

**Protocol for the PATHOME Study: A Cohort Study on Urban Societal Development and the Ecology of Enteric Disease Transmission among Infants, Domestic Animals, and the Environment**

Kelly K Baker<sup>1\*</sup>, Sheillah Simiyu<sup>2</sup>, Phylis Busienei<sup>2</sup>, Fanta D Gutema<sup>1</sup>, Bonphace Okoth<sup>2</sup>, John Agira<sup>2</sup>, Christine S Amondi<sup>2</sup>, Abdhalah Ziraba<sup>3</sup>, Alexis G Kapanka<sup>1</sup>, Abisola Osinuga<sup>4</sup>, Collins Ouma<sup>5</sup>, Daniel K Sewell<sup>6</sup>, Sabin Gaire<sup>6</sup>, Innocent K Tumwebaze, Blessing Mberu<sup>2</sup>

**Affiliations**

<sup>1\*</sup> Corresponding author: Kelly K. Baker, PhD, University of Iowa College of Public health, Department of Occupational and Environmental Health, 145 Riverside Dr. S316, Iowa City, Iowa 52246, [Kelly-k-baker@uiowa.edu](mailto:Kelly-k-baker@uiowa.edu), 410-375-3615

<sup>2</sup> Sheillah Simiyu, PhD, African Population and Health Research Center, Division of Population Dynamics and Urbanization, [ssimiyu@aphrc.org](mailto:ssimiyu@aphrc.org)

<sup>2</sup> Phylis J. Busienei, African Population and Health Research Center, Division of Population Dynamics and Urbanization, [pbusienei@aphrc.org](mailto:pbusienei@aphrc.org)

<sup>1</sup> Fanta D Gutema, PhD, University of Iowa College of Public health, Department of Occupational and Environmental Health, [fantadesissa-gutema@uiowa.edu](mailto:fantadesissa-gutema@uiowa.edu);

<sup>2</sup> Bonphace Okoth, African Population and Health Research Center, Division of Population Dynamics and Urbanization, [bonphaceo@gmail.com](mailto:bonphaceo@gmail.com)

<sup>2</sup> John Agira, African Population and Health Research Center, Division of Population Dynamics and Urbanization, [agiradenge@gmail.com](mailto:agiradenge@gmail.com)

<sup>2</sup> Christine Sharon Amondi, African Population and Health Research Center, Division of Population Dynamics and Urbanization, [amondiogalo632@gmail.com](mailto:amondiogalo632@gmail.com)

<sup>3</sup> Abdhalah Ziraba, PhD, African Population and Health Research Center, Division of Health and Wellbeing, [aziraba@aphrc.org](mailto:aziraba@aphrc.org)

<sup>1</sup> Alexis G Kapanka, University of Iowa College of Public health, Department of Occupational and Environmental Health, [alexis-kapanka@uiowa.edu](mailto:alexis-kapanka@uiowa.edu)

<sup>4</sup> Abisola Modupe Osinuga, University of North Carolina, [abisola\\_osinuga@med.unc.edu](mailto:abisola_osinuga@med.unc.edu)

<sup>5</sup> Collins Ouma, PhD, Maseno University, Department of Biomedical Sciences and Technology, [collinouma@yahoo.com](mailto:collinouma@yahoo.com)

<sup>6</sup> Daniel Sewell, PhD, University of Iowa College of Public health, Department of Biostatistics, [daniel-sewell@uiowa.edu](mailto:daniel-sewell@uiowa.edu)

<sup>6</sup> Sabin Gaire, University of Iowa College of Public health, Department of Biostatistics, [sabin-gaire@uiowa.edu](mailto:sabin-gaire@uiowa.edu)

<sup>2</sup> Innocent K Tumwebaze, African Population and Health Research Center, Division of Population Dynamics and Urbanization, [itumwebaze@aphrc.org](mailto:itumwebaze@aphrc.org)

<sup>2</sup> Blessing Mberu, African Population and Health Research Center, Division of Population Dynamics and Urbanization, [bمبرu@aphrc.org](mailto:bمبرu@aphrc.org)

Keywords: gastrointestinal infection, diarrhea, One Health, Environmental Exposure, Global Health

1    **1. ABSTRACT**

2    **Introduction.** Global morbidity from enteric infections and diarrhea remains high in children  
3    in low- and middle-income countries (LMICs), despite significant investment over recent  
4    decades in health systems and global water and sanitation infrastructure. Other types of  
5    societal changes may be required to reduce disease burden. Ecological research on the  
6    influence of household and neighborhood societal development on pathogen transmission  
7    dynamics between humans, animals, and the environment could identify more effective  
8    strategies for preventing enteric infections.

9    **Methods and analysis.** The “enteric pathome” - i.e., the communities of viral, bacterial, and  
10    parasitic pathogens transmitted from human and animal feces through the environment is  
11    taxonomically complex in high burden settings. This integrated cohort-exposure assessment  
12    study leverages natural socio-economic spectrums of development to study how pathome  
13    complexity is influenced by household and neighborhood development. We are enrolling  
14    under 12-month-old children in low- and middle-income neighborhoods of two Kenyan cities  
15    (Nairobi and Kisumu) into a “short-cohort” study involving repeat testing of child feces for  
16    enteric pathogens. A mid-study exposure assessment documenting infrastructural,  
17    behavioral, spatial, climate, environmental, and zoonotic factors characterizes pathogen  
18    exposure pathways in household and neighborhood settings. These data will be used to  
19    inform and validate statistical and agent-based models that identify individual or combined  
20    intervention strategies for reducing multi-pathogen transmission between humans, animals,  
21    and environment in urban Kenya.

22    **Ethics and dissemination.** The protocols for human subjects’ research were approved by  
23    Institutional Review Boards at the University of Iowa (ID - 202004606) and AMREF Health  
24    Africa (ID - ESRC P887/2020), and a national permit was obtained from the Kenya National  
25    Commission for Science Technology and Innovation (NACOSTI) (ID# P/21/8441). The study  
26    was registered on Clinicaltrials.gov (Identifier: NCT05322655). Protocols for research on

27 animals were approved by the University of Iowa Animal Care and Use Committee (ID  
28 0042302).

29

### 30 **STRENGTHS AND LIMITATIONS OF THE STUDY**

31 • This cohort-exposure assessment study will provide new evidence on the nature of  
32 household and neighborhood developmental strategies that are most effective at  
33 preventing critical multi-pathogen transmission pathways among humans, animals,  
34 and environment in cities of low- and middle-income countries.

35 • We study middle class households and neighborhoods to test counterfactual  
36 theories about meeting basic developmental standards to reduce pathogen  
37 transmission.

38 • Our data collection uses objective methods to comprehensively document  
39 socioeconomic, weather, infrastructural, spatial, behavioral, environmental,  
40 zoonotic, and human data, including use of both selective culture and molecular  
41 methods to characterize pathogen community patterns.

42 • The observational study design is vulnerable to unmeasured confounders.

43 • The living conditions in middle-class households and neighborhoods may not offset  
44 hygiene conditions in the overall urban environment enough to alter enteric  
45 pathogen transmission patterns.

46

47

## 48 Introduction

### 49 Study Rationale:

50 In the past half century, advances in access to health care and early childhood vaccines  
51 and therapies both in high income countries (HICs) and in low-to-middle income countries  
52 (LMICs) have successfully reduced global mortality from diarrheal diseases. According to  
53 Global Burden of Disease (GBD) data, diarrhea mortality in children under five years  
54 decreased by 59.3% (from 173.3 per 100,000 to 70.6 per 100,000) between 2000 and 2016,  
55 largely in South Asia and sub-Saharan Africa where fatality rates were highest [1]. However,  
56 reductions in diarrhea incidence have lagged with only a 12.7% reduction (from 2.0 per  
57 child-year to 1.75 per child-year) over the same 16 years, and large differences remain  
58 between incidence in children in HICs (0.58 per child year) versus regions like Latin America  
59 and sub-Saharan Africa (2.82 cases and 2.37 cases per child-year, respectively). Typhoid and  
60 paratyphoid enteric fever cases in endemic areas also remain high at 14.3 million cases per  
61 year [2]. Asymptomatic infection prevalence is even higher than symptomatic rates [3], and  
62 symptomatic and asymptomatic infections can elevate a child's susceptibility to co-infection  
63 [4], enteric dysfunction and malnutrition [5, 6], and long-term cognitive and developmental  
64 stunting [7, 8]. Greater investment in disease prevention and control programs that protect  
65 children from pathogen exposure will be needed to accelerate the reduction in global  
66 enteric disease burden and achieve health equity between HIC and LMIC populations.

67 At the end of the 19<sup>th</sup> century, typhoid fever, cholera, and diarrhea epidemics routinely  
68 plagued cities like London and Philadelphia in today's HICs [9]. Annualized death rates were  
69 quite similar to rates in LMICs today, although with stronger seasonal patterns [10]. Death  
70 rates dramatically decreased in the late 19<sup>th</sup> to early 20<sup>th</sup> centuries, even as urban  
71 populations were growing rapidly and well before the development of vaccines [11].  
72 Development of municipal piped water and wastewater systems contributed to these health  
73 transitions, although inference from these studies is limited by the lack of model adjustment

74 for confounding from other causal pathways (food safety, handwashing, housing quality)  
75 [12-14]. Water, Sanitation and Hygiene (WASH) has occupied a central strategy in global  
76 development policy over the last fifty years for reducing global diarrhea burden and resulted  
77 in impressive increases between 1990 and 2015 in global population with access to  
78 household drinking water and to a lesser extent sanitation infrastructure. These investments  
79 have improved quality of life globally, but as stated above have not generated expected  
80 declines in diarrheal incidence in LMICs [15]. Randomized controlled studies reporting  
81 relatively small benefits from household-level improvements in WASH and recent  
82 demographic studies raise questions as to what other conditions are required for disease  
83 control [16-20]. Achieving population-wide water and sanitation coverage, infrastructure  
84 maintenance, and chlorination of a water supply may be necessary, but not sufficient for  
85 disease control [19-24].

86 Control of enteric pathogen transmission may require more wide sweeping overhauls  
87 of societal sanitary conditions beyond drinking water and sanitation systems. The building of  
88 water and sanitation systems in HIC cities was just one piece of a much larger societal  
89 reformation that included parallel improvements in health systems, housing quality, waste  
90 disposal laws, awareness of personal hygiene, growth of a middle class, and better nutrition  
91 [25]. The scientists who were directly observing health transitions between 1850 and 1950  
92 proposed diverse theories for the collapse of enteric disease epidemics. Some studies noted  
93 improvements in water sources in Glasgow and London correlated poorly with observed  
94 declines in diarrhea rates over time, with major decreases in disease occurring decades after  
95 improved water sources and decades before sewerage [10, 26]. Contamination of milk  
96 sources was considered a cause of typhoid and scarlet fever outbreaks [27, 28]. Widespread  
97 regulation of milk pasteurization also emerged in the late 19<sup>th</sup> century and has been  
98 attributed with decreases in diarrhea case rates in Philadelphia and New York City in the  
99 1900's [29, 30]. Others suggested sharing of unsanitary living conditions rather than

100 sewerage problems caused typhoid fever cases [31]. One comprehensive study examined  
101 housing quality, shared living conditions, food supply and type of water and toilet in early  
102 1900's Mansfield, England, which had a seasonal diarrhea prevalence of 10% despite high  
103 sewerage sanitation and water supply coverage [32]. It supported the premise that shared  
104 housing and unhygienic household habits contributed most to diarrhea prevalence and  
105 noted that middle class households in otherwise crowded areas with poor neighborhood  
106 hygiene were not protected from infection.

