Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Robert Giordano Date

Socio-Economic Risk Factors of Pathogenic Enterobacteria Infection and Antibiotic Resistance in Ranomafana Commune, Madagascar

By

Robert Giordano
Masters of Public Health

Environmental Health

Thomas Gillespie, PhD
Committee Chair

Paige Tolbert, PhD
Committee Member

Socio-Economic Risk Factors for Pathogenic Enterobacteria Infection and Antibiotic Resistance in the Ranomafana Commune, Madagascar

By

Robert Giordano

Bachelor of Science
Emory University
2015

Thesis Committee Chair: Thomas Gillespie, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Environmental Health
2016

Abstract

Socio-Economic Risk Factors of Pathogenic Enterobacteria Infection and Antibiotic Resistance in the Ranomafana Commune, Madagascar

By Robert Giordano

Background

Diarrheal diseases represent a primary cause of mortality and morbidity among individuals – particularly those under the age of 5 – who live in rural areas of low-income nations characterized by limited access to safe and reliable water and sanitation infrastructure. Although global health efforts, such as the United Nation's Millennium Project, have had a substantial impact on the prevalence of, and death rates associated with, communicable disease, the progressive emergence of bacterial resistance to antibiotics threatens the effectiveness of therapeutic drug treatment against the pathogens responsible.

Methodology

From July 1st – July 14th, 2014, 20 households were randomly selected within each of four geographically distinct census tracts within the Ranomafana Commune of Madagascar, to participate in the study. Within each household, in-person interviews were administered to assess: demographic information, individual health status, common hygiene practices, water usage/treatment, antibiotic usage, and household economics. All survey participants were asked to provide a fresh fecal specimen for molecular analysis using single polymerase chain reaction (PCR) assays to detect for diarrheal pathogen infection and the absence/presence of genes that encode for antibiotic resistance (AR).

Seventy-six households elected to partake in the study and 65% of participants agreed to provide a fecal specimen resulting in a study population of 225 characterized individuals. This cross-sectional study describes the current prevalence of four globally significant diarrheal pathogens and five genetic elements that encode resistance to four classes of antibiotic.

Principle Findings

The weighted prevalence of enteric virulence markers among study participants were: 22.04% for *Shigella* spp., 5.4% for *Salmonella* spp., 0.35% for *Vibrio cholerae*, and 7.11% for *Yersinia* spp. No significant association was found between enteric infection and: sex, access to household latrine, primary water treatment methods, or experiencing diarrhea within four weeks prior to the study. However, children aged 5-14 had a marked increase in the odds of enteric infection when compared to individuals under the age of five.

The weighted prevalence of genes encoding AR among study participants were: 42.11% for β-lactam resistance genes, 54.29% for *sul1* sulfonamide resistance, 31.77% for *sul2* sulfonamide resistance, 50.47% for *strAB* aminoglycoside resistance, and 43.41% for *dfrA7* trimethoprim resistance. No significant associations were found between AR carriage and: sex, cluster, or – curiously – antibiotic usage. However, the odds of harboring any AR encoding gene were nearly 4.27 times higher among individuals belonging to households defined as being in the bottom 80% of the household wealth index compared to those individual belonging to a household in the top 20%. Significant associations were found between *Shigella* spp. infection and harboring β-lactam and aminoglycoside resistance genes. Additionally, a significant association was found between *Yersinia* spp. infection and harboring β-lactam resistance genes.

Conclusions

The high prevalence of *Shigella* spp. observed in this region highlights the need to introduce a targeted intervention that actively works to disrupt the traditional transmission pathways of *Shigella*. Additionally, the high prevalence of divergent genes encoding antibiotic resistance in this region suggests that clinical treatment of bacterial diarrheal infections with conventional antibiotics may prove ineffective. In addition, current patterns of antibiotic usage in this region may facilitate the evolution of novel patterns of antibiotic resistance. Further, the results presented demonstrate the need to establish health and hygiene sensitization programs that reach out to those of lower wealth standing, emphasizing the significance of proper antibiotic usage and WASH practices.

Socio-Economic Risk Factors of Pathogenic Enterobacteria Infection and Antibiotic Resistance in the Ranomafana Commune, Madagascar

By

Robert Giordano

Bachelor of Science
Emory University
2015

Thesis Committee Chair: Thomas Gillespie, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Environmental Health
2016

Acknowledgements
This project would not have been possible if it were not for the in-country guidance and logistical support of the Malagasy Institute for the Conservation of Tropical Environments (MICET), particularly that of the director, Benjamin Andriamihaja. I would like to extend my sincere thanks and gratitude to the support personnel and staff of the Centre ValBio for providing in-country laboratory accommodations and a home during my time spent in Ranomafana.

Permission for this research was provided by the Government of Madagascar. Funding for this work was provided largely by a Gates Grand Challenges Explorations Grant with additional support from the Jim and Robin Herrnstein Foundation, the Bliska Laboratory at Stony Brook University, the Emory Global Health Institute, James G. Lester Research Grant Program, the Scholarly Inquiry and Research at Emory Grant, and the Emory College International Research Experience for Science Students Award.

A thank you to our project partners: Harvard School of Medicine, Department of Global Health and Social Medicine; Stony Brook University; L' Institut National De La Statistique De Madagascar (INSTAT) and PIVOT.

A special thank you to: Thomas Gillespie PhD, MSc, James Bliska PhD, Matthew Bonds PhD, Michele Parsons PhD, Cassidy Rist DVM, MPH and Rebecca Bridges for advising and guiding me throughout all stages of this project and the thesis process, Katie Smith MPH, Adria Narharimehefa, and Ashley Stumvoll for their help, collaboration, and company during sample collection and analysis in Madagascar.

Table of Contents

Introduction

Globally, acute diarrheal, lower respiratory, and other common communicable disease infections accounted for nearly 10% of the total mortality in 2013 [1]. Although global health efforts such as the United Nation's Millennium Project have had a substantial impact on the prevalence and death rates associated with communicable disease – in 1990, the mortality rate for the above was 20.64% – the progressive emergence of bacterial resistance to common/available antibiotics threatens the effectiveness of therapeutic drug treatment against the pathogens responsible [1, 2]. Unsurprisingly, global dissemination of antibiotic resistance (AR) is a topic of concern for many major global health organizations [2]; from an ecological perspective, AR is unavoidable and has subsequently been documented in every class of available antibiotic drugs [3]. Despite the growing interest, few studies have evaluated the relationship between the prevalence of resistance and potential socio-economic risk factors in a community setting [2].

Background

The Global Significance of Diarrheal Pathogens and Antibiotic Resistance

Diarrheal pathogens account for nearly 2.2 million deaths annually, mostly among children in Africa and South Asia [4]. There is a well-known association between frequent gastrointestinal infection at a young age and gut dysfunction; over time, poor absorption of nutrients will lead to a weakened immune response to pathogens and oral vaccines, stunted growth, and impaired cognitive development [4]. Those who are frequently exposed to enteric pathogens are subjected to diminished opportunities throughout life and thus have limited socio-economic mobility, perpetuating the classic poverty trap [5].

The historical effectiveness of antibiotics has allowed for the creation of a false belief among some patients that most/all infections require antibiotic therapy. Across the globe – developed nations included – prescriber understanding of the pharmacology of respective antimicrobials, differential diagnosis and pathogen epidemiology is highly variable from region to region, accounting for one of the largest factors responsible for improper antimicrobial use [6]. Two independent studies conducted in China and Vietnam respectively found that 63% of antimicrobials prescribed were found to be inappropriate and more than 70% were prescribed inadequate dosage [7, 8].

