

1 **Geospatial Analysis of Multilevel Socio-environmental Factors Impacting the**  
2 ***Campylobacter* Burden Among Infants in Rural Eastern Ethiopia: A One Health**  
3 **Perspective**

4  
5 Xiaolong Li<sup>1,2</sup>, Dehao Chen<sup>1,2</sup>, Song Liang<sup>1,2</sup>, Jemal Yousuf Hassen<sup>3</sup>, Sarah L.  
6 McKune<sup>1,4,6</sup>, Arie H. Havelaar<sup>2,5,6</sup>, Jason K. Blackburn<sup>2,7\*</sup> on behalf of the CAGED  
7 Research Team

8 <sup>1</sup>Department of Environmental and Global Health, College of Public Health and Health  
9 Professions, University of Florida, Gainesville, FL, USA

10 <sup>2</sup>Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA

11 <sup>3</sup>Haramaya University, Dire Dawa, Ethiopia

12 <sup>4</sup>Center for African Studies, University of Florida, Gainesville, FL, USA

13 <sup>5</sup>Department of Animal Sciences, Institute of Food and Agricultural Sciences, University  
14 of Florida, Gainesville, FL, USA

15 <sup>6</sup>Global Food Systems Institute, University of Florida, Gainesville, FL, USA

16 <sup>7</sup>Spatial Epidemiology and Ecology Research Laboratory, Department of Geography,  
17 College of Liberal Arts and Sciences, University of Florida, Gainesville, FL, USA

18 \*Corresponding author: Jason K. Blackburn (jkblackburn@ufl.edu)

19 **Key words:** *Campylobacter* infection; geospatial analysis; infants; Ethiopia; One Health

20

21 **Abstract**

22 Increasing attention has focused on health outcomes of *Campylobacter* infections  
23 among children under five years in low-resource settings. Recent evidence suggests  
24 colonization of *Campylobacter* species contributes to environmental enteric dysfunction,  
25 malnutrition, and growth faltering in young children. *Campylobacter* species are  
26 zoonotic, and factors from humans, animals, and the environment are involved in  
27 transmission. Few studies have assessed geospatial effects of environmental factors  
28 along with human and animal factors on *Campylobacter* infections. Here, we leveraged  
29 *Campylobacter* Genomics and Environmental Enteric Dysfunction (CAGED) project  
30 data to model multiple socio-environmental factors on *Campylobacter* burden among  
31 infants in Eastern Ethiopia. Stool samples from 106 infants were collected monthly from  
32 birth through the first year of life (December 2020 – June 2022). Genus-specific  
33 Taqman real-time PCR was performed to detect and quantify *Campylobacter* spp. and  
34 calculate cumulative *Campylobacter* burden for each child as the outcome variable.  
35 Thirteen regional environmental covariates describing topography, climate, vegetation,  
36 soil, and human population density were combined with household demographics,  
37 livelihoods/wealth, livestock ownership, and child-animal interactions as explanatory  
38 variables. We dichotomized all continuous outcome and explanatory variables and built  
39 logistic regression models for the first and second half of the infant's first year of life.  
40 Infants being female, living in households with cattle, reported to have physical contact  
41 with animals, or reported to have mouthed soil or animal feces had increased odds of  
42 higher cumulative *Campylobacter* burden in Eastern Ethiopia. Future interventions

- 43 should focus on infant-specific transmission pathways and create adequate separation  
44 of domestic animals from humans to prevent potential fecal exposures.  
45

## 46 Introduction

47 *Campylobacter* species are zoonotic bacteria, and the animal reservoirs include  
48 poultry, free-living birds, ruminants such as cattle and goats, as well as other warm-  
49 blooded animals <sup>1,2</sup>. Common symptoms of *Campylobacter* infection include fever,  
50 diarrhea, and abdominal pain <sup>3</sup>. In addition to symptomatic infections, asymptomatic  
51 *Campylobacter* cases are common, specifically among children and adults in low- and  
52 middle-income countries (LMICs) <sup>4,5</sup>. Evidence from studies in low-resource settings has  
53 shown that colonization of *Campylobacter* species acquired at the early stage of child  
54 development potentially contributes to environmental enteric dysfunction (EED),  
55 malnutrition, and growth faltering in children <sup>4,6,7</sup>, which draws increasing attention to the  
56 long-term health effects of *Campylobacter* infections among children under five years of  
57 age (CU5) in LMICs.

58 *Campylobacter jejuni* and *C. coli* have been reported as the two most common  
59 species that cause illness in humans <sup>3</sup>. Traditional methods selectively isolating these  
60 two species have resulted in a large body of knowledge specific to their disease burden  
61 and clinical manifestations <sup>8</sup>. However, several *Campylobacter* species other than *C.*  
62 *jejuni/coli* (non- *C. jejuni/coli* species), such as *C. concisus*, *C. lari*, *C. upsaliensis*, and  
63 *C. ureolyticus*, have been recognized as “emerging species” and have shown increasing  
64 clinical importance <sup>9</sup>. An increasing proportion of non-*C. jejuni/coli* species, including a  
65 new species "*Candidatus Campylobacter infans*" (*C. infans*), were isolated from child  
66 stool samples collected in low-resource settings, and several studies have revealed  
67 their potential linkages with child stunting <sup>7,10-13</sup>. This evidence shifts our attention

68 towards the genus *Campylobacter* as a whole, rather than focusing solely on *C.*  
69 *jejuni/coli* in the context of child health.

70 The classic “F-diagram” (fluids, fields, flies, fingers, fomites, and food) from WHO  
71 indicates that adequate sanitation and hygiene can be effective in interrupting the fecal-  
72 oral transmission routes of diarrheic pathogens, including *Campylobacter*, in low-  
73 resource settings<sup>14,15</sup>. Driven by this idea, water, sanitation, and hygiene (WaSH)  
74 interventions, including improved pit latrine, hand-washing stations, liquid soap, and  
75 point-of-use water chlorination, have been developed and implemented in LMICs<sup>16,17</sup>.  
76 Substantial evidence suggests that adequate WaSH contributes to reducing the risk of  
77 childhood diarrheal disease<sup>18,19</sup>. However, several recent large-scale randomized  
78 controlled trials have found that WaSH interventions did not significantly reduce the  
79 occurrence of diarrhea or other enteric infections<sup>20-22</sup>. One possible explanation is that  
80 traditional WaSH interventions generally target routes of children’s exposure to human  
81 feces but fail to pay adequate attention to children’s exposure to animal feces<sup>15,23</sup>. Risk  
82 factors for exposing children to animal feces might include livestock ownership,  
83 proximity to livestock, cohabitation of humans and animals, etc.

84 The first year of a child's life is marked by remarkable developmental milestones,  
85 and the differentiation between the first and second six months brings about major  
86 changes in both biological risk and environmental interaction. In the initial six months,  
87 breastfeeding plays a crucial role in offering infants essential nutrients and immune  
88 protection against pathogens<sup>24</sup>. However, as the infant transitions into the second six  
89 months, the introduction of solid foods broadens the child's nutritional spectrum but also  
90 amplifies their exposure to contaminants and pathogens through the food pathway<sup>25</sup>.

91 Meanwhile, the infant becomes more mobile, gaining the ability to explore their  
92 surroundings independently. This newfound mobility empowers infants to explore their  
93 environment actively through touching and tasting<sup>26</sup>, potentially increasing exposure to  
94 pathogens via the fecal-oral pathway. This developmental transition underscores the  
95 dynamic effects of age on infant's biological risk and environmental interaction.

96 Moreover, other environmental factors could impact the non-food transmission  
97 route of *Campylobacter* to humans directly or via animal reservoirs indirectly<sup>27,28</sup>.  
98 Studies conducted in European countries, the United States, and Canada found  
99 campylobacteriosis incidence to be correlated with climatic variables, including  
100 temperature and precipitation<sup>29-33</sup>. In addition to weather and climate factors,  
101 Sanderson et al. also examined the potential impacts of hydrology and landscape  
102 features, like soil type and land use, on the rates of human *Campylobacter* cases in the  
103 UK<sup>34</sup>. This study showed that an increased risk of *Campylobacter* infections was  
104 associated with periods of high surface-water flow and catchment areas with  
105 cattle/sheep grazing on stagnogley soils. Another similar study conducted in New  
106 Zealand linked the risk of infection to a high dairy cattle density<sup>35</sup>. However, these kinds  
107 of studies that investigate the influence of environmental factors have been  
108 predominantly conducted in developed countries with a focus on *C. jejuni/coli*. There is  
109 a critical knowledge gap in environmental effects on *Campylobacter* infections in LMICs.

110 Although multiple factors are involved in the transmission pathways of  
111 *Campylobacter*, few studies have assessed the combined effects of environmental  
112 covariates with other relevant factors from different domains including humans and  
113 animals. Since environmental data often comprise spatial information such as land

114 cover, precipitation, and topography, geospatial analysis is needed to address the  
115 inherent differences in data types between environmental factors and human/animal  
116 factors by integrating them within a spatial framework (e.g., spatial regression). This  
117 capability to analyze data at different levels, from individual households to broader  
118 environmental landscapes, enhances the depth of risk factor identification. Coupled with  
119 the One Health perspective, this kind of analysis will help unravel the complex interplay  
120 between human, animal, and environmental factors involved in *Campylobacter*  
121 transmissions among children in a low-resource setting.

122         The longitudinal study of the *Campylobacter* Genomics and Environmental  
123 Enteric Dysfunction (CAGED) project conducted in rural Eastern Ethiopia aimed to  
124 examine the association between *Campylobacter* infection, related reservoirs, and child  
125 health outcomes<sup>36</sup>. It collected not only the child fecal samples but also environmental  
126 (e.g., soil and drinking water) and livestock samples. A household questionnaire  
127 included components on people (e.g., demographics, livelihoods and wealth), livestock,  
128 and the interactions between people, livestock, and the environment. Combined with the  
129 child and environmental samples, these data provide an opportunity to investigate the  
130 potential combined effects of these risk factors on *Campylobacter* infection in infants.