107        Evolutionary biology evidence suggests that the establishment of permanent  
108 agricultural population centers around 10,000 years ago triggered pre-industrial  
109 epidemiological transitions that *increased* pathogen adaptation to humans and multi-  
110 parasitism in human populations. Migration and urbanization further concentrated animals,  
111 humans, and their waste into relatively small areas, increased pathogen host adaptation,  
112 and increased rates of within-species and across-species transmission of many pathogenic  
113 species [33]. Over time the number of households in HICs choosing to own animals and the  
114 types of animals own has shifted as a byproduct of changing lifestyles, technology, and  
115 limited space and need for animal rearing. The importance of animals to urban industry has  
116 also shifted. Brownlee and Young (1922) reported strong consensus at scientific society  
117 meetings with the theory that the replacement of horses and gravel road surfaces with  
118 motorized vehicles and tarmac prior to 1910, reduced visible animal feces in the streets and  
119 coincided strongly in time with disease declines. Strong temporal correlation between  
120 decreases in annual issued licenses for horse-drawn vehicles in English and Welsh cities and  
121 declines in diarrheal deaths provides support for that premise [34]. Fly control may have  
122 further enhanced the community hygiene-driven declines in typhoid fever and diarrheal  
123 deaths [35]. Changes in contact between humans, work animals, and other domestic species  
124 that prevented zoonotic transmission could also explain early epidemiological transitions.

125 Elimination of cholera, typhoid fever, and other enteric disease may have required a  
126 combination of societal developmental changes that collectively attained a standard of  
127 environmental hygiene capable of preventing transmission of infectious agents– essentially  
128 a microbial community effect. Unfortunately, evidence on which aspects of development  
129 had the greatest impact on enteric disease ecology during historical transitions cannot be  
130 disentangled because records on the timing, rate, and scale of coverage of most societal  
131 development activities, and personal adoption of hygiene behaviors are fragmented and  
132 poorly documented [36]. Which persons benefitted from those improved living conditions is  
133 also poorly recorded, requiring studies to rely upon aggregated health records until the  
134 middle of the 20<sup>th</sup> century [14]. Records on outcomes are predominantly death records for  
135 cholera and typhoid due to inconsistent health care seeking and recording of non-fatal cases  
136 and investigations of other enteric causes of diarrheal deaths. Contemporary studies  
137 examining societal development and multi-decade enteric disease rates in countries that  
138 have recently undergone epidemiological transitions would be more culturally relevant, but  
139 such studies are also lacking for many of the same data availability challenges.

140 Objectives:

141 A study combining Planetary Health and One Health approaches to examining ongoing  
142 health transitions in high burden countries of the 21<sup>st</sup> century could better inform which  
143 developmental factors most influence enteric disease prevention and control for prioritizing  
144 health program policy. To address this knowledge gap, we describe the scientific rationale,  
145 hypotheses, and design for the Pathogen Transmission and Health Outcome Models of  
146 Enteric Disease (PATHOME) study supported by the National Institutes of Health (NIH)  
147 through the Evolution and Ecology of Infectious Disease funding mechanism (1 R01  
148 TW011795, years 2020-2025). The objectives of PATHOME are to (1) characterize  
149 interactions among enteric pathogenic agents in infants, domestic animals, and the  
150 environment and the dynamic interactions between infants, caregivers, animals, and



151 environmental materials across space and seasons; and (2) develop statistical and  
152 computational methods to examine this complex disease system and predict which social  
153 and environmental developmental improvements best prevent multi-pathogen transmission  
154 in urbanizing areas of a high disease burden country. Our overall hypothesis is that joint  
155 modeling of “enteric pathome” agents (i.e. the microbial communities of viral, bacterial, and  
156 protozoan pathogens transmitted by human and animal feces) across households and  
157 neighborhoods that represent contrasts in urban societal development will show that  
158 development leads to lower pathogen-specific detection frequencies, and thus evolution of  
159 the pathome from complex to simple microbial community structure distributions in  
160 humans, animals, and the environment (Figure 1). Our study design is motivated by several  
161 fundamental science theories (Table 1).

162 **Table 1. Study Design Strategies and Scientific Rationale for the PATHOME Study**

Study Design Strategy	Rationale
Recruit from neighborhoods representing low to moderate levels of societal development to test counterfactual hypotheses that high thresholds of basic development can reduce pathogen transmission	The rapid growth in middle-class urban households in LMICs globally can be leveraged to compare enteric pathogen exposure and health patterns across a socioeconomic developmental spectrum to identify the conditions that most influence enteric disease transmission. While increases in household socioeconomic status are known to decrease the risk of pediatric diarrhea in LMICs, most studies have been based in low-income populations. Evidence from middle class neighborhoods testing is critical for identifying tipping points in transmission and accurately predicting potential program impact on health outcomes.
Apply a One Health approach that includes zoonotic and environmental health to identify all sources of pathogen transmission	Domestic and sylvatic animals are susceptible to infection from many of the same enteric pathogens and may have more widespread carriage than humans due to less frequent disease surveillance and treatment. Wildlife are considered major drivers of novel emerging zoonotic disease outbreaks, but human interactions with domestic species may be more important in sustaining endemic zoonotic burden because the higher frequency of human-animal interaction increases human contact with zoonotic vectors. Keeping animals for food, pest control, and companionship are common practices in most urban and peri-urban LMICs settings and these animals are also often allowed to freely roam and forage for food in household and neighborhood areas which can spread feces with pathogens across the environment [37-39]. Improvements in sanitation coverage does not reduce animal feces pollution in the environment so data collection on safe feces management must account for infection prevalence in animals, animal hygiene practices, and human-animal contact. A similar study to this one in Nairobi detected human-animal <i>E. coli</i> genotype sharing but found sharing between humans to be more common [40]. Pathogen sharing is also influenced by exposure dose, type and effectiveness of microbial virulence genes,

	and host anti-pathogen immunity, so within- and between-host pathogen sharing patterns may not be the same as for non-infectious <i>E. coli</i> .
Rely upon a microbial ecology approach to detect pathogen diversity at the human-environment-animal interface to identify pathogen sources and transmission dynamics	Studies in urban Kenya and elsewhere have demonstrated taxonomic diversity in enteric pathogens in humans [3, 18, 41], animals, and environment [39, 42, 43] in areas where open defecation, dumping of untreated sewage, and unsanitary animal management practices are common [44]. Taxonomic diversity varies by exposure pathway, and multi-pathogen exposure risks may be elevated in infants living in underdeveloped houses with dirt floors, shared latrines, and domestic animals. Evidence on pathogen diversity in middle-class neighborhoods is unknown, although theoretically should be low if moderate to high thresholds of neighborhood infrastructure and hygiene have been met. Identifying intervention strategies that best reduce multi-pathogen transmission could enable advocacy for improved investment in high impact developmental and public health programs.
Use a community-based enrollment design and objective measures of pathogen presence or absence to detect transmission dynamics that are unconfounded by host type, health status, and health care utilization behaviors	Reliance upon diarrhea symptoms as an indicator of exposure can lead to biased study designs that presume health care seeking cases are exposed while asymptomatic controls are presumed unexposed. The relationships between ingestion of a pathogen, gut infection, and progression to symptomatic status are influenced by many factors, such as type of pathogen, infectious dose, host immunity, comorbidities, and breastfeeding. Asymptomatic carriage of enteric pathogens is common in endemic areas [3], and these individuals can still pose a transmission risk to others, but typically go undetected due to lack of symptoms. Evidence from cities in the United States in the 1930s and 1940 suggests that city development decreased both symptomatic and asymptomatic <i>Shigella</i> infections [9]. Combining community-based sampling with objective measures of pathogen presence in infants, animals, and environmental samples avoids misclassification of exposed asymptomatic subjects as unexposed.
Enroll children between 0 to 12 months of age to improve detection of transmission dynamics	A high pre-existing pathogen detection prevalence makes it more challenging to distinguish between persistent or pre-existing infections versus new infections and increases the chances that pathogen detection in children is confounded by pre-existing host immunity or disease conditions. Large studies of enteric pathogen detection in under- five children consistently show that detection and co-detection of enteric pathogens in endemic areas climbs rapidly within the first months of life [45, 46]. Study of children between birth and twelve months leverages the dynamic increases in prevalence rate in this age group for identifying early causes of population disease transmission.
Use a short-cohort, repeat measures design that matches a natural timeline of exposure and infection to improve detection of transmission dynamics	For most enteric pathogens, the timeline between known exposure and confirmed infection is between 1 to 5 days. The timeline between infection and recovery varies from a few days to a few weeks. Pathogen carriage and diarrhea lasting more than 14 days is considered persistent and unusual. A cohort study that spaces human health assessment by 30 days or more cannot detect exposure-outcome relationships or confirm that concurrent pathogen detections over a month or more are persistent carriage versus new exposure events. Given known temporal pathogen-host disease trends, the dynamic rate of infections in infants in Kenya, and the evidence of widespread pathogen contamination in low-income urban neighborhoods, a 14-day short cohort design with repeat diagnostic testing for enteric pathogens in infants could detect

	transmission events and monitor their progression. Assessing pathogen presence in infants at enrollment can improve screening for post-enrollment pathogen exposures.
<b>Apply a Planetary Health approach to characterize household and community socio-economic and infrastructure development on enteric pathogen transmission</b>	Natural societal development involves improvements in many living conditions that are inherently correlated, like water and sanitation access, flooring, ability to buy safer foods and soap for handwashing, and possibly changes in zoonotic exposures. The rate with which households and neighborhood improve their circumstances can vary creating spatial variability in when and where hygiene conditions are achieved. If human and animal feces containing enteric pathogens are discarded in drains or lots in the neighborhood, it can be carried by human and animal foot traffic or flood water into the home impacting household hygiene conditions and exposure of children [43, 46, 47]. Children also can be exposed to neighborhood soil, water, objects, and animal feces containing enteric pathogens during play or movement in residential environments outside the household [42, 43]. Understanding how changes in socio-economic and built environment conditions influence population health dynamics in the present could improve the future sustainability and health effectiveness of urban planning policies [48]. In the context of enteric disease control, evidence is needed on how household and community development and the spatial exposures of infants influences the collapse of enteric pathogen transmission.
<b>Select enteric pathogen indicators that are likely to occur dynamically over space and time in the study population for discriminatory power</b>	Pathogens selected for characterization of pathome transmission were chosen based upon evidence that (1) population prevalence of the pathogen in diarrheal and asymptomatic infants changes between 0 and 12 months of age in Kenya [3]; (2) pathogen detection in humans, animals, and/or environment is a sensitive indicator of unhygienic conditions; (3) domestic animals and environmental reservoirs (e.g., soil and water) contribute to pathogen transmission; (4) methods exist or can be adapted to isolate organisms and validate pathogen presence in human, animal, and environmental materials.

