Distribution Agreement

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

Signature:

_____   _____

                                                                                    Date

Family Activity Space Overlap as a Risk Factor for

Vector-Borne Disease Transmission

By

Masato Yoshihara

Master of Public Health

Global Environmental Health

_____

Gonzalo Vazquez-Prokopec, PhD

Committee Chair

_____

Paige Tolbert, PhD

Committee Member

Family Activity Space Overlap as a Risk Factor for

Vector-Borne Disease Transmission

By

Masato Yoshihara

B.S.

University of California, San Diego

2016

Thesis Committee Chair: Gonzalo Vazquez-Prokopec, PhD

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Global Environmental Health

2018

Abstract

Family Activity Space Overlap as a Risk Factor for
Vector-Borne Disease Transmission
By Masato Yoshihara

Purpose: Socially structured human movement is key to *Aedes aegypti*-borne disease (ABD) transmission since it increases the human-mosquito interactions within a narrow space-time window for diseases to be transmitted between the two hosts. Recent theoretical explorations suggest that activity space overlap, specifically amongst those with close social distances (i.e. family members), can drive ABD transmission.

Methods: Using movement survey data collected from families living in three cities in the Yucatan State of Mexico, we aimed to identify family member activity space overlap as a risk factor for ABD transmission. This was accomplished using linear regression models to identify the correlation between the frequency of shared locations amongst families containing members that have acute ABD infection or history of ABD infection with the percent ABD-positive member count in household.

Results: As predicted by mathematical models, empirical analyses indicate that family members have increased risk of ABD contraction with higher degrees of shared activity space with a family member who is either acutely infected with an ABD or has previous ABD history.

Conclusions: These results suggest that there may be an ideal ratio of family activity space overlap that facilitates ABD transmission and persistence within an urban environment. These findings have implications for developing more accurate disease models to aid resource-limited public health programs to efficiently implement vector and disease control.

Family Activity Space Overlap as a Risk Factor for

Vector-Borne Disease Transmission

By

Masato Yoshihara

B.S.

University of California, San Diego

2016

Thesis Committee Chair: Gonzalo Vazquez-Prokopec, PhD

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master in Public Health

in Global Environmental Health

2018

Table of Contents:

Introduction:

Social interactions play a significant role in the transmission of communicable diseases, since they increase the space-time interactions necessary for pathogens to pass from a host to a susceptible individual [1]. Identifying human social interactions via contact tracing has been a standard procedure in addressing directly-transmitted infections such as sexually transmitted diseases, water-borne illnesses, and air-borne illnesses [2-5]. Mosquito-borne pathogens, on the other hand, differ since they are indirectly transmitted between individuals via a pathogen-carrying vector. Mosquito dispersal and human movement operate on the same spatial and temporal level which obscures the role of socially structured movement [6-8]. Recent evidence suggests that this may not be the case for Dengue (DENV), Chikungunya (CHIKV), and Zika (ZIKV) viruses. These arboviruses are predominantly found in the tropics and subtropics and are transmitted by day-biting mosquitos in the *Aedes* genus, predominantly by *Ae. aegypti* [9]. The day biting nature and limited flight range of *Ae. aegypti* (<100 meters) suggests that human movement to locations that they frequent within their social network (i.e. social interactions) may drive *Aedes*-borne disease (ABD; i.e. DENV, CHIKV, ZIKV) transmission dynamics in urban settings [10-13]. This observation implies that spatiotemporal dynamics of ABD's are similar to that of directly transmitted diseases since the vector is found within homes or other locations that individuals may frequent [1, 14].

Transmission dynamics of ABD's vary widely from individual to individual due to coupled functional heterogeneities between humans, mosquitos, and the local environment [15]. Mosquito heterogeneities, particularly *Ae. aegypti* dispersion, have been incorporated into recent ABD prediction models and have been shown to have a diffusive effect on activity space, a human source of heterogeneity for ABD transmission [1]. An individual's activity space is critical to determining an individual's exposure to *Ae. aegypti* across space and time. Variation in human-mosquito contacts influence how rapidly and efficiently the ABD can disseminate within a community and contribute to a sustained epidemic [16]. Consequently, such contact heterogeneities with *Ae. aegypti* resulting from variability in activity space overlap have explained why house-to-house movement drives dengue virus transmission [8] and why spatiotemporally unstructured movement routines can facilitate pathogen persistence [17, 18].

Current understanding of human mobility is limited at the individual level. A recent study from

Iquitos, Peru, used a seasonally forced transmission model to determine that individuals sharing a location, specifically houses, within a DENV-infected individual's activity space significantly increased risk of ABD infection [8]. The houses that participants visited in the study were of relatives and friends, and overlapping in residential locations were not random [8]. The findings from this study may be explained by the social proximity theory. The theory postulates that individuals with closer social distances, such as family members or friends, have a higher probability of overlapping frequented locations and therefore have a higher infection risk [1]. Furthermore, the theory states that risk decreases as social distance increases until it reaches a baseline infection risk of a situation in which an infectious mosquito travels between houses [1]. Social proximity defines the frequency of overlapping activity space, influences infection patterns of ABD's, and ultimately raises questions whether family-wide sharing of activity space can facilitate ABD transmission.

Studies testing the social proximity theory have been limited to empirical evidence for directly communicable diseases such as influenza, or theoretical modeling for ABD's [1, 19-22]. None have linked the implications of family-wide activity space overlap as a potential risk factor for ABD transmission. Here, we present the findings from our study which utilized semi-structured interviews in the Yucatán state of Mexico to examine overlap of family member activity space as a potential risk factor for vector-borne disease transmission. Data collected in this study was used to quantify movement characteristics of residents in the state. Using these characteristics, we determined shared locations amongst family members to produce linear regression models that associated the frequency of shared locations with the prevalence of acute ABD infection or history of ABD infection. The findings from this study will aid resource-limited public health programs to better develop ABD surveillance models and efficiently implement vector and disease control.

Methods:

<u>Hypothesis</u>

Vector-borne disease transmission heterogeneities are heavily driven by variances in socially driven human mobility behaviors. While previous studies have identified the significance of these factors for individuals, attempts to understand the impact of shared human activity space amongst family members have largely remained theoretical. We hypothesize that overlapping activity space amongst family members with a family member with previous history with ABD

infection or with acute infection is correlated with the number of individuals with previous history with ABD infection or with acute infection in the household. We also hypothesize that retrospective family members will overlap in locations that are associated with further social distances (i.e. marketplaces and healthcare centers) whereas prospective family members will overlap in locations that are associated with closer social distances (i.e. residential locations and churches).

<u>Study Area</u>

The study site was situated around Mérida, the capital of the Mexican state of Yucatán. Merida is the largest city in the State and of the Yucatán peninsula; it has a population of approximately 890,000 according to the 2015 census reports. Merida is surrounded by smaller municipalities such as Ticul, a small town of 34,000 persons located 100 km south of Merida, and Progreso, a popular resort destination for local tourists that has a population of 54,000 individuals (Figure 1). These three cities were chosen due to varying dengue transmission rates; Merida experiences the highest dengue transmission, followed by Ticul and Progreso [23]. Weather is consistent throughout the three cities. Temperatures range from 29°C in December to 34°C in July, and the rainy season from May through October coincides with the peak of the dengue transmission season between July and November. These conditions make Mérida a significant contributor to dengue virus transmission in the Yucatán peninsula region [24].