While the global health implications from AR are substantial, there has been a lackluster surveillance effort in the rural communities within developing nations [9]. As resistance to antimicrobials is a natural phenomenon, all antibiotics have the capacity to select for resistant strains of pathogens [6, 10]. Antibiotic-producing bacteria occur naturally in the environment and are ubiquitous throughout terrestrial and aquatic fauna and flora [11, 12]. The mixing of environmental strains with exogenous anthropogenic derived strains under clinical antimicrobial use or environmental pollutants create selective conditions that give rise to new resistant strains [13]. Resistant strains can emerge rapidly as a result of horizontal gene transfer of mobile genetic elements. Further, for many microorganisms, the fitness cost associated with harboring genes which encode resistance were found to have a negligible impact on their dissemination throughout a given

environment, as subsequent mutations/evolutions frequently result in the amelioration of potential negative costs [14].

Factors including antimicrobial agent abundance and cost, average duration of drug therapy, the route of administration, and the average dose interval could all play a central role in AR incidence fluctuations within a community [2, 6]. It has been estimated that at least 30% of all medicines sold in Africa are counterfeits, with antimicrobial agents being the most popular [10]. Counterfeit drugs often have a lower than stated dose, which not only promote resistance from weakened effectiveness, but more importantly, foster the diffusion of resistant pathogens into the environment when sanitation practices are poor [15]. Drugs that were properly manufactured but improperly stored may impact the potency and the effectiveness of the antibiotics. Tropical conditions are unfavorable for antimicrobial agents, requiring expensive electrical equipment, a constant electrical supply, an efficient supply chain and expert pharmaceutical handling; these expectations are often unmet. In the absence of a proper health care system that informs/prescribes usages and adequate sanitation infrastructure, there is a great potential for enhanced transmission rates of AR enteric bacteria, thus leading to an increase in morbidity from the decreased effectiveness of treatment [4, 16-18].

The Island Nation of Madagascar

Madagascar is characterized by high rates of enteric disease, intensified by poor living conditions and inadequate sanitation infrastructure. As a result, diarrheal disease has become the leading cause of mortality in children under five and the second leading factor for increased morbidity across all age groups, second only to malaria [19]. Over the past decade, moreover, the population has steadily increased to 22.4 million people and is projected to rise to 52.8 million by mid-2050 [20]. Seventy-seven percent of the Malagasy population live in rural communities and primarily rely on primitive and highly unsustainable slash-and-burn (Tavy) rice production for sustenance [20]. Typically, the Tavy process is repeated on a given plot of land every four to six years until the soil is exhausted of nutrients and left to be colonized by scrub vegetation or alien grasses,

wherein a new plot of forested land is converted via Tavy for harvesting, encouraging high levels of deforestation and anthropogenic disturbance [21]. This process has led to the loss of more than 90% of the original Malagasy forest-cover [22]. Although the Malagasy population density is, at its current state, relatively low (38 inhabitants per sq. kilometer), the projected population growth will inevitably lead to higher levels of anthropogenic disturbance within the environment, thus allowing for more frequent interactions between humans, livestock, and wildlife [20, 23].

Enteric pathogens, moreover, pose an elevated threat to the Malagasy, who depend on an underdeveloped healthcare system. In 2013, the measured density of hospitals/physicians in Madagascar was 0.5 per 10,000 and 0.16 per 100,000 respectively [24]. In comparison, the 2013 WHO African Region average for hospitals/physicians density was 2.7 per 10,000 and 0.8 per 100,000 respectively. Further, only 14% of the rural population frequently use WASH facilities – annually accounting for over 5,000 diarrheal deaths [18]. As a result, diarrheal disease has become the second leading cause of mortality associated with communicable disease across all age groups, responsible for nearly 8% of all recorded deaths in 2013 [1]. Further, a 2011 pilot study conducted within three rural villages of the Ranomafana Commune found that 77% of all humans included in the study tested positive – using PCR analysis – for virulence markers associated with *Shigella* spp., *Salmonella enterica*, *Vibrio cholera* or *Yersinia pseudotuburculosis* [25]. While it is likely that these results are being influenced by asymptomatic carriage, the results highlight the potential importance of these pathogens in explaining patterns of enteric disease within this region [26]. This pilot study also demonstrated that of those who took medications within four weeks of experiencing diarrhea or diarrhea like symptoms, 76% of participants reported antibiotic/anti-parasitic usage, 78% had used anti-inflammatories, and 52% used both to some degree [25]. This level of antibiotic usage in a country where the health care system is virtually absent and the pharmaceutical industry is unregulated is concerning given the potential of large-scale drug misusage/handling and the subsequent risk of increased dissemination of AR enteric bacteria. Furthermore, these enteric pathogens, particularly *Salmonella* spp., exhibit a high potential for zoonotic transmission, which has

been shown to exacerbate outbreaks [27-29].

The main objective for the following cross-sectional study was to describe the prevalence of both enteric infections using genetic virulence markers (*Shigella* spp. [*ipaH*], *Salmonella* spp. [*invA*], *Vibrio cholera* [*ctxA*], or *Yersinia* spp. [*yadA*]) and genes associated with AR for four common and available antibiotic classes (β-lactam [bla_{TEM}], sulfonamides [*sul1, sul2*], aminoglycoside [*strAB*], and trimethoprim [*dfrA7*]) in the population of Ranomafana – a commune of approximately 10,000 people, located within the Ifanadiana District. Patterns of enteric and AR infection and co-infection were evaluated for potential associations with demographic, socio-economic, and behavioral risk factors. Methods were adapted from the laboratory and surveillance standards recommended by the World Health Organization, the Centers for Disease Control and Prevention, and the Clinical and Laboratory Standards Institute [30-32].

The reality of the matter is simple: if the effectiveness of an antimicrobial is to be retained, effective epidemiological surveillance is essential. These results shed light on some of the underlying association that exists between enteric infection, AR, and socio-economic factors within a rural, tropical, community setting.

Methods

Hypotheses

1. The prevalence of enteric infection is expected to be high due to low levels of access to proper hygiene and sanitation infrastructure.
2. The prevalence of AR is expected to be high given low literacy rates, restricted access to proper healthcare facilities and medical practitioners, and unrestricted pharmaceutical industry.
3. A negative association is expected between household wealth and both enteric infection and AR, given that positive associations are often found between increased income and higher education levels, and access to health care/sanitation infrastructure in developing nations.
4. Antibiotic usage will have the strongest association with the observed prevalence and diversity of AR, given antibiotic use is completely unregulated in the Ranomafana Commune.

Ethical Statement

All protocols for the following were reviewed and approved by both the regional health officials stationed in Ifanadiana, Madagascar, and the Emory University Internal Review Board. All adult study participants (≥ 15 years of age) provided verbal and written consent for both themselves and their children prior to specimen collection and survey administration. In situations where written consent was unattainable due to illiteracy, oral consent was recorded on the survey sheets by the native interpreter conducting the interview. All forms of data collection were performed by L' Institut National De La Statistique De Madagascar (INSTAT) in Malagasy, the native language of Madagascar. Participants were anonymously given unique identifiers, such that the subsequent dataset provided to the investigators was void of all individual identifiers.

Study Area and Population

The commune of Ranomafana, Madagascar (47°18' 40 – 47°37' E and 21°2' – 21°25' S) is located within the westernmost portion of the Ifanadiana district in southeastern Madagascar [25]. The commune is located within the humid eastern rainforest corridor, a small strip of land that extends in a north-south direction for a distance of about 200 km along the eastern escarpment of Madagascar [33]. The commune has an average elevation of 466 m and an average annual temperature of 23.9 °C. About 1700 mm of precipitation falls annually, mostly from November to March [25]. The periphery of Ranomafana National Park (RNP) – a 41,600 hectare National Park, established in 1991 and has since declared a UNESCO World Heritage site – define the commune's northern and southern borders. The vast majority of lands surrounding the populated regions of the commune are agricultural lands, consisting largely of rice, cassava, and banana subsistence farms [25].

The commune was further divided into eight geographically distinct census tracts or 'clusters' for the Demographic and Health Survey in Madagascar conducted between 2008-09. This study focuses on the four clusters previously described in a baseline study performed by INSTAT for PIVOT – a new healthcare NGO established in Ranomafana [34] – earlier in 2014, where all households within the four clusters were enumerated.