## 131 **Materials and Methods**

### 132 **Study Design and Protocol**

133         A detailed description of the CAGED study design and protocol can be found  
134 elsewhere<sup>36</sup>. Briefly, a total of 106 infants from 10 kebeles (the smallest administrative  
135 unit) of Haramaya woreda, Eastern Hararghe Zone, Oromia, Ethiopia (Figure 1) were  
136 enrolled at birth and followed until approximately 13 months of age. Written informed  
137 consent was obtained from both parents of each participating child in the local language

138 (Afan Oromo). Household information, including demographics, livelihoods and wealth,  
139 livestock ownership, and child health and nutrition status, was collected through  
140 household surveys at baseline and endline, along with short surveys conducted  
141 monthly. During the study period, child stool samples were collected monthly and tested  
142 for *Campylobacter* species. In addition, fecal samples from mothers and siblings of the  
143 enrolled children, feces from livestock (i.e., chicken, cattle, goat, and sheep), and  
144 environmental samples (soil and drinking water) were collected biannually. DNA  
145 extraction, genus-specific Taqman real-time PCR, and species-specific Sybr Green  
146 real-time PCR were performed afterwards to detect, quantify, and characterize  
147 *Campylobacter* spp. in these samples <sup>37</sup>.

#### 148 **Outcome variable**

149 We calculated the cumulative burden of *Campylobacter* infection for each  
150 enrolled child as the outcome variable of interest derived from the Ct values of their  
151 fecal samples collected over time. A genus-specific standard curve for expected  
152 bacterial load (log genome copies per 50 ng of DNA) against the Ct values was first  
153 generated using the 16S Taqman approach <sup>38</sup>. One ml normalized bacterial culture  
154 cocktail (including *Campylobacter jejuni*, *Campylobacter coli*, *Campylobacter*  
155 *hyointestinalis*, *Campylobacter lari* and *Campylobacter fetus*) was used for DNA  
156 extraction, and 2 µl extracted DNA for qPCR afterwards. Nuclease free water and  
157 *Salmonella* genomic DNA were used as negative controls. We tested up to 10  
158 *Campylobacter* DNA concentrations and repeated the experiment three times. Then the  
159 Ct values of the tested stool samples were converted to the expected bacterial load  
160 using this standard curve. The cumulative *Campylobacter* burden was defined as the  
161 average of the expected bacterial loads from available stool samples for each child <sup>37</sup>.



## 162 **Explanatory Variables**

### 163 **Household surveys**

164 Human and animal data were derived from household surveys. Demographic  
165 data including child sex, mother's age, and mother's education level were selected from  
166 the baseline survey, along with ownership of livestock (i.e., cattle, goat, sheep, and  
167 chicken) and assets (Table 1). To quantify all the livestock kept in a household, a  
168 composite metric, tropical livestock unit (TLU) was calculated based on the number of  
169 each species of livestock recorded in the baseline household survey<sup>39</sup>.

170 In the monthly short surveys, we selected variables that reflect interactions  
171 between the target child and animals or the environment (e.g., physical contact with  
172 animals, crawling in areas with animal droppings, and mouthing of soil or animal feces),  
173 diet and nutrition (e.g., consumption of animal source food [ASF]), feeding practices  
174 (e.g., pre-lacteal feeding, introduction of complementary foods, and consumption of any  
175 solid food in the past 24 h), and use of antibiotics (i.e., being treated with antibiotics in  
176 the past month) (Table 2). In addition, two composite variables, minimum dietary  
177 diversity (MDD) and household food insecurity access score (HFIAS), were generated  
178 according to their respective literature<sup>40,41</sup>.

### 179 **Environmental variables**

180 Thirteen environmental covariates used for ecological niche modeling of the  
181 genus *Campylobacter* in a previous analysis<sup>42</sup> were also included in this study (Table  
182 3). Corresponding values in grids within which the 106 enrolled households were  
183 located were extracted from the raster layer of each environmental covariate using the  
184 *raster* package<sup>43</sup> in R.

## 185 **Statistical Analysis**

186           Given that feeding practices (exclusive breastfeeding vs. complementary  
187 feeding) and the motor ability of infants are most often quite different between the first  
188 and second half of infants' first year of life, related risk factors such as probability of  
189 exposure to livestock and contaminated environments are likely to dynamically change  
190 over time. Child age is also an important confounding factor for *Campylobacter*  
191 infections<sup>4,39</sup>. To be consistent with our previous analysis about the prevalence of  
192 *Campylobacter* by age groups<sup>42</sup>, we split the whole study period into two parts, with a  
193 cut-off of 177 days of child age, which reflects the boundary between the first two and  
194 second two age groups classified in that analysis.

195           The outcome variable, cumulative *Campylobacter* burden, and all explanatory  
196 variables derived from the short household surveys were calculated separately for the  
197 two periods. We took the average of the monthly short survey data for each selected  
198 short survey variable, which resulted in an average value for the numeric variables or a  
199 proportion of being "1" for the binary variables (ordinal variable MDD was dichotomized  
200 using a generally accepted cut-off of  $< 5$  and  $\geq 5$ ) during each period. Given the  
201 relatively small sample size, a median split approach was applied to dichotomize all  
202 continuous variables to improve the model robustness following a previous study<sup>44</sup>.

203           The purposeful selection of covariates approach was employed in this study to  
204 choose the candidate variables and determine which to include in the final model<sup>45,46</sup>.  
205 The likelihood ratio test from logistic regression was performed for each explanatory  
206 variable with the outcome variable. The univariate analysis was conducted for the two  
207 time periods separately with the same pool of explanatory variables. For the first period,  
208 two variables, consumption of ASF in the past 24 h and MDD, were excluded from the

209 univariate analysis, as these questions were not asked until a child reached six months  
210 of age. Given 105 of 106 children didn't have vitamin-A supplements in the first period,  
211 the variable vitamin-A supplementation was also excluded from the first period  
212 univariate analysis. Any variables with a  $p$ -value less than 0.25 were selected as  
213 candidates for the following multivariate analysis.

214 A multivariate logistic regression model was built first with all candidate variables,  
215 followed by an iterative variable selection process. In each iteration, the covariate with  
216 the maximum  $p$ -value in the model was identified first. If the covariate was not  
217 significant at the 0.1 alpha level, it was removed from the model temporarily, and a  
218 reduced model was fit to compare the change of coefficients ( $\Delta\beta$ ) without that covariate.  
219 If the maximum  $\Delta\beta$  was greater than 20%, which indicates a potential confounding  
220 effect of the excluded covariate with one or more of the remaining variables in the  
221 model, the excluded covariate was added back to the model, and the next covariate  
222 with the maximum  $p$ -value was evaluated. Otherwise, this temporarily excluded  
223 covariate was deleted from the pool of candidates, and a new multivariate model was  
224 built to start another round of evaluations. After multiple iterations, the final model with  
225 significant covariates and potential confounders was obtained.

226 Nagelkerke  $R^2$ <sup>47</sup>, a so-called pseudo- $R^2$  measure, and Hosmer & Lemeshow  
227 goodness-of-fit test<sup>48</sup> were used to test the model fit. For a sample size less than 200  
228 and percentage of success in the outcome variable ranging from 38 – 62%, benchmark  
229 values between 0.32 and 0.58 of Nagelkerke  $R^2$  indicate good fit of the model<sup>49</sup>. The  
230 Hosmer & Lemeshow test is a hypothesis test and evaluates if the expected event  
231 frequencies from the logistic regression model match the observed event frequencies in

232 subgroups. The area under the receiver operating characteristic (ROC) curve (AUC)  
233 was used to assess the model performance in discriminating the positive results from  
234 the negatives<sup>50</sup>. AUC values range from 0 to 1, and, empirically, values between 0.7  
235 and 0.9 are a sign of good predictive performance. Values greater than 0.9 represent an  
236 excellent performance<sup>51</sup>.

237 We first fitted a multivariate model using the data collected from the first period  
238 and tested if it was a good fit for the second period data. If not, a separate model was  
239 built for the second period, and effects of the covariates between two periods were  
240 evaluated. All the statistical analyses were performed using R version 4.1.1<sup>52</sup>.

#### 241 **Spatial Autocorrelation Test**

242 Spatial autocorrelation refers to the correlation within variables across different  
243 spatial units<sup>53</sup>. If the values of a particular variable in nearby locations tend to be  
244 similar, a positive spatial autocorrelation exists in this variable, while negative spatial  
245 autocorrelation occurs when the values of a variable are more dissimilar than expected  
246 with their spatial neighbors. If spatial autocorrelation exists in the residuals of a  
247 regression model, it violates the assumption of independent errors and may result in an  
248 underestimation of the standard errors of the coefficient estimates of the model<sup>54</sup>. To  
249 test the spatial autocorrelation in the residuals of the logistic models fit in this study, we  
250 performed Moran's I test on the residuals using *spdep* package in R<sup>55</sup>.

### 251 **Results**

#### 252 **Demographics and Socio-economic Status**

253 Among the 106 enrolled children, 51.9% were male and 48.1% were female  
254 (Table 1). The mother's age at baseline ranged from 17 to 43 years, with a median age  
255 of 27 years. A high proportion of mothers reported not attending school at any level;

256 only 28% had some primary education. For livestock ownership, 49%, 57%, 45%, and  
257 50% of households owned cattle, goat, sheep, and chicken, respectively. Accordingly,  
258 the indicator of tropical livestock units (TLU) kept by the household ranged from 0 to  
259 5.54, with a median value of 0.62.

## 260 **Cumulative *Campylobacter* Burden**

261 The calculated cumulative *Campylobacter* burden for infants in the first period  
262 ranged from 0.77 to 3.75 log genome copies per 50 ng of DNA with a median value of  
263 2.10, while the minimum and maximum of cumulative *Campylobacter* burden for the  
264 second period were 1.95 and 5.06 log genome copies per 50 ng of DNA, respectively.  
265 The *Campylobacter* burden in the second half year of life ( $M = 3.52$ ,  $SD = 0.75$ ) was  
266 significantly higher than in the first half year of life ( $M = 2.14$ ,  $SD = 0.62$ );  $t(210) = 14.7$ ,  
267  $p < .01$  (Figure 2).

## 268 **Univariate Analysis**

269 Based on the  $p$ -value of likelihood ratio test with a cut-off of 0.25, 13 and 8  
270 variables were selected as candidate variables for the first and second period,  
271 respectively (Table 2). Among the candidates, three variables, namely, cattle  
272 ownership, sheep ownership, and consumption of any solid food in the past 24h, were  
273 included in the multivariate analysis for both periods. Though child sex was not  
274 significant at the alpha level of 0.25 in the first period, we still included it in the  
275 multivariate analysis given the confounding effect of child sex in *Campylobacter*  
276 infection reported in previous studies<sup>4,7,39,56</sup>.

