Spatio-Temporal Trends in Typhoid Fever Incidence in Kibera (2007-2015)

Gabriella Veytsel

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Public Health

University of Washington
2018

Committee:
Grace John-Stewart
Adam Akullian
Jon Wakefield

Program Authorized to Offer Degree:
Public Health

University of Washington

Abstract

Spatio-Temporal Trends in Typhoid Incidence in Kibera (2007-2015)

Gabriella Veytsel

Chair of Supervisory Committee:
Dr. Grace John-Stewart
Department of Epidemiology

Background:

Typhoid fever is a bloodstream infection caused by the bacterium *Salmonella* typhi and transmitted either person-to-person through direct fecal-oral transmission or via environmental factors.

Methods:

We utilized household and clinical data from an infectious disease surveillance system in Kibera, from 2007-2015 to assess spatial and temporal patterns of typhoid fever risk in two age groups (children under 10 years of age and individuals 10 and older). We calculated incidence rate differences (IRDs) within clusters before and after a major Water, Sanitation, and Hygiene (WASH) intervention and utilized regression methods to compare risk between clusters in two age groups.

Results:

All clusters had declines in typhoid incidence following the WASH implementation, most with statistically significant declines. While we found that the risk between clusters was not significantly different in children under 10 years of age, we identified one cluster that exhibited elevation in risk compared to the reference in individuals 10 years of age and older (IRR = 2.80, p-value = 0.02).

Conclusion:

Residual risk of typhoid fever following the WASH intervention suggests that other environmental risk factors may play an important role in these areas.

Introduction:

Typhoid fever is a bloodstream infection caused by the bacterium *Salmonella typhi*. Overcrowding and poor drinking water quality, underdeveloped sanitation infrastructure, and hygiene are associated with an increased risk of typhoid transmission [1]. In some urban areas of sub-Saharan Africa, the incidence of typhoid fever is 247 cases per 100,000 person-years, comparable to high burden regions in Asia [2]. However, there are limited data on the epidemiology and transmission dynamics of typhoid fever in Africa [3]. Though contaminated food and water are thought to be the major risk factors [4], other environmental factors, including proximity to water bodies [5], [6], [7], residency in low elevation areas [8] [9], close contact with typhoid cases [7], flooding [10], poor sanitation [11], and large household size [5], may contribute to increased risk. In addition, the contribution of direct fecal-oral transmission versus indirect environmental transmission [9].

We mapped the spatial distribution of typhoid fever incidence over time to gain a preliminary understanding of the spatio-temporal variation in typhoid risk in an urban informal settlement in Nairobi, Kenya. Our primary aim was to identify clusters that exhibited significant reductions in typhoid fever risk over time amongst children under 10 years of age and individuals 10 and older; hence, we tested the null hypothesis that typhoid fever incidence did not decline significantly within clusters before and after a major WASH intervention. Our secondary aim was to identify clusters within each age group that exhibited either significantly elevated or reduced typhoid fever risk; we hypothesized that typhoid risk for children under 10 years of age would be higher in downstream clusters compared to upstream clusters.

The results of this study will be used to inform model parameters in a dynamic mathematical model that will be applied to identify the most likely drivers of the reduction in typhoid incidence, develop targeted interventions and allocate resources, and direct environmental sampling. The generated area-specific risk estimates will also be used to generate hypotheses as to environmental drivers of risk, including open drainage lines, water kiosk and latrine sites, streams, and elevation.

Methods:

Study site and setting

This study was conducted in Kibera, an urban informal settlement in Nairobi, Kenya, between January 2007 and December 2015. The surveillance area covers two of the 12 neighborhoods in Kibera; the study site is 0.4 km^2 and households are geographically grouped into 10 clusters [12]. Cases were defined as individuals who presented with fever ≥ 38°C and received a positive *S*. typhi blood culture at the Tabitha Clinic.

Data sources

The Kenya Medical Research Institute-Centers for Disease Control and Prevention collaboration (KEMRI-CDC) has been conducting active population-based surveillance for febrile illness, pneumonia, diarrheal disease, and jaundice in Kibera since 2005. The surveillance methods have been previously described [12]; briefly, community interviewers followed a cohort of about 50,000 individuals biweekly until April 2015, after which home visits were reduced to biannually. With an open study population, individuals are at risk for varying periods of time; therefore, person-days are accumulated for individuals until acquisition of typhoid fever, death, outmigration, internal migration, refusal, termination, and inability to trace (when individuals are missing for more than 120 days).