Data Collection and Surveys

Household and individual surveys were administered on July 1st – July 14th 2014 by two, four-member teams from INSTAT within four of the eight clusters that comprise the Ranomafana Commune. Household eligibility criteria were based on prior participation in a two-step randomized survey of 1520 households within the Ifanadiana District, conducted earlier that year. Within each cluster, a random sample of 20 eligible households was drawn from the population without replacement, for a base study population of 80 households. Once consent was received, the male and female head-of-household (HOH) were asked to participate in in-person interviews that were designed to assess: demographic information, individual health status, common hygiene practices,

water usage/treatment, antibiotic usage, and household economics. Distinct surveys were created for men, women, and children, where female HOH's were asked to answered questions about any child that was ≤5 years old. Potential risk factors associated with: diarrheal disease, enteric infection, and the development of AR were assessed in all surveys with multiple redundancies, such to reduce the risk of exposure misclassification and general information bias.

Specimen Collection and Transport

All survey participants, regardless of age, were asked to provide a fresh fecal specimen for the molecular analysis of diarrheal pathogen infection and the absence/presence of genes that encode for AR. Specimen cups were distributed by INSTAT to all consenting individuals within a household, and INSTAT members were trained to provide ample instruction of proper collection protocols designed to insure freshly voided specimens were aseptically collected [31]. All fecal specimens collected were linked to survey data via a unique identification number at the point of collection.

Following collection, approximately one milliliter of fecal material was extracted from the center of each specimen, homogenized, and stored in an equal volume of RNAlater® nucleic acid stabilizing buffer (Qiagen, Valencia, CA) at -20°C until transport to the United States. Specimen preparation and preservation was at the Centre ValBio Research Laboratory, in Ranomafana, Madagascar within 24 hours from the time of collection.

Molecular Methods

Total nucleic acid was extracted from all fecal specimens preserved in RNAlater using the FastDNA SPIN Kit (MP Biomedicals, Solon, OH), following the manufacturer recommended procedures. The concentration of DNA was quantified using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and subsequently standardized at 150 $ng/\mu l$ using DNase/Pyrogen-Free Water. Extractions were stored at -20°C until later analysis [35].

Using single PCR assays, DNA extractions were screened for the presence of five target genes associated with enteric AR (*strAB, bla*$_{TEM,}$ *sul1, sul2, dfrA7*) and four virulence markers associated with *Shigella* spp., *Salmonella* spp, *Vibrio cholera*, and *Yersinia* spp. (*iaph*, *invA, ctxA, yadA*) [36-39]. Additionally, each extraction was screened for the presence of the bacterial 16s rRNA gene, using universal primers described in [40], such to evaluate if PCR was inhibited in any given specimen. All primers were synthesized by Eurofilms MWG Operon (Stony Brook University, Stony Brook, NY) and are listed in Appendix I.

Each PCR reaction was run in a 1.5 micro-centrifuge tube and consisted of 25 μl of Platinum PCR mix (Invitrogen, Life Technologies, Grand Island, NY), 2.0 μl of of template, and 0.5 μl of a 25 μM mix of each primer set. Sterile distilled water was used as a negative control. Positive strains of *V. cholerae* and *S. flexneri* strains were obtained from American Type Culture Collection (ATCC, Manassas, VA). The 32,777 strains in the Bliska Laboratory collection were used as a positive control for *Y. pseudotuberculosis* (Stony Brook University, Stony Brook, NY). A strain of *Salmonella enterica* serotype Enteriditis – which exhibited both the phenotypic and genotypic characteristics of encoded resistance genes in question – was obtained from the National Center for Emerging and Zoonotic Infectious Diseases (The Centers for Disease Control and Prevention, Atlanta, GA) and used as the positive control for *invA* and all resistance markers. Genomic DNA was extracted from control isolates using the boiling method. All primers were tested on each of the positive control strains to confirm sensitivity and specificity. Ten microliters of each PCR product was analyzed with gel electrophoresed using 1.2% Agarose Gels. Gels were stained with 3 5 μl of GelRed Nucleic Acid Gel Stain (Biotium, Hayward, CA). Gel images were captured under UV exposure using a Gel Doc illumination system (Bio-Rad, Hercules, CA).

Data Analysis and Statistical Methods

Given that a fecal specimen was necessary to categorize the outcomes of interest for this study, eligible study participants were restricted to those that provided a specimen. Raw data were entered using Microsoft Excel (Redmond, WA) and subsequently analyzed using SAS Version 9.4 (SAS Institute, Cary, NC).

Individuals were considered infected with an enteric pathogen (*Shigella* spp., *Salmonella* spp. *Vibero cholerae,* or *Yersinia* spp.) if their respective fecal specimen tested PCR positive for any one of the four species-specific virulence markers of interest (*iahp, invA, ctxA,* or *yadA*). Further, individuals were considered carriers for a given AR gene if their respective fecal specimen tested PCR positive for any one of the five resistance genes in question (bla_{TEM}. *sul1*, *sul2*, *strAB*, or *dfra7*). Survey data from INSTAT were evaluated for missing or implausible values using frequency procedures for categorical variables and univariate procedures for continuous variables. Following this preliminary analysis, it was determined that missing survey values were not completely at random, given that the distribution of non-response was differential with respect to an individual's age. Household Wealth Index, a continuous variable calculated by INSTAT using Principal Component Analysis and traditional household wealth indicators used in the Malagasy Demographic Health Survey (Appendix II), was categorized into a binary variable (0,1) such that the odds of a respective outcome could be compared between the bottom 80% of the population and the top 20%.

All subsequent analyses were preformed using SAS Survey Procedures (i.e. SURVEYMEANS, SURVEYFREQ, SURVEYLOGISTIC), which employ Taylor linear approximation for variance estimation to adjust for the clustered sample design. Household weights, calculated by INSTAT (Appendix II), were applied to provide unbiased effect estimates at the commune level.

Descriptive statistics for outcome variables, exposures variables, and covariates were calculated using the entire study population when possible. Prevalence was determined for each study

outcome using weighted frequencies and were subsequently analyzed for correlations in positivity for all pathogen/enteric-specific pairwise comparison using a Bonferroni adjusted McNemar's test of marginal homogeneity (α=0.0125). A Wald Chi-square test (α=0.05) was used to quantify potential bivariate associations between the study outcomes, risk factors, and potential covariates.

Simple logistic regression models were created to evaluate the crude odds ratios (ORs) between study outcomes and risk factors of interest, as well as the quantify strength of association between all potential confounders/effect modifiers (α=0.1). Variable selection for subsequent multivariable analysis was primarily informed by the literature review, as well as the bivariate results. Likelihood ratio tests were used to test the significance of potential interaction terms included in multivariate models, while potential confounders were evaluated through a comparison between unadjusted and adjusted ORs.

Results

Population Characteristics

Seventy-six households elected to partake in the study and 65% of participants agreed to provide a fecal specimen resulting in a study population of 225 characterized individuals (Table 1). Gender distribution among participants was approximately evenly split between males (Weighted n=109.87, CI_{95}: 90.38,129.38), and females (Weighted n=110.72, CI_{95}: 87.62, 133.82). With respect to age, the distribution among participants was 13% children under the age of five years (Weighted n=28.48, CI_{95}: 17.67, 38.28), 26% children age 5-14 years (Weighted n=57.74, CI_{95}: 41.28, 74.20), 49% adults age 15-49 years (Weighted n=106.52, CI_{95}: 88.00, 125.05), and 12% adults 50 years and older (Weighted n=26.26, CI_{95}: 16.50, 36.01).