<u>Study Design</u>

Human Subjects (Ethics) Approval

Protocols for processing and analyzing data were approved by Emory University's ethics committee under protocol ID: IRB00088659. The protocol was also approved by the Ethics and Research Committee from the O´Horan General Hospital from the state Ministry of Health, Register No. CEI-0-34-1-14. See Appendix I.

Semi-Structured Interviews (SSI)

The goal of the SSI is to capture temporal and positional information of a participant's activity space. The SSI's were divided into two groups: Retrospective movement surveys (RMS's) and prospective movement surveys (PMS's). RMS's were conducted to a reference case, and his or her family members, who was diagnosed as acutely infected with any ABD. RMS's contain details

of places visited on a routine basis over the previous two weeks from the interview date. Conversely, PMS's were conducted to a reference case along with family members, who was not diagnosed with ZIKV, CHIKV, nor DENV by the time of the interview. PMS's contain details of locations they frequent routinely, without a pre-defined time window of occurrence. When respondents were not familiar with the exact address of a location, they were asked by the interviewer to call someone at the location to retrieve the address; the interviewer assisted the respondents with maps and calendars of public events when necessary. If the address was still unknown, the participant was asked to give the best estimate of the address by providing street intersections, nearby landmarks, or the establishment name. The results from the surveys were recorded and later geocoded for further analysis [25, 26].

Methods of Data Collection

Participants in our study were enrolled as part of Familias sin Dengue (FSD), a dengue surveillance and monitoring project run by Universidad Autonoma de Yucatán (UADY) in conjunction with the Yucatán State public health department.

Reference cases were identified through FSD surveillance. A random selection of elementary, middle, and high schools in the three cities followed by another random selection was conducted to identify minors (i.e. individuals 18 years old or younger); this was performed to reduce selection bias. Information, such as home address, phone number, and ages, were collected to help derive a proportional number of students from each grade from the same school and to contact the family for consent. Blood samples were taken once informed consent paperwork was signed, and seroprevalence results were directly delivered to the child's family. Adult reference cases (i.e. over 19 years old) were identified through primary care clinic visits. Qualifying adults were randomly chosen with respect to age and sex distribution of the locality. Blood samples were taken once informed consent paperwork was signed, and seroprevalence results were reported by phone or home visits. Both minor and adult patients were recruited for the SSIs when seroprevalence results were reported. These individuals were identified either as positive retrospective (i.e. acutely infected with zika, chikungunya, and dengue) or as positive prospective (i.e. not acutely infected, but has previous infection with mosquito-borne illnesses). It was possible to have multiple reference cases per family. When a family consented to partake in the study, the family was matched with a similar family within close vicinity.

Inclusion criteria for retrospective interviews were: 1) The participants must live in the same household of the index reference, 2) must have given consent to be surveyed and if a minor, consent from a responsible adult, 3) be diagnosed with ABD during the period in which the survey is administered, and 4) not hospitalized due to the infection. Inclusion criteria for prospective interviews were: 1) The participants must live in the same household of the index reference, 2) must have given consent to be surveyed and if a minor, consent from a responsible adult, and 3) not be diagnosed with ABD during the period in which the survey is administered. Prospective and retrospective interviews were conducted once the selected families agreed to participate in the study.

A dataset describing an individual's activity space was conglomerated after the conclusion of the interview process. Each row denoted the family code, location type, location address, duration and frequency of visits to each location, and the location's GPS coordinates that he or she visits on a routine basis. Another dataset included the home address associated with the individual along with the time spent at home each day of the week.

Statistical Analyses

Exploratory analyses were conducted using R Studio (version 1.0.153) and cumulative probability distributions were produced using MatLab R2017a (version 9.2.0.556344). Families were defined as a group of individuals sharing the same home address that contains two family members or more. Age groups were divided as follows: 0-9, 10-17, 18-33, 34-44, and 45+. Locations were defined as places visited outside the participant's home identified in the SSI's. Location types were categorized as follows: education, church, marketplace, residence, restaurants, health, work, and other locations. A Kolmogrov-Smirnov Test was performed to determine whether the cumulative probability distributions were significantly different from one another. Both time spent at a given location and time spent at home were parameterized as hours spent at a location per month. Google Earth was used to identify GPS coordinates of the households and the reported locations. When locations could not be identified based on the addresses provided, higher level addresses were recorded as a proxy. Linear regression was used to model the relationship between family member activity space overlap and the prevalence of positive prospective or positive retrospective individuals within the family.

Results:

<u>Overall Descriptive Statistics</u>

A total of 962 individuals ages 1 or older (Merida: 630 (65.5%), Ticul: 88 (9.1%), Progreso: 244 (25.4%)) and 229 families (Merida: 155 (67.7%), Ticul: 57 (24.8%), Progreso: 17 (7.4%)) were recruited for this study as part of the long-term prospective longitudinal cohort between June 2016 and May 2017. Distribution of ages were similar amongst the three cities (Figure 2).

<u>Prospective Cohort Descriptive Statistics</u>

There were 852 prospective interviews (Merida: 562 (66.0%), Ticul: 58 (6.8%), Progreso: 228(26.8%)) were conducted. This was equivalent to 208 family prospective interviews (Merida: 140 (67.3%), Ticul: 55 (26.4%), Progreso: 13 (6.3%)). Average number +/- SD of family members in a household in the prospective group was 3.75 +/- 1.36 (Merida: 3.60 +/- 1.29, Ticul: 3.92 +/- 1.89, Progreso: 4.091 +/- 1.35). 62.5% of families in the prospective cohort had at least one positive member.

<u>Retrospective Cohort Descriptive Statistics</u>

There were 113 individual retrospective interviews (Merida: 66 (58.4%), Ticul: 30 (26.5%), Progreso: 12 (10.6%)). This was equivalent to 21 family retrospective interviews (Merida: 15 (71.4%), Ticul: 2 (9.5%), Progreso: 4 (19.0%)), and the average number of family members in a household in the retrospective group was 3.182 +/- 2.72 (Merida: 2.67 +/- 0.62, Ticul: 5.50 +/- 6.35, 2.00 +/- 0). Only 36.3% of the families in the retrospective cohort had at least one positive member.