Statistical analysis

We conducted a retrospective cohort study including 48,728 individuals followed until the first occurrence of culture-confirmed typhoid fever. Typhoid fever is a rare disease in this population with 302 cases confirmed in the study period (2007-2015). Study visits were excluded if they occurred outside of the study period, had missing person-time, or were missing either latitude and longitude coordinates or cluster assignment. Crude maps of cases and non-cases were generated for children under 10 years of age and individuals 10 and over using GPS coordinates of households. Point-level data was aggregated into population clusters (N=10) defined by the original study protocol and Standardized Incidence Ratios (SIRs) were calculated at the cluster-level to estimate the relative risk of typhoid fever within each cluster compared to the entire region. Given evidence of different routes of transmission between children less than 10 years of age and individuals 10 and over, all analyses were stratified by age (<10 and 10+).

Recently, there has been a sharp decline in typhoid fever cases in Kibera, beginning 2012; that same year, a major Water, Sanitation, and Hygiene (WASH) intervention was initiated in Kibera. We were interested in comparing the incidence rates before and after this public health intervention; hence, we divided the study period into 2007-2012 and

2013-2015. To determine if there was a significant reduction in typhoid incidence between these time periods within all clusters, we calculated age-stratified incidence rate differences (IRDs) for each cluster by subtracting the incidence rates in time period 2 from time period 1. Associated p-values were calculated via Pearson's chi-square test for trend.

Crude Incidence Rates:

The overall crude incidence rate was calculated by dividing the number of observed cases by the person-years accumulated in the entire study period. The age-specific incidence rates were calculated by dividing the number of observed cases in each age group by the person-years accumulated in each age group during the entire study period. Age-stratified cluster-specific incidence rates were calculated for each time period and Fisher's exact mid-p confidence intervals were generated from OpenEpi.

Quasi-Poisson Loglinear Model

In order to test the null hypothesis that typhoid fever risk is the same in all clusters regardless of year, we used an age-stratified Quasi-Poisson generalized linear model (GLM) to account for overdispersion with a log link function and an offset for expected number of cases. We included year as a fixed effect with 2007 as a reference in order to reduce potential omitted variable bias—by including year as a categorical variable in this regression, we control for observable and unobservable differences in environmental and demographic factors across time. Thus, we tested for an association between the probability of becoming a typhoid fever case and cluster residence, after adjustment for year.

$$E[Y_i \mid \beta] = \beta_0 + \log(PY_i) + \beta_1(x_1)$$

Where $\log(PY_i)$ is an offset for person-years, Y_i is the number of cases, and x_1 is cluster.

Results

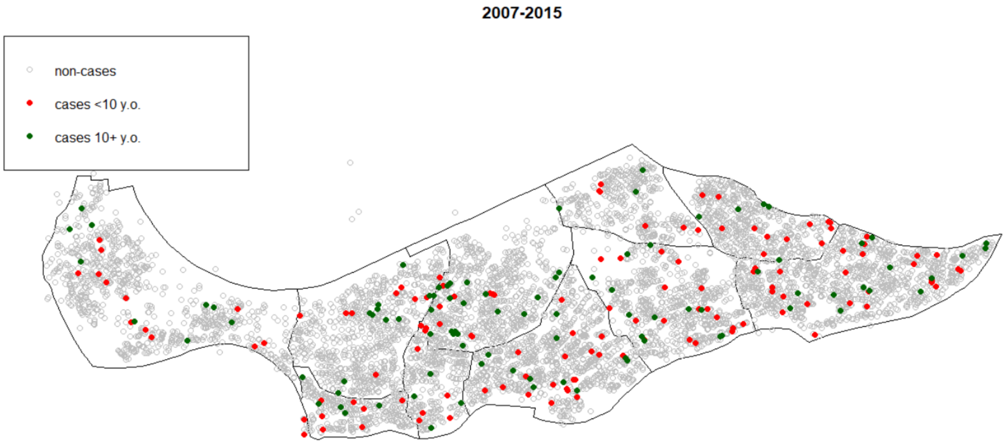

Figure 1: Household residence of cases and non-cases in Kibera, 2007-2015

Between January 1st, 2007 and December 31st, 2015, 306 cases of culture-confirmed typhoid fever were reported in Kibera. Of those, 302 had confirmed residence information and household GPS coordinates and were included in this study (Figure 1). Demographic and environmental characteristics of the study population are displayed in Table 1 by cluster. Of note, the population density is lowest in cluster 1 (132,300 persons per km$^{2)}$ and highest in cluster 4 (171,127.4 persons per km^2); additionally, cluster 10 is the most downstream at an altitude of 1717.7 meters.