163

## 164 **METHODS AND ANALYSIS**

### 165 **1. Study setting**

166 The study is situated in low-income and middle-income communities in two different  
167 Kenyan cities, Nairobi (population 4.4 million persons) and Kisumu (population ~610,000 in  
168 2019) where our team has performed prior health research [43, 45, 49-60]. The prevalence  
169 of diarrhea in children in Kenya is high with most cases occurring in children under 3 years of  
170 age, and roughly 25% of children are severely stunted [61]. However, diarrhea prevalence is  
171 much lower in Kenya's rapidly growing middle class [59], indicating that epidemiological  
172 transitions in enteric disease risk are occurring along low- to middle-income neighborhood  
173 gradients. Communities in each city were selected based upon similarity in population's

174 ethnic and cultural characteristics but with distinct differences in overall community  
175 developmental conditions. Environmental hygiene indicators also differ by neighborhood  
176 socioeconomic status and thus are expected to be different in the two target cities. Private  
177 toilets are the norm in middle-income neighborhoods, while the majority of households in  
178 low-income areas rely upon shared latrines for disposal of human feces and a quarter of  
179 households dispose of feces in open defecation sites, drains, or waste dumps [62]. Low-  
180 income households [46, 62] and neighborhoods [39] are heavily polluted with indicator  
181 bacteria and enteric pathogens, while contamination of middle-income neighborhoods is  
182 unknown. Half of households in Kisumu keep domestic animals [37], while animal ownership  
183 is less common in low- and middle- income Nairobi neighborhoods. While development or  
184 the lack thereof may be uniform across communities, in urban settings there may be outlier  
185 households that differ from their neighbors, such as relatively wealthier households in low-  
186 income communities or relatively poorer households in middle-income communities. The  
187 communities selected for this study offer the opportunity to enroll these types of  
188 households and study how a spectrum of household conditions might influence transmission  
189 dynamics in the context of broader community conditions.

## 190 **2. Study Design**

191 To capture exposure-outcome relationships reflecting pathogen transmission dynamics,  
192 we are conducting a prospective 14-day cohort epidemiological study with repeat health  
193 outcome assessment and integrated exposure assessment data collection. We are recruiting  
194 32 children in 4 age groups (<3 months, 3 to <6 months, 6 to <9 months, 9 to 12 months) on  
195 a rolling basis over 2 consecutive years across wet and dry seasons (12 months) in low-  
196 income and middle-income neighborhoods of Kisumu and Nairobi for a total of 248  
197 households with complete data assuming a 10% attrition rate (Figure 2). Enrollment in  
198 Nairobi spans November 2021 to November 2022, with the last participant follow-up in  
199 December 2022. Enrollment in Kisumu spans late March 2023 to February 2024 with the last

200 participant follow-up expected in March 2024. Participant recruitment follows a  
201 standardized schedule of sixteen households over a four-week period, eight each from the  
202 low- and middle-income neighborhood, followed by a one week break in enrollment for  
203 neighborhood-level observations and environmental sampling. This five-week cycle is  
204 repeated 12 times or until sample size requirements for each neighborhood are met.

### 205 **3. Eligibility, Participant Recruitment, and Enrollment**

206 Eligibility is defined as infants between 0 and 12 months of age, as verified by birth  
207 registration card. Infants with disabilities that would impact normal behavior are excluded.  
208 Study of infant age groups also provides an opportunity to examine pathogen transmission  
209 dynamics related to development of motor skills that allow infants to explore their  
210 environment [61]. Participants are identified from study neighborhoods by extracting infant  
211 age and the primary caregiver's name and contact information from monthly surveillance  
212 records collected routinely by Kenyan Ministry of Health (MOH) Community Health  
213 Volunteers (CHVs) in each neighborhood. Per their standard practice, CHVs approach all  
214 households within their area of work with pregnant women or young children to ensure they  
215 receive health information and have support to access health care. Recruitment data is  
216 compiled as separate lists for the low- and middle-income neighborhoods of each city. When  
217 CHVs identify potentially eligible households, they notify study staff of the African  
218 Population Health and Research Center (APHRC) who follow up with households in the  
219 company of the CHV to verify information about age-eligible children.

220 The information sheet/form explains the study activities and the purpose for them, as  
221 well as risk and benefits, costs for participation (none), right to withdraw, and  
222 compensation. The adult/primary caregiver is informed about all study activities by the field  
223 staff who read the information on the form to the caregiver in English or Swahili, in the  
224 presence of their CHV as a witness, and answers any questions they may have. Consent  
225 forms are signed by the caregiver with signature or thumb print and by their CHV witness if

226 the caregiver is not in a position to sign themselves. A copy of the information sheet and  
227 consent form is provided for the caregiver's permanent records. After informed consent is  
228 completed, the study team schedules a time to begin data collection and provides the  
229 caregiver with a package of size-appropriate diapers with instructions to collect infant feces  
230 specimens over the subsequent two weeks. Collection of a baseline or Day 1 feces sample is  
231 required for continued enrollment in the study. Should consent or an enrollment feces  
232 sample not be provided, the research team and CHV identifies another infant of same age  
233 group and neighborhood from the list for recruitment. Biological sex of the infant is not an  
234 enrollment criterion and does not influence recruitment although we expect the sex of the  
235 study population to be representative of natural sex distributions in urban Kenya.  
236 Households are provided one pack of diapers and flour and sugar to make porridge (items  
237 identified by caregivers in a pilot study as being highly desired that directly benefit the child)  
238 as compensation for their time at the end of enrollment (approximately 6 USD).

#### 239 **4. Study Outcomes**

240 The primary outcomes of the epidemiological study are detection and co-detection of  
241 specific types of enteric pathogens in infant feces during the enrollment period. The three  
242 approaches for examining pathogen detection are:

- 243 1. Pathogen prevalence - the proportion of infants with a specific type of enteric virus,  
244 bacteria, or parasite detected in their feces at enrollment (Day 1).
- 245 2. Pathogen incidence –the proportion of infants who have a specific type of enteric virus,  
246 bacteria, or parasite detected in their feces in the 13 days after enrollment that was not  
247 present in their feces at enrollment (Day 2 to 14).
- 248 3. Pathogen diversity – the number of unique types of enteric virus, bacteria, or parasite  
249 detected in infant feces at enrollment (Day 1).

250 The study is evaluating the following secondary outcomes:

- 251 1. 14-day prevalence of prospective self-reported diarrheal symptoms - proportion of  
252 caregivers who report an enrolled infant had 3 or more loose feces, with or without  
253 blood or mucus, in any 24-hour period during the enrollment period (Day 1 to 14).
- 254 2. 7-day prevalence of self-reported diarrheal and other symptoms prior to enrollment -  
255 proportion of caregivers reporting at enrollment that in the past 7 days the enrolled  
256 infant had symptoms of: malaise, vomiting, difficulty in breathing, unusual lack of  
257 appetite or willingness to take liquids, fever, runny nose, cough, rash (pre-Day 1).
- 258 3. Preterm birth prevalence - proportion of infants born before 37 weeks gestational age  
259 (Day 1).
- 260 4. Stunting prevalence - proportion of infants whose length-for-age z-score (centimeters) is  
261 -2 or more standard deviations away from the international standard during the  
262 enrollment period (Day 3 and 7).
- 263 5. Wasting prevalence - proportion of infants whose weight-for-age z-score (kilograms) is -  
264 2 or more standard deviations away from the international standard (Day 3 and 7).
- 265 6. Mid-upper arm circumference (Day 3 and 7).

266 The study is evaluating the following Intermediate Pathogen Exposure Outcomes:

- 267 1. Frequency of child exposure behaviors - number of infant contacts with caregivers, other  
268 children, domestic animals, and the environment per unit of time (Day 2, 3)
- 269 2. Presence and concentration of *E. coli* and pathogenic sub-types, *Salmonella enterica*,  
270 *Shigella* spp., *Campylobacter jejuni/coli*, and *Listeria monocytogenes* in household (Day  
271 2) and public (monthly) environments.
- 272 3. Prevalence and concentration of enteric viruses, bacteria, and protozoan pathogens in  
273 feces of domestic animals (Day 2).

#### 274 5. **Sample Size Determination**

275 Sample size calculations aimed to ensure enough observations to reliably estimate the  
276 model parameters and limit the set of plausible parameter values to a reasonably narrow

277 range. Towards this, we considered the set of models to be used to examine the  
278 relationship between exposures and multi-pathogen infections and ensured that we  
279 achieved 11 to 21 observations per parameter, depending on the model.

## 280 **6. Data Collection**

281 PATHOME methods are summarized below and in Figure 3, with detailed protocols provided  
282 sequentially in Supplemental Materials and at <https://pathomelab.github.io/>.

### 283 **6.1. Household Data Collection**

284 Over the 14-day enrollment window households participate in a series of data collection  
285 activities as shown in **Error! Reference source not found.**3. On Day 1, our field staff collect  
286 the baseline diaper with infant feces and conduct a survey on socioeconomic conditions,  
287 behaviors, and health in the household. Caregivers are also given a 14-day self-reported  
288 diarrhea symptom calendar that spans enrollment to study completion to record diarrheal  
289 symptoms for the infant. On Day 2, field staff return for the first day of exposure assessment  
290 activities involving a five-hour period of structured observation, the placement of a  
291 geotracking device on infants and their domestic animals (if available), and collection of  
292 domestic animal feces and household environmental samples. The caregiver is reminded to  
293 continue using the diapers to collect infant feces and to record symptoms, if any, on the  
294 calendar. The geotracker is left on infants and animals for a 24-hour period spanning the  
295 morning of Day 2 to mid-afternoon on Day 3. On Day 3, field staff repeat 5-hour structured  
296 observation, collect a Day 3 feces specimen, document anthropometric measurements of  
297 the participating infant, and retrieve the geotracking devices just before departure. Field  
298 staff return to the household on Days 5, 7, and 14 to collect three more feces specimens (5  
299 feces samples total). On day 7, the second anthropometric measurements are also  
300 documented. A second household soil/swab is also collected on this day. On Day 14, the  
301 enumerators return to the households to collect the diarrhea calendar, the last infant stool  
302 sample and provide caregivers the compensation gift.



303 **6.1.1. Household Survey**

304 On Day 1, caregivers of infants are asked to respond to a 45-minute survey that  
305 documents demographics of persons living in the household, number and type of caregivers,  
306 wealth assets, household and compound infrastructure including access to drinking water  
307 sources and latrines, ownership and management of domestic animals, breastfeeding and  
308 feeding practices, method of handling infant and animal feces, handwashing practices, 7-day  
309 history of diarrhea and other symptoms in the infant and other household members and  
310 related health care utilization or self-administered treatments (e.g. antibiotics).  
311 Developmental indicators were chosen for their ability to serve as culturally relevant and  
312 discriminatory measures of upward socioeconomic mobility in Kenya.

313 **6.1.2. Prospective 14-day diarrhea calendar**

314 The caregiver is given a 14-day pictorial diary with images of a sick and healthy infant  
315 and asked to mark any days in the next two weeks when the infant experiences 3 or more  
316 loose water or bloody feces in a 24-hour period. Enumerators and CHVs review the calendar  
317 on scheduled visits and prompt the caregiver to complete the record. On Day 14, the  
318 calendar is replaced with a gift of a 12-month calendar.