<u>Descriptive Statistics of Prospective Human Movement</u>

Interview data was analyzed to understand the landscape of routine movement of individuals living in these three cities. Figure 3 shows household locations and destinations visited by individuals on a routine basis. Most reported locations that were routinely visited were inside the city of residence (Merida: 94.5%, Ticul: 88.16%, Progreso: 93.5%) (Figure 4). Most frequented city for Ticul (39.47%) and Progreso (50.94%) residents was Merida. Uman was the most frequented city for Merida residents (11.36%). A nonsignificant relationship was identified between distance traveled yearly and percent of locations visited outside of Merida (Figure 5). Residences (24.49%), followed by schools (15.79%) and marketplaces (15.46%), were the most frequently visited

location types (Figure 6). The relative frequency of visits to each location type differed by age group. Those 0-9 years old routinely visited residential and educational locations (25.5% and 25.2% respectively), while those 45 and older routinely visited residential locations and market (28.6% and 26.0% respectively) (Figure 7). The kernel density estimate distribution for the number of locations shows that most individuals routinely visited at most 5 locations for all three categories of interest (by city, age group, and gender) (Figure 8).

Human Movement Characteristics Analysis

Several facets of human movement (i.e. distance traveled, time spent at a location away from home, time spent at home) were visualized via cumulative probability distributions to describe movement characteristics. Probability of distance traveled outside an individual's home $P(\Delta D)$ followed an exponential decay (Figure 9). Distributions for distance traveled were significantly different by age groups for all comparisons ($p < 0.001$) apart from those made between 18-33 and 34-44 year-olds; 18-33 and 45+ year-olds; and 34-44 and 45+ year-olds (Table 1a). Those between the ages of 0-9 traveled the least, whereas individuals between 18-33 years old traveled the furthest (Figure 9). Distributions for distances traveled were significantly different by gender ($p < 0.001$), and males traveled further than females do (Table 1b, Figure 9). Furthermore, distributions for distances traveled were significantly different by city of residence ($p < 0.001$), and those from Merida traveled the furthest, while those from Ticul traveled the closest (Table 1c, Figure 9). The cumulative probability of time spent at a given location $P(\Delta T)$ followed an exponential decay as well (Figure 10). Distributions for time spent at a location were significantly different by age groups for all comparisons ($p < 0.001$) apart from those made between 18-33 and 34-44 year-olds; 18-33 and 45+ year-olds; and 34-44 and 45+ year-olds (Table 2a). 10-17 year-olds spent the most time at a given location, while those over 45 years old spent the least time at a given location (Figure 10). Distributions for time spent at a location were significantly different by gender ($p < 0.001$, Table 2b); males spent more time at a given location than women do (Figure 10). Distributions for time spent at a location by city of residence were significantly different apart from the Merida-Ticul comparison (Table 2c); people living in Ticul spent the most amount of time at a given location, while those living in Progreso spent the least amount of time at a given location. Time spent at different location types did not significantly differ from each other (Table 2d). As with the $P(\Delta D)$ and $P(\Delta T)$ distributions, the cumulative probability distribution of time residents spent at home $P(\Delta Th)$ followed an exponential decay (Figure 11). Time residents spent at home

by age group were also significantly different by age groups for all comparisons apart from those made between 18-33 and 34-44 year-olds; 18-33 and 45+ year-olds; and 34-44 and 45+ year-olds (Table 3a). 0-9 year-olds spent time at home the longest, whereas 34-44 year-olds spent time at home the least (Figure 11). Distributions for time spent at home were significantly different by gender; women spent more time at home than men did (Table 3b, Figure 11). Distributions for time spent at home by city were significantly different apart from the Merida-Progreso (Table 3c). Those from Progreso spent the most time at home, whereas those from Merida spent the least time at home (Figure 11).

Activity Space Overlap Analysis

Most families overlapped in their activity space less than 50% of the time (Figure 12). In the prospective cohort, 23 out of 43 families (53.5%) that had over 50% of family activity space overlap had at least one positive prospective individual. In the retrospective cohort, 2 out of 6 families (33.3%) that had over 50% of family activity space overlap had at least one positive retrospective individual. The location types overlapped most frequently amongst family members in the prospective cohorts were residential (25.55%), educational (17.91%), and marketplaces (15.9%). On the other hand, the location types overlapped most frequently amongst family members in the retrospective cohorts were educational (24.52%), residential (22.64%), and marketplaces (16.98%). A linear regression was used to model the relationship between percent family overlap with positive case members and percent positive case count within the household for both cohorts. The relationship between percent family overlap with index case members and percent index case count within household was significant for both prospective and retrospective groups despite the medium to low correlation coefficient (prospective: $r^2 = 0.032$, $p < 0.001$; retrospective: $r^2 = 0.532$, $p < 0.001$) (Figure 13). Family size was tested as an additional variable in the linear regression to determine the impact of the variable in further predicting percent positive prospective and positive retrospective count within household. The family size variable was not significant ($p > 0.1$) in either models (Table 4a). In addition, the interaction term was also not significant ($p > 0.1$), and the inclusion of the interaction term reduced the correlation coefficient in the prospective model, but increased the coefficient in the retrospective model (Table 4b).

Discussion:

The distribution of human movement variables analyzed in this study were consistent with those

found in a similar study conducted in Iquitos [6]. In both settings, the probability distributions distance traveled, time spent at a location away from home, and time spent at home followed an exponential decay. This could possibly be explained by a distance decay, in which the probability of an event occurring decreases as distance increases [27]. However, the decay of distance traveled differed between the two sites. In Iquitos, 81% of all reported movement occurred within 1 km of the house of residence. In our study, nearly 35% of all reported movement occurred within 1 km of the house of residence. This may be attributed to the context of the study sites. Iquitos is a geographically isolated port city located in the Amazon forest, whereas Merida is much larger in size and, as the capital city of Yucatan, has highways connecting other towns in the State. Despite the infrastructure available for travel, the participants in our study mainly exhibited in-city movement more than out-of-town movement. This represents a standard metropolitan network movement pattern, in which most movement activity takes place in the city, with out-of-town activity limited to special occasions. Often, individuals will travel to the urban hub for such instances, as seen in our study results in which those living in Progreso and Ticul were most likely to visit Merida if they were to leave city limits. This is highly possible since Merida is the capital of the Yucatan State and has a concentration of government and healthcare services as well as other leisure activities. This source-sink movement pattern is in agreement with findings from Cairns, Australia, where there was high mobility from smaller coastal towns to Cairns, the most populous city in the region [28].

The overlap of activity spaces amongst family members in the prospective cohort at residential, educational, and marketplace settings was consistent with our hypothesis, given that family members have a higher probability of spending time together outside of their shared home at a relative's or friend's house. Educational and marketplace settings were also within expectation; however, it was more likely for one adult to accompany his or her child to each location. Therefore, while there was an overlap in activity space amongst family members, the number of family members participating in the overlap at educational locations and marketplaces were lower than that of residential settings. Overlap of activity spaces in educational, residential, and marketplace locations amongst family members in the retrospective cohort slightly deviates from our hypothesis. We expected family members to visit locations associated with locations where they may meet acquaintances or the public at-large (i.e. those with greater social distances) to meet their health and daily minimum needs. Observations in both cohorts may be due to the greater number

9

of educational, residential, and marketplace locations visited in the total cohort as compared to any other location types, as shown in Figure 6.