	Cluster									
Characteristics	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Follow-up (P-Y)	30,002.1	19,944.4	5429.9	5,656.9	15849.5	16,692.5	10,292.9	18,932.7	9,760.7	19,594.0
Population density/km²	132,300	165,606	164,052	171,127.4	135,722.2	151,939.9	151,157.8	162,140.9	161,153.4	153,360.4
Median elevation, m	1744.7	1754.5	1732.1	1740.68	1756.2	1738.0	1747.1	1731.7	1731.0	1717.7
Median household size	4.62	4.33	3.95	4.1	4.06	4.21	4.76	4.33	4.43	4.4
Proportion of person-years in age group 0-9, %	55	50	57	61	53	54	53	52	53	53
Median age, year	20	21	21	20	21	21	20	21	20	20

Table 1: Descriptive statistics of study participants by cluster, 2007-2015

The overall crude incidence rate during 2007-2015 was 1.99 cases per 1,000 person-years (range 0.19-4.38). The age-specific rates for children under 10 years of age was 4.75 cases per 1,000 person-years, compared to 1.03 cases per 1,000 person-years in individuals 10 years of age and older. Thus, children under 10 had nearly 5 times the rate of typhoid fever compared to individuals 10 and older (IRR = 4.59, p-value = 2.2E-16). Incidence rates for both age groups fluctuated in a similar pattern throughout the study period, approaching an all-time low in 2015 (Figure 2).

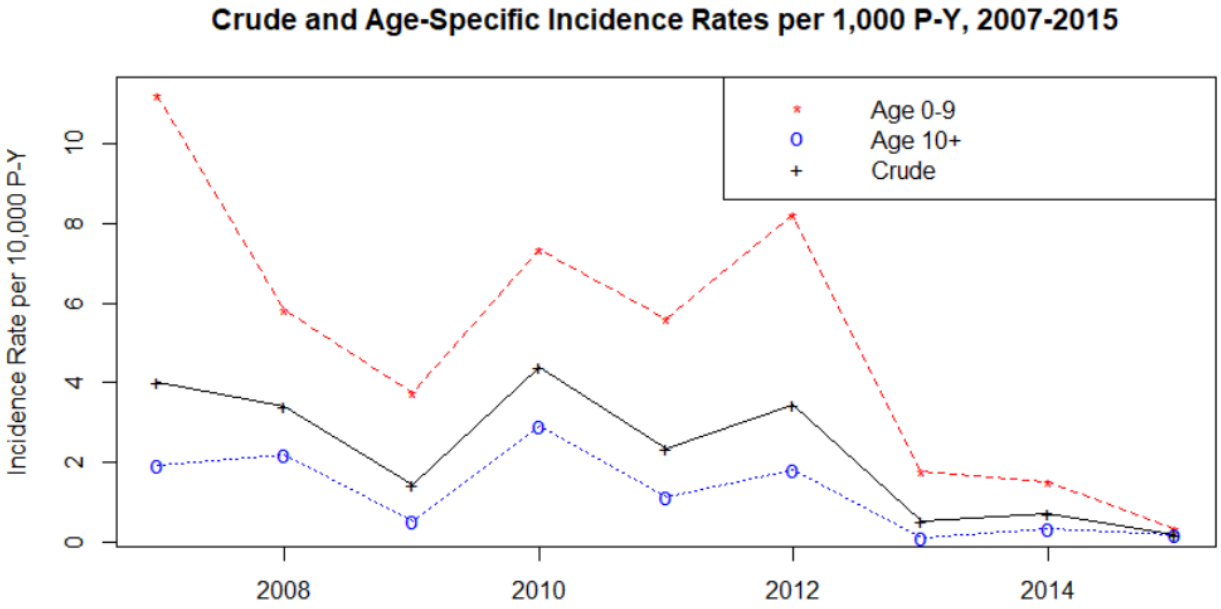

Figure 2: Crude and age-stratified incidence rate by year, 2007-2015

Cluster-specific typhoid fever incidence rates for all age groups during the study period are shown in Figure 3. The lowest incidence rate was reported in cluster 1 and 7 (IR = 1.20 and 1.36 cases per 1,000 person-years, respectively) and the highest incidence rates were reported in clusters 5 and 3 (IR = 2.78 and 2.76 cases per 1,000 person-years, respectively). Compared to the overall incidence in the region, cluster 1 had significantly lower risk of typhoid fever (IRR = 0.62, 95% CI: 0.41, 0.94), while cluster 5 had a significantly higher risk of typhoid fever (IRR = 1.73, 95% CI: 1.13, 2.65).