319 **6.1.3. Infant feces Collection**

320 Five feces samples are collected per subject [Days 1, 3, 5, 7, and 14, yielding a total of  
321 1,240 infant feces samples (248 infants \* five feces samples)]. If staff return as scheduled  
322 but no feces is available, they are allowed to return anytime in the next 24 hours (e.g., Day 4  
323 for Day 3) to fulfill the scheduled feces sample plan. Infant feces is collected by providing the  
324 caregiver a pack of disposable diapers sized by the infant's weight, and asking her to use  
325 them immediately and for the duration of the 14-day study, providing any containing infant  
326 feces to the study team [55]. Diapers are used to collect infant feces to protect the feces  
327 from contamination by soil (e.g., scooped from ground) or feces of other children (e.g.,  
328 potties in use by multiple children). If the infant defecates, the caregiver places it in a

329 provided sterile, Ziploc bag and stores it in a safe, shaded place. The field team collects any  
330 diapers with feces during visits in a cooler with ice packs and transfers them to a supervisor  
331 for transport to a central lab for analysis (APHRC lab in Nairobi and Maseno University lab in  
332 Kisumu).

#### 333 **6.1.4. *Infant and Domestic Animal Spatial Geotracking***

334 Data on the spatial-temporal movement of infants and domestic animals between  
335 household and public areas, and frequency and duration of co-localization that could lead to  
336 infant-animal interactions is measured by placing GPS mobile data loggers on infants and  
337 animals for a 24-hour period [63, 64]. The detailed development and validation of the  
338 geotracking protocol for infants and animals will be summarized in a separate manuscript. In  
339 brief, on Day 2 enumerators place a small (28 x 11-inch, 1.23 ounce) GPS Tractive Pet  
340 Tracker (Tractive Co., Austria) linked to an online real-time logging account on the infant.  
341 Enumerators work jointly with caregivers to identify an unobtrusive place for attaching the  
342 geotracker to infants, typically a snug child-sized wrist or ankle band or in a pocket of the  
343 infant's clothing. Caregivers are assured that the devices do not record audio or visual  
344 information [65]. They are asked to move the location of the geotracker as necessary to  
345 ensure child comfort, minimize concerns of theft or visibility, changes in clothes, or other  
346 concerns, while keeping the tracker within arm's reach of the infant [66, 67]. This tracker is  
347 placed prior to structured observation to allow enumerators to monitor infant discomfort or  
348 reactivity over the five-hour observation window and assist with placement fit strategies if  
349 necessary. This also provides parallel observation and satellite-based location data for  
350 examining the validity of methods for recording human behavior. The geotrackers remain  
351 placed for 24 hours, although if the caregiver chooses to remove it from the infant for  
352 sleeping, she is asked to place the device next to, but out of reach of the infant and then  
353 replace the device when the infant wakes up.

354 An additional one or two geotrackers are attached to harnesses (chicken/duck, cat) and  
355 collars (dogs, goats, sheep, cow) of domestic animals belonging to the enrolled household.  
356 Enumerators monitor the animal for species-specific signs of distress for 15 minutes, and  
357 then turn the device on, which begins real-time upload of date, time, and longitude and  
358 latitude data to the web. If the animal does not adjust to the harness, then the harness is  
359 removed, and another animal, if available, is selected for observation. If the household owns  
360 multiple animals, staff prioritize placing geotrackers on two different species that most  
361 commonly come into contact with the infant over two animals of the same species or  
362 animals kept apart from infants and household areas. Geotrackers are set to transmit  
363 location data to satellites in real-time at five-minute intervals. The geotrackers are  
364 recovered after 24 hours on Day 3 for cleaning, recharging and reuse. Data is extracted from  
365 the Tractive server via an api and merged by geotracker ID with other household data for  
366 cleaning and summarization of movement and location of the wearer. A selected number of  
367 caregivers will be selected from the study sites for in depth interviews about their  
368 experiences with the geo tracking exercise. These caregivers will be randomly selected from  
369 households where tracking has been done for both animals and infants. The tool will contain  
370 questions about their experiences with the tracking, experiences of the infants and animals,  
371 placement of the tracker, their fears/concerns, and recommendations for improvement.

372 **6.1.5. Structured Observation of Infant and Caregiver Behaviors.**

373 After placing the geotrackers, enumerators begin 5 hours of structured observation of  
374 the enrolled child and their caregiver using a pre-piloted tool developed by our team using  
375 the Livetrak (developed by Stanford University) app for tablet data collection [42, 45, 65].  
376 Tools document where the child is taken (public and household settings), what the infant is  
377 doing (e.g., sleeping, crawling, standing), and the rate of the child touching and mouthing  
378 soil, objects, food, drinking water, surface water, animals or their feces, and hands or other  
379 body parts of other children and caregivers. Tools also document the frequency of mitigating

380 behaviors performed for infants, such as cleaning after defecation and infant feces disposal,  
381 cleaning face or hands, and hand washing [42]. In cases of multiple caregivers for children,  
382 the caregiver who is interacting with the child at any given time is recorded to capture  
383 differences in behavior by caregiver [54]. Data will be summarized by frequency of different  
384 observations as well as the sequence of behaviors.

#### 385 **6.1.6. Domestic animal feces collection**

386 Up to two specimens of animal feces representing different animal species are collected  
387 in each household, equating to 496 animal feces samples (248 households \* 2 animal feces  
388 per household). Animal feces could include feces observed in animal enclosures if a  
389 household-owned animals or feces on the ground that may have been from personal or  
390 neighbor-owned animals. Enumerators will attempt to sample feces on Day 2, at the same  
391 time as observations, geotracking, and environmental sample collection. However, if no  
392 feces are observed on Day 2, enumerators can inspect the household on subsequent days  
393 when they return to collect infant feces and then collect samples. If no animal feces are  
394 observed over the 14-day time period, then the household will be considered  
395 uncontaminated by domestic animal feces. Feces specimens are collected into sterile  
396 WhirlPak bags, from the center of the pile to avoid contamination by soil, surface water, or  
397 other animal feces. These samples are stored in a cooler on ice packs and transported, as  
398 stated above, to the central lab for analysis.

#### 399 **6.1.7. Household Environmental Samples**

400 In each participating household, we collect two soil samples (Day 2 and 7), one  
401 caregiver hand rinse, one toy rinse, and in Kisumu only, one drinking water sample. Soil  
402 samples are about a 30 gram composite of soil from different locations in the household or a  
403 swab of a 30X30 cm squared floor space [68]. Toy rinses aim to measure the amount of  
404 microbial contamination that can be transferred onto household fomites in 24 hours. This is  
405 measured by providing the caregiver an alcohol sterilized infant toy, such as teething ring or

406 rattle, on Day 1 for the infant to play with for 24 hours [69]. This infant toy is retrieved 24  
407 hours later on Day 2 into a sterile barcode labeled bag for transport to the lab for microbial  
408 testing. An identical replacement toy is given to the infant for permanent keeping. At the  
409 end of behavioral observation on Day 2, one caregiver hand rinse is collected by submerging  
410 each hand into a barcode-labeled sterile one-liter WhirlPak bag with Phosphate Buffered  
411 Saline (PBS) and massaging all surfaces of the hand for 30 second before switching to the  
412 next hand. In total, we expect to collect 496 household soils, 248 toy rinses, and 248 hand  
413 rinses across all neighborhoods and cities, plus up to 124 drinking water samples (when  
414 water is consumed directly or indirectly via food preparation) in Kisumu. All samples are  
415 stored on ice packs and transported to the APHRC (Nairobi) or Maseno University (Kisumu)  
416 lab for processing and microbial analysis on the day of collection.

#### 417 **6.1.8. Infant birth status and anthropometry data**

418 Preterm birth status is recorded by extracting gestational week of age from the infant's  
419 national birth registration booklet if possible, using caregiver self-report if the booklet is not  
420 available. Infant weight at the point of enrollment is measured using a digital scale. Infant  
421 length is measured using length boards per standard WHO guidelines [70]. These  
422 measurements are transformed into Z scores and categorized as wasted and stunted based  
423 upon whether their Z score is <2 standard deviations below international standard for  
424 weight and length for an infant of the same gender and weeks of age. If the infant's booklet  
425 contains a record on infant weight and length at birth, this information is also recorded and  
426 used to calculate the absolute difference in weight and length from birth to enrollment age  
427 per week for comparison of average growth rates by city, neighborhood status, and other  
428 factors.

#### 429 **6.2. Neighborhood Data Collection**

##### 430 **6.2.1. Environmental hygiene observations and soil and water collection**

431 One neighborhood location is identified every four weeks in each study neighborhood  
432 using feedback from field staff as to locations where infants enrolled in the past four-week  
433 cycle were observed outside of their private households. If no infants left their homes during  
434 structured observation in the last four weeks, we consult the geotracker data to identify  
435 locations where the tracker placed upon the infant was found outside of household  
436 property. Repeating this process each month (12 months) in two neighborhoods in two cities  
437 results in 48 neighborhood areas where enrolled infants may have been exposed to  
438 neighborhood soil and water across a seasonal year. The designated locations are  
439 characterized by recording the types of urban infrastructure present, the presence of  
440 observed human waste disposal (diapers, flying toilets, individual piles, emptied buckets),  
441 and the presence of animal feces. Infrastructure includes small to large wastewater drains,  
442 paved roads, residences, electricity lines, community parks, businesses, and waste dumps.  
443 During neighborhood environmental observations, eight soil samples and up to four surface  
444 water samples are collected at each site to assess for pathogen contamination. The eight soil  
445 samples and four water samples are collected approximately 2m apart as previously  
446 described [39].

#### 447 *6.2.2. Seasonality and weather data*

448 Climate events like precipitation can cause small to large scale flooding that carries  
449 pathogens across the environment and contributes to their persistence in the environment.  
450 Middle class neighborhoods may have better infrastructure for preventing and managing  
451 run-off, so land-based precipitation estimates are likely inaccurate for measuring differences  
452 in urban flood exposure. A supplemental grant to the parent grant is supporting the  
453 development of strategies to generate neighborhood-specific satellite data on cumulative  
454 precipitation and flooding from multiple satellite-based hydrometeorological online platforms  
455 over the 12 month time of data collection in each city for assessment of seasonal  
456 contributions to pathogen detection patterns.

## 457 7. Sample Processing and Microbial Assays

458 The overall sample processing strategy is shown below and in Figure 4 with detailed  
459 methods in Supplemental Materials.

### 460 11.1 Feces Processing

461 In a biosafety cabinet, a 300 mg portion of each feces specimen is transferred from the  
462 diaper into DNA/RNA Shield™ Collection Tubes containing ceramic beads and Shield cell lysis  
463 buffer (Zymo Research Corp, California, USA), spiked with a 10<sup>5</sup> concentration of an extrinsic  
464 control for monitoring extraction efficiency, and extracted using the ZymoBIOMICS Quick  
465 DNA/RNA kit (Zymo Research Corp, California, USA). An extrinsic negative control tube of  
466 Shield buffer is processed twice a week to monitor for potential laboratory contamination.  
467 Approximately half of the Zymo tube is filled with feces and vortexed with an equal amount  
468 of Cary Blair preservation media (HiMedia, India) and 80% glycerol and stored in a  
469 biorepository in an ultralow freezer at the central laboratories. Some frozen feces samples  
470 are cultured for specific bacterial pathogens using the protocols described below to isolate  
471 individual bacteria for further validation of molecular assays.