The median income among participating households was MGA 1,614,975 (~US$573) per year. Approximately 80% of study participants reported having access to any form of latrine (improved or unimproved) at their place of residence (Weighted n=175.63, CI_{95}: 141.38, 209.89). Curiously, nearly 80% of study participants reported 'No Treatment' when asked how they treated their primary source of water (Weighted n=178.30, CI_{95}: 140.51, 216.08). Sixteen percent of study participants reported experiencing a bout of diarrhea, defined as the passage of three or more loose or liquid stools per day (or more frequent passage than is normal for the individual), within 4 weeks prior to being surveyed (Weighted n=20.55, CI_{95}: 12.29, 28.81). Of the individuals who responded to antibiotic usage questions, 56% responded as having taken some form of antibiotic within 6 weeks prior to being surveyed (Weighted n=63.35, CI_{95}: 48.55, 78.15)

Prevalence of Enteric Infection and Antibiotic Resistance

The weighted prevalence of enteric virulence markers among study participants were: 22.04% for *Shigella* spp. (CI_{95}: 16.22, 27.85), 5.4% for *Salmonella* spp. (CI_{95}: 2.03, 8.77), 0.35% for *Vibrio cholerae* (CI_{95}: 0.00, 1.05), and 7.11% for *Yersinia* spp. (CI_{95}: 3.63, 10.59) (Table 2). *Shigella* spp. virulence markers were found at a much higher frequency when compared to *Salmonella* spp.

(McNemar's: 23.46, df=1, p<0.0001), and *Yersinia* spp. (McNemar's: 22.26, df=1, p<0.0001). Further, the frequency of the *Salmonella* spp. virulence marker was not found to differ significantly when compared to *Yersinia* spp. (McNemar's: 0.58, df=1, p=0.445).

In regards to genes encoding AR among study participants, weighted prevalence among study participants were: 42.11% for β-lactam resistance genes (CI_{95}: 34.93, 49.29), 54.29% for *sul1* sulfonamide resistance (CI_{95}: 46.91, 61.66), 31.77% for *sul2* sulfonamide resistance (CI_{95}: 25.23, 38.30), 50.47% for aminoglycoside resistance genes (CI_{95}: 43.75, 57.20), and 43.41% for trimethoprim resistance genes (CI_{95}: 35.84, 50.99) (Table 2). A significant difference between gene frequencies was found when comparing sulfonamide resistance (*sul1* or *sul2*) to: β-lactam resistance (McNemar's: 18.73, df=1, p<0.0001), aminoglycoside resistance (McNemar's: 34.19 df=1, p<0.0001), and trimethoprim resistance (McNemar's: 15.32, df=1, p<0.0001). All other two-way comparisons between resistance genes were not significant.

Risk Factors for Enteric Infection

Seven multivariable logistic regression models were constructed to evaluate associations between infection with any enteric pathogen and: sex, age, cluster, household wealth index, access to household latrine, primary water treatment methods, and experiencing diarrhea within four weeks prior to the study (Table 3). No significant association was found between enteric infection and: sex, access to household latrine, primary water treatment methods, or experiencing diarrhea within four weeks prior to the study.

However, when compared to individuals who were under five years of age, the odds of infection among individuals ages five to fourteen were nearly 2.87 times higher than odds of infection for the reference group (CI_{95}: 1.03, 8.03, Wald χ^2 p=0.044). Further, when compared to cluster 57, the odds of infection among individuals in cluster 54 were nearly 6.20 times higher than odds of infection for the reference group – after adjusting for age and household wealth index (CI_{95}: 2.42, 15.90, Wald χ^2 p<0.0001).

Risk Factors for Antibiotic Resistance

Five multivariable logistic regression models were constructed to evaluate the associations between carriage of any AR encoding gene and: sex, age, household wealth index, cluster, and any antibiotic used within the six months prior to the study (Table 4). No significant associations were found between AR carriage and: sex, cluster, or – curiously – antibiotic usage.

A significant association was found, however, when comparing individuals aged 5-14 and 15-49 to those <5 years old. The odds of AR carriage among participants 5-14 were 4.92 times higher (CI_{95}: 1.66, 14.61, Wald χ^2 p=0.004) compared to participants less than five years of age. Further, the odds of AR carriage among participants 15-49 were 4.15 times higher (CI_{95}: 1.39, 12.44, Wald χ^2 p=0.011) compared to participants less than five years of age. Additionally, the odds of AR carriage among participants in the bottom 80% of the household wealth index were 4.27 times greater (CI_{95}: 1.56, 11.70, Wald χ^2 p=0.005) compared to participants in the top 20% of the household wealth index.

Odds of Antibiotic Resistance Carriage given Enteric Infection

Finally, twelve multivariable logistic regression models – controlling for co-infection – were constructed to evaluate the odds of carrying of specific AR encoding genes given infection with either *Shigella* spp., *Salmonella* spp., or *Yersinia* spp.; *Vibrio cholerae* was left out of the analysis given its negligible prevalence (Table 5). Significant associations were found between *Shigella* spp. infection and harboring β-lactam resistance genes (aOR: 4.72, CI_{95}: 2.03, 11.01, Wald χ^2 p<0.0001) and aminoglycoside resistance genes (aOR: 7.48, CI_{95}: 3.29, 17.01, Wald χ^2 p<0.0001). Additionally, a significant association was found between *Yersinia* spp. infection and harboring β-lactam resistance genes (aOR: 4.73, CI_{95}: 1.24, 18.03, Wald χ^2 p<0.0001).

Discussion

Few studies have examined the eco-epidemiology of enteric pathogens and AR within a rural, tropical community setting. Using well-defined and validated molecular methods, this cross-sectional study sought to determine the prevalence of four globally important enteric pathogens and identify potential risk factors for infection amongst the human population of the Ranomafana Commune, Madagascar [4, 41]. Further, given the well-documented need for AR surveillance within developing nations, this study attempted to evaluate the prevalence of AR, and identify potential risk factors for harboring genetic elements that encode resistance to four commonly used and available classes of antibiotics [2, 25].

The Prevalence of Enteric Infection and Associated Risk Factors

Of the four pathogens investigated, the prevalence of infection (Table 2) with *Shigella* spp. was the highest (22.04%), followed by *Yersinia* spp. (7.11%), *Salmonella* spp. (5.40%), and *Vibrio cholerae* (0.35%). These results serve to validate and extend one of the major conclusions presented in the 2013 case-control GEMS study, which found that the majority of cases of moderate-to-severe diarrhea in children under 5 could be attributed to four pathogens: *Shigella*, rotavirus, *Cryptosporidium*, and enterotoxigenic *Escherichia coli* producing heat-stable toxin (ST-ETEC) [41]. It should be noted, however, that the *iaph* gene used to identify *Shigella* infections is not completely specific to only *Shigella*. Enteroinvasive *Escherichia coli* (EIEC) strains have been found to also express the toxin encoded by *iaph*, thus it is possible that some of the bacteria detected were indeed EIEC, not *Shigella*. However, the symptoms of an EIEC infection is nearly identical to Shigellosis and thus presents an equal threat to the health and well being of the inhabitants of Ranomafana [42].

That being said, these results suggest that an intervention, targeted on disrupting the transmission of *Shigella*/EIEC could potentially reduce the overall mortality and morbidity attributed to diarrheal diseases within the Ranomafana commune. Further, in comparison to a previous study of enteric infections in this region [24], although infection still remains high, there has been an overall

reduction in enteric infections over the last three years. While there are a number of plausible explanations for this apparent reduction, it is likely that the reductions stem from improved WASH practices among individuals and, generally speaking, improved water quality. Unfortunately, due to study limitations: such as relatively small sample size, inconsistent survey notation, and high survey item non-response; the impact of factors could not comprehensively be evaluated with high validity. Preliminary bivariate analysis made evident that these data were undoubtedly vulnerable to over-stratification, subsequently resulting in unstable and non-valid ORs. As a result, key risk factors for enteric infection (i.e. access to health care, household access to an improved latrine, hand washing practices, and livestock practices) could not be evaluated. Additionally a condensed outcome of infection, defined by infection with any enteric pathogen, was created allowing for a more statistically robust analysis. Further, given the intrinsic limitations of the cross-sectional design, the temporality between exposure and disease outcome cannot validly be established, and thus restricts the degree to which study outcomes can be interpreted.