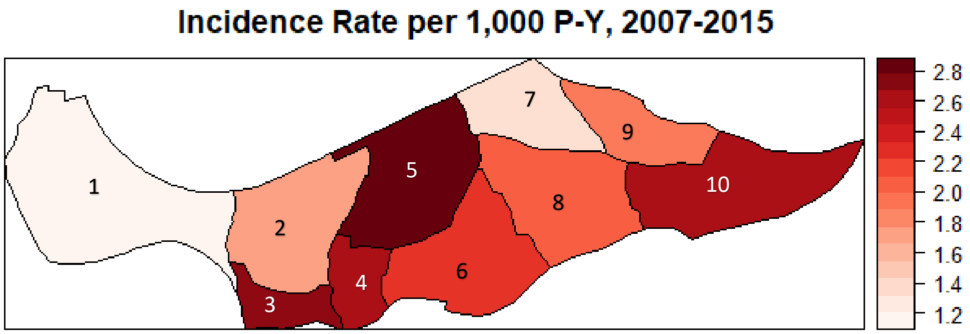

Figure 3: Incidence rate by cluster, 2007-2015

	Age 0-9									
	2007-2012				2013-2015					
	Cases	Person-Years	IR_1*	95% CI	Cases	Person-Years	IR_2*	95% CI	IRD**	p-value
Cluster 1	21	4,467.45	4.70	2.99, 7.06	1	3,049.30	0.33	0.16, 16.17	**4.37***	**0.00057**
Cluster 2	15	2,834.78	5.29	3.08, 8.53	3	1,711.35	1.75	0.45, 4.77	3.54	0.066
Cluster 3	10	960.32	10.41	5.29, 18.56	0	533.04	0	--	**10.41***	**0.018**
Cluster 4	9	1,028.32	8.75	4.27, 16.06	0	586.64	0	--	**8.75***	**0.023**
Cluster 5	22	2,576.71	8.54	5.49, 12.71	1	1,481.49	0.67	0.03, 3.33	**7.87***	**0.0013**
Cluster 6	22	2,858.24	7.70	4.95, 11.46	1	1,560.56	0.64	0.03, 3.16	**7.06***	**0.0018**
Cluster 7	9	1,824.69	4.93	2.41, 9.05	0	675.05	0	--	4.93	0.068
Cluster 8	23	3,134.72	7.34	4.76, 10.84	1	1,784.12	0.56	0.28, 2.76	**7.78***	**0.0004**
Cluster 9	11	1,697.13	6.48	3.41, 11.27	2	920.08	2.17	0.36, 7.18	4.31	0.14
Cluster 10	28	3,521.77	7.95	5.39, 11.34	6	1,780.84	3.37	1.37, 7.01	**4.58***	**0.048**
Total	170	24,904.13	6.83	5.86, 7.91	15	14,082.47	1.07	0.62, 1.72	**5.76***	**1.85E-15**

Table 2: Incidence rates per 1,000 person-years over time in ages 0-9, by cluster
*Incidence Rate per 1,000 P-Y
**Incidence Rate Difference $(IR_1 – IR_2)$

	Age 10+									
	2007-2012				2013-2015					
	Observed Cases	Person-Years	IR_1*	95% CI	Observed Cases	Person-Years	IR_2*	95% CI	IRD**	p-value
Cluster 1	13	11,843.89	1.10	0.61, 1.83	1	10,641.43	0.09	0.005, 0.46	**1.01***	**0.0026**
Cluster 2	12	7,534.06	1.59	0.86, 2.71	3	7,864.23	0.38	0.97, 10.38	**1.21***	**0.016**
Cluster 3	5	2,000.42	2.50	0.92, 5.54	0	1,936.14	0	--	**2.50***	**0.028**
Cluster 4	6	2,199.34	2.73	1.11, 5.67	0	1,842.55	0	--	**2.73***	**0.025**
Cluster 5	21	6,349.59	3.31	2.10, 4.97	0	5,441.73	0	--	**3.31***	**2.18E-5**
Cluster 6	13	7,091.81	1.83	1.02, 3.06	1	5,181.86	0.19	0.01, 0.95	**1.64***	**0.0078**
Cluster 7	5	4,855.14	1.03	0.37, 2.28	0	2,937.99	0	--	1.03	0.082
Cluster 8	14	8,222.40	1.70	0.97, 2.79	0	5,791.47	0	--	**1.70***	**0.0017**
Cluster 9	6	4,123.68	1.46	0.59, 3.03	0	3,019.82	0	--	**1.46***	**0.036**
Cluster 10	13	8,501.79	1.53	0.85, 2.55	4	5,789.56	0.69	0.22, 1.67	0.84*	0.15
Total	108	62,722.12	1.72	1.42, 2.07	9	50,446.78	0.18	0.09, 0.33	**1.54***	**9.69E-16**