## 277 **Multivariate Analysis**

278 *Period One*

279           The final logistic regression model for the first period showed that contact with  
280 animals, mouthing of soil or animal feces, and drinking from a bottle with nipple in the  
281 past 24h were statistically significant at the 0.05 level (Table 4). The direction of the  
282 estimated coefficients indicated children in the first period who had more physical  
283 contact with animals and more mouthing of soil or animal feces had greater odds of high  
284 cumulative *Campylobacter* burden, while drinking from a bottle with nipple was found to  
285 be a protective factor for high *Campylobacter* burden. Sheep ownership was marginally  
286 significant in the model and had a negative coefficient estimate. Through the purposeful  
287 selection of covariates process, child sex, pre-lacteal feeding, elevation, and soil pH  
288 were identified as potential confounding factors for *Campylobacter* burden.

289           The Nagelkerke  $R^2$  of the logistic regression model was 0.40, which fell into the  
290 benchmark values (0.32 – 0.58) for good model fit (Table 4). Hosmer & Lemeshow test  
291 also showed there was no evidence of poor model fit. The model also had a relatively  
292 good predictive performance with an AUC of 0.79 (95% CI: 0.70 – 0.88) (Figure 3A).  
293 Moran's I statistics showed that no spatial autocorrelation existed in the model  
294 residuals, suggesting there was no need to consider a spatial regression model to  
295 account for the spatial autocorrelation.

296           To test if one single model can be obtained for both periods, the second period  
297 data were plugged into the model for the first period as a testing data set. Hosmer &  
298 Lemeshow test checked the difference between observed data and predicted values  
299 and showed this difference was significant ( $p = 0.00$ ), suggesting a poor fit of the model  
300 to the second period data. It was also supported by a low AUC of 0.57 (95% CI: 0.46 –  
301 0.69) (Figure 3B).

302 *Period two*

303 Logistic regression models were then fitted using the candidate variables  
304 selected from the univariate analysis for the second period, and the final model included  
305 eight variables with child sex, cattle ownership, and population density (per 100m<sup>2</sup>)  
306 being statistically significant at the 0.05 level (Table 5). Children being female and living  
307 in households that kept cattle had greater odds of high cumulative *Campylobacter*  
308 burden, while population density showed protective effects on *Campylobacter* burden.  
309 Household food insecurity access score, proportion of clay particles in soil, consumption  
310 of any solid food in the past 24h, and sheep ownership were identified as potential  
311 confounding factors for *Campylobacter* burden in the second period.

312 The Nagelkerke R<sup>2</sup> for the second model was 0.34, and Hosmer & Lemeshow  
313 test showed there was no evidence of poor model fit (Table 5). This model had a good  
314 predictive performance with an AUC of 0.80 (95% CI: 0.71 – 0.89) (Figure 3C). Again,  
315 there was no spatial autocorrelation existing in the model residuals.

## 316 **Discussion**

317 From a One Health perspective, this study identified potential factors involved in  
318 the *Campylobacter* transmission pathways between humans, animals, and the  
319 environment in a low-resource setting. Child age-specific behaviors and other factors  
320 that may increase the risk of children's exposure to animal feces are largely missing  
321 from previous interventions that aimed to mitigate the burden of *Campylobacter*  
322 infection or other enteric illnesses among infants and young children in LMICs<sup>57</sup>. Risk  
323 factors of *Campylobacter* infections that have been commonly identified in previous  
324 studies conducted among CU5 in low-resource settings are primarily limited to maternal

325 educational level, feeding practices, WaSH indicators, and ownership of domestic  
326 animals<sup>4,58-60</sup>. However, fewer studies have delved into infant-specific behaviors and  
327 the interactions between infants and livestock, which could also contribute to the risk of  
328 *Campylobacter* infections<sup>61</sup>. Considering the dynamic effects of age, particularly  
329 distinguishing between infants aged less than six months and those aged six to twelve  
330 months, on biological risk and environmental interaction, our study introduces a novel  
331 approach. By fitting models and identifying risk factors separately for the first and  
332 second six months of infancy, we aim to capture the nuanced variations in infection  
333 dynamics during these critical developmental stages. This innovative methodology,  
334 coupled with the exploration of previously overlooked factors using longitudinal data,  
335 sets our study apart from existing research on *Campylobacter* infection risk factors.

336         At an early stage of life, infants and young children put non-food objects in their  
337 mouths as one way to explore the surrounding environment<sup>62</sup>. However, this  
338 exploratory behavior could put children at higher risk of contracting zoonotic pathogens  
339 if they live in an environment contaminated by livestock and poultry feces. Our results  
340 identify mouthing of soil or animal feces as a significant risk factor contributing to a  
341 higher cumulative *Campylobacter* burden during the first half of the first year of life. This  
342 association can be attributed to the high prevalence of *Campylobacter* spp. in the feces  
343 of animals, including domestic livestock and poultry. Results from the laboratory showed  
344 that the prevalence of *Campylobacter* in the fecal samples of cattle, sheep, goats, and  
345 chicken collected from the CAGED longitudinal study were 99%, 98%, 99%, and 93%,  
346 respectively<sup>37</sup>. In rural areas of LMICs, infants and young children are frequently placed  
347 on the ground sharing space with free-ranging livestock<sup>57</sup>. As animals defecate in the



348 homestead, or sometimes even inside the homes, enteric pathogens harbored in  
349 livestock feces or the soil contaminated by the feces may be ingested by infants through  
350 their routine mouthing behaviors while they are on the floor. Ingestion of soil and  
351 livestock feces is therefore a point of exposure specific to infant and young child  
352 behavior. This infant-specific transmission pathway needs to be considered in the future  
353 design of intervention strategies for preventing children from fecal exposure. This  
354 finding echoes the collective conclusion from the WASH Benefits and the Sanitation,  
355 Hygiene, Infant Nutrition Efficacy (SHINE) trials; that more effective interventions are  
356 needed to reduce the exposure to fecal contaminations in the domestic environments  
357 other than the traditional WaSH interventions<sup>63</sup>.

358 Similarly, physical contact with animals was identified as another risk factor  
359 existing in the child-livestock interaction in the first period (before six months of age)  
360 that increases the exposure of children to animal feces. It is very common that people in  
361 rural areas share living and sleeping quarters with their livestock, which frequently  
362 exposes young children in the household to direct contact with livestock. Keeping  
363 animals inside human living spaces has been associated with *Campylobacter*-positive  
364 child stools in another study in rural Ethiopia<sup>57</sup>, highlighting the need to design  
365 interventions creating adequate separation of domestic animals from the human living  
366 spaces to block the transmission pathway through direct contact.

367 The risks and benefits of raising livestock in smallholder families on child health  
368 are intertwined<sup>23</sup>. On one hand, livestock raised in the household can provide animal  
369 source foods, which are seen as the best source of nutrition-rich food for infants and  
370 young children. Livestock production is also a source of household income. Increased

371 income from livestock sales can grant families more purchasing power for food, which  
372 further helps improve child nutritional status. However, livestock ownership can  
373 potentially increase the risk of children's exposure to animal feces, as discussed above.  
374 In this study, we didn't find a significant association between consumption of animal  
375 source food and *Campylobacter* burden, but cattle ownership was identified as a  
376 significant factor associated with increased odds of having higher cumulative  
377 *Campylobacter* burden during the second half of the first year of life. In our study  
378 population, when cattle were kept either inside the homes or on the homestead but  
379 outside the home during the day, 100% (27/27) and 93.3% (28/30) of households,  
380 respectively, confined those cattle. During the night, almost all cattle were kept and tied  
381 inside the home. Regarding the purposes of raising cattle, households reported selling  
382 livestock for income (44.2%; 23/52) and consumption of milk (51.9%; 27/52) as the two  
383 major purposes. These data support other findings indicating a strong cultural focus on  
384 milk consumption in this area, which may correspond with an underlying risk in the  
385 consumption of raw/unpasteurized milk which has been associated with reported  
386 outbreaks of campylobacteriosis in HICs<sup>64,65</sup>. A recent Ethiopian study showed a higher  
387 prevalence of *C. jejuni* (16%) in raw milk compared to other dairy products<sup>66</sup>. Though  
388 consumption of raw milk was not identified in our models as a significant risk factor of  
389 *Campylobacter* burden in our study population, further work is still needed to unpack  
390 this association among children in low-resource settings.

391         Appropriate infant and young child feeding practices could improve child  
392 nutritional status, growth, and development<sup>67</sup>. For infant aged from 0 – 5 months,  
393 exclusive breastfeeding is strongly recommended by WHO, while, bottle feeding, using

394 a bottle with a nipple/teat to feed any liquid or semi-solid food, is discouraged at this  
395 early stage of life due to global concerns that include excessive weight gain, iron  
396 depletion, etc. <sup>68</sup>. In addition, a bottle with a nipple is more likely to be contaminated in  
397 low-resource settings where inadequate cleaning and disinfection of bottles is more  
398 common, which increases the risk of enteric infections <sup>69</sup>. Lengerh et al. showed that  
399 bottle feeding was associated with increased odds of *Campylobacter* infection among  
400 diarrheic children in northwest Ethiopia <sup>60</sup>. However, in this study, drinking from a bottle  
401 with nipple was shown to be a protective factor for higher cumulative *Campylobacter*  
402 burden in the first time period. Our results should not be interpreted as an  
403 encouragement of bottle feeding. One potential explanation could be that this practice is  
404 linked to other socio-economic factors that contribute to reducing the risk of  
405 *Campylobacter* infections. A previous study using Ethiopian Demographic and Health  
406 Surveys data (EDHS) to examine the determinants of bottle feeding suggested that  
407 women who had a higher education, came from a richer household, and lived in urban  
408 areas were more likely to bottle feed <sup>68</sup>. The role of mother's education level and  
409 household wealth status in *Campylobacter* burden was not clear in our analysis with a  
410 relatively small sample size, and further work is needed to unpack these potential links.  
411 Influences of environmental and climatic factors have been rarely reported in  
412 studies investigating risk factors of *Campylobacter* infections among children in LMICs.  
413 Here, we included several environmental covariates previously used to model the  
414 prevalence of enteric diseases globally and regionally. Due to the small geography of  
415 our study area, most environmental covariates have less heterogeneity to show signals  
416 in their association with *Campylobacter* infections. This is one of the limitations in this

417 study. Evaluating these environmental effects on *Campylobacter* infections at a larger  
418 scale in LMICs would be a future direction.