This study did, however, identify important relational associations between enteric infection and age/household locale within the Ranomafana commune (Table 3). Individuals aged 5-14 had a increase odds of enteric infection when compared to individuals under the age of 5. This result carries with it a slew of curious interpretations, given that children under the age of 5 have historically been shown to have the highest risk of mortality and morbidity from diarrheal disease [4, 18, 41, 43]. These results suggest that while the current methods in place for reducing mortality and morbidity of diarrheal disease have been successful among children less than five years, they lose effectiveness among school-aged children. As such, it may be useful for future studies to evaluate the effectiveness of health and hygiene sensitization programs – specifically when stressing the importance of WASH based transmission control measures – in place among school-aged children, such that future educational interventions have a more targeted and informed approach for establishing future programs and reforming those already in place.

Further, it was found that the odds of enteric infection for individuals living within cluster 54, a region of Ranomafana in close proximity to the Centre ValBio research station and along the only major roadway, were nearly 6.2 times higher compared to individuals living in cluster 57 (the reference group), the most rural region of Ranomafana. This result conflicts with patterns seen elsewhere in regards to the risk of enteric infection along an infrastructural gradient [4, 25, 41]. While this study falls short of identifying the potential factors at play in this situation, the scope of possible explanations are vast. Given the regions close proximity to Centre ValBio – a well-established institution in Ranomafana, where many health-based services/interventions originate – one possibility could be a general water source contamination to the potential existence combined with a false sense of health security. As such, future studies should systematically investigate why individuals that live within this region are subject to a marked increase in the odds of enteric infection when compared to others in Ranomafana.

Curiously, household wealth index did not prove to be a significant risk factor for enteric infections, which is surprising given that wealthier individuals, generally speaking, have greater access to improved water sources, sanitation infrastructure, healthcare, and educational programs [5, 41]. This lack of significance suggests that the transmission of enteric pathogens may perhaps be more of a systemic infrastructural issue affecting the Ranomafana commune as a whole, rather than only those poorest individuals. As such, future studies should work to identify the potential reservoirs for enteric pathogens, like *Shigella*, *Salmonella*, and *Yersinia*, and evaluate possible down-stream transmission pathways that could produce this 'universal predisposition' for exposure and subsequent infection.

The Prevalence of Antibiotic Resistance and Associated Risk Factors

This was the first study of its kind to evaluate the prevalence of AR encoding genes in this region of Madagascar. Of the five AR encoding genes evaluated (Table 2), the prevalence of sulfonamide (*sul1*) and aminoglycosides (*strAB*) resistance encoding genes were found to be the highest – both above 50% – followed by trimethoprim (43.41%) and finally β-lactam (42.11%). These results make sense, given that Co-trimoxazole, Gentamicin, and Amoxicillin are all cheaply and readily available in the two pharmacies of Ranomafana, and, more importantly, could be found and purchased in most shops scattered along the major-roadway that transects the commune.

It is difficult to firmly establish just how elevated these results are, given the widespread overuse of antibiotics globally and the lack of community setting surveillance [2]. These results, however, suggest that gut flora within the human population of the Ranomafana commune could serve as a reservoir for AR, thus creating the potential for horizontal transfer of mobile genetic elements carrying antimicrobial resistance genes from commensal to pathogenic bacteria during times of infection, potentially reducing the effectiveness of previously dependable treatments [44, 45].

Once again due to study limitations stemming from a relatively small sample size, inconsistent survey notation, and high survey item non-response – particularly in regards to the questions concerning antibiotic usage practice (55% response rate) – the analysis of risk factors associated with harboring said AR encoding genes was limited in scope, and the definition of AR had to be condensed to those harboring any one of the genes in question such that statistically valid estimates could be calculated. After accounting and adjusting for these limitations, however, this study was able to identify two significant and highly relevant associations (Table 4).

First, it was found that among individuals aged 5-14 and 15-45, the respective odds of harboring any AR encoding gene was 4.92 and 4.15 times higher then the odds for those individuals under five years of age. These results suggest that access to antibiotics may be increased among these two age categories or, more likely, these individuals are less likely to fully complete the recommended/prescribed dose regimen. As such, future studies should focus on taking a more in-

depth and targeted approach to analyzing how and why antibiotics are being used within these age categories, such that educational programs can be more effective in conveying the significance of proper usage, as it relates to health and wellbeing on both the individual and community levels.

Additionally, a rather strong association was found between harboring any AR encoding gene and the wealth status of a given household. The odds of harboring any AR encoding gene were nearly 4.27 times higher among individuals belonging to households defined as being in the bottom 80% of the household wealth index compared to those individual belonging to a household in the top 20% (Table 4). One possible explanation for this result could be that low literacy/education rates among these individuals are preventing them from reading or fully understanding the directions on proper use/dosage. Another possible explanation could be that individuals within this wealth class cannot afford to purchase a complete dosage regimen, thereby encouraging AR development through incomplete infection clearage [2, 7, 46]. This hypothesis is further bolstered given that most pharmacies and shops that sell antibiotics will sell individual pills to customers (Patricia Wright, PhD, personal communication). This is particularly concerning considering that strong associations were found between infection with *Shigella* and harboring genes encoding for β-lactam and aminoglycoside resistance (Table 5).

Given the above, it is essential that health organizations, pharmacies, and clinics begin taking a more active stance in antimicrobial resistance monitoring and begin to establish education programs that reach these at-risk individuals such to encourage more judicious antibiotic usage practices aimed at minimizing the spread of resistance and maintaining the effectiveness of antibiotic treatments [47, 48].

Conclusions and Recommendations

Although the above utilized a cross-sectional design and is thus limited in how results can be interpreted, interesting and rather useful conclusions can be drawn and subsequently extended to the Ranomafana community and PIVOT, a health care systems strengthening NGO incubated by Partners in Health and the Centre ValBio that supports the existing health care system of the Malagasy Ministry of Health for the Ifanadiana District. While the significance of *Shigella* infections have previously been established both globally [41], and in Ranomafana [25], the results presented here further validate the need to introduce a targeted intervention that actively works to disrupt the traditional transmission pathways of *Shigella*, which potentially could greatly reduce the overall health burden attributed to diarrheal disease in Ranomafana.

Additionally, the AR results present useful surveillance information that can be used to inform clinical prescription practices, especially in light of the results presented in Table 5, which potentially suggest that β-lactam and aminoglycoside class antibiotics may have decreased effectiveness in treating suspected *Shigellosis* infections. Further, the results presented demonstrate the need to establish educational programs that reach out to those of a lower wealth standing, emphasizing the significance of proper antibiotic usage and WASH practices. Based on the hypothesis that those in the bottom 80% of the wealth distribution are incorrectly self-administering antibiotics – or not fully completing recommended dosage regimens – PIVOT should work with community leaders and pharmacists to educate them on the recommended administration of antibiotics to maintain their effectiveness over time.