Table 3: Incidence rates per 1,000 person-years over time in ages 10+, by cluster
*Incidence Rate per 1,000 P-Y
**Incidence Rate Difference $(IR_1 – IR_2)$

Tables 2 and 3 show the incidence rate of typhoid fever in age groups 0-9 and 10+ between 2007-2012 and 2013-2015. For children under 10 years of age, the incidence rate is highest in cluster 3 in the first time period (IR = 10.42, 95% CI: 5.29, 18.56), yet this cluster undergoes the largest reduction in the entire region in the second time period (IRD = 10.42, p-value = 0.00057), thus shifting the highest incidence to cluster 10 (IR = 3.37, 95% CI: 1.37, 7.01). Similarly, in individuals 10 years of age and older, cluster 5 exhibits the highest incidence in the first time period, yet it experiences the greatest incidence rate difference (IRD = 3.31, p-value = 2.18E-5), changing the highest incidence to cluster 10 (IR = 0.69, 95% CI: 0.22, 1.67). In children under 10 years of age in clusters 2, 7, and 9, typhoid fever risk decreased in the second time period but either had a trend towards or did not attain statistical significance. Similarly, in individuals over 10 years of age residing in clusters 7 and 10, typhoid risk declined albeit without statistically significant differences, perhaps due to insufficient power.

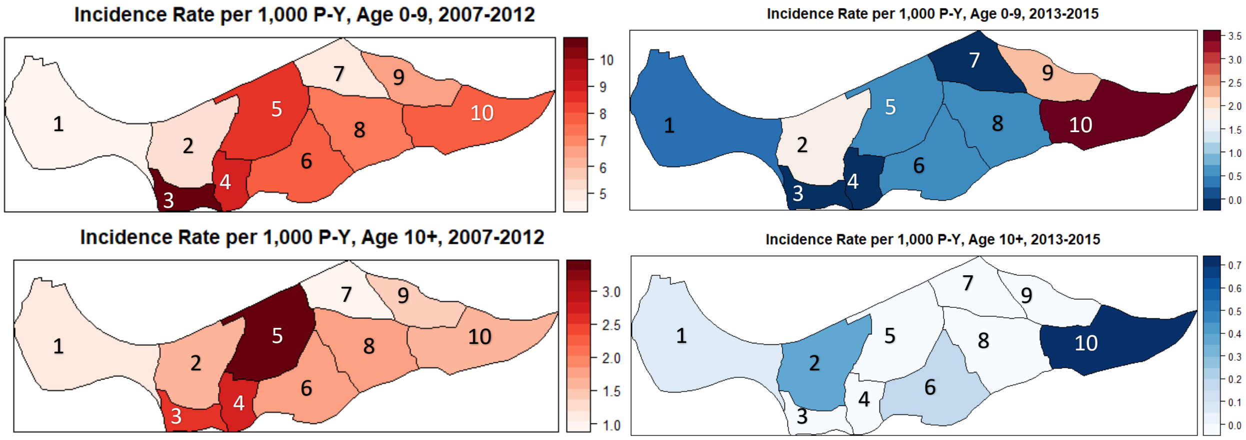

Figure 4: Incidence rates in ages 0-9 and 10+ over time

After adjustment for year, cluster 5 exhibited an elevated risk of typhoid fever for individuals 10 years of age and older (IRR = 2.80, p-value = 0.02). Model inclusion of elevation, household size, and population density covariates had a trend for statistical significance and improved model fit (IRR = 2.14, p-value = 0.07) (Appendix). We were unable to reject the null that risk is homogenous in the study area among children under 10 years of age. Additionally, we identified 2009 and 2013-2015 as study years that underwent a significant reduction in typhoid fever incidence for both age groups compared to 2007 (Table 4).