### 472 11.2 Quantitative molecular detection of enteric pathogens

473 DNA and RNA from infant and animal feces is analyzed as previously described using  
474 microfluidic TaqMan Array cards containing primers and probes for Norovirus GI and GII,  
475 adenovirus 40/41, *Sapovirus*, *Enterovirus*, Rotavirus, *Salmonella enterica*, *Shigella/EIEC*  
476 *backbone*, *Shigella/EIEC plasmid*, *Campylobacter jejuni/coli*, *Listeria monocytogenes*,  
477 *Clostridium difficile*, *Helicobacter pylori*, enterotoxigenic *Escherichia coli* (ETEC) LT/ST,  
478 enteropathogenic *E. coli* (EPEC), enteroaggregative *E. coli* (EAEC), and Shiga-like toxin-  
479 producing *E. coli* (STEC) stx1/stx2, *Enterocytozoon bienersi*, *Giardia*, *Cryptosporidium spp.*,  
480 *Entamoeba histolytica*, and the MS2 extrinsic control [3, 41, 71]. The TaqMan assays are run  
481 on a QuantiStudio 12K Flex Real-Time PCR System (ThermoFisher, Chicago, IL).

482 Prior to using the TaqMan Array cards, the assays are validated using ten-fold serial  
483 dilutions ( $10^0$  to  $10^5$  genes per sample) of positive controls. Bacterial positive controls are  
484 prepared by culture of bacterial strains maintained at the University of Iowa. Other positive  
485 controls were prepared using qBlocks gene fragments (IDT, Coralville, IA, USA) with  
486 pathogen-specific sequence inserts targeted by the primer and probes. One molecular water  
487 PCR control for every box of 25 cards and two negative extraction controls per 100 samples  
488 is included to verify no contamination is introduced during sample preparation. Two field  
489 process negative controls per week are collected on days when samples are placed into  
490 Zymo tubes. These results are examined to verify MS2 extrinsic controls are positive but  
491 other genes are negative. Any feces samples processed on the same day as controls with  
492 false positive amplification are repeated by reisolating and extracting feces from the  
493 biorepository at the central labs. Inhibition is assessed by comparing cycle threshold (Ct) of  
494 MS2 to known concentration spiked into feces prior to extraction, plus calculating 260:280  
495 and 230:260 ratios on a Nanodrop. The recovery efficiency is calculated by dividing the  
496 quantity of MS2 by the quantity of MS2 spiked into the sample prior to concentration.

### 497 *11.3 Selective Detection and Quantification of Environmental Bacteria*

498 Presence and concentration of *Salmonella*, *Shigella*, *Campylobacter*, *Listeria*, and *E. coli*  
499 spp. in household soils, hand rinses, toy rinses, drinking water (Kisumu only), and  
500 neighborhood soil and water is measured using methods adapted from ISO and FDA food  
501 microbiology guidelines and relevant literature, and a Most Probable Number method that  
502 involves primary and secondary pre-enrichment and selective culture of serially diluted  
503 sample volumes.

504 First, three different volumes of soil, hand rinse, toy rinse, and water are measured into  
505 Buffered peptone water pre-enrichment broth (BPW, Himedia, India). Except for public  
506 surface water samples, all samples are subjected to 1:10, 1:100, and 1: 1000 dilutions of the  
507 original sample by measuring 25, 2.5 and 0.25 (g/ml) of samples into 225, 250 and 250 ml



508 BPW, respectively. For water samples, assuming high likely chance of harboring diverse and  
509 high bacterial load, first a 10-fold dilution (1ml sample: 9ml BPW ratio) of public water  
510 surface sample is made from which subsequent serial dilutions (1:100 and 1:1000) are  
511 performed by serially transferring 1ml from the first dilution into tubes containing 9ml of  
512 BPW to yield 1:100, and 1:1000 dilutions. All the three dilutions are incubated at 37°C for  
513 24h to recover sub-lethally injured bacteria.

514       Second, a 100 µL aliquot from each of the three BPW primary enrichment culture is  
515 transferred into 10 mL of selective secondary enrichment media using Rappaport Vassiliadis  
516 Soya Peptone Broth (RVS, Himedia, India) for *Salmonella* and *Shigella* spp., *E.coli* broth  
517 (Oxoid, UK) for *E. coli* spp., Bolton Broth (BB, Oxoid, UK) for *Campylobacter* spp., and Half  
518 Fraser Broth (HFB, Oxoid, UK) for *Listeria monocytogenes*. These are incubated at a growth  
519 temperature recommended for each pathogen (e.g. *Salmonella* spp. at 41.5°C for 24h).

520       Third, a loopful of each ten-fold sample dilution in secondary enrichment broth culture  
521 is streaked on selective agar and grown overnight to isolate specific strains, including  
522 streaking RVS on modified XLD (mXLD, Oxoid, UK) for *Salmonella* and *Shigella* spp., *E. coli*  
523 broth on Eosin methylene blue or MacConkey agar (Oxoid, UK) for *E. coli*, BB on mCCDA  
524 (Oxoid, UK) for *Campylobacter* spp., and FB on ALOA (Oxoid, UK) for *Listeria monocytogenes*.  
525 These are incubated at 37°C for 24h. Incubation of *Campylobacter* spp. occurs in anaerobic  
526 chambers with gas packs (Mitsubishi gas Chemical America Inc., Japan) except for the  
527 primary enrichment.

528       Fourth, up to ten representative colonies of each distinct phenotype are collected from  
529 each plate, sub-cultured into tryptic soya broth (Himedia, India) and incubated at 37°C for  
530 24h and then preserved as glycerol stocks at 1:2 ratio for future re-isolation and analysis. For  
531 DNA extraction, a loopful of aliquot from the glycerol stocks are sub-cultured into Tryptic  
532 Soy agar/selective agar and then five to ten representative colonies are transferred into 100  
533 µL molecular grade water, boiled for 10 minutes, centrifuged to precipitate cell debris, and

534 the supernatant transferred to a new sterile labeled Eppendorf tube. DNA samples from  
535 presumptive positive bacteria are stored at -20°C until pathogen status can be confirmed by  
536 qRT-PCR. Identity of each type of presumptive enteric bacteria phenotype is verified by  
537 polymerase chain reaction assays for gene indicators of pathogen-specific identity of  
538 isolated colonies. Concentration of presumptive bacterial species is determined by  
539 examining which sample cultures across the three volumes of environmental sample tested  
540 for each bacterial protocols contained in presumptive positive colonies using Most Probable  
541 Number (MPN) method as described in Appendix 2 of the Food and Drug Administration  
542 Bacteriology Analytical Manual [72].

#### 543 *11.4 qRT-PCR Verification of environmental bacteria species*

544 Pathogen status of presumptive *Salmonella*, *Shigella*, *Campylobacter*, *Listeria*, and *E.*  
545 *coli* spp. isolated from household soil, hands, toys, and neighborhood soil and water is  
546 tested by qRT-PCR of DNA isolated from bacterial colonies growing on tryptic soya  
547 agar/selective agar. The PCR is conducted by distributing 18 µL of mix (10 µL of master mix,  
548 1 µL of TaqMan assay containing specific primers and probes and 7 µL of nucleic acid-free  
549 water) first and then, 2ul of DNA template of the bacterial isolate to each well in a 0.2 ml 96-  
550 well PCR plate, alongside PCR negative controls. The qRT-PCR primers and probes and the  
551 PCR cycling conditions are identical to those used on TAC to ensure comparability of  
552 detection information.

#### 553 **14. Biorepository**

554 A biorepository generated through this study is preserving all human and animal feces  
555 specimens, bacterial isolates from selective culture of environmental samples, and the non-  
556 selective buffered peptone broth primary enrichment cultures of those environmental  
557 samples for validation of results as well as for future research. Primary stool specimens and  
558 bacterial isolates from the environment are stored in an ultralow freezer (-76°C) at the

559 stated central labs and DNA/RNA/cDNA isolated from all specimens is stored in an ultralow  
560 freezer (-76°C) at the University of Iowa.

#### 561 **14. Analysis**

562 Primary data analysis has been described on [clinicaltrials.gov](https://clinicaltrials.gov), and will involve  
563 characterizing and modeling differences in socioeconomic, infrastructure, behavioral,  
564 spatial, environmental, zoonotic, and human data per our study design: by city and low  
565 versus middle-income neighborhood type, and by age group. The form of analysis and  
566 presentation will be determined by data type and distribution (e.g., principal component  
567 analysis of wealth, diversity and network mapping of pathogen data). Estimates of pathogen  
568 exposure dose and diversity in pathogen type per exposure pathway will be performed as  
569 previously described [43]. Many statistical methods and agent-based models are being  
570 developed newly through this study and will be publicly disseminated via publications, open  
571 data sharing resources, and on our own websites.

#### 572 **15. Community engagement and Translational Science**

573 Community and stakeholder engagement by the research team occurs prior to study  
574 onset in each city to gather feedback from community leaders, health workers, policy  
575 makers, and other academic and non-governmental stakeholders in each city. Suggestions  
576 for improvement or concerns about risks to participants are used to improve study  
577 protocols. Feedback also contributes to a sense of ownership of the study by local policy  
578 makers and practitioners, ensuring that results will be received and potentially implemented  
579 and sustained after the study to address local health challenges. Community and  
580 stakeholder engagement is repeated at the end of the study to share initial learning from  
581 the studies, and a final end-of-study meeting at the end of the statistical modeling and  
582 virtual lab design stage to disseminate final recommendations. To ensure research is quickly  
583 and effectively translated into policy, the research team is conducting a landscaping review  
584 of Government of Kenya policies across sectors, including urban housing and development,

585 health, and agricultural to identify opportunities for improving existing policies,  
586 implementing policies that address gaps, or linking policies between sectors for improved  
587 planning and monitoring.

588 **What this study will contribute**

589 This study will be able to identify the most significant pathways of enteric pathogen  
590 transmission in urban low-to-middle income settings in Kenya. Based on the new learning  
591 we will be able recommended new or modified strategies for combining interventions to  
592 expedite enteric disease control for settings like urban Kenya. Unique elements of the study  
593 as described in our Rationale include a focus on patho-microbial community (PATHOME)  
594 dynamics for examining pathogen transmission patterns across exposure pathways and  
595 rigorous microbiological tools for characterizing the pathome. We recruit middle class  
596 households with high basic standards of development and use a 14-day repeat sampling  
597 cohort design with integrated household and neighborhood microbial exposure assessment.  
598 Our combined behavioral observation and newly developed geotracking methods provide  
599 rigorous new evidence about spatial dimensions of exposure, in terms of where infants  
600 experience environmental exposures and frequency and duration of localized interaction  
601 with domestic animals. Our Planetary Health-One Health integrated approach includes  
602 documentation of urban socio-economic development as well as domestic animal health  
603 and their potential to serve as zoonotic vectors through direct or indirect interaction with  
604 humans. The data generated through this cohort study will be used to construct new  
605 statistical and agent-based models to identify spatial-temporal dynamics in enteric pathogen  
606 transmission in LMICs and the role of societal development in preventing transmission. An  
607 interactive virtual laboratory will be developed to predict how different improvements in  
608 household and neighborhood development alone and in combination could influence  
609 endemic enteric disease burden in Kenya or other urban settings. The overarching goal of  
610 this virtual laboratory is to generate targeted recommendations for improving policy and

611 practice, especially in terms of preventive options, and provide a system for refining those  
612 interventions over time in the context of ongoing societal changes. We will publicly share  
613 the large datasets we collect, thereby providing an invaluable asset to other modeling  
614 groups working on reducing incidence rates of enteric pathogen infections in children in  
615 LMICs.