References

1. Institute for Health Metrics and Evaluation (IHME). GBD Compare. Seattle, WA: IHME, University of Washington, **2015**.
2. World Health O. Antimicrobial Resistance: Global Report on Surveillance, **2014**.
3. Harbarth S, Samore MH. Antimicrobial resistance determinants and future control. Emerging infectious diseases **2005**; 11:794-801.
4. Wardlaw T, Salama P, Brocklehurst C, Chopra M, Mason E. Diarrhoea: why children are still dying and what can be done. Lancet **2010**; 375:870-2.
5. Bonds MH, Keenan DC, Rohani P, Sachs JD. Poverty trap formed by the ecology of infectious diseases. Proceedings of the Royal Society of London B: Biological Sciences **2010**; 277:1185-92.
6. World Health O. WHO global strategy for containment of antimicrobial resistance. World Health Organization, **2001**.
7. Hui L LX, Zeng XJ, Dai YH, Foy HM. . Patterns and determinants of use of antibiotics for acute respiratory tract infection in children in China. Pediatr Infect Dis J **1997**; 16:560–4.
8. Chalker J PN. Combating the growth of resistance to antibiotics: antibiotic dose as an indicator for rational drug use. ICIUM. Chang Mai, **1997**.
9. Jacobson v. Massachusetts. Us. Vol. 197: Supreme Court, **1905**:11.
10. World Health O. Counterfeit medicines: an update on estimates. Geneva: WHO International Medical Products Anti-Counterfeiting Task Force **2006**.
11. Martinez JL. Antibiotics and antibiotic resistance genes in natural environments. Science **2008**; 321:365-7.
12. Riesenfeld CS, Goodman RM, Handelsman J. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. Environmental Microbiology **2004**; 6:981-9.
13. Wellington EMH, Boxall ABA, Cross P, et al. The role of the natural environment in the emergence of antibiotic resistance in Gram-negative bacteria. Lancet Infectious Diseases **2013**; 13:155-65.
14. Andersson DI, Levin BR. The biological cost of antibiotic resistance. Current opinion in microbiology **1999**; 2:489-93.
15. Hart CA, Kariuki S. Antimicrobial resistance in developing countries. BMJ: British Medical Journal **1998**; 317:647.
16. Measham AR, Alleyne G, Mills A, et al. Disease control priorities in developing countries. Washington, DC: World Bank and Oxford University Press, **2006**.
17. Okeke IN, Aboderin OA, Byarugaba DK, Ojo KK, Opintan JA. Growing problem of multidrug-resistant enteric pathogens in Africa. Emerging infectious diseases **2007**; 13.
18. World Health Organization U-W. UN-Water Global Analysis and Assessment of Sanitation and Drinking-Water (GLAAS) 2014 - Report. In: Organization WH, ed. Investing in Water and Sanitation: Increasing Access, Reducing Inequalities, **2014**.
19. Evaluation IfHma. HME Data Visualizations: GDB Compare **2013**.
20. Bureau PR. 2014 World Population Data Sheet. In: Bureau PR, ed, **2014**.
21. Kull CA. Isle of fire: the political ecology of landscape burning in Madagascar. Vol. 245. University of Chicago Press, **2004**.
22. Harper GJ, Steininger MK, Tucker CJ, Juhn D, Hawkins F. Fifty years of deforestation and forest fragmentation in Madagascar. Environmental Conservation **2007**; 34:325-33.
23. Bublitz DC, Wright PC, Rasambainarivo FT, Arrigo-Nelson SJ, Bodager JR, Gillespie TR. Pathogenic enterobacteria in lemurs associated with anthropogenic disturbance. Am J Primatol **2014**.
24. World Health Organization. World Health Statistics 2015. World Health Organization, **2015**.
25. Bublitz DC, Wright PC, Bodager JR, Rasambainarivo FT, Bliska JB, Gillespie TR. Epidemiology of pathogenic enterobacteria in humans, livestock, and peridomestic rodents in rural Madagascar. PloS one **2014**; 9:e101456.

26. Frickmann H, Schwarz NG, Rakotozandrindrainy R, May J, Hagen RM. PCR for enteric pathogens in high-prevalence settings. What does a positive signal tell us? Infectious diseases (London, England) **2015**; 47:491-8.
27. Chen Y, Liang W, Yang S, et al. Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: clinical analysis and characterisation of viral genome. The Lancet **2013**; 381:1916-25.
28. Daszak P, Zambrana-Torrelio C, Bogich TL, et al. Interdisciplinary approaches to understanding disease emergence: The past, present, and future drivers of Nipah virus emergence. Proceedings of the National Academy of Sciences of the United States of America **2013**; 110:3681-8.
29. Garten RJ, Davis CT, Russell CA, et al. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. Science **2009**; 325:197-201.
30. World Health O. Surveillance standards for antimicrobial resistance: World Health Organization, **2001**.
31. Prevention CfDCa. Laboratory Methods for the Diagnosis of Epidemic Dysentery and Cholera. Atlanta, Georgia: CDC, **1999**.
32. CLSI. Performance Standards for Antimicrobial Disk Susceptibility Tests; Approved Standard. Eleventh ed. Wayne, PA: Clinical and Laboratory Standards Institute, **2012**.
33. Wright PC, Andriamihaja B, Terborgh J, Schaik Cv, Davenport L, Madhu R. Making a rain forest national park work in Madagascar: Ranomafana National Park and its long-term research commitment. Making parks work: Strategies for preserving tropical nature **2002**:112-36.
34. PIVOT: About. Available at: http://pivotworks.org/what-we-do/mission/.
35. Da Silva AJ, Bornay-Llinares FJ, Moura IN, Slemenda SB, Tuttle JL, Pieniazek NJ. Fast and reliable extraction of protozoan parasite DNA from fecal specimens. Molecular Diagnosis **1999**; 4:57-64.
36. Chen S, Zhao S, White DG, et al. Characterization of multiple-antimicrobial-resistant Salmonella serovars isolated from retail meats. Applied and environmental microbiology **2004**; 70:1-7.
37. Rasheed JK, Tenover FC. Detection and Characterization of Antimicrobial Resistance Genes in Pathogenic Bacteria*. Manual of Clinical Microbiology, 10th Edition: American Society of Microbiology, **2011**.
38. Thoerner P, Kingombe CB, Bögli-Stuber K, et al. PCR detection of virulence genes in Yersinia enterocolitica and Yersinia pseudotuberculosis and investigation of virulence gene distribution. Applied and environmental microbiology **2003**; 69:1810-6.
39. Wang RF, Cao WW, Cerniglia CE. A universal protocol for PCR detection of 13 species of foodborne pathogens in foods. Journal of applied microbiology **1997**; 83:727-36.
40. Gulati AS, Shanahan MT, Arthur JC, et al. Mouse background strain profoundly influences Paneth cell function and intestinal microbial composition. PloS one **2012**; 7:e32403.
41. Kotloff KL, Nataro JP, Blackwelder WC, 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. The Lancet **2013**; 382:209-22.
42. Van den Beld M, Reubsaet F. Differentiation between Shigella, enteroinvasive Escherichia coli (EIEC) and noninvasive Escherichia coli. European journal of clinical microbiology & infectious diseases **2012**; 31:899-904.
43. Jones KE, Patel NG, Levy MA, et al. Global trends in emerging infectious diseases. Nature **2008**; 451:990-3.
44. Wellington EM, Boxall AB, Cross P, et al. The role of the natural environment in the emergence of antibiotic resistance in Gram-negative bacteria. The Lancet infectious diseases **2013**; 13:155-65.
45. Bartoloni A, Benedetti M, Pallecchi L, et al. Evaluation of a rapid screening method for detection of antimicrobial resistance in the commensal microbiota of the gut. Transactions of the Royal Society of Tropical Medicine and Hygiene **2006**; 100:119-25.

46. Bell BG, Schellevis F, Stobberingh E, Goossens H, Pringle M. A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance. BMC Infectious Diseases **2014**; 14:13.
47. Bennish ML, Khan WA. What the Future Holds for Resistance in Developing Countries. **2010** (Sosa AJ, Byarugaba DK, AmabileCuevas CF, Hsueh PR, Kariuki S, Okeke IN, eds. Antimicrobial Resistance in Developing Countries).
48. Moeller AH, Shilts M, Li YY, et al. SIV-Induced Instability of the Chimpanzee Gut Microbiome. Cell Host & Microbe **2013**; 14:340-5.
49. Gwatkin DR, Rutstein S, Johnson K, Suliman E, Wagstaff A. Socio-economic differences in health nutrition and population. Madagascar: 1997. **2007**.

Tables

Table 1: Cluster Population Size and Individual Characteristics of Study Population of Ranomafana Commune, Ifanadiana District, Madagascar.