		Quasi-Poisson Loglinear Model					
		Age 0-9			Age 10+		
		IRR**	95% CI	p-value	IRR**	95% CI	p-value
Cluster	1	Ref	--	--	Ref	--	--
	2	1.26	0.51, 3.07	0.61	1.60	0.65, 3.93	0.30
	3	2.11	0.66, 5.82	0.17	2.01	0.49, 6.39	0.27
	4	1.79	0.53, 5.08	0.31	2.26	0.63, 6.81	0.17
	5	1.79	0.77, 4.17	0.17	**2.80**	**1.25, 6.58**	**0.02**
	6	1.67	0.72, 3.89	0.23	1.65	0.66, 4.12	0.28
	7	1.04	0.31, 2.96	0.95	0.85	0.21, 2.69	0.79
	8	1.55	0.68, 3.58	0.30	1.42	0.57, 3.54	0.44
	9	1.45	0.51, 3.77	0.46	1.25	0.35, 3.77	0.70
	10	1.93	0.91, 4.27	0.09	1.68	0.71, 4.07	0.24
Year	2007	Ref	--	--	Ref	--	--
	2008	0.42	0.25, 1.04	0.07	1.08	0.51, 2.26	0.85
	2009	**0.34**	**0.13, 0.77**	**0.02**	**0.26**	**0.07, 0.77**	**0.03**
	2010	0.65	0.33, 1.28	0.21	1.48	0.74, 3.01	0.27
	2011	0.50	0.23, 1.02	0.06	0.57	0.24, 1.30	0.19
	2012	0.73	0.37, 1.43	0.36	00.93	0.44, 1.95	0.84
	2013	**0.15**	**0.04, 0.43**	**0.002**	**0.04**	**0.001, 0.27**	**0.01**
	2014	**0.14**	**0.03, 0.40**	**0.002**	**0.17**	**0.03, 0.60**	**0.02**
	2015	**0.03**	**0.002, 0.15**	**0.001**	**0.08**	**0.021, 0.25**	**0.0001**
Model Fit	Dispersion parameter	2.01			1.47		
	Null deviance (df)	265.31 (89)			212.02 (89)		
	Residual deviance (df)	152.66 (72)			97.11 (72)		

Table 4: Quasi-Poisson Loglinear Model
**Incidence Rate Ratio

Discussion:

For our primary aim, we found that typhoid fever risk declined in all clusters post-intervention, although for a few the incidence risk decline did not attain statistical significance. Given the small study area, it is unlikely that an unequal distribution of water kiosks is responsible for the lack of significant reduction in risk in these clusters; rather, this suggests either that statistical power was limited or that risk factors or transmission routes in these areas persisted despite a water and sanitation intervention. A recent study found that soil ingestion is associated with diarrheal illnesses in children 5 years of age and younger in Kibera; thus, interventions aimed at improving clean water access, latrines, and handwashing are not likely to remove all exposures in children [13].

For our secondary aim, we were unable to reject the null hypothesis that risk is homogenous in this region for children under 10 years of age. Cluster 5 exhibited significantly elevated risk compared to the reference in individuals 10 years of age and older. Given that we still had a trend for significance upon model inclusion of several demographic and environmental variables, cluster 5 may be an area that would benefit from vaccination efforts.

Our results from the regression model are contrary to a recent study on geographic patterns in typhoid risk in Kibera that found elevated risk of typhoid fever in low elevation areas among children under 10 years of age [8]. Low elevation areas aggregate fecal waste from upstream sources via runoff of surface waters and may function as environmental reservoirs for water-related infectious disease in children [14], [15], [16], [17], [18], [19]. Environmental transmission may largely account for the disproportionately high burden of typhoid incidence in children in Kibera [8]. The inclusion of environmental risk factors such as geographical proximity to streams, latrines, water kiosks, and open drainage sites may improve model fit.

There are many hypotheses for the observed reduction in typhoid incidence after 2012, including improved access to medical care and earlier case detection, and enhancements in water and sanitation. For instance, in the same time frame, Shining Hope for Communities (SHOFCO) launched its Clean Water, Sanitation, & Hygiene (WASH) program, which has been scaled up every year since 2012. Currently, SHOFCO's intervention package includes 10 water kiosks that supply clean and low-cost water, 46 community pit latrines, and hygiene education [20].