## 616 ETHICS AND DISSEMINATION

617 This study conducts research on both human and animal health outcomes. The  
618 protocols for human subjects' research for the PATHOME study were approved by  
619 Institutional Review Boards at the University of Iowa (ID - 202004606), AMREF Health Africa  
620 (ID - ESRC P887/2020), and a permit obtained from the Kenyan National Commission for  
621 Science Technology and Innovation (ID# P/21/8441). The study is also registered on  
622 Clinicaltrials.gov (Identifier: NCT05322655). Protocols for research on canines, felines,  
623 avians, and ruminants were approved by the UI Animal Care and Use Committee (ID  
624 0042302) in accordance with the Guide for the Care and Use of Laboratory Animals, NIH  
625 Publication No. 85-23 (2010). Details on human and animal research ethical considerations  
626 and protocols are reported in Supplemental Material. The entire research team took  
627 refresher training on ethical principles on protecting human research participants.

628

## 629 REFERENCES

- 630 1. Collaborators, G.B.D.D.D., *Estimates of the global, regional, and national*  
631 *morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a*  
632 *systematic analysis for the Global Burden of Disease Study 2016*. *Lancet*  
633 *Infect Dis*, 2018. **18**(11): p. 1211-1228.
- 634 2. Typhoid, G.B.D. and C. Paratyphoid, *The global burden of typhoid and*  
635 *paratyphoid fevers: a systematic analysis for the Global Burden of Disease*  
636 *Study 2017*. *Lancet Infect Dis*, 2019. **19**(4): p. 369-381.
- 637 3. Platts-Mills, J.A., et al., *Pathogen-specific burdens of community diarrhoea*  
638 *in developing countries: a multisite birth cohort study (MAL-ED)*. *Lancet*  
639 *Glob Health*, 2015. **3**(9): p. e564-75.
- 640 4. Fischer Walker, C.L., et al., *Diarrhea incidence in low- and middle-income*  
641 *countries in 1990 and 2010: a systematic review*. *BMC Public Health*, 2012.  
642 **12**: p. 220.

- 643 5. Petri Jr, W.A., et al., *Enteric infections, diarrhea, and their impact on*  
644 *function and development*. The Journal of clinical investigation, 2008.  
645 **118**(4): p. 1277.
- 646 6. Humphrey, J.H., *Child undernutrition, tropical enteropathy, toilets, and*  
647 *handwashing*. Lancet, 2009. **374**(9694): p. 1032-5.
- 648 7. Checkley, W., et al., *Multi-country analysis of the effects of diarrhoea on*  
649 *childhood stunting*. International journal of epidemiology, 2008. **37**(4): p.  
650 816-830.
- 651 8. Rogawski McQuade, E.T., et al., *Impact of Shigella infections and*  
652 *inflammation early in life on child growth and school-aged cognitive*  
653 *outcomes: Findings from three birth cohorts over eight years*. PLoS Negl  
654 Trop Dis, 2022. **16**(9): p. e0010722.
- 655 9. Hardy, A.V. and J. Watt, *Studies of the acute diarrheal diseases;*  
656 *epidemiology*. Public Health Rep (1896), 1948. **63**(12): p. 363-78.
- 657 10. Brownlee, J. and M. Young, *The Epidemiology of Summer Diarrhoea*. Proc R  
658 Soc Med, 1922. **15**(Sect Epidemiol State Med): p. 55-74.
- 659 11. Anderson, D.M., D.I. Rees, and T. Wang, *The phenomenon of summer*  
660 *diarrhea and its waning, 1910-1930*. Explor Econ Hist, 2020. **78**.
- 661 12. Alsan, M. and C. Goldin, *Watersheds in Child Mortality: The Role of Effective*  
662 *Water and Sewerage Infrastructure, 1880 to 1920*. J Polit Econ, 2019.  
663 **127**(2): p. 586-638.
- 664 13. Cain, L., *Death and spending: urban mortality and municipal expenditure on*  
665 *sanitation*. Annales De Demographie Historique, 2001. **1**: p. 139-154.
- 666 14. Cutler, D. and G. Miller, *The role of public health improvements in health*  
667 *advances: the twentieth-century United States*. Demography, 2005. **42**(1):  
668 p. 1-22.
- 669 15. Collaborators, G.B.D.C.o.D., *Global, regional, and national age-sex specific*  
670 *mortality for 264 causes of death, 1980-2016: a systematic analysis for the*  
671 *Global Burden of Disease Study 2016*. Lancet, 2017. **390**(10100): p. 1151-  
672 1210.
- 673 16. Luby, S.P., et al., *Effects of water quality, sanitation, handwashing, and*  
674 *nutritional interventions on diarrhoea and child growth in rural*  
675 *Bangladesh: a cluster randomised controlled trial*. Lancet Glob Health,  
676 2018. **6**(3): p. e302-e315.
- 677 17. Null, C., et al., *Effects of water quality, sanitation, handwashing, and*  
678 *nutritional interventions on diarrhoea and child growth in rural Kenya: a*  
679 *cluster-randomised controlled trial*. Lancet Glob Health, 2018. **6**(3): p.  
680 e316-e329.
- 681 18. Rogawski McQuade, E.T., et al., *Impact of water quality, sanitation,*  
682 *handwashing, and nutritional interventions on enteric infections in rural*  
683 *Zimbabwe: the Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial*. J  
684 Infect Dis, 2019.
- 685 19. Wolf, J., et al., *Impact of drinking water, sanitation and handwashing with*  
686 *soap on childhood diarrhoeal disease: updated meta-analysis and meta-*  
687 *regression*. Trop Med Int Health, 2018. **23**(5): p. 508-525.
- 688 20. Headey, D. and G. Palloni, *Water, Sanitation, and Child Health: Evidence*  
689 *From Subnational Panel Data in 59 Countries*. Demography, 2019. **56**(2):  
690 p. 729-752.

- 691 21. Barreto, M.L., et al., *Impact of a citywide sanitation program in Northeast*  
692 *Brazil on intestinal parasites infection in young children*. Environ Health  
693 *Perspect*, 2010. **118**(11): p. 1637-42.
- 694 22. Barreto, M.L., et al., *Effect of city-wide sanitation programme on reduction*  
695 *in rate of childhood diarrhoea in northeast Brazil: assessment by two cohort*  
696 *studies*. Lancet, 2007. **370**(9599): p. 1622-8.
- 697 23. Jung, Y.T., et al., *Effects of neighbourhood and household sanitation*  
698 *conditions on diarrhea morbidity: Systematic review and meta-analysis*.  
699 PLoS One, 2017. **12**(3): p. e0173808.
- 700 24. Bhalotra, S.R., et al., *Urban Water Disinfection And Mortality Decline In*  
701 *Developing Countries*, N.B.o.E. Research, Editor. 2018: Cambridge, MA.
- 702 25. McKeown, R.E., *The Epidemiologic Transition: Changing Patterns of*  
703 *Mortality and Population Dynamics*. Am J Lifestyle Med, 2009. **3**(1 Suppl):  
704 p. 19S-26S.
- 705 26. Russell, J.B., *The Evolution of the Function of Public Health Administration*.  
706 1895.
- 707 27. Russell, J.B., *On Certain Epidemic Outbreaks of Enteric Fever in April, 1880,*  
708 *Traced to Contamination of Milk, Being a Report Presented to the Police*  
709 *Board of Glasgow*. Glasgow Med J, 1880. **14**(8): p. 102-125.
- 710 28. Russell, J.B. and A.K. Chalmers, *On an Outbreak of Scarlet Fever in Glasgow*  
711 *Connected with an Epidemic Teat Eruption on Milch Cows at Jaapston*.  
712 Glasgow Med J, 1893. **39**(1): p. 1-22.
- 713 29. Russell, H.L., *Typhoid Fever Disseminated through the Milk Supply*. Science,  
714 1895. **2**(47): p. 682-3.
- 715 30. *Milk and Other Food Poisoning and Epidemic Diarrhoea*. Br Med J, 1903.  
716 **1**(2199): p. 456-9.
- 717 31. Russell, J.B., *On the Distribution of Enteric Fever in Glasgow*. Glasgow Med  
718 J, 1868. **1**(1): p. 59-67.
- 719 32. Peters, O.H., *Observations upon the Natural History of Epidemic Diarrhoea*.  
720 J Hyg (Lond), 1910. **10**(4): p. 602-777.
- 721 33. Gibb, R., et al., *Zoonotic host diversity increases in human-dominated*  
722 *ecosystems*. Nature, 2020. **584**(7821): p. 398-402.
- 723 34. Graham-Smith, G.S., *The Relation of the Decline in the Number of Horse-*  
724 *drawn Vehicles, and consequently of the Urban Breeding-Grounds of Flies,*  
725 *to the Fall in the Summer Diarrhoea Death-rate*. J Hyg (Lond), 1929. **29**(2):  
726 p. 132-8.
- 727 35. Nash, J.T.C., *Remarks on the Seasonal Incidence of Diarrhoea, and the Part*  
728 *Played by Flies*. Hospital (Lond 1886), 1903. **33**(854): p. 317.
- 729 36. Sanders, J.W., et al., *The epidemiological transition: the current status of*  
730 *infectious diseases in the developed world versus the developing world*. Sci  
731 Prog, 2008. **91**(Pt 1): p. 1-37.
- 732 37. Barnes, A.N., J. Mumma, and O. Cumming, *Role, ownership and presence of*  
733 *domestic animals in peri-urban households of Kisumu, Kenya*. Zoonoses  
734 Public Health, 2018. **65**(1): p. 202-214.
- 735 38. Jenkins, M.W., et al., *Identifying human and livestock sources of fecal*  
736 *contamination in Kenya with host-specific Bacteroidales assays*. Water Res,  
737 2009. **43**(19): p. 4956-66.