The results of the linear regression were as expected for both cohorts. However, the association between the variables were low, particularly for the prospective cohort. While the correlation coefficient was greater for the retrospective cohort, the lack of data prevented us from making any definitive conclusions as to whether overlap with an acutely infected family member was an indicator for presence of an ABD-infected member. The family size variable was weaker than expected within the linear regression model. We believed that increased household size would increase the likelihood of an individual bringing an ABD into the household. Unfortunately, the limited data in our study prevented us from exploring further potential explanatory variables for the presence of a positive individual.

Other limitations impact the validity of this study. Recall bias and reporting bias, in addition to input errors, may have affected the data collected through the SSI. We are confident that the data collected in the study accurately represents the locations that the participants frequent since they are part of their routine for the prospective cohort participants and for the retrospective cohort, the locations are visited within the last two weeks. Furthermore, we are confident with this collection method since previous studies have reported the high correlation between the data collected between SSI's and GPS data [6, 29]. While there were 208 prospective families, there were only 21 retrospective families in the study. This limit in the families in the retrospective cohort prevented us from analyzing movement characteristics and producing meaningful data. Finally, we were limited in higher-level social network data that allowed us to determine social distances between family members and their relatives along with their friends. This data, if available, would have allowed us to directly relate the findings of our study to that of Reiner et al. [1].

Conclusion and Recommendations:

Further investigation of high overlap locations amongst the index case and family members is crucial in better understanding the transmission heterogeneities of *Ae. Aegypti*-borne diseases. Several studies using empirical evidence have highlighted the influence of activity space overlap on driving dengue infection patterns in urban environments such as Merida. However, these studies have focused on house-to-house transmission of the arbovirus[8, 28]. This study has attempted to

fill in this knowledge gap by empirically testing the social proximity theory with movement data collected from three cities in the Yucatan state of Mexico. In short, our study has shown that family members, who have short social distances, have increased risk of *Ae. Aegypti*-borne disease infection with a greater degree of activity space overlap with other family members who have history of ABD's or are acutely infected with the arboviruses. While the association was far from perfect, our work lays foundation for future work to analyze other factors that may further explain the association of familial activity space overlap as an indicator for ABD risk. Exposure to mosquito-borne diseases can occur anywhere infectious mosquitos are present [30-32]. While *Aedes* mosquitos prefer to aggregate near or within homes, they have been found in other location types, including schools, churches, and factories [32, 33]. Non-residential locations can possibly be a gateway for a family member to contract ABD's and then introduce the virus into the family at the household. This is a key consideration since there may be an ideal ratio of family-wide activity space overlap that allows for ABD's to propagate efficiently. Too little of overlap and the risk of ABD transmission between family members is decreased. On the other hand, too great of overlap and the ABD is not introduced to a diversified social network, making sustained propagation more difficult within the approximately week-long viremic period amongst the infected. With this in mind, more empirical studies must be performed to 1) further support the social proximity thesis in other urban settings and to 2) further understand the implication of family-wide shared activity space outside the household. A more comprehensive definition of overlapping activity space can better inform us of *Ae. Aegypti*-borne disease transmission dynamics to aid disease prevention and vector control programs.

References:

1. Reiner Jr, R.C., S.T. Stoddard, and T.W. Scott, *Socially structured human movement shapes dengue transmission despite the diffusive effect of mosquito dispersal.* Epidemics, 2014. **6**: p. 30-36.

2. Snow, J., *On the mode of communication of cholera.* 1855: John Churchill.

3. Dennis, A.M., et al., *Integration of Contact Tracing and Phylogenetics in an Investigation of Acute HIV Infection.* Sexually Transmitted Diseases, 2017.

4. Eames, K.T., et al., *Assessing the role of contact tracing in a suspected H7N2 influenza A outbreak in humans in Wales.* BMC infectious diseases, 2010. **10**(1): p. 141.

5. Bjorn-Mortensen, K., et al., *Extent of transmission captured by contact tracing in a tuberculosis high endemic setting.* European Respiratory Journal, 2017. **49**(3): p. 1601851.

6. Vazquez-Prokopec, G.M., et al., *Using GPS technology to quantify human mobility, dynamic contacts and infectious disease dynamics in a resource-poor urban environment.* PloS one, 2013. **8**(4): p. e58802.

7. Vazquez-Prokopec, G.M., et al., *Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus.* International journal of health geographics, 2009. **8**(1): p. 68.

8. Stoddard, S.T., et al., *House-to-house human movement drives dengue virus transmission.* Proceedings of the National Academy of Sciences, 2013. **110**(3): p. 994-999.

9. Kuno, G., *Review of the factors modulating dengue transmission.* Epidemiologic reviews, 1995. **17**(2): p. 321-335.

10. Harrington, L.C., et al., *Dispersal of the dengue vector Aedes aegypti within and between rural communities.* The American journal of tropical medicine and hygiene, 2005. **72**(2): p. 209-220.

11. Getis, A., et al., *Characteristics of the spatial pattern of the dengue vector, Aedes aegypti, in Iquitos, Peru*, in *Perspectives on Spatial Data Analysis*. 2010, Springer. p. 203-225.

12. MACIEL- DE- FREITAS, R., et al., *Influence of the spatial distribution of human hosts and large size containers on the dispersal of the mosquito Aedes aegypti within the first gonotrophic cycle.* Medical and veterinary entomology, 2010. **24**(1): p. 74-82.

13. Reiter, P., *Oviposition, dispersal, and survival in Aedes aegypti: implications for the efficacy of control strategies.* Vector-Borne and Zoonotic Diseases, 2007. **7**(2): p. 261-273.

14. May, R.M., *Network structure and the biology of populations.* Trends in Ecology &

Evolution, 2006. **21**(7): p. 394-399.

15.   Vazquez-Prokopec, G.M., et al., *Coupled heterogeneities and their impact on parasite transmission and control.* Trends in parasitology, 2016. **32**(5): p. 356-367.

16.   Perkins, T.A., et al., *Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission.* PLoS computational biology, 2013. **9**(12): p. e1003327.

17.   Adams, B. and D.D. Kapan, *Man bites mosquito: understanding the contribution of human movement to vector-borne disease dynamics.* PloS one, 2009. **4**(8): p. e6763.

18.   Stoddard, S.T., et al., *The role of human movement in the transmission of vector-borne pathogens.* PLoS neglected tropical diseases, 2009. **3**(7): p. e481.

19.   Danon, L., T. House, and M.J. Keeling, *The role of routine versus random movements on the spread of disease in Great Britain.* Epidemics, 2009. **1**(4): p. 250-258.

20.   Keeling, M.J., et al., *Individual identity and movement networks for disease metapopulations.* Proceedings of the National Academy of Sciences, 2010. **107**(19): p. 8866-8870.

21.   Salathé, M., et al., *A high-resolution human contact network for infectious disease transmission.* Proceedings of the National Academy of Sciences, 2010. **107**(51): p. 22020-22025.

22.   Salathé, M. and J.H. Jones, *Dynamics and control of diseases in networks with community structure.* PLoS computational biology, 2010. **6**(4): p. e1000736.