This study has several limitations. First, the potential for misclassification of the outcome exists as the study cannot exclude paratyphoid fever in clinical cases reported as typhoid fever, an illness that presents with similar symptoms, yet is milder compared to typhoid fever; importantly, typhoid vaccines do not protect against paratyphoid fever. Thus, we may be overestimating the rate of typhoid fever. Secondly, the choice of cluster boundaries may not be entirely discriminatory—individuals from different clusters that are close in distance may face more similar exposures than individuals who are far apart in the same cluster. Thirdly, incidence rates were not adjusted to account for patients who visited the Tabitha study clinic and did not have a blood culture done despite meeting the typhoid case definition (based on fever and respiratory criteria) or patients who met the illness criteria who visited a clinic other than the study clinic and therefore did not have their blood drawn. Thus, we may be underestimating the true burden of typhoid fever illness as people who live near the study clinic may be more likely to visit it rather than an alternative clinic. Fourthly, there may be omitted variable bias present in the fixed effects model if the demographic and environmental variables we did not include are time-variant within clusters (i.e. the scale-up of WASH interventions varies between clusters over time). Finally, the study involved multiple comparisons, potentially resulting in the rejection of a null hypothesis that should not have been rejected.

One of the strengths of this study was the inclusion of person-time to account for differing amounts of risk amongst individuals due to migration. A second strength was the stratification by age that allowed us to compare typhoid fever risk between age groups, thus potentially identifying a cluster that in which children under 10 years of age and individuals 10 years of age and older underwent different routes of transmission. Previous studies on typhoid fever in Kibera focused on 2-3 year periods before a major public health intervention was initiated [8] [2]; thus, the third strength of our study is the comparison of typhoid fever incidence before and after 2012. Lastly, the inclusion of time as a fixed effect in our regression model allowed us to include unmeasured covariates that varied with study year.

Without a definitive explanation for this reduction in typhoid incidence, it is unclear if this trend will continue or if vaccination interventions should be initiated. Mounting drug resistance adds urgency to this question: a 2012 study found that 77.8% of S. typhi isolates were multi-drug resistant in Kibera. Importantly, 7% had decreased susceptibility or resistance to nalidixic acid; there is increasing evidence that nalidixic acid-resistant S. enterica infections have reduced susceptibility to fluoroquinolones, which could result in increases in typhoid illness duration, rate of treatment failure, and treatment cost [12].

References

1. *Unsafe drinking-water, sanitation and waste management*. WHO 2016 2016-08-04 10:18:41; Available from: http://www.who.int/sustainable-development/cities/health-risks/water-sanitation/en/.
2. Breiman, R.F., et al., *Population-Based Incidence of Typhoid Fever in an Urban Informal Settlement and a Rural Area in Kenya: Implications for Typhoid Vaccine Use in Africa.* PLoS One, 2012. **7**(1): p. e29119.
3. John Wain, R.S.H., Matthew L Mikoleit, Karen H Keddy, R Leon Ochiai, *Typhoid fever.* Lancet, 2015. **385**: p. 1136–45.
4. Dewan, A.M., et al., *Typhoid Fever and Its Association with Environmental Factors in the Dhaka Metropolitan Area of Bangladesh: A Spatial and Time-Series Approach.* PLoS Neglected Tropical Diseases, 2013. **7**(1): p. e1998.
5. Corner, R.J., A.M. Dewan, and M. Hashizume, *Modelling typhoid risk in Dhaka Metropolitan Area of Bangladesh: the role of socio-economic and environmental factors.* Int J Health Geogr, 2013. **12**(13).
6. Sur, D., et al., *Comparisons of predictors for typhoid and paratyphoid fever in Kolkata, India.* BMC Public Health, 2007. **7**(289).
7. Tran, H.H., et al., *Risk factors associated with typhoid fever in Son La province, northern Vietnam.* Transactions of The Royal Society of Tropical Medicine and Hygiene, 2005. **99**(11): p. 819-826.
8. Akullian, A., et al., *Environmental Transmission of Typhoid Fever in an Urban Slum.* PLoS Neglected Tropical Diseases, 2015. **9**(12): p. e0004212.
9. Baker, S., et al., *Combined high-resolution genotyping and geospatial analysis reveals modes of endemic urban typhoid fever transmission.* Open Biology, 2011. **1**(2): p. 110008.
10. Vollaard, A.M., et al., *Risk Factors for Typhoid and Paratyphoid Fever in Jakarta, Indonesia.* JAMA, 2004. **291**(21): p. 2607-2615.
11. Karkey, A., et al., *The Burden and Characteristics of Enteric Fever at a Healthcare Facility in a Densely Populated Area of Kathmandu.* PLoS One, 2010. **5**(11): p. e13988.
12. Feikin, D.R., et al., *The Burden of Common Infectious Disease Syndromes at the Clinic and Household Level from Population-Based Surveillance in Rural and Urban Kenya.* PLoS One, 2011. **6**(1): p. e16085.
13. Bauza V , O.R., Nguyen TH, Guest JS, *Soil Ingestion is Associated with Child Diarrhea in an Urban Slum of Nairobi, Kenya. - PubMed - NCBI.* Am J Trop Med Hyg., 2017. **96**(3): p. 569-575.
14. Dorner, S.M., et al., *Hydrologic modeling of pathogen fate and transport.* Environ Sci Technol, 2006. **40**(15): p. 4746-53.
15. Ferguson, C., et al., *Fate and Transport of Surface Water Pathogens in Watersheds.* http://dx.doi.org/10.1080/10643380390814497, 2010. **33**(3): p. 299-361.
16. Mawdsley, J.L., et al., *Pathogens in livestock waste, their potential for movement through soil and environmental pollution.* Applied Soil Ecology, 1995. **2**(1): p. 1-15.
17. Pachepsky, Y.A., et al., *Transport and fate of manure-borne pathogens: Modeling perspective.* Agricultural Water Management, 2006. **86**(1): p. 81-92.
18. Trask, J.R., et al., *Overland and near-surface transport of Cryptosporidium parvum from vegetated and nonvegetated surfaces.* J Environ Qual, 2004. **33**(3): p. 984-93.
19. Tyrrel, S.F. and J.N. Quinton, *Overland flow transport of pathogens from agricultural land receiving faecal wastes.* J Appl Microbiol, 2003. **94 Suppl**: p. 87s-93s.
20. *Clean Water, Sanitation & Hygiene (WASH)*. 2018 [cited 2018 5/30/2018]; Available from: https://www.shofco.org/approach/wash/.