- 738 39. Baker, K.K., et al., *Fecal Fingerprints of Enteric Pathogen Contamination in*  
739 *Public Environments of Kisumu, Kenya associated with Human Sanitation*  
740 *Conditions and Domestic Animals*. Environ Sci Technol, 2018.
- 741 40. Muloj, D.M., et al., *Population genomics of Escherichia coli in livestock-*  
742 *keeping households across a rapidly developing urban landscape*. Nat  
743 Microbiol, 2022. **7**(4): p. 581-589.
- 744 41. Liu, J., et al., *Use of quantitative molecular diagnostic methods to identify*  
745 *causes of diarrhoea in children: a reanalysis of the GEMS case-control study*.  
746 Lancet, 2016. **388**(10051): p. 1291-301.
- 747 42. Medgyesi, D.N., et al., *Where Children Play: Young child exposure to*  
748 *environmental hazards during play in public areas in a transitioning*  
749 *internally displaced persons community in Haiti*. International Journal of  
750 Environmental Research and Public Health, 2018. **15**(8): p. 1682.
- 751 43. Medgyesi, D., et al., *The landscape of enteric pathogen exposure of young*  
752 *children in public domains of low-income, urban Kenya: The influence of*  
753 *exposure pathway and spatial range of play on multi-pathogen exposure*  
754 *risks*. PLoS Negl Trop Dis, 2019. **13**(3): p. e0007292.
- 755 44. Cairncross, S., et al., *The public and domestic domains in the transmission of*  
756 *disease*. Trop Med Int Health, 1996. **1**(1): p. 27-34.
- 757 45. Davis, E., et al., *Oral Contact Events and Caregiver Hand Hygiene:*  
758 *Implications for Fecal-Oral Exposure to Enteric Pathogens among Infants 3-*  
759 *9 Months Living in Informal, Peri-Urban Communities in Kisumu, Kenya*. Int  
760 J Environ Res Public Health, 2018. **15**(2).
- 761 46. Tsai, K., et al., *Enteric Pathogen Diversity in Infant Foods in Low-Income*  
762 *Neighborhoods of Kisumu, Kenya*. Int J Environ Res Public Health, 2019.  
763 **16**(3).
- 764 47. Wagner, E.G. and J.N. Lanoix, *Excreta disposal for rural areas and small*  
765 *communities*. Excreta Disposal for Rural Areas and Small Communities,  
766 1958.
- 767 48. Pineo, H., et al., *Building a Methodological Foundation for Impactful Urban*  
768 *Planetary Health Science*. J Urban Health, 2021. **98**(3): p. 442-452.
- 769 49. Amugsi, D.A., T.N. Haregu, and B.U. Mberu, *Levels and determinants of*  
770 *perceived health risk from solid wastes among communities living near to*  
771 *dumpsites in Kenya*. Int J Environ Health Res, 2019: p. 1-12.
- 772 50. Aseyo, R.E., et al., *Realities and experiences of community health volunteers*  
773 *as agents for behaviour change: evidence from an informal urban*  
774 *settlement in Kisumu, Kenya*. Hum Resour Health, 2018. **16**(1): p. 53.
- 775 51. Simiyu, S., M. Swilling, and S. Cairncross, *Decision-making on shared*  
776 *sanitation in the informal settlements of Kisumu, Kenya*. Int J Environ  
777 Health Res, 2017. **27**(5): p. 377-393.
- 778 52. Simiyu, S., et al., *Determinants of quality of shared sanitation facilities in*  
779 *informal settlements: case study of Kisumu, Kenya*. BMC Public Health,  
780 2017. **17**(1): p. 68.
- 781 53. Simiyu, S., et al., *Estimating the Cost and Payment for Sanitation in the*  
782 *Informal Settlements of Kisumu, Kenya: A Cross Sectional Study*. Int J  
783 Environ Res Public Health, 2017. **14**(1).
- 784 54. Mumma, J.A.O., et al., *Infant food hygiene and childcare practices in context:*  
785 *Findings from an urban informal settlement in Kenya*. American Journal of  
786 Tropical Medicine and Hygiene, In Press.



- 787 55. Mumma, J.A.O., et al., *The Safe Start Trial A Cluster Randomized Controlled*  
788 *Trial for the Effect of an Infant Hygiene Intervention on Enteric Infections*  
789 *and Diarrhoea in Low-Income Informal Settlements of Kisumu, Kenya*. BMC  
790 Infectious Diseases, In Press.
- 791 56. Chege, M. and B. Mberu, *Improving solid waste management practices and*  
792 *addressing associated health risks in Nairobi and Mombasa.*, in *Urban*  
793 *Africa Risk and Knowledge Briefing No. 11*. 2018, African Population and  
794 Health Research Center: Nairobi, Kenya.
- 795 57. Ezeh, A., et al., *The history, geography, and sociology of slums and the*  
796 *health problems of people who live in slums*. Lancet, 2017. **389**(10068): p.  
797 547-558.
- 798 58. Mberu, B., et al., *Impact of Solid Waste Management on Health: A*  
799 *Biomedical Study of Solid Waste Workers at Dandora Dumpsite, Nairobi,*  
800 *Kenya*, in *Urban Africa Risk Knowledge Working Papers*. 2019, African  
801 Population and Health Research Centre: Nairobi, Kenya.
- 802 59. Mberu, B.U., et al., *Health and health-related indicators in slum, rural, and*  
803 *urban communities: a comparative analysis*. Glob Health Action, 2016. **9**: p.  
804 33163.
- 805 60. Ziraba, A.K., T.N. Haregu, and B. Mberu, *A review and framework for*  
806 *understanding the potential impact of poor solid waste management on*  
807 *health in developing countries*. Arch Public Health, 2016. **74**: p. 55.
- 808 61. Statistics, K.N.B.o., *Nyanza Province Multiple Indicator Cluster Survey 2011*.  
809 2013, Kenyan National Bureau of Statistics: Nairobi, Kenya.
- 810 62. Barnes, A.N., et al., *The association between domestic animal presence and*  
811 *ownership and household drinking water contamination among peri-urban*  
812 *communities of Kisumu, Kenya*. PLoS One, 2018. **13**(6): p. e0197587.
- 813 63. Ngure, F.M., et al., *Formative research on hygiene behaviors and geophagy*  
814 *among infants and young children and implications of exposure to fecal*  
815 *bacteria*. Am J Trop Med Hyg, 2013. **89**(4): p. 709-16.
- 816 64. Levy, K., et al., *Untangling the Impacts of Climate Change on Waterborne*  
817 *Diseases: a Systematic Review of Relationships between Diarrheal Diseases*  
818 *and Temperature, Rainfall, Flooding, and Drought*. Environ Sci Technol,  
819 2016. **50**(10): p. 4905-22.
- 820 65. Kwong, L.H., et al., *Hand- and Object-Mouthing of Rural Bangladeshi*  
821 *Children 3-18 Months Old*. Int J Environ Res Public Health, 2016. **13**(6).
- 822 66. Cumming, O., et al., *The implications of three major new trials for the effect*  
823 *of water, sanitation and hygiene on childhood diarrhea and stunting: a*  
824 *consensus statement*. BMC Med, 2019. **17**(1): p. 173.
- 825 67. Arnold, B.F., et al., *Implications of WASH Benefits trials for water and*  
826 *sanitation - Authors' reply*. Lancet Glob Health, 2018. **6**(6): p. e616-e617.
- 827 68. Ritter, R.L., et al., *Within-Compound Versus Public Latrine Access and Child*  
828 *Feces Disposal Practices in Low-Income Neighborhoods of Accra, Ghana*. Am  
829 J Trop Med Hyg, 2018.
- 830 69. Torondel, B., et al., *Laboratory development and field testing of sentinel*  
831 *toys to assess environmental faecal exposure of young children in rural*  
832 *India*. Trans R Soc Trop Med Hyg, 2015. **109**(6): p. 386-92.
- 833 70. *Recommendations for data collection, analysis and reporting on*  
834 *anthropometric indicators in children under 5 years old*. 2019, World

835 Health Organization and the United Nations Children’s Fund (UNICEF):  
836 Geneva.  
837 71. Steiner, K.L., et al., *Species of Cryptosporidia Causing Subclinical Infection*  
838 *Associated with Growth Faltering in Rural and Urban Bangladesh- a Birth*  
839 *Cohort Study*. Clin Infect Dis, 2018.  
840 72. *Bacteriological Analytical Manual*, C.f.F.S.a.A. Nutrition, Editor. Continuous  
841 updating: Silver Spring, MD.

842

#### 843 **AUTHOR’S CONTRIBUTIONS**

844 KKB and DKS conceived of the study and secured funding. KKB drafted the epidemiological  
845 study design, with input from DKS, SS, AZ, and BM. SS and PJB designed field protocols and  
846 survey, behavioral, diarrhea calendar, and anthropometry tools, with support from AZ, AO,  
847 and KKB. KKB and FDG designed field sampling and laboratory protocols, with input from BO,  
848 JA, CSA, AGK, and CO. BO, JA, CA, and CO implemented culture-based laboratory protocols,  
849 and FDG and AK implemented molecular and genomic protocols. PJB, AMO, and KKB  
850 designed geotracking protocols. PJB, SS, SG, and DKS oversaw data management and quality  
851 control. AZ, SS, and CO facilitated human subjects’ and animal research approvals and  
852 stakeholder engagement in Kenya. KKB oversaw human subjects and animal research  
853 protocol approvals in Iowa. BM conducted policy landscape analysis. KKB and FDG oversaw  
854 enteric pathogen analysis approach. DKS oversaw the statistical and agent-based analysis of  
855 the epidemiological data. KKB wrote the manuscript, with input from all authors.

856

#### 857 **FUNDING STATEMENT**

858 The PATHOME study is funded by the National Institutes of Health Fogarty Institute Grant  
859 Number 01TW011795 to University of Iowa. The first draft of the proposal was submitted to  
860 the National Science Foundation Evolution and Ecology of Infectious Disease mechanism in  
861 November 2019 with notice of award and study launch occurring in July 2020. The content is  
862 solely the responsibility of the authors and does not necessarily represent the official views

863 of the National Institutes of Health. The funders had no role in study design, data collection  
864 and analysis, decision to publish or preparation of the manuscript.

865

866 **COMPETING INTERESTS STATEMENT**

867 The authors declare no conflicts of interests

868

869 **WORD COUNT:** 7,089

870

871 **Figure 1.** Conceptual Model of the Enteric PATHOME Study. Societal development of  
872 household water, latrines, flooring, animal health, and hygiene conditions and neighborhood  
873 infrastructure (drainage) and hygiene (animal penning, waste regulations) conditions  
874 triggers an evolutionary change in the enteric PATHOME from taxonomically complex to  
875 simple microbial community detection patterns in humans, animals, and the environment.

876

877 **Figure 2.** Low- and middle-income study neighborhoods in Nairobi and Kisumu, Kenya,  
878 where households with 0-to-12-month of age children are enrolled into the PATHOME  
879 study. Yellow bars represent 100 meters distance. Image licensed through Vectorstock  
880 Standard License.

881

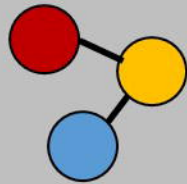
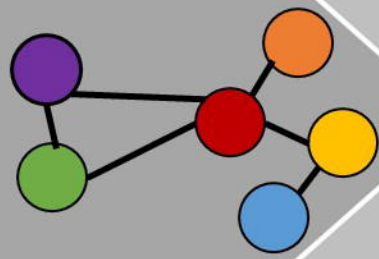
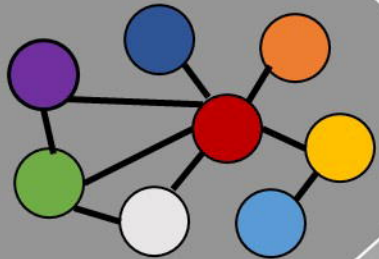
882 **Figure 3.** Study design and timeline for PATHOME household data and biological specimen  
883 collection. Symbols matching data collection activity placed under days when activities are  
884 implemented. Survey = self-reported data on household conditions. Diaper = provision of  
885 diapers for feces collection. Infant = days of diaper/feces collection. Chicken = animal feces  
886 collected. Binoculars = structured observation days. Geotracker = 24-hour time window  
887 where spatial movement is monitored. Toy ball = days of environmental sample collection.  
888 Infant on length board = anthropometric data. Calendar = window of time for prospective  
889 self-reporting of infant diarrhea symptoms. Photo credit for tape measure image: ShareAlike  
890 3.0 Unported (CC BY-SA 3.0). Image of child licensed through Vectorstock Standard License.  
891 Other images taken by authors or are available through a Creative Commons license.