23.   Pavía-Ruz, N., et al., *Seroprevalence of Dengue Antibodies in Three Urban Settings in Yucatan, Mexico.* 2018.

24.   Dantés, H.G., J.A. Farfán-Ale, and E. Sarti, *Epidemiological trends of dengue disease in Mexico (2000–2011): a systematic literature search and analysis.* PLoS neglected tropical diseases, 2014. **8**(11): p. e3158.

25.   Paz-Soldan, V.A., et al., *Assessing and maximizing the acceptability of global positioning system device use for studying the role of human movement in dengue virus transmission in Iquitos, Peru.* The American journal of tropical medicine and hygiene, 2010. **82**(4): p. 723-730.

26.   Paz-Soldan, V.A., et al., *Strengths and weaknesses of Global Positioning System (GPS) data-loggers and semi-structured interviews for capturing fine-scale human mobility: findings from Iquitos, Peru.* PLoS neglected tropical diseases, 2014. **8**(6): p. e2888.

27.   Brockmann, D., L. Hufnagel, and T. Geisel, *The scaling laws of human travel.* Nature,

2006. **439**(7075): p. 462.

28.   Vazquez-Prokopec, G.M., et al., *Combining contact tracing with targeted indoor residual spraying significantly reduces dengue transmission.* Science advances, 2017. **3**(2): p. e1602024.

29.   Perkins, T.A., et al., *Calling in sick: impacts of fever on intra-urban human mobility.* Proc. R. Soc. B, 2016. **283**(1834): p. 20160390.

30.   Adams, B., et al., *Cross-protective immunity can account for the alternating epidemic pattern of dengue virus serotypes circulating in Bangkok.* Proceedings of the National Academy of Sciences, 2006. **103**(38): p. 14234-14239.

31.   Honório, N.A., et al., *Spatial evaluation and modeling of dengue seroprevalence and vector density in Rio de Janeiro, Brazil.* PLoS neglected tropical diseases, 2009. **3**(11): p. e545.

32.   García-Rejón, J.E., et al., *Mosquito infestation and dengue virus infection in Aedes aegypti females in schools in Merida, Mexico.* The American journal of tropical medicine and hygiene, 2011. **84**(3): p. 489-496.

33.   Zarate-Nahon, E.A., et al., *Aedes aegypti mosquitoes at nonresidential sites might be related to transmission of dengue virus in Monterrey, Northeastern Mexico.* Southwestern Entomologist, 2013. **38**(3): p. 465-476.
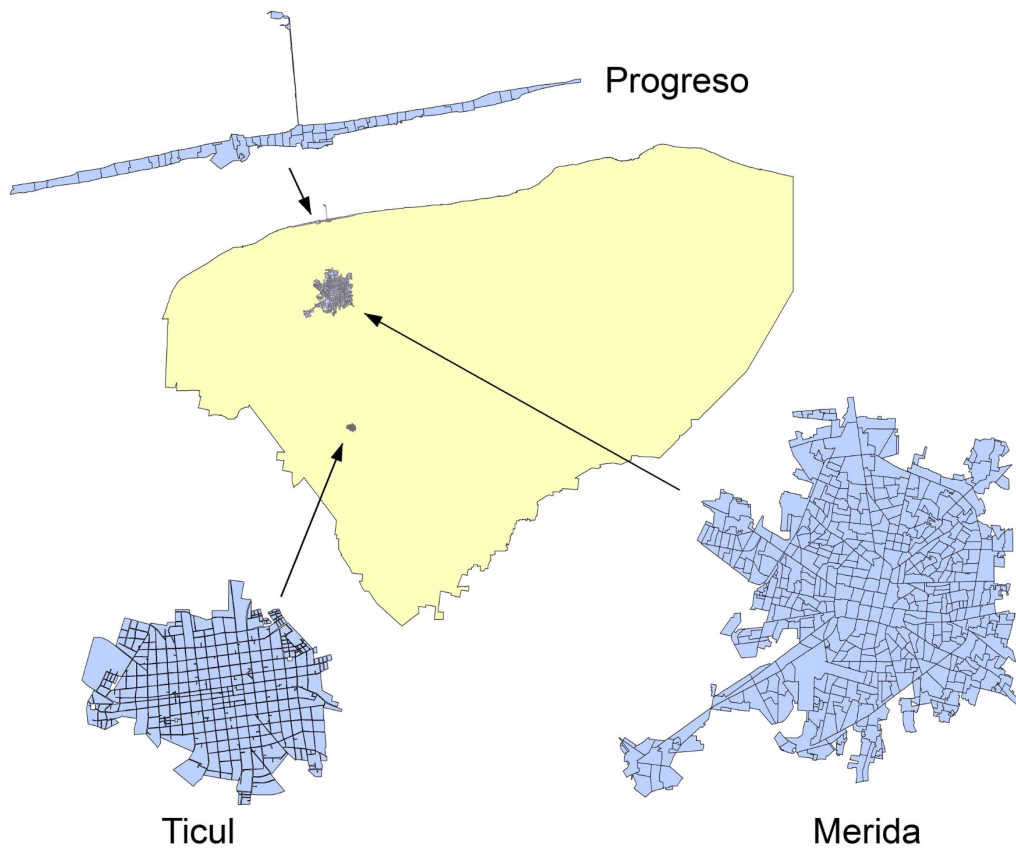
Tables and Figures:

Figure 1: Map of Yucatan State with cities analyzed in study.
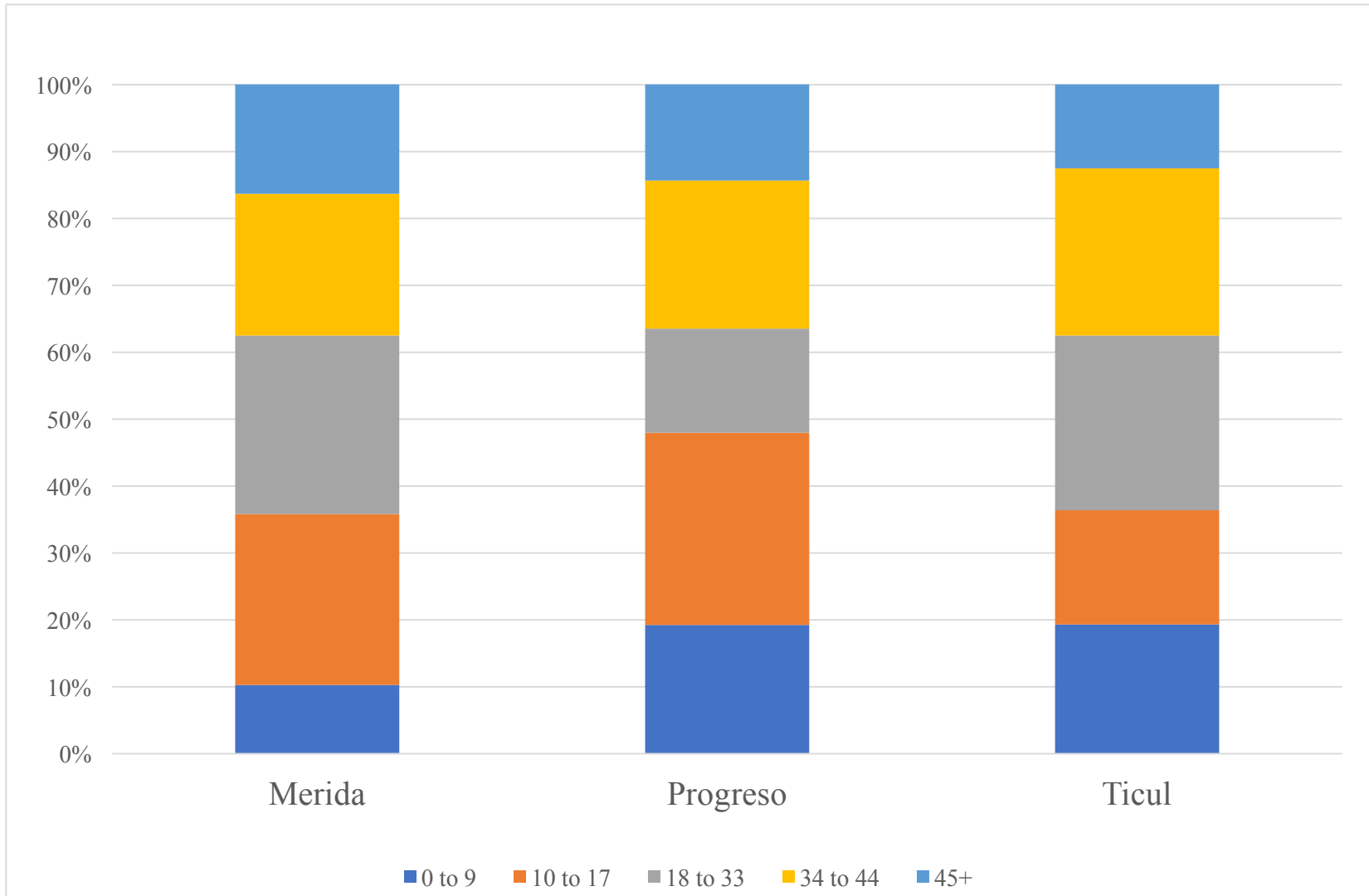
Figure 2: Age Structure of Study Participants. Distribution of ages appear to be similar amongst the three cities.
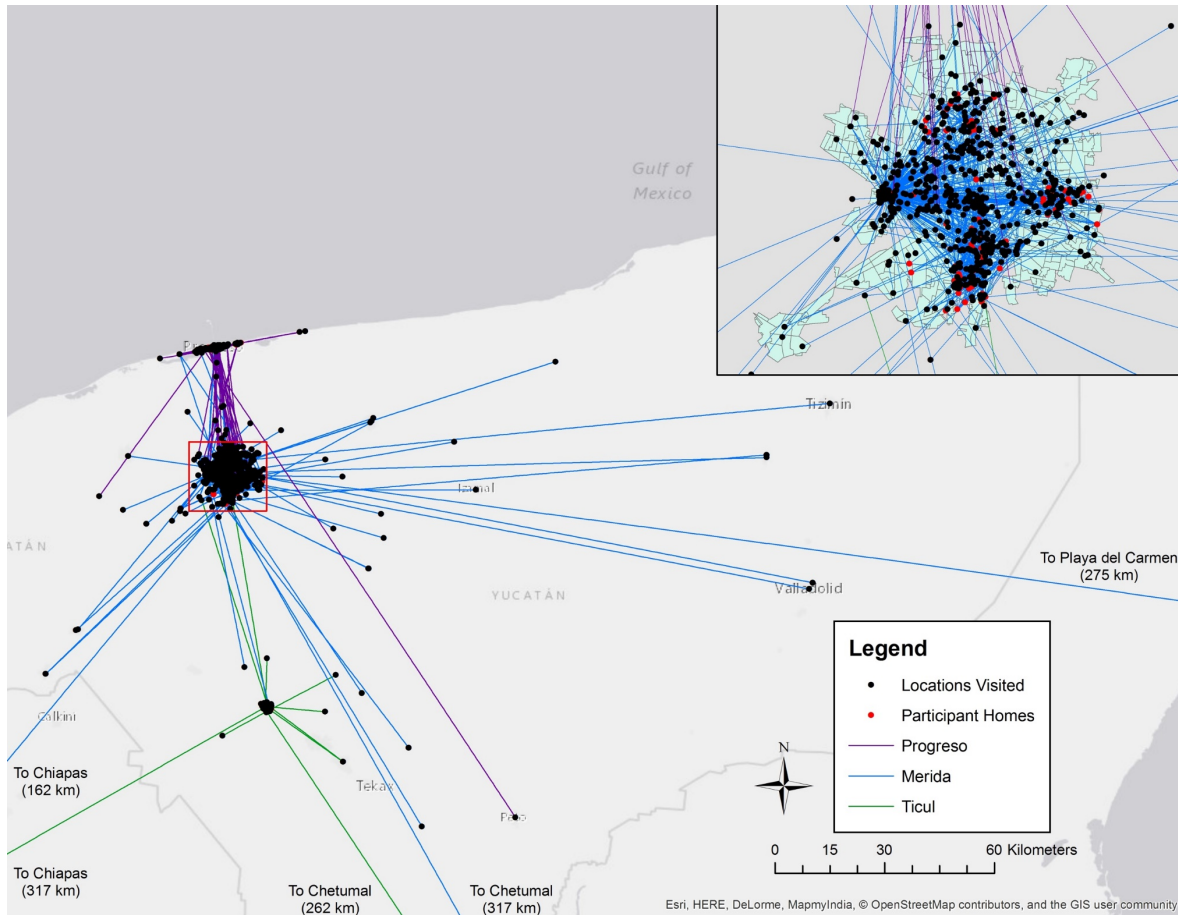
Figure 3: Map of Yucatan State with locations visited in relation to participant homes. Lines differ in color based on the participant's city of residence. Locations traveled to that were not captured within the map boundaries were labeled along the edges along with the distance from individual's residence.
