eschenchia coli*. *Infect Immun*. 2007;75(1):252–9.
264. Graham DY, Estes MK, Gentry LO. Double-Blind Comparison of Bismuth Subsalicylate and Placebo in the Prevention and Treatment of Enterotoxigenic *Escherichia coli*-Induced Diarrhea in Volunteers. *Gastroenterology* [Internet]. 1983;85(5):1017–22. Available from: [http://dx.doi.org/10.1016/S0016-5085\(83\)80066-3](http://dx.doi.org/10.1016/S0016-5085(83)80066-3)
265. Levine MM, Nalin DR, Hoover DL, Bergquist EJ, Hornick RB, Young CR. Immunity to enterotoxigenic *Escherichia coli*. *Infect Immun*. 1979;23(3):729–36.
266. Clements ML, Levine MM, Black RE, Robins-Browne RM, Cisneros LA, Drusano GL, et al. *Lactobacillus* prophylaxis for diarrhea due to enterotoxigenic *Escherichia coli*. *Antimicrob Agents Chemother*. 1981;20(1):104–8.
267. Dupont HL, Hornick RB, Snyder MJ, Libonati JP, Formal SB, Gangarosa EJ. Immunity in shigellosis. II. protection induced by oral live vaccine or primary infection. *J Infect Dis*. 1972;125(1):12–6.
268. Crockett CS, Haas CN, Fazil A, Rose JB, Gerba CP. Prevalence of shigellosis: In the U.S. consistency with dose-response information. *Int J Food Microbiol*. 1996;30(1–2):87–99.
269. Lechevallier MW, Norton WD, Lee RG. Occurrence of *Giardia* and *Cryptosporidium* spp . in Surface Water Supplies. 1991;57(9):2610–6.
270. Schmidt PJ, Emelko MB, Thompson ME. Recognizing Structural Nonidentifiability: When Experiments Do Not Provide Information About Important Parameters and Misleading Models Can Still Have Great Fit. *Risk Anal* [Internet]. 2019;0(0):risa.13386. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.13386>
271. Majuru B, Suhrcke M, Hunter PR. How do households respond to unreliable water supplies? a systematic review. *Int J Environ Res Public Health*. 2016;13(12).
272. Bivins A, Beetsch N, Majuru B, Montgomery M, Sumner T, Brown J. Selecting Household Water Treatment Options on the Basis of World Health Organization Performance Testing Protocols. *Environ Sci Technol* [Internet]. 2019 May 7;53(9):5043–51. Available from: <http://pubs.acs.org/doi/10.1021/acs.est.8b05682>
273. Bain R, Cronk R, Wright J, Yang H, Slaymaker T, Bartram J. Fecal Contamination of Drinking-Water in Low- and Middle-Income Countries: A Systematic Review

- and Meta-Analysis. Hunter PR, editor. PLoS Med [Internet]. 2014 May 6;11(5):e1001644. Available from: <https://dx.plos.org/10.1371/journal.pmed.1001644>
274. Bain R, Gundry S, Wright J, Yang H, Pedley S, Bartram J. Accounting for water quality in monitoring access to safe drinking-water as part of the Millennium Development Goals: lessons from five countries. Bull World Health Organ [Internet]. 2012;90(3):228–35. Available from: <http://www.who.int/bulletin/volumes/90/3/11-094284.pdf>
 275. UNICEF WHO (WHO) & WASH Post-2015: Proposed indicators for drinking water, sanitation and hygiene. World Heal Organ [Internet]. 2015;1–8. Available from: <http://www.wssinfo.org/%0Ahttp://www.unwater.org/gemi/en/%0Ahttp://www.wssinfo.org/post-2015-monitoring/>
 276. Ayoub G, Malaeb L. Impact of intermittent water supply on water quality in Lebanon. Int J Environ Pollut. 2006;26(4):379–97.
 277. Eshcol J, Mahapatra P, Keshapagu S. Is fecal contamination of drinking water after collection associated with household water handling and hygiene practices? A study of urban slum households in Hyderabad, India. J Water Health [Internet]. 2009 Mar;7(1):145–54. Available from: <https://iwaponline.com/jwh/article/7/1/145/9917/Is-fecal-contamination-of-drinking-water-after>
 278. Alexandra D, Renwick V. The Effects of an Intermittent Piped Water Network and Storage Practices On Household Water Quality in Tamale , Ghana. 2013.
 279. Schlueter V, Schmolke S, Stark K, Hess G, Ofenloch-Haehnle B, Engel AM. Reverse transcription-PCR detection of hepatitis G virus. J Clin Microbiol. 1996;34(11):2660–4.
 280. Rutledge RG, Stewart D. Assessing the performance capabilities of LRE-based assays for absolute quantitative real-time PCR. PLoS One. 2010;5(3).
 281. Boxus M, Letellier C, Kerkhofs P. Real Time RT-PCR for the detection and quantitation of bovine respiratory syncytial virus. J Virol Methods. 2005;125(2):125–30.
 282. Wang J, O’Keefe J, Orr D, Loth L, Banks M, Wakeley P, et al. Validation of a real-time PCR assay for the detection of bovine herpesvirus 1 in bovine semen. J Virol Methods. 2007;144(1–2):103–8.

VITA

Aaron William Bivins was born on July 4, 1984 in Atlanta, Georgia. His interest in engineering began as a young boy when he would spend hours building cities with millions of Legos. He earned his Bachelor of Science in Civil Engineering from Georgia Tech in 2007. In August 2012 after five years of engineering consulting and earning his professional engineering license, Aaron returned to Georgia Tech and earned his Master of Science in Environmental Engineering. He enjoyed learning so much that he continued for his PhD in Environmental Engineering. He hopes to spend his career training and mentoring engineers to overcome the constraints of delivering clean drinking water in under-resourced settings. When not in the lab, Aaron can be found running the mountainous trails of north Georgia or fly fishing for trout in north Georgia's cold streams. He still lives in Atlanta with his wife Katherine and cat Lily.