892

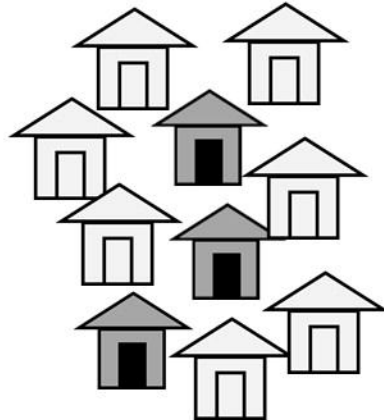
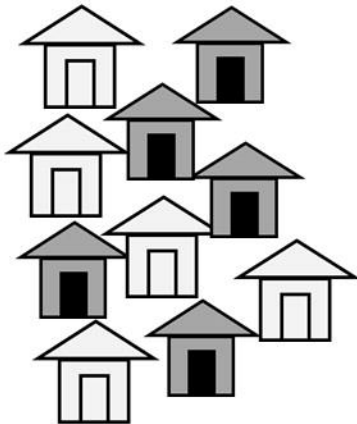
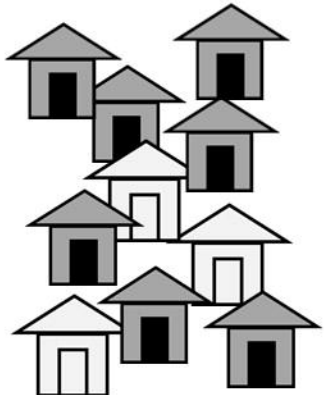
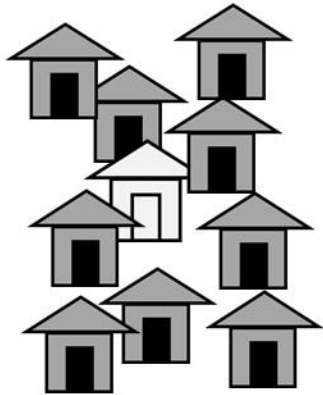
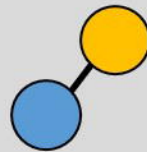
893 **Figure 4.** Flow diagram of the strategy for recovery, detection, quantification, and  
894 confirmation of enteric pathogens in human and animal feces and environmental fomites.  
895 Buffered Peptone water (BPW); Rappaport Vassiliadis Soya Peptone Broth (RVS); Half Fraser

- 896 Broth (HFB); Bolton Broth (BB); modified Xylose Lysine Deoxycholate (mXLD); quantitative  
897 reverse transcription polymerase chain reaction (qRT-PCR); multiplex PCR (mPCR).

**HIGH PATHOME DIVERSITY**



**LOW PATHOME DIVERSITY**



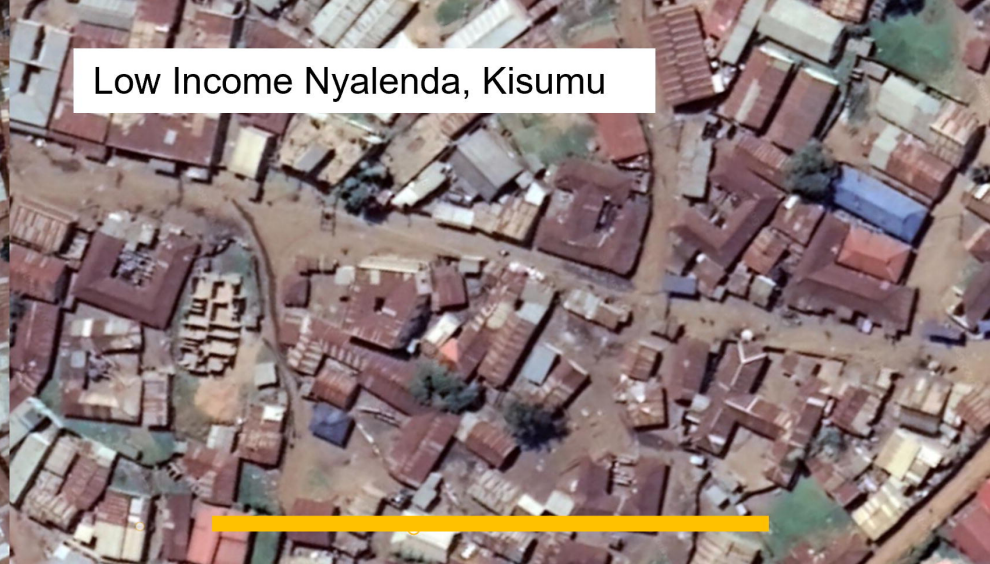
**LOW LEVEL OF  
URBAN DEVELOPMENT**

**HIGH LEVEL OF  
URBAN DEVELOPMENT**

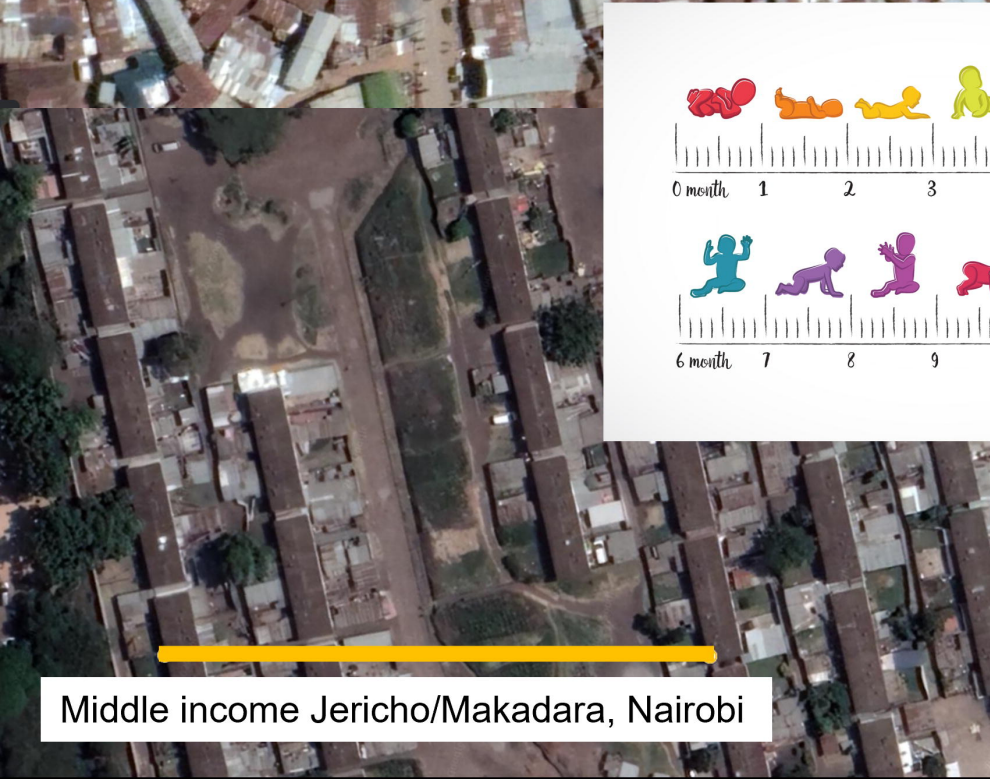
Low-income Kibera, Nairobi



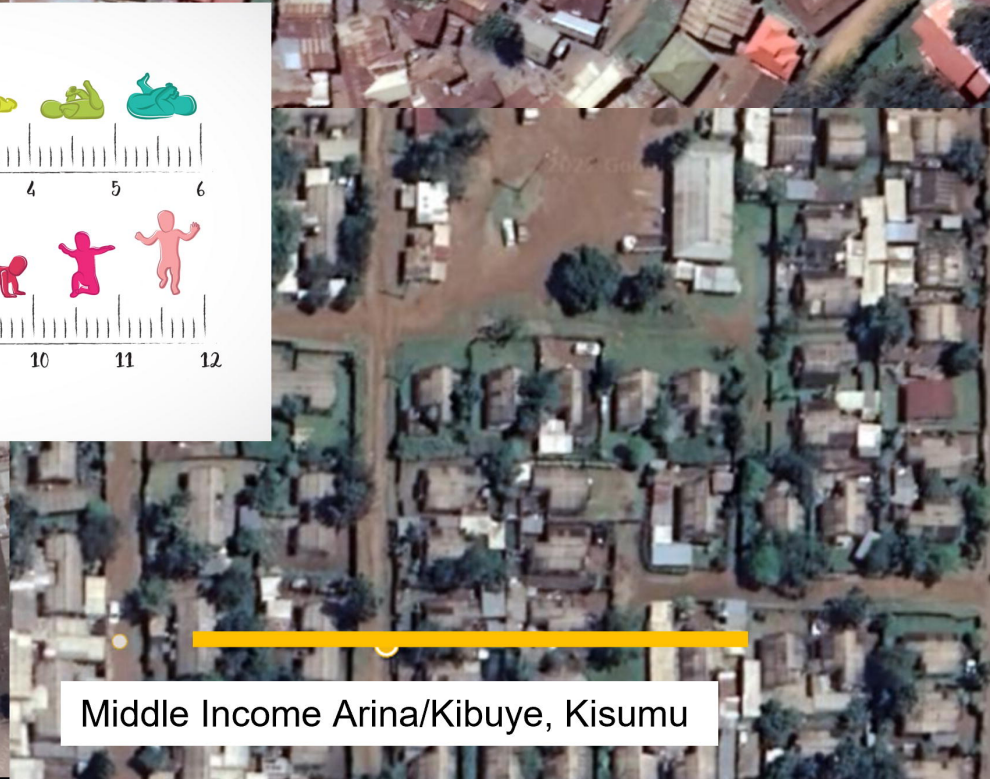
Low Income Nyalenda, Kisumu
























Middle income Jericho/Makadara, Nairobi



Middle Income Arina/Kibuye, Kisumu



	Enrollment	Day 1	Day 2	Day 3	Day 4	Day 7	Day 14
Survey							
Infant feces samples							
Structured Observation							
Infant and animal geotracking							
Animal feces samples							
Household environmental samples							
Length, weight, MUAC							
14-day self-reported diarrhea diary							



Infant (n<=1,240) and  
Animal Feces (n<=496)



Neighborhood soil (n=320)  
and surface water (n<=80)



Household soil (n=496), toys (n=248)  
and caregiver hands (n=248)



24 hour BPW pre-enrichment of 3 ten-fold dilutions  
of sample at 37°C

24 hour BPW pre-enrichment of 3 ten-fold dilutions  
of sample at 37°C

-80°C  
repository

DNA/RNA  
extraction

multi-pathogen  
qRT-PCR

Pathogen-specific  
isolation/validation

DNA/RNA  
extraction

-80°C  
repository

multi-pathogen TAC  
qRT-PCR

Pathogen-specific  
isolation/validation

Enrich RVS  
broth

Selective  
XLD/MPN

-80°C  
repository

DNA  
extraction

mPCR *Shigella* and/  
or *Salmonella*

Enrich *E. coli*  
broth

Selective  
MAC/MPN

-80°C  
repository

DNA  
extraction

mPCR *E. coli*

Enrich HFB  
broth

Selective  
PALCAM/MPN

-80°C  
repository

DNA  
extraction

PCR *hlyA*

Enrich BB

Selective  
mCCDA/MPN

-80°C  
repository

DNA  
extraction

PCR *cadF*

Enrich RVS  
broth

Selective  
XLD/MPN

-80°C  
repository

DNA  
extraction

mPCR *Shigella*  
and/or *Salmonella*

Enrich *E. coli*  
broth

Selective  
MAC/MPN

-80°C  
repository

DNA  
extraction

mPCR *E. coli*