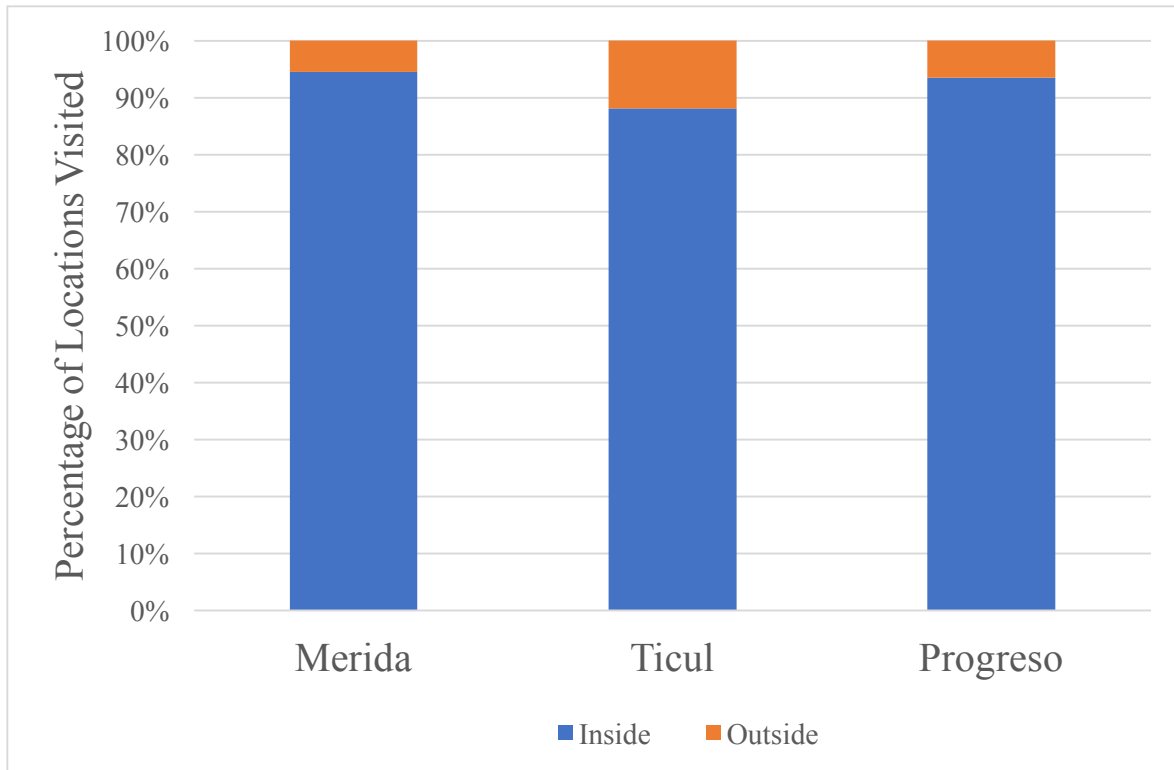
Figure 4: Percentage of locations traveled within city of residence. Most movement activity occurs within city of residence. This is typical of urban movement patterns, in which out-of-town movements occur only for special occasions.
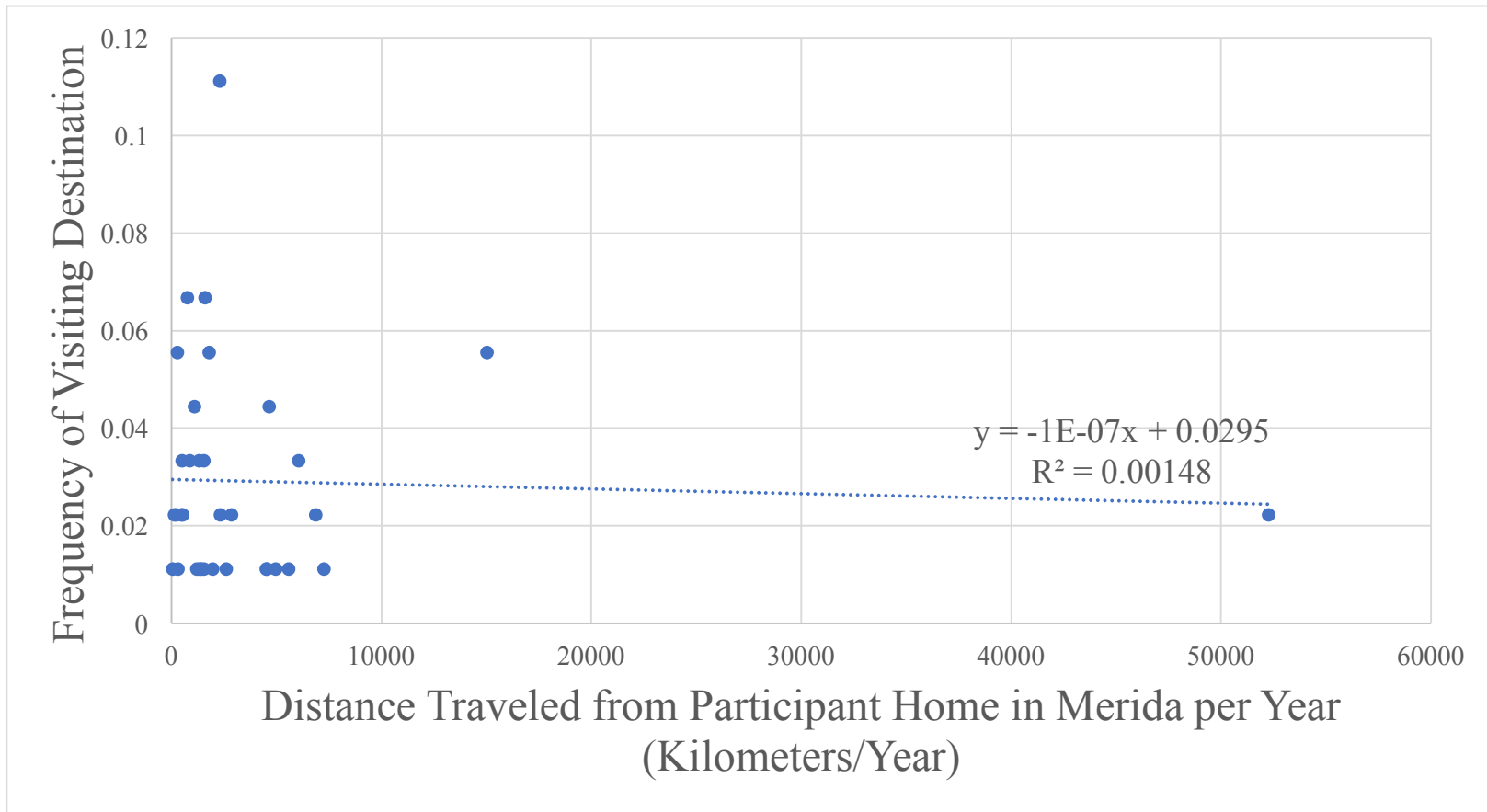
Figure 5: Likelihood of traveling outside of Merida. Each point in the graph represents a city visited by a participant residing in Merida. Distance traveled from participant home in Merida per year was calculated by averaging the distances traveled by a participant to the destination. 90 locations out of 1645 total locations visited by participants in the prospective cohort were located outside of Merida. Frequency of visiting destination was derived by dividing the number of instances that a participant traveled to the destination divided by the total (90).
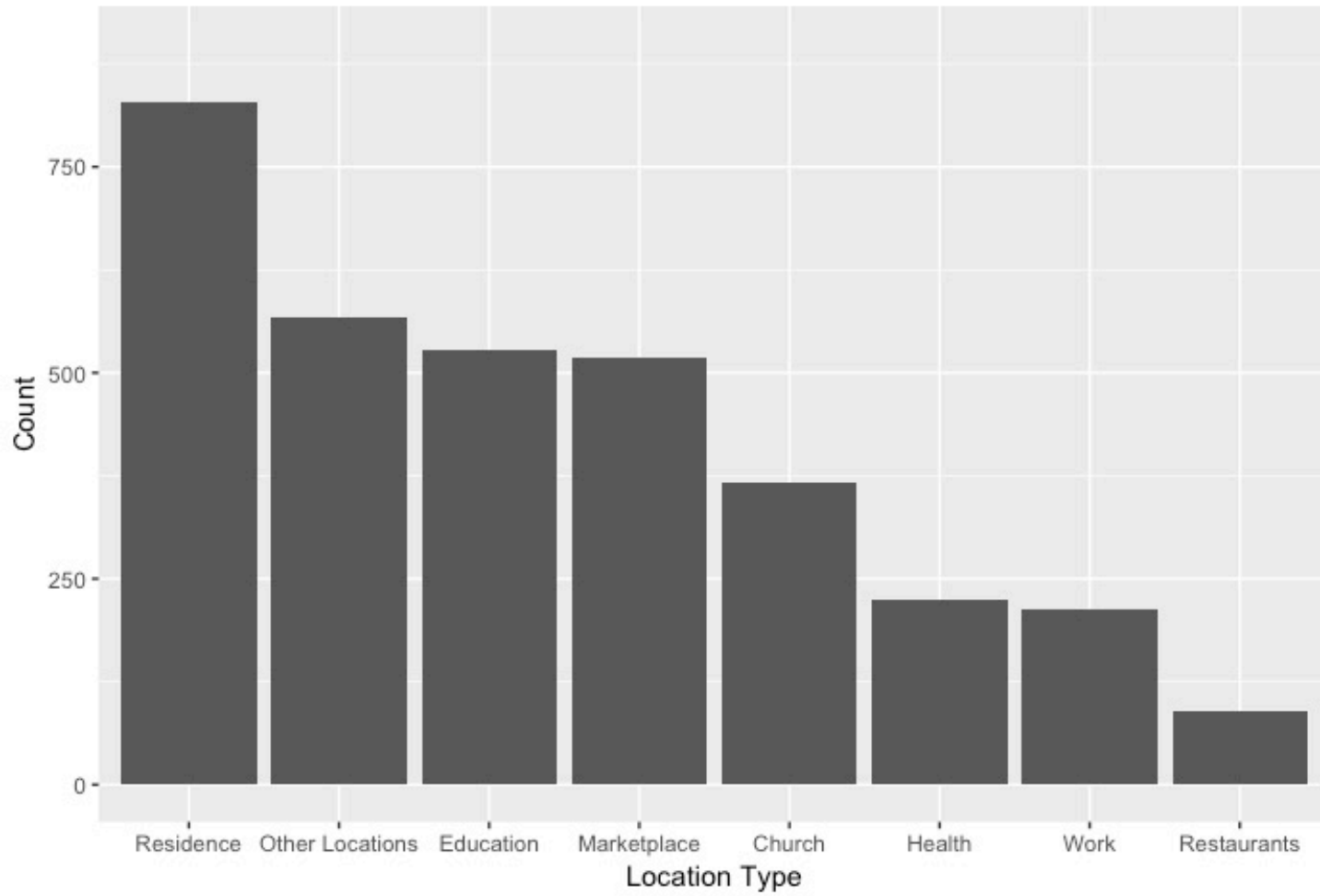
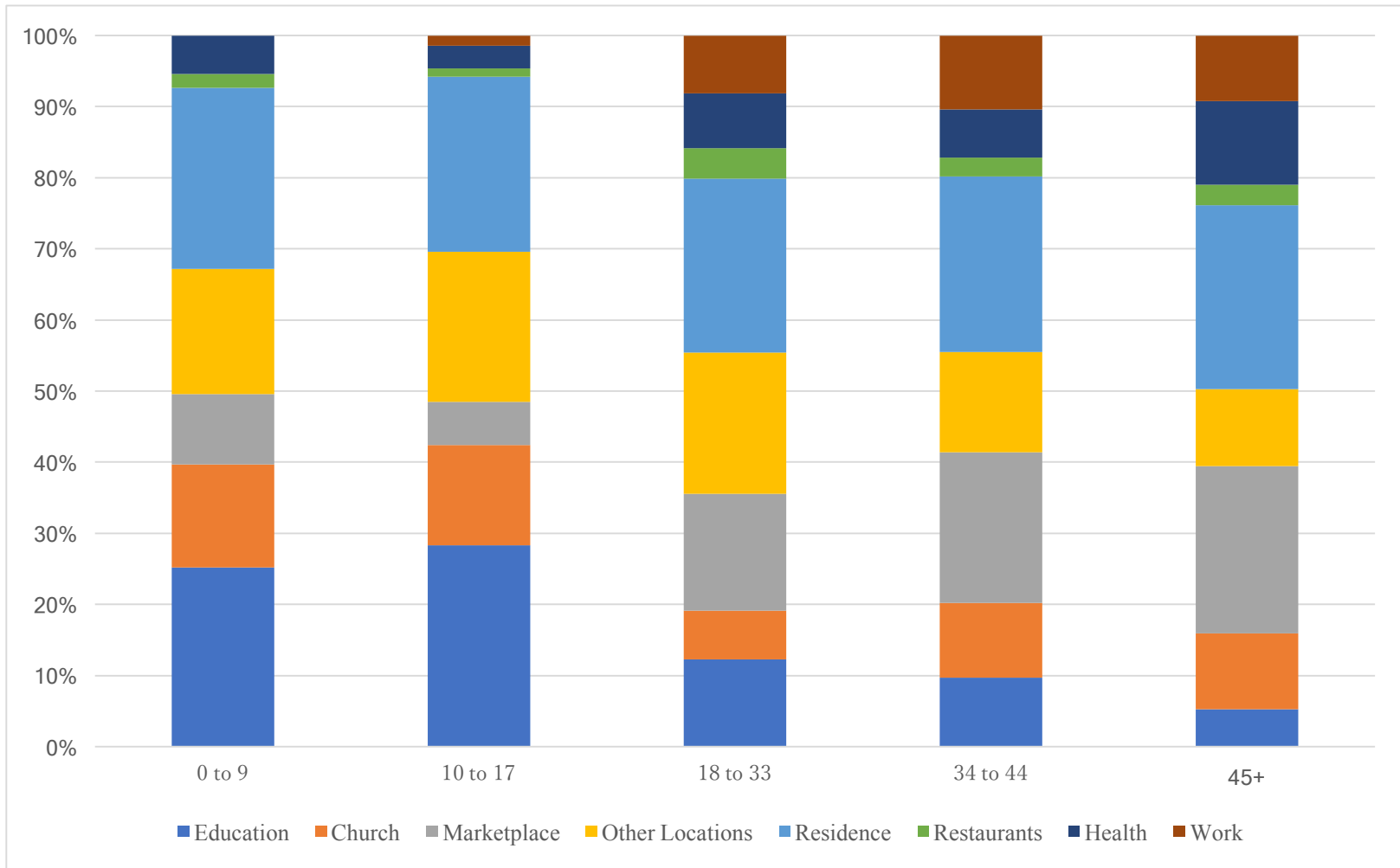**Figure 6: Location Type Visited by Participants.**