Characteristic	n[1]	Frequency /Median	Weighted Frequency[2]	95% Weighted Confidence Limits
Sex	225			
Male		113	109.87	90.38 – 129.38
Female		112	110.72	87.62 – 133.82
Age, years	223			
<5		25	28.48	17.67 – 38.28
5-14		65	57.74	41.28 – 74.20
15-49		105	106.52	88.00 – 125.05
>49		28	26.26	16.50 – 36.01
Household Latrine Available	225			
Yes		171	175.63	141.38 – 209.89
No		54	44.97	23.05 – 66.88
Primary Water Treatment Method	225			
Boiled		35	32.64	13.09 – 52.19
Nothing		179	178.30	140.51 – 216.08
Other[3]		11	9.66	0.00 – 19.88
Diarrhea (Past 4 Weeks)	124			
Yes		19	20.55	12.29 – 28.81
No		105	111.45	87.56 – 135.34
Antibiotic Usage (Past 6 Months)	113			
Yes		59	63.35	48.55 – 78.15
No		54	50.38	38.20 – 62.55
Household Wealth Index[5]	225			
Top 20%		42	47.52	19.79 – 75.26
Bottom 80%		183	220.60	135.68 – 210.48
Median Yearly Income[4]	225	1,614,075	-	-

[1]Total n varies due to incomplete notation for some survey questions
[2] Adjusted for Complex Cluster Sampling and Non-Response
[3]Includes Mechanical Filtration and Chemical Treatment
[4]Income reported in Malagasy Ariary (1 USD = 2,860 MGA)
[5]See Appendix II for Wealth Index Calculation Methods

Table 2: Weighted Prevalence of Enteric Pathogen Virulence Markers and Antibiotic Resistant Genes Detected in Fecal Specimens within the Study Population of Ranomafana Commune, Ifanadiana District, Madagascar.

Pathogen Virulence Marker	**n**	**Weighted Frequency[1]**	**Weighted Prevalence[1]**	**95% Weighted Confidence Limits[1]**
iaph	225	48.62	22.04	16.22 – 27.85
invA	225	11.91	5.40	2.03 – 8.77
ctxA	225	0.78	0.35	0.00 – 1.05
yadA	225	15.69	7.11	3.63 – 10.59
Antibiotic Resistance Gene	**n**	**Weighted Frequency[1]**	**Weighted Prevalence[1]**	**95% Weighted Confidence Limits[1]**
bla_{TEM}	225	92.89	42.11	34.93 – 49.29
sul1	225	119.75	54.29	46.91 – 61.66
sul2	225	70.08	31.77	25.23 – 38.30
sul1 and *sul2*	225	55.78	25.29	19.01 – 31.57
strAB	225	111.35	50.47	43.75 – 57.20
dfrA7	225	95.78	43.41	35.84 – 50.99
16s rRNA Control Gene	**n**	**Positive**	**Negative**	
U16s	225	225	0	

[1]Adjusted for Complex Cluster Sampling

Table 3: Risk Factors for Infection with Any Enteric Pathogen within the Study Population of Ranomafana Commune, Ifanadiana District, Madagascar

Variable	n[1]	OR[2]	95% Wald Confidence Limits[2]	Wald χ^2 p-value[1,2]
Sex	225			
Male		Ref	Ref	Ref
Female		0.81	0.41 – 1.59	0.533
Age, Years	223			
<5		Ref	Ref	Ref
5-14		2.87	1.03 – 8.03	0.044*
15-49		1.28	0.49 – 3.29	0.616
>49		0.939	0.23 – 3.91	0.931
Household Wealth Index	225			
Top 20%		Ref	Ref	Ref
Bottom 80%		2.23	0.77 – 6.47	0.140
Cluster[3]	223			
54		6.20	2.42 – 15.90	0.0001*
55		0.937	0.29 – 3.00	0.913
56		0.77	0.22 – 2.72	0.690
57		Ref	Ref	Ref
Household Latrine Available[4]	225			
Yes		Ref	Ref	Ref
No		1.24	0.50 – 3.06	0.645
Primary Water Treatment Method[5]	225			
Boiled		Ref	Ref	Ref
Nothing		0.63	0.17 – 2.37	0.500
Other[6]		0.92	0.12 – 6.945	0.932
Diarrhea (Past 4 Weeks)	124			
Yes		1.12	0.39 – 3.16	0.838
No		Ref	Ref	Ref

[1]Total n varies due to incomplete notation for some survey questions
[2]Adjusted for Complex Cluster Sampling and Non-Response
[3]Adjusted for Age and Household Wealth Index
[4]Adjusting for Household Wealth Index
[5]Adjusting for Cluster
[6]Includes Mechanical Filtration and Chemical Treatment
*Significant at 0.05 α level

Table 4: Risk Factors for Carriage of Any Antibiotic Resistance Gene within the Study Population of Ranomafana Commune, Ifanadiana District, Madagascar

Variable	n[1]	OR[2]	95% Wald Confidence Limits[2]	Wald χ^2 p-value[1,2]
Sex	225			
Male		Ref	Ref	Ref
Female		0.79	0.29 – 2.11	0.632
Age, Years	223			
<5		Ref	Ref	Ref
5-14		4.92	1.66 – 14.61	0.004*
15-49		4.15	1.39 – 12.44	0.011*
>49		1.38	0.40 – 4.77	0.616
Household Wealth Index	225			
Top 20%		Ref	Ref	Ref
Bottom 80%		4.27	1.56 – 11.70	0.005*
Cluster[3]	223			
54		4.27	0.70 – 25.92	0.115
55		1.38	0.36 – 5.24	0.637
56		0.92	0.32 – 2.62	0.877
57		Ref	Ref	Ref
Antibiotic Use (Past 6 Months)[4]	113			
Yes		0.362	0.09 – 1.52	0.162
No		Ref	Ref	Ref

[1]Total n varies due to incomplete notation for some survey questions
[2]Adjusted for Complex Cluster Sampling and Non-Response
[3]Adjusted for Age and Household Wealth Index
[4]Adjusting for Cluster and Household Wealth Index
*Significant at 0.05 α level

Table 5: Bivariate Associations between Infection with an Enteric Pathogen and Harboring a Gene encoding for Antibiotic Resistance within the Study Population of Ranomafana Commune, Ifanadiana District, Madagascar.

Pathogen Virulence Marker	n	OR[1]	95% Wald Confidence Limits[1]	Wald χ^2 p-value[1,2]
iaph[3]	225			
bla$_{TEM}$		4.72	2.03 – 11.01	<0.0001*
sul 1 or *sul2*		1.66	0.76 – 3.64	0.202
strAB		7.48	3.29 – 17.01	<0.0001*
dfrA7		1.24	0.56 – 2.74	0.597
invA[4]	225			
bla$_{TEM}$		0.68	0.24 – 1.96	0.475
sul 1 or *sul2*		0.59	0.20 – 1.79	0.352
strAB		2.54	0.58 – 11.01	0.214
dfrA7		1.18	0.28 – 5.00	0.820
yadA[5]	225			
bla$_{TEM}$		4.73	1.24 – 18.03	<0.0001*
sul 1 or *sul2*		0.66	0.23 – 1.87	0.434
strAB		1.48	0.55 – 4.04	0.439
dfrA7		0.82	0.28 – 2.41	0.717

[1]Adjusted for Complex Cluster Sampling
[2]p-value adjusted for multiple comparisons using Bonferroni Correction
[3]Adjusted for *invA* and *yadA*
[4]Adjusted for *iaph* and *yadA*
[5]Adjusted for *iaph* and *yadA*
*Significant at 0.0125 α level

Appendix I: Oligonucleotide primers sequences used in PCR assays for identification of antimicrobial resistance genes and pathogenic virulence markers