419 We tested the spatial autocorrelation in the model residuals during the model-  
420 building process, a step that has seldom been taken in previous work on identifying risk  
421 factors of *Campylobacter* infections. We did so not only because we included  
422 environmental covariates in the models, but to consider the potential spatial effects (i.e.,  
423 spatial dependence and spatial heterogeneity) introduced by georeferenced data <sup>70</sup>.  
424 Neglecting the spatial effects could lead to an inflation of variance in regression  
425 estimates and, consequently, a less reliable regression model <sup>71</sup>. Therefore, it would be  
426 appropriate for future studies to consider including the spatial autocorrelation test as a  
427 core component of the modeling-building process to ensure the reliability and accuracy  
428 of the regression model.

429 Another innovative aspect of this study involves the adoption of bacterial load-  
430 based cumulative burden, departing from traditional prevalence measures when dealing  
431 with longitudinal data. By calculating the cumulative burden, we gain insight into the  
432 persistent impact of *Campylobacter* infections over time, offering a nuanced perspective  
433 on disease dynamics. This method enables a comprehensive assessment of  
434 *Campylobacter* burden over a certain period and holds promise for advancing  
435 longitudinal studies on *Campylobacter* infections.

436 In conclusion, our study reveals that factors involved in the interactions between  
437 human (infants), livestock, and home environment impacted the presence of higher  
438 cumulative *Campylobacter* burden among infants in eastern Ethiopia. For young infants  
439 aged less than 6 months, being reported to have physical contact with animals and

440 have mouthing of soil or animal feces were identified as risk factors of higher  
441 *Campylobacter* burden. And drinking from a bottle with nipple was shown to be a  
442 protective factor. This result requires additional research to understand and should not  
443 be interpreted to encourage bottle feeding; additional information is required to  
444 understand whether there is a direct causal mechanism or whether underlying factors or  
445 confounders, such as socio-economic status or overall household hygiene, might  
446 explain the finding. While older infants (aged between 6 and 12 months) being female  
447 and living in households with cattle had increased odds of higher *Campylobacter*  
448 burden. High population density (potentially linked to urban residency) was identified as  
449 a protective factor for this age group. Future interventions should pay more attention to  
450 the infant-specific transmission pathway and create adequate separation of domestic  
451 animals from humans to prevent infants and young children from potential fecal  
452 exposures.

### 453 **Acknowledgments**

454 This work is a result of the CAGED Research Team whose members include: Amanda  
455 Evelyn Ojeda, Arie H. Havelaar, Abadir Jemal Seran, Abdulmuen Mohammed Ibrahim,  
456 Bahar Mummed Hassen, Belisa Usmael Ahmedo, Cyrus Saleem, Dehao Chen, Efraim  
457 Ali Yusuf, Gireesh Rajashekara, Getnet Yimer, Ibsa A. Ahmed, Ibsa Aliyi Usmane,  
458 Jafer Kedir Amin, Jason K. Blackburn, Jemal Y. Hassen, Kedir A. Hassen, Kunuza  
459 Adem Umer, Karah Mechlowitz, Kedir Teji Roba, Loic Deblais, Mussie Bhrane, Mark J.  
460 Manary, Mawardi M. Dawid, Mahammad Mahammad Usmail, Nigel P. French, Nur  
461 Shaikh, Nitya Singh, Sarah L. McKune, Wondwossen A. Gebreyes, Xiaolong Li,  
462 Yenenesh Demisie Weldesenbet, Yang Yang, Zelalem Hailu Mekuria.

463 **Financial Support**

464 This project is funded by the United States Agency for International Development  
465 Bureau for Food Security under Agreement #AID-OAA-L-15-00003 as part of Feed the  
466 Future Innovation Lab for Livestock Systems, and by the Bill & Melinda Gates  
467 Foundation OPP#1175487. Under the grant conditions of the Foundation, a Creative  
468 Commons Attribution 4.0 Generic License has already been assigned to the Author  
469 Accepted Manuscript version that might arise from this submission. Any opinions,  
470 findings, conclusions or recommendations expressed here are those of the authors  
471 alone. Research reported in this publication was supported by the University of Florida  
472 Clinical and Translational Science Institute, which was supported in part by the NIH  
473 National Center for Advancing Translational Sciences under award number  
474 UL1TR001427.

475 **Authors contributions**

476 XL, SL, SLM, AHH and JKB conceptualized and designed the study; XL and DC curated  
477 and analyzed data; JYH, SLM, and AHH managed and coordinated the research activity  
478 in this project; XL drafted the initial manuscript; and SL, DC, SLM, AHH, JKB reviewed  
479 and critically revised the manuscript. All authors approved the final manuscript as  
480 submitted and agreed to be accountable for all aspects of the work.

481 **Authors contact information**

482 Xiaolong Li, Department of Environmental and Global Health, College of Public Health  
483 and Health Professions, University of Florida, Gainesville, FL, Email:  
484 [xiaolong.li@ufl.edu](mailto:xiaolong.li@ufl.edu); Dehao Chen, Department of Environmental and Global Health,  
485 College of Public Health and Health Professions, University of Florida, Gainesville, FL,  
486 Email: [dehao.chen@emory.edu](mailto:dehao.chen@emory.edu); Song Liang, Department of Environmental and Global

487 Health, College of Public Health and Health Professions, University of Florida,  
488 Gainesville, FL, Email: [songliang@umass.edu](mailto:songliang@umass.edu); Jemal Yousuf Hassen, Haramaya  
489 University, Dire Dawa, Ethiopia, Email: [jemaly2001@yahoo.com](mailto:jemaly2001@yahoo.com); Sarah L. McKune,  
490 Department of Environmental and Global Health, College of Public Health and Health  
491 Professions, University of Florida, Gainesville, FL, Email: [smckune@ufl.edu](mailto:smckune@ufl.edu); Arie H.  
492 Havelaar, Department of Animal Sciences, Institute of Food and Agricultural Sciences,  
493 University of Florida, Gainesville, FL, Email: [ariehavelaar@ufl.edu](mailto:ariehavelaar@ufl.edu); Jason K. Blackburn,  
494 Spatial Epidemiology and Ecology Research Laboratory, Department of Geography,  
495 College of Liberal Arts and Sciences, University of Florida, Gainesville, FL, Email:  
496 [jkblackburn@ufl.edu](mailto:jkblackburn@ufl.edu).

#### 497 **Availability of data and materials**

498 Deidentified individual participant data will be made available through Dataverse  
499 (<https://dataverse.org/>) after December 31, 2024.

#### 500 **References**

- 501 1. Sahin O, Fitzgerald C, Stroika S, Zhao S, Sippy RJ, Kwan P, Plummer PJ, Han J,  
502 Yaeger MJ, Zhang Q., 2012. Molecular Evidence for Zoonotic Transmission of an  
503 Emergent, Highly Pathogenic *Campylobacter jejuni* Clone in the United States. *J*  
504 *Clin Microbiol* 50: 680
- 505 2. Chlebicz A, Śliżewska K., 2018. Campylobacteriosis, Salmonellosis, Yersiniosis,  
506 and Listeriosis as Zoonotic Foodborne Diseases: A Review. *Int J Environ Res*  
507 *Public Health* 15
- 508 3. Kaakoush NO, Castaño-Rodríguez N, Mitchell HM, Man SM., 2015. Global  
509 epidemiology of campylobacter infection. *Clin Microbiol Rev* 28: 687–720
- 510 4. Amour C, Gratz J, Mduma ER, Svensen E, Rogawski ET, McGrath M, Seidman

- 511 JC, McCormick BJJ, Shrestha S, Samie A, Mahfuz M, Qureshi S, Hotwani A,  
512 Babji S, Trigoso DR, et al., 2016. Epidemiology and Impact of Campylobacter  
513 Infection in Children in 8 Low-Resource Settings: Results From the MAL-ED  
514 Study. *Clin Infect Dis* 63: 1171–1179
- 515 5. Platts-Mills JA, Kosek M., 2014. Update on the burden of Campylobacter in  
516 developing countries
- 517 6. Lee G, Pan W, Peñataro Yori P, Paredes Olortegui M, Tilley D, Gregory M,  
518 Oberhelman R, Burga R, Chavez CB, Kosek M., 2013. Symptomatic and  
519 Asymptomatic Campylobacter Infections Associated with Reduced Growth in  
520 Peruvian Children. *PLoS Negl Trop Dis*
- 521 7. Haque MA, Platts-Mills JA, Mduma E, Bodhidatta L, Bessong P, Shakoor S, Kang  
522 G, Kosek MN, Lima AAM, Shrestha SK, Alam MA, Havt A, Samie A, Guerrant RL,  
523 Lang D, et al., 2019. Determinants of Campylobacter infection and association  
524 with growth and enteric inflammation in children under 2 years of age in low-  
525 resource settings. *Sci Rep* 9
- 526 8. François R, Yori PP, Rouhani S, Siguas Salas M, Paredes Olortegui M, Rengifo  
527 Trigoso D, Pisanic N, Burga R, Meza R, Meza Sanchez G, Gregory MJ, Houpt  
528 ER, Platts-Mills JA, Kosek MN., 2018. The other Campylobacters: Not innocent  
529 bystanders in endemic diarrhea and dysentery in children in low-income settings.  
530 *PLoS Negl Trop Dis* 12: e0006200
- 531 9. Man SM., 2011. The clinical importance of emerging Campylobacter species. *Nat*  
532 *Rev Gastroenterol Hepatol* 2011 8: 669–685
- 533 10. Terefe Y, Deblais L, Ghanem M, Helmy YA, Mummmed B, Chen D, Singh N,



- 534 Ahyong V, Kalantar K, Yimer G, Yousuf Hassen J, Mohammed A, McKune SL,  
535 Manary MJ, Ordiz MI, et al., 2020. Co-occurrence of Campylobacter Species in  
536 Children From Eastern Ethiopia, and Their Association With Environmental  
537 Enteric Dysfunction, Diarrhea, and Host Microbiome. *Front Public Heal* 8: 99
- 538 11. Bian X, Garber JM, Cooper KK, Huynh S, Jones J, Mills MK, Rafala D, Nasrin D,  
539 Kotloff KL, Parker CT, Tennant SM, Miller WG, Szymanski CM., 2020.  
540 Campylobacter Abundance in Breastfed Infants and Identification of a New  
541 Species in the Global Enterics Multicenter Study. *mSphere* 5: e00735-19
- 542 12. Parker CT, Schiaffino F, Huynh S, Paredes Olortegui M, Peñataro Yori P, Garcia  
543 Bardales PF, Vasquez TP, Curico Huansi GE, Villanueva KM, Shapiama Lopez W  
544 V., Cooper KK, Kosek MN., 2022. Shotgun metagenomics of fecal samples from  
545 children in Peru reveals frequent complex co-infections with multiple  
546 Campylobacter species. *PLoS Negl Trop Dis* 16: 1–17
- 547 13. Garcia Bardales PF, Schiaffino F, Huynh S, Olortegui MP, Yori PP, Vasquez TP,  
548 Villanueva KM, Curico Huansi GE, Shapiama Lopez W V., Cooper KK, Parker CT,  
549 Kosek MN., 2022. “Candidatus Campylobacter infans” detection is not associated  
550 with diarrhea in children under the age of 2 in Peru. *PLoS Negl Trop Dis* 16: 1–12
- 551 14. WAGNER EG, LANOIX JN., 1958. Excreta disposal for rural areas and small  
552 communities. *Monogr Ser World Health Organ*
- 553 15. Penakalapati G, Swarthout J, Delahoy MJ, McAliley L, Wodnik B, Levy K,  
554 Freeman MC., 2017. Exposure to Animal Feces and Human Health: A Systematic  
555 Review and Proposed Research Priorities. *Environ Sci Technol* 51: 11537–11552
- 556 16. Sclar GD, Penakalapati G, Amato HK, Garn J V., Alexander K, Freeman MC,