Appendix

$$E[Yi \mid \beta] = \beta 0 + \log(PYi) + \beta 1(x1) + \beta 2(x2) + \beta 3(x3) + \beta 4(x4)$$

Where log(PYi) is an offset for person-years, Yi is the number of cases, x1 is cluster, x2 is average household size, x3 is average elevation, and x4 is population density.

		Quasi-Poisson Loglinear Model					
		Age 0-9			Age 10+		
		IRR**	95% CI	p-value	IRR**	95% CI	p-value
Cluster	1	Ref	--	--	Ref	--	--
	2	1.23	0.52, 2.87	0.63	1.36	0.56, 3.30	0.49
	3	1.80	0.62, 4.65	0.25	1.66	0.43, 5.06	0.41
	4	1.35	0.44, 3.64	0.56	1.76	0.51, 5.19	0.33
	5	1.35	0.61, 2.99	0.46	2.14	0.95, 4.97	0.07
	6	1.43	0.65, 3.14	0.38	1.33	0.54, 3.27	0.53
	7	1.02	0.34, 2.68	0.97	0.80	0.21, 2.43	0.71
	8	1.57	0.73, 3.42	0.26	1.27	0.53, 3.07	0.59
	9	1.60	0.62, 3.88	0.31	1.26	0.37, 3.63	0.68
	10	1.70	0.84, 3.55	0.15	1.49	0.64, 3.52	0.35
Year	2007	Ref	--	--	Ref	--	--
	2008	0.55	0.28, 1.05	0.07	1.10	0.53, 2.26	0.80
	2009	**0.41**	**0.17, 0.90**	**0.04**	**0.29***	0.08, 0.81	0.04
	2010	0.81	0.43, 1.53	0.51	1.62	0.82, 3.24	0.17
	2011	0.57	0.28, 1.12	0.11	0.59	0.25, 1.30	0.20
	2012	0.79	0.42, 1.49	0.47	0.97	0.48, 0.28	0.94
	2013	**0.16**	**0.05, 0.43**	**0.001**	**0.05**	0.001, 0.28	0.01
	2014	**0.15**	**0.04, 0.42**	**0.002**	**0.18**	**0.03, 0.60**	**0.02**
	2015	**0.03**	**0.002, 0.13**	**0.0004**	**0.08**	**0.02, 0.24**	**0.00006**
Explanatory variables	Average household size	1.90	0.79, 4.62	0.16	1.89	0.64, 5.62	0.25
	Average elevation (m)	**0.98**	**0.96, 1.00**	**0.019**	**0.98**	**0.96, 1.00**	**0.04**
	Population density (persons/km^2)	**1.00**	**1.00, 1.00**	**0.01**	1.00	1.00, 1.00	0.72
Model Fit	Dispersion parameter	1.67			1.33		
	Null deviance (df)	265.31 (89)			212.02 (89)		
	Residual deviance (df)	126.00 (69)			88.67 (69)		

Table 5: Quasi-Poisson Loglinear Model
**Incidence Rate Ratio