Antimicrobial Class	Target Gene	Oligonucleotide Primer Sequence		Amplicon Size (bp)	Ta (°C)	Reference
		Forward Primer (5'-3')	Reverse Primer (5'-3')			
Aminoglycosides	*strAB*	**GACGAGGACAAGAGTACGCC**	**TAGCTAGATCGCGTTGCTCC**	1155	58	This Study
β-Lactams	*bla*$_{TEM}$	**ATGAGTATTCAACATTTCCG**	**CTGACAGTTACCAATGCTTA**	867	49	[37]
Sulfonamide	*sul1*	**TCACCGAGGACTCCTTCTTC**	**CAGTCCGCCTCAGCAATATC**	331	55	[36]
	sul2	**CCTGTTTCGTCCGACACAGA**	**GAAGCGCAGCCGCAATTCAT**	435	55	
Trimethoprim	*dfrA7*	**TCGCTTTGCAAGAACTATCGAA**	**CACCTTCAACCTCAACGTGAAC**	129	55	This Study
Pathogen Genus	**Target Gene**	**Oligonucleotide Primer Sequence**		**Amplicon Size (bp)**	**Ta (°C)**	**Reference**
		Forward Primer (5'-3')	**Reverse Primer (5'-3')**			
Shigella spp.	*iaph*	**CTTGACCGCCTTTCCGATAC**	**CAGCCACCCTCTGAGAGTA**	610	55	[39]
Salmonella spp.	*invA*	**TATCGCCACGTTCGGGCAA**	**TCGCACCGTCAAAGGAACC**	275	55	
Vibrio cholerae	*ctxA*	**GGCAGATTCTAGACCTCCT**	**TCGATGATCTTGGAGCATTC**	563	55	
Yersinia spp.	*yadA*	**CTTCAGATACTGGTGTCGCTGT**	**ATGCCTGACTAGAGCGATATCC**	681	55	[38]
PCR Control	**Target Gene**	**Oligonucleotide Primer Sequence**		**Amplicon Size (bp)**	**Ta (°C)**	**Reference**
		Forward Primer (5'-3')	**Reverse Primer (5'-3')**			
Total Bacteria 16s rRNA Gene	U16s	**GTGSTGCAYGGYTGTCGTCA**	**ACGTCRTCCMCACCTTCCTC**	146	53	[40]

Appendix II: Weighting Procedures and Wealth Index Explanation Provided by the Malagasy Institute of Statistics (INSTAT)

Survey Context:

These procedures were used with PIVOT baseline survey for the whole district of Ifanadiana. At this level stratum was each commune of the district, 80 clusters were sampled.

Another survey was conducted at 4 clusters in the Commune of Ranomafana. Sample design weights for household in these clusters are the same as used in PIVOT baseline survey. Final and normalized weights are equal to sample design weight updated by nonresponse during fieldwork. Considering this sampling procedure, we could not say if these 4 clusters are statistically representative of Commune of Ranomafana.

Weighting Procedures:

In order for the sample estimates from the basic survey to be representative of the population, it is necessary to multiply the data by a sampling weight. The basic weight for each sample household would be equal to the inverse of its overall probability of selection, calculated by multiplying the probabilities at each sampling stage. A household weight will be attached to each sample household record in the data files; in addition, woman weights and child weights will be attached to the corresponding data files. The sampling probabilities at each stage of selection will be maintained in an Excel spreadsheet with information from the sampling frame for all the sample clusters so that the corresponding weights can be calculated. Following the fieldwork it will be necessary to enter in this spreadsheet the total number of households listed and the final number of completed household interviews in each sample cluster.

Based on the proposed sample design, the overall probability of selection for the sample households can be expressed as follows:

$$P_{hij} = \frac{z_h \times N_{hi}}{N_h} \times P_{Shij} \times \frac{m_{hij}}{M'_{hij}}$$

Where:

- p_{hij} = probability of selection for the sample households in the jth sample segment in the ith sample cluster in stratum (commune) h
- z_h = number of sample cluster (or segments) selected in stratum h
- N_{hi} = total population in the frame for the ith sample cluster in stratum h
- N_h = total population in the sampling frame for stratum h (cumulated measures of size)
- p_{Shij} = probability of selecting the jth sample segment in the ith sample cluster in stratum h
- m_{hij} = number of sample households selected in the jth sample segment in the ith sample cluster in stratum h
- M'_{hij} = total number of households listed in the jth sample segment in the ith sample cluster in stratum h

In the case of the sample cluster that are not divided into segments, p_{Shij} = 1. For the remaining (large) sample cluster, the formula for p_{Shij} will depend on whether the sample segment within the cluster is selected with PPS or equal probability.

The basic household weight is calculated as the inverse of this probability of selection. Based on the previous expression for the probability, the weight can be simplified as follows:

$$W_{hij} = \frac{N_h \times M'_{hij}}{z_h \times N_{hi} \times P_{2hij} \times m_{hij}}$$

Where:

- W_{hij} = basic weight for the sample households in the j[th] sample segment in the i[th] sample cluster in stratum h

If m_{hij} is constant for each segment (for example, 20 households), the sample will be approximately self-weighting within each stratum. The variability in the weights within each stratum depends on the correlation between the population in the frame for the cluster and the number of households listed in the sample segment (multiplied by the number of segments in the cluster).

It is also important to adjust the basic weights for the sample households to take into account the nonresponse in each sample cluster. Since the weights will be calculated at the level of the sample cluster, it would be advantageous to adjust the weights at this level. The final weight (W'_{hij}) for the sample households in the i[th] sample cluster in stratum h can be expressed as follows:

$$W'_{hij} = W_{hij} \times \frac{m'_{hij}}{m''_{hij}}$$

Where:

- m'_{hij} = total number of valid (occupied) sample households selected in the j[th] sample segment in the i[th] sample cluster in stratum h
- m''_{hij} = number of sample households with completed interviews in the j[th] sample segment in the i[th] sample cluster in stratum h

Following the adjustment of the household weights for nonresponse, these weights are generally normalized (standardized) in data files so that relative weights can be used for the analysis of the survey data. In this way the sum of the relative weights will be equal to the number of sample households. The household weights were normalized by dividing each weight by the average weight at the whole-district level – that is, the sum of the weights for all sample households divided by the number of sample households). Therefore the relative weights will have a mean value of 1.

Given that sometimes it is not possible to complete a woman questionnaire for each woman identified in the household roster, it is also necessary to have a separate woman weight with an additional nonresponse adjustment factor applied to the household weight. The woman weight can be expressed as follows:

$$W_{whij} = W'_{hij} \times \frac{w_{hij}}{w'_{hij}}$$

Where:

- W_{whij} = adjusted weight for data in woman questionnaires for the j[th] sample segment the i[th] sample cluster in stratum h
- w_{hij} = total number of women age 15 to 49 years identified in the questionnaire roster for all sample households in the j[th] sample segment in the i[th] sample cluster in stratum h
- w'_{hij} = number of women with completed interviews for all sample households in the j[th] sample segment i[th] sample cluster in stratum h

There will also be cases where men questionnaire is not completed for all eligible men in some sample households. Therefore it is necessary to have a separate men weight with an additional

nonresponse adjustment factor applied to the household weight. The men weight can be expressed as follows:

$$W_{chij} = W'_{hij} \times \frac{c_{hij}}{c'_{hij}}$$

Where:

- W_{chij} = adjusted weight for data in child questionnaires for the jth sample segment in the ith sample cluster in stratum h
- c_{hij} = total number of children under 5 years identified in the questionnaire roster for all sample households in the jth sample segment in the ith sample cluster in stratum h
- c'_{hij} = number of women with completed interviews for all sample households in the jth sample segment in the ith sample cluster in stratum h

Wealth index

Wealth index calculations are exactly the same as in DHS and MICS surveys. The wealth index is a background characteristic that is used throughout the report as a proxy for long-term standard of living of the household. It is based on the data on the household's ownership of consumer goods; dwelling characteristics; type of drinking water source; toilet facilities; and other characteristics that are related to a household's socio-economic status. To construct the index, each of these assets was assigned a weight (factor score) generated through principal component analysis, and the resulting asset scores were standardized in relation to a standard normal distribution with a mean of zero and standard deviation of one [49].

Each household was then assigned a score for each asset, and the scores were summed for each household. Individuals were ranked according to the total score of the household in which they resided. The sample was then divided into quintiles from one (poorest) to five (richest). A single asset index was developed on the basis of data from the entire.