- 557 Boisson S, Medlicott KO, Clasen T., 2016. Assessing the impact of sanitation on  
558 indicators of fecal exposure along principal transmission pathways: A systematic  
559 review
- 560 17. Dangour AD, Watson L, Cumming O, Boisson S, Che Y, Velleman Y, Cavill S,  
561 Allen E, Uauy R., 2013. Interventions to improve water quality and supply,  
562 sanitation and hygiene practices, and their effects on the nutritional status of  
563 children
- 564 18. Wolf J, Hunter PR, Freeman MC, Cumming O, Clasen T, Bartram J, Higgins JPT,  
565 Johnston R, Medlicott K, Boisson S, Prüss-Ustün A., 2018. Impact of drinking  
566 water, sanitation and handwashing with soap on childhood diarrhoeal disease:  
567 updated meta-analysis and meta-regression
- 568 19. Freeman MC, Garn J V., Sclar GD, Boisson S, Medlicott K, Alexander KT,  
569 Penakalapati G, Anderson D, Mahtani AG, Grimes JET, Rehfuess EA, Clasen  
570 TF., 2017. The impact of sanitation on infectious disease and nutritional status: A  
571 systematic review and meta-analysis
- 572 20. Clasen T, Boisson S, Routray P, Torondel B, Bell M, Cumming O, Ensink J,  
573 Freeman M, Jenkins M, Odagiri M, Ray S, Sinha A, Suar M, Schmidt WP., 2014.  
574 Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted  
575 helminth infection, and child malnutrition in Odisha, India: A cluster-randomised  
576 trial. *Lancet Glob Heal*
- 577 21. Null C, Stewart CP, Pickering AJ, Dentz HN, Arnold BF, Arnold CD, Benjamin-  
578 Chung J, Clasen T, Dewey KG, Fernald LCH, Hubbard AE, Kariger P, Lin A, Luby  
579 SP, Mertens A, et al., 2018. Effects of water quality, sanitation, handwashing, and

- 580 nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-  
581 randomised controlled trial. *Lancet Glob Heal* 6: e316–e329
- 582 22. Pickering AJ, Djebbari H, Lopez C, Coulibaly M, Alzua ML., 2015. Effect of a  
583 community-led sanitation intervention on child diarrhoea and child growth in rural  
584 Mali: A cluster-randomised controlled trial. *Lancet Glob Heal*
- 585 23. Chen D, Mechlowitz K, Li X, Schaefer N, Havelaar AH, McKune SL., 2021.  
586 Benefits and Risks of Smallholder Livestock Production on Child Nutrition in Low-  
587 and Middle-Income Countries. *Front Nutr* 8
- 588 24. Hanson LÅ, Wiedermann U, Ashraf R, Zaman S, Adlerberth I, Dahlgren U, Wold  
589 A, Jalil F., 1996. Effects of breastfeeding on the baby and on its immune system.  
590 *Food Nutr Bull* 17: 384–389
- 591 25. Oriá RB, Murray-Kolb LE, Scharf RJ, Pendergast LL, Lang DR, Kolling GL,  
592 Guerrant RL., 2016. Early-life enteric infections: Relation between chronic  
593 systemic inflammation and poor cognition in children. *Nutr Rev* 74: 374–386
- 594 26. Gibson E., 1988. Exploratory Behavior In The Development Of Perceiving, Acting,  
595 And The Acquiring Of Knowledge. *Annu Rev Psychol* 39: 1–41
- 596 27. Bronowski C, James CE, Winstanley C., 2014. Role of environmental survival in  
597 transmission of *Campylobacter jejuni*. *FEMS Microbiol Lett* 356: 8–19
- 598 28. Whiley H, van den Akker B, Giglio S, Bentham R., 2013. The Role of  
599 Environmental Reservoirs in Human Campylobacteriosis. *Int J Environ Res Public*  
600 *Health* 10: 5886
- 601 29. Arsenault J, Berke O, Michel P, Ravel A, Gosselin P., 2012. Environmental and  
602 demographic risk factors for campylobacteriosis: Do various geographical scales

- 603 tell the same story? *BMC Infect Dis* 12: 1–12
- 604 30. Louis VR, Gillespie IA, O'Brien SJ, Russek-Cohen E, Pearson AD, Colwell RR.,  
605 2005. Temperature-Driven *Campylobacter* Seasonality in England and Wales.  
606 *Appl Environ Microbiol* 71: 85–92
- 607 31. Kuhn KG, Nygård KM, Guzman-Herrador B, Sunde LS, Rimhanen-Finne R,  
608 Trönnberg L, Jepsen MR, Ruuhela R, Wong WK, Ethelberg S., 2020.  
609 *Campylobacter* infections expected to increase due to climate change in Northern  
610 Europe. *Sci Rep* 10: 1–9
- 611 32. Soneja S, Jiang C, Romeo Upperman C, Murtugudde R, Mitchell CS, Blythe D,  
612 Sapkota AR, Sapkota A., 2016. Extreme precipitation events and increased risk of  
613 campylobacteriosis in Maryland, U.S.A. *Environ Res* 149: 216–221
- 614 33. Weisent J, Seaver W, Odoi A, Rohrbach B., 2014. The importance of climatic  
615 factors and outliers in predicting regional monthly campylobacteriosis risk in  
616 Georgia, USA. *Int J Biometeorol* 58: 1865–1878
- 617 34. Sanderson RA, Maas JA, Blain AP, Gorton R, Ward J, O'Brien SJ, Hunter PR,  
618 Rushton SP., 2018. Spatio-temporal models to determine association between  
619 *Campylobacter* cases and environment. *Int J Epidemiol* 47: 202–216
- 620 35. Spencer SEF, Marshall J, Pirie R, Campbell D, Baker MG, French NP., 2012. The  
621 spatial and temporal determinants of campylobacteriosis notifications in New  
622 Zealand, 2001-2007. *Epidemiol Infect* 140: 1663–1677
- 623 36. Havelaar AH, Brhane M, Ahmed IA, Kedir J, Chen D, Deblais L, French N,  
624 Gebreyes WA, Hassen JY, Li X, Manary MJ, Mekuria Z, Ibrahim AM, Mammed B,  
625 Ojeda A, et al., 2022. Unravelling the reservoirs for colonisation of infants with

- 626           Campylobacter spp. in rural Ethiopia: protocol for a longitudinal study during a  
627           global pandemic and political tensions. *BMJ Open* 12: 61311
- 628   37.   Deblais L, Ojeda A, Brhane M, Mummed B, Hassen KA, Ahmedo BU,  
629           Weldesebet YD, Amin JK, Ahmed IA, Usmane IA, Yusuf EA, Seran AJ, Abraham  
630           FI, Game HT, Mummed BA, et al., 2023. Prevalence and Load of the  
631           Campylobacter Genus in Infants and Associated Household Contacts in Rural  
632           Eastern Ethiopia: a Longitudinal Study from the Campylobacter Genomics and  
633           Environmental Enteric Dysfunction (CAGED) Project. *Appl Environ Microbiol* 89:  
634           e00424-23
- 635   38.   Platts-Mills JA, Liu J, Gratz J, Mduma E, Amour C, Swai N, Taniuchi M, Begum S,  
636           Yori PP, Tilley DH, Lee G, Shen Z, Whary MT, Fox JG, McGrath M, et al., 2014.  
637           Detection of Campylobacter in stool and determination of significance by culture,  
638           enzyme immunoassay, and PCR in developing countries. *J Clin Microbiol*
- 639   39.   Chen D, McKune SL, Singh N, Yousuf Hassen J, Gebreyes W, Manary MJ,  
640           Bardosh K, Yang Y, Diaz N, Mohammed A, Terefe Y, Roba KT, Ketema M,  
641           Ameha N, Assefa N, et al., 2021. Campylobacter Colonization, Environmental  
642           Enteric Dysfunction, Stunting, and Associated Risk Factors Among Young  
643           Children in Rural Ethiopia: A Cross-Sectional Study From the Campylobacter  
644           Genomics and Environmental Enteric Dysfunction (CAGED) Project. *Front Public*  
645           *Heal* 8: 1043
- 646   40.   World Health Organization., 2017. Global Nutrition Monitoring Framework:  
647           operational guidance for tracking progress in meeting targets for 2025. Available  
648           at: <https://apps.who.int/iris/bitstream/handle/10665/259904/9789241513609->

- 649 eng.pdf;jsessionid=5B7CD35139464EA9E9214B4F68A81B5E?sequence=1.
- 650 Accessed. 2017
- 651 41. Coates J, Swindale A, Bilinsky P., 2007. Household Food Insecurity Access Scale  
652 (HFIAS) for measurement of food access: indicator guide. *Washington, DC FHI*  
653 *360/FNATA*
- 654 42. Li X., 2023. *Applying Geospatial Approaches to Studying the Epidemiology of*  
655 *Enteric Illnesses at Different Scales*. University of Florida
- 656 43. Hijmans RJ., 2024. raster: Geographic analysis and modeling with raster data.  
657 Available at: <https://rspatial.org/raster>. Accessed. 2024
- 658 44. Chen D, Mckune SL, Yang Y, Usmane IA, Ahmed A, Amin JK, Ibrahim AM, Jemal  
659 A, Shaik N, Ojeda A, Hassen BM, Deblais L, Ahmedo BU, Hassen KA, Bhrane M,  
660 et al., 2024. Campylobacter colonization and undernutrition in infants in rural  
661 Eastern Ethiopia: a longitudinal community-based birth cohort study. *medRxiv*: 1–  
662 44
- 663 45. Bursac Z, Gauss CH, Williams DK, Hosmer DW., 2008. Purposeful selection of  
664 variables in logistic regression. *Source Code Biol Med* 3: 17
- 665 46. Hosmer DW, Lemeshow S, Sturdivant RX., 2013. *Applied Logistic Regression:*  
666 *Third Edition*
- 667 47. Nagelkerke NJD., 1991. A note on a general definition of the coefficient of  
668 determination
- 669 48. Hosmer DW, Lemeshow S., 1980. Goodness of fit tests for the multiple logistic  
670 regression model. *Commun Stat - Theory Methods* 9: 1043–1069
- 671 49. Hemmert GAJ, Schons LM, Wieseke J, Schimmelpfennig H., 2018. Log-

- 672 likelihood-based Pseudo-R<sup>2</sup> in Logistic Regression: Deriving Sample-sensitive  
673 Benchmarks. *Sociol Methods Res* 47: 507–531
- 674 50. Hanley JA, McNeil BJ., 1982. The meaning and use of the area under a receiver  
675 operating characteristic (ROC) curve. *Radiology* 143: 29–36
- 676 51. Li H, Chen Y, Deng S, Chen M, Fang T, Tan H., 2019. Eigenvector Spatial  
677 Filtering-Based Logistic Regression for Landslide Susceptibility Assessment.  
678 *ISPRS Int J Geo-Information* 8: 332
- 679 52. R Core Team., 2021. R: A language and environment for statistical computing. R  
680 Foundation for Statistical Computing, Vienna, Austria.
- 681 53. Getis A., 2007. Reflections on spatial autocorrelation. *Reg Sci Urban Econ* 37:  
682 491–6
- 683 54. Lichstein JW, Simons TR, Shiner SA, Franzreb KE., 2002. Spatial autocorrelation  
684 and autoregressive models in ecology. *Ecol Monogr* 72: 445–463
- 685 55. Bivand R., 2022. R Packages for Analyzing Spatial Data: A Comparative Case  
686 Study with Areal Data. *Geogr Anal* 54: 488–518
- 687 56. Strachan NJC, Watson RO, Novik V, Hofreuter D, Ogden ID, Galán JE., 2008.  
688 Sexual dimorphism in campylobacteriosis. *Epidemiol Infect* 136: 1492–5
- 689 57. Budge S, Barnett M, Hutchings P, Parker A, Tyrrel S, Hassard F, Garbutt C,  
690 Moges M, Woldemedhin F, Jemal M., 2020. Risk factors and transmission  
691 pathways associated with infant *Campylobacter* spp. Prevalence and malnutrition:  
692 A formative study in rural Ethiopia. *PLoS One* 15: 1–17
- 693 58. Haque MA, Platts-Mills JA, Mduma E, Bodhidatta L, Bessong P, Shakoob S, Kang  
694 G, Kosek MN, Lima AAM, Shrestha SK, Alam MA, Havt A, Samie A, Guerrant RL,

- 695 Lang D, et al., 2019. Determinants of Campylobacter infection and association  
696 with growth and enteric inflammation in children under 2 years of age in low-  
697 resource settings. *Sci Rep*
- 698 59. Diriba K, Awulachew E, Anja A., 2021. Prevalence and associated factor of  
699 Campylobacter species among less than 5-year-old children in Ethiopia: a  
700 systematic review and meta-analysis. *Eur J Med Res* 26: 2
- 701 60. Lengerh A, Moges F, Unakal C, Anagaw B., 2013. Prevalence, associated risk  
702 factors and antimicrobial susceptibility pattern of Campylobacter species among  
703 under five diarrheic children at Gondar University Hospital, Northwest Ethiopia.  
704 *BMC Pediatr* 13: 82
- 705 61. Budge S, Barnett M, Hutchings P, Parker A, Tyrrel S, Hassard F, Garbutt C,  
706 Moges M, Woldemedhin F, Jemal M., 2020. Risk factors and transmission  
707 pathways associated with infant Campylobacter spp. prevalence and malnutrition:  
708 A formative study in rural Ethiopia. *PLoS One* 15: e0232541
- 709 62. Ruff HA, Saltarelli LM, Capozzoli M, Dubiner K., 1992. The Differentiation of  
710 Activity in Infants' Exploration of Objects. *Dev Psychol* 28: 851–861
- 711 63. Pickering AJ, Null C, Winch PJ, Mangwadu G, Arnold BF, Prendergast AJ, Njenga  
712 SM, Rahman M, Ntozini R, Benjamin-Chung J, Stewart CP, Huda TMN, Moulton  
713 LH, Colford JM, Luby SP, et al., 2019. The WASH Benefits and SHINE trials:  
714 interpretation of WASH intervention effects on linear growth and diarrhoea. *Lancet*  
715 *Glob Heal* 7: e1139–e1146
- 716 64. Wulsten IF, Galeev A, Stingl K., 2020. Underestimated Survival of Campylobacter  
717 in Raw Milk Highlighted by Viability Real-Time PCR and Growth Recovery. *Front*



- 718            *Microbiol* 11: 1107
- 719    65.    Davys G, Marshall JC, Fayaz A, Weir RP, Benschop J., 2020. Campylobacteriosis  
720            associated with the consumption of unpasteurised milk: Findings from a sentinel  
721            surveillance site. *Epidemiol Infect* 148: e16
- 722    66.    Admasie A, Eshetu A, Tessema TS, Vipham J, Kovac J, Zewdu A., 2023.  
723            Prevalence of Campylobacter species and associated risk factors for  
724            contamination of dairy products collected in a dry season from major milk sheds in  
725            Ethiopia. *Food Microbiol* 109: 104145
- 726    67.    World Health Organization., 2008. *Indicators for Assessing Infant and Young*  
727            *Child Feeding Practices: Conclusions of a Consensus Meeting Held 6-8*  
728            *November 2007 in Washington DC, USA*
- 729    68.    Belay DG, Getnet M, Akalu Y, Diress M, Gela YY, Getahun AB, Bitew DA, Terefe  
730            B, Belsti Y., 2022. Spatial distribution and determinants of bottle feeding among  
731            children 0-23 months in Ethiopia: spatial and multi-level analysis based on 2016  
732            EDHS. *BMC Pediatr* 22: 120
- 733    69.    Rothstein JD, Mendoza AL, Cabrera LZ, Pachas J, Calderón M, Pajuelo MJ,  
734            Caulfield LE, Winch PJ, Gilman RH., 2019. Household contamination of baby  
735            bottles and opportunities to improve bottle hygiene in Peri-Urban Lima, Peru. *Am*  
736            *J Trop Med Hyg* 100: 988–997
- 737    70.    Anselin L, Getis A., 1992. Spatial statistical analysis and geographic information  
738            systems. *Ann Reg Sci* 26: 19–33
- 739    71.    Bertazzon S, Olson S, Knudtson M., 2010. A spatial analysis of the demographic  
740            and socio-economic variables associated with cardiovascular disease in calgary

741 (Canada). *Appl Spat Anal Policy* 3: 1–23

742

743

744 **Tables**

745 Table 1. Demographics and socio-economic status of the households enrolled in this  
746 study.

Variable	N = 106 <sup>\$</sup>
<b>Sex</b>	
Female	51 (48%)
Male	55 (52%)
<b>Mother's age</b>	
Unknown	1
<b>Mother's education</b>	
No primary education	76 (72%)
Some primary education	29 (28%)
Unknown	1
<b>Livestock ownership</b>	
Cattle	52 (49%)
Goat	60 (57%)
Sheep	48 (45%)
Chicken	53 (50%)
Tropical livestock unit	0.62 (0.20, 1.40)

747 <sup>\$</sup>n (%); Median (IQR)

748

749 Table 2. Variables used for regression modeling and univariate analysis results.

Variable	First period		Second period	
	OR (CI) <sup>&amp;</sup>	<i>p</i>	OR (CI)	<i>p</i>
Child sex	0.8 (0.4, 1.7)	0.56	2.3 (1.1, 5.1)	<b>0.03</b>
Mother's age	1.3 (0.6, 2.8)	0.55	0.8 (0.4, 1.7)	0.55
Mother's education	1.2 (0.5, 2.9)	0.49	1.1 (0.4, 2.6)	0.49
Cattle ownership	0.6 (0.3, 1.4)	<b>0.24*</b>	2.1 (1.0, 4.7)	<b>0.05</b>
Goat ownership	1.4 (0.6, 3.0)	0.43	0.9 (0.4, 1.9)	0.70
Sheep ownership	0.6 (0.3, 1.4)	<b>0.24</b>	0.5 (0.2, 1.2)	<b>0.12</b>
Chicken ownership	0.8 (0.4, 1.7)	0.56	1.5 (0.7, 3.2)	0.33
Tropical livestock unit	1.1 (0.5, 2.3)	0.85	1.3 (0.6, 2.7)	0.56
Asset	1.3 (0.6, 3.0)	0.44	0.8 (0.4, 1.8)	0.70
Pre-lacteal feeding	2.3 (1.0, 5.5)	<b>0.06</b>	NA	NA
Time of introduction of complementary foods (days)	0.6 (0.3, 1.3)	<b>0.17</b>	NA	NA
Drinking from a bottle with nipple in the past 24h	0.6 (0.3, 1.3)	<b>0.17</b>	1.0 (0.5, 2.1)	1
Consumption of unpasteurized animal milk in the past 24h	1.0 (0.2, 4.4)	1	1.4 (0.6, 2.9)	0.44
Consumption of any solid food in the past 24h	4.5 (1.1, 31.1)	<b>0.04</b>	1.8 (0.9, 4.0)	<b>0.12</b>
Consumption of animal source food in the past 24h <sup>§</sup>	NA	NA	0.9 (0.4, 1.9)	0.70
Minimum dietary diversity <sup>§</sup>	NA	NA	0.8 (0.4, 1.7)	0.56
Household food insecurity access score	1.3 (0.6, 2.7)	0.56	2.3 (1.1, 5.1)	<b>0.03</b>
Vitamin-A supplementation in the past month <sup>#</sup>	NA	NA	1.2 (0.5, 2.6)	0.69
Being treated with antibiotics in the past month	1.6 (0.7, 3.7)	<b>0.22</b>	0.8 (0.4, 1.7)	0.55
Contact with animals	3.2 (1.4, 7.7)	<b>0.01</b>	1.3 (0.6, 2.7)	0.56
Crawling in areas with animal feces	3.0 (0.8, 14.2)	<b>0.11</b>	1.5 (0.7, 3.2)	0.33
Mouthing of soil or animal feces	12.1 (2.2, 226)	<b>0.00</b>	1.5 (0.7, 3.2)	0.33
Preventive actions (mother) for mouthing behavior	1.1 (0.5, 2.4)	0.77	0.9 (0.4, 2.0)	0.84
Elevation (m)	0.6 (0.3, 1.3)	<b>0.17</b>	0.7 (0.3, 1.5)	0.33

Table 2. Continued

Variable	First period		Second period	
	OR (CI)	<i>p</i>	OR (CI)	<i>p</i>
Maximum daily land surface temperature, 2021 (°C)	1.3 (0.6, 2.7)	0.56	1.1 (0.5, 2.3)	0.85
Mean daily land surface temperature, 2021 (°C)	0.9 (0.4, 2.0)	0.85	0.8 (0.4, 1.7)	0.56
Minimum daily land surface temperature, 2021 (°C)	1.3 (0.6, 2.7)	0.56	1.5 (0.7, 3.2)	0.33
Maximum 16-day NDVI <sup>‡</sup> , 2021	1.5 (0.7, 3.2)	0.33	0.9 (0.4, 2.0)	0.85
Mean 16-day NDVI, 2021	1.3 (0.6, 2.7)	0.56	0.7 (0.3, 1.5)	0.33
Minimum 16-day NDVI, 2021	0.8 (0.4, 1.7)	0.56	0.5 (0.2, 1.1)	<b>0.08</b>
Population count 100m x 100m, 2020	1.2 (0.5, 2.5)	0.70	0.3 (0.1, 0.6)	<b>0.00</b>
Slope (degree)	0.4 (0.2, 0.9)	<b>0.03</b>	1.3 (0.6, 2.7)	0.56
Proportion of clay particles in the fine earth fraction (g/kg)	1.3 (0.6, 2.7)	0.56	2.3 (1.1 to 5.1)	<b>0.03</b>
Soil organic carbon content in the fine earth fraction (0.1g/kg)	0.9 (0.4, 1.8)	0.70	1.2 (0.5, 2.5)	0.70
Soil pH	0.5 (0.2, 1.3)	<b>0.16</b>	0.8 (0.3, 2.0)	0.64

750 & Odds ratio (confidence interval)

751 \* Bold script represents the *p*-value is less than 0.25, and the corresponding variable  
752 was included as a candidate for the multivariate analysis.

753 \$ Data were not available for the first period as the minimum dietary diversity (MDD)  
754 survey was administered after 6 months of age.

755 # Only one child had vitamin-A supplements for the first period. This variable was  
756 excluded from the univariate analysis for the first period.

757 ‡ Normalized difference vegetation index

758

759 Table 3. Environmental variables used in this study and their sources.

Variable	Median	Data source
Elevation (m)	2083.0	WorldPop ( <a href="https://www.worldpop.org/">https://www.worldpop.org/</a> )
Maximum daily land surface temperature, 2021 (°C)	44.18	MODIS Land Surface Temperature /Emissivity Daily (MOD11A1) Version 6.1
Mean daily land surface temperature, 2021 (°C)	31.17	MODIS Land Surface Temperature /Emissivity Daily (MOD11A1) Version 6.1
Minimum daily land surface temperature, 2021 (°C)	16.58	MODIS Land Surface Temperature /Emissivity Daily (MOD11A1) Version 6.1
Maximum 16-day NDVI*, 2021	0.68	MODIS Vegetation Indices (MOD13Q1) Version 6.1
Mean 16-day NDVI, 2021	0.48	MODIS Vegetation Indices (MOD13Q1) Version 6.1
Minimum 16-day NDVI, 2021	0.30	MODIS Vegetation Indices (MOD13Q1) Version 6.1
Population count 100m x 100m, 2020	6	WorldPop ( <a href="https://www.worldpop.org/">https://www.worldpop.org/</a> )
Slope (degree)	5.0	WorldPop ( <a href="https://www.worldpop.org/">https://www.worldpop.org/</a> )
Proportion of clay particles in the fine earth fraction (g/kg)	421.5	SoilGrids version 2.0 ( <a href="https://soilgrids.org/">https://soilgrids.org/</a> )
Soil organic carbon content in the fine earth fraction (0.1g/kg)	328	SoilGrids version 2.0 ( <a href="https://soilgrids.org/">https://soilgrids.org/</a> )
Soil pH	7.4	SoilGrids version 2.0 ( <a href="https://soilgrids.org/">https://soilgrids.org/</a> )

760 \* Normalized difference vegetation index  
761

762 Table 4. Logistic regression model for the cumulative *Campylobacter* burden in the first  
763 period.

Characteristic	OR <sup>^</sup>	95% CI <sup>*</sup>	p-value
Child sex	0.59	(0.22, 1051)	0.3
Sheep ownership	0.40	(0.15, 1.01)	0.06
Pre-lacteal feeding	1.77	(0.66, 4.87)	0.3
Contact with animals	3.13	(1.11, 9.6)	<b>0.04</b> <sup>§</sup>
Mouthing of soil or animal feces	12.8	(1.80, 272)	<b>0.03</b>
Drinking from a bottle with nipple in the past 24h	0.35	(0.12, 0.94)	<b>0.04</b>
Elevation	0.61	(0.23, 1.60)	0.30
Soil pH	0.43	(0.14, 1.29)	0.14
Nagelkerke R <sup>2</sup>	0.40		
Hosmer & Lemeshow test			0.42
AUC <sup>#</sup>	0.79	(0.70, 0.88)	
Moran's I statistic	-0.06		0.75

764 <sup>^</sup>Odds ratio

765 <sup>\*</sup>Confidence interval

766 <sup>§</sup>Boldness represents the p-value is less than 0.05.

767 <sup>#</sup>Area under the receiver operating characteristic curve

768

769 Table 5. Logistic regression model for the cumulative *Campylobacter* burden in the  
770 second period.

Characteristic	OR <sup>^</sup>	95% CI*	p-value
Child sex	2.85	(1.15, 7.47)	<b>0.027</b> <sup>\$</sup>
Cattle ownership	2.87	(1.14, 7.70)	<b>0.029</b>
Sheep ownership	0.49	(0.91, 1.21)	0.12
household food insecurity access score	1.78	(0.69, 4.59)	0.2
Consumption of any solid food in the past 24h	2.01	(0.79, 5.30)	0.15
Minimum 16-day NDVI, 2021	0.43	(0.17, 1.05)	0.069
Population count 100m x 100m, 2020	0.37	(0.15, 0.92)	<b>0.033</b>
Proportion of clay particles in the fine earth fraction	2.12	(0.84, 5.54)	0.12
Nagelkerke R <sup>2</sup>	0.34		
Hosmer & Lemeshow test			0.47
AUC <sup>#</sup>	0.80	(0.71, 0.89)	
Moran's I statistic	0.014		0.37

771 <sup>^</sup>Odds ratio

772 <sup>\*</sup>Confidence interval

773 <sup>\$</sup>Bold script represents the p-value is less than 0.05.

774 <sup>#</sup>Area under the receiver operating characteristic curve

775



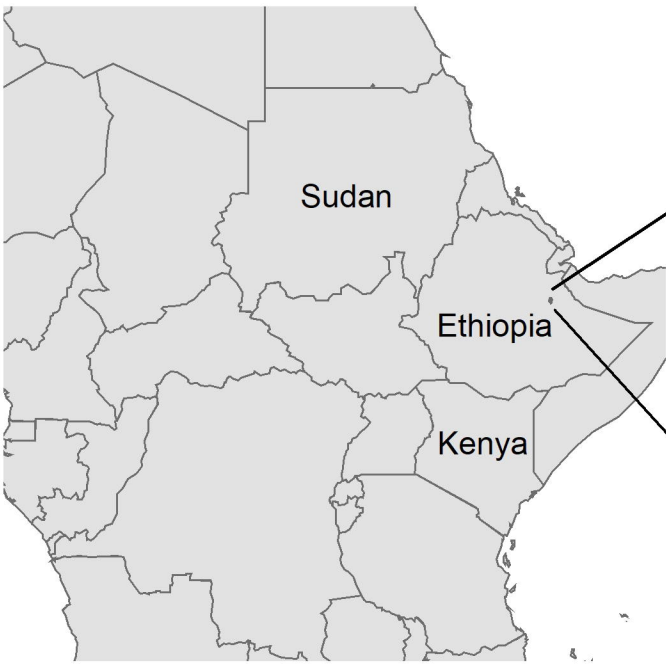
776 **Figure Legends**

777 **Figure 1.** Study area and enrolled households in the *Campylobacter* Genomics and  
778 Environmental Enteric Dysfunction (CAGED) project.

779 **Figure 2.** Frequency distribution of the cumulative *Campylobacter* burden for the first  
780 and second period in this study.

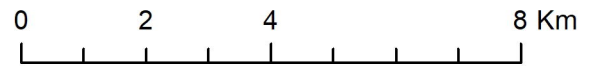
781 **Figure 3.** Receiver operating curves (ROC) for logistic regression models fit in this  
782 study. A: model fit with the first period data; B: test of the fit of the first period model to  
783 the second period data; C: model fit with the second period data.

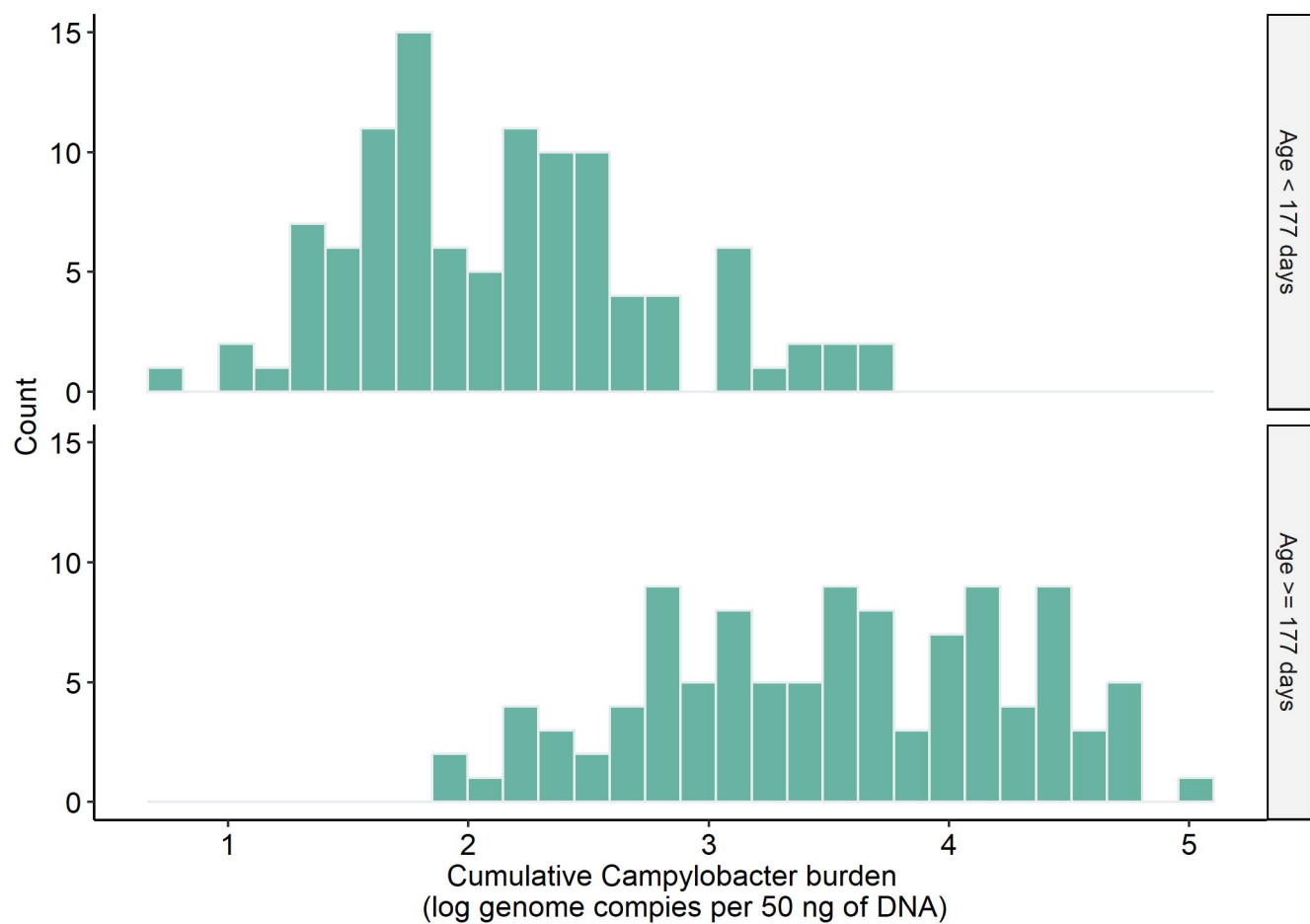
784



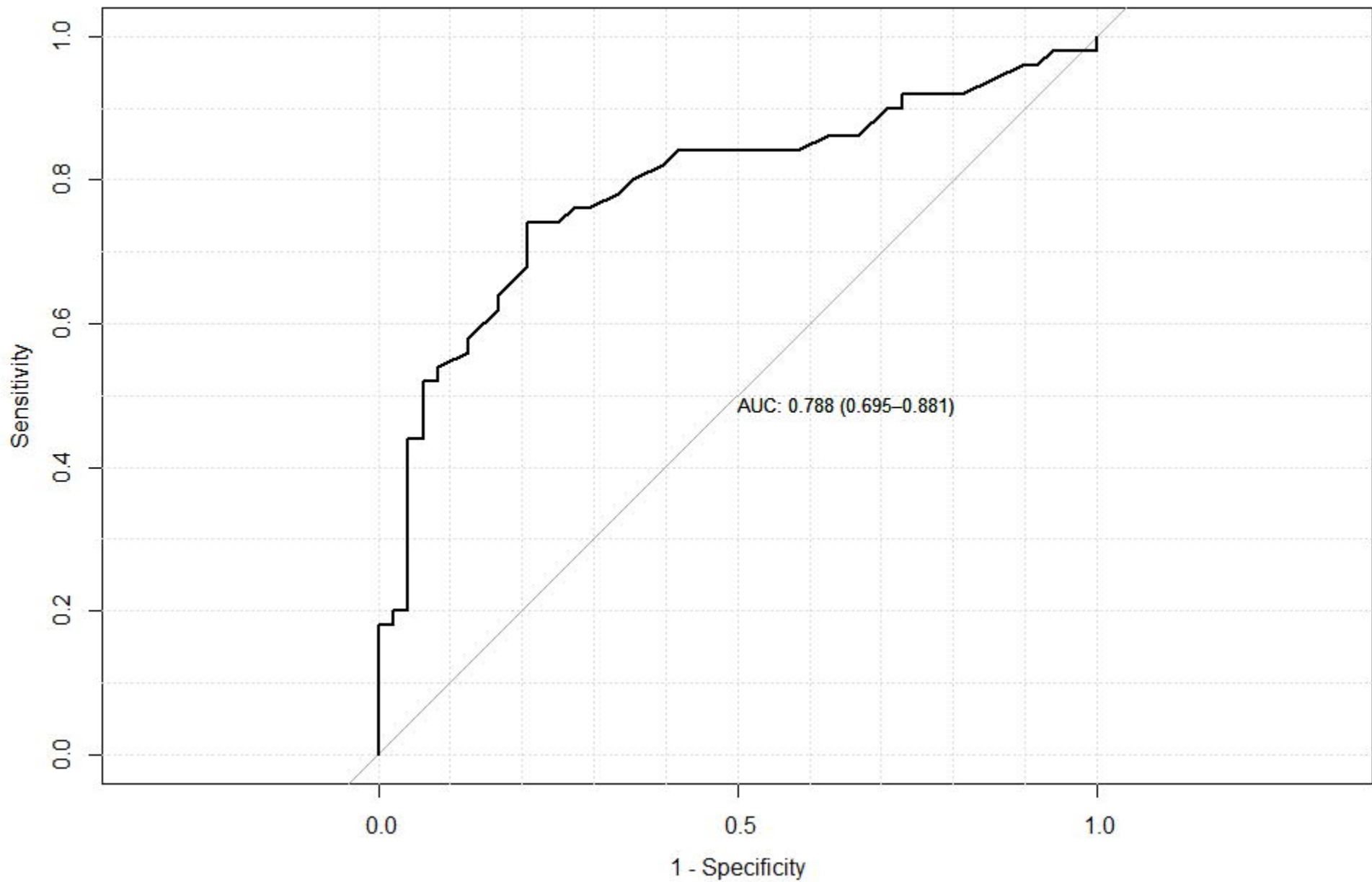
**Legend**

- Enrolled households

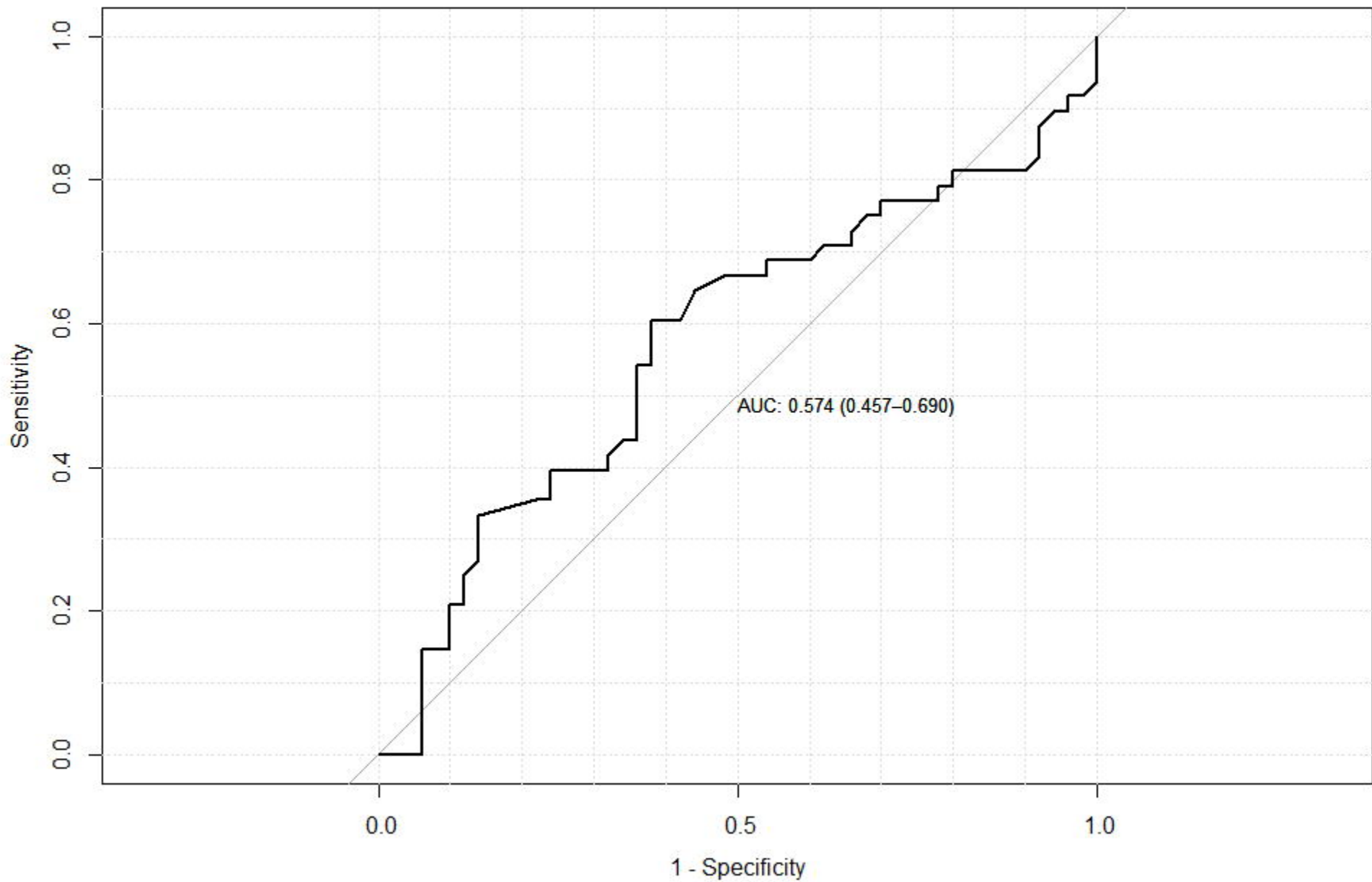




# Receiver Operating Curve



# Receiver Operating Curve



# Receiver Operating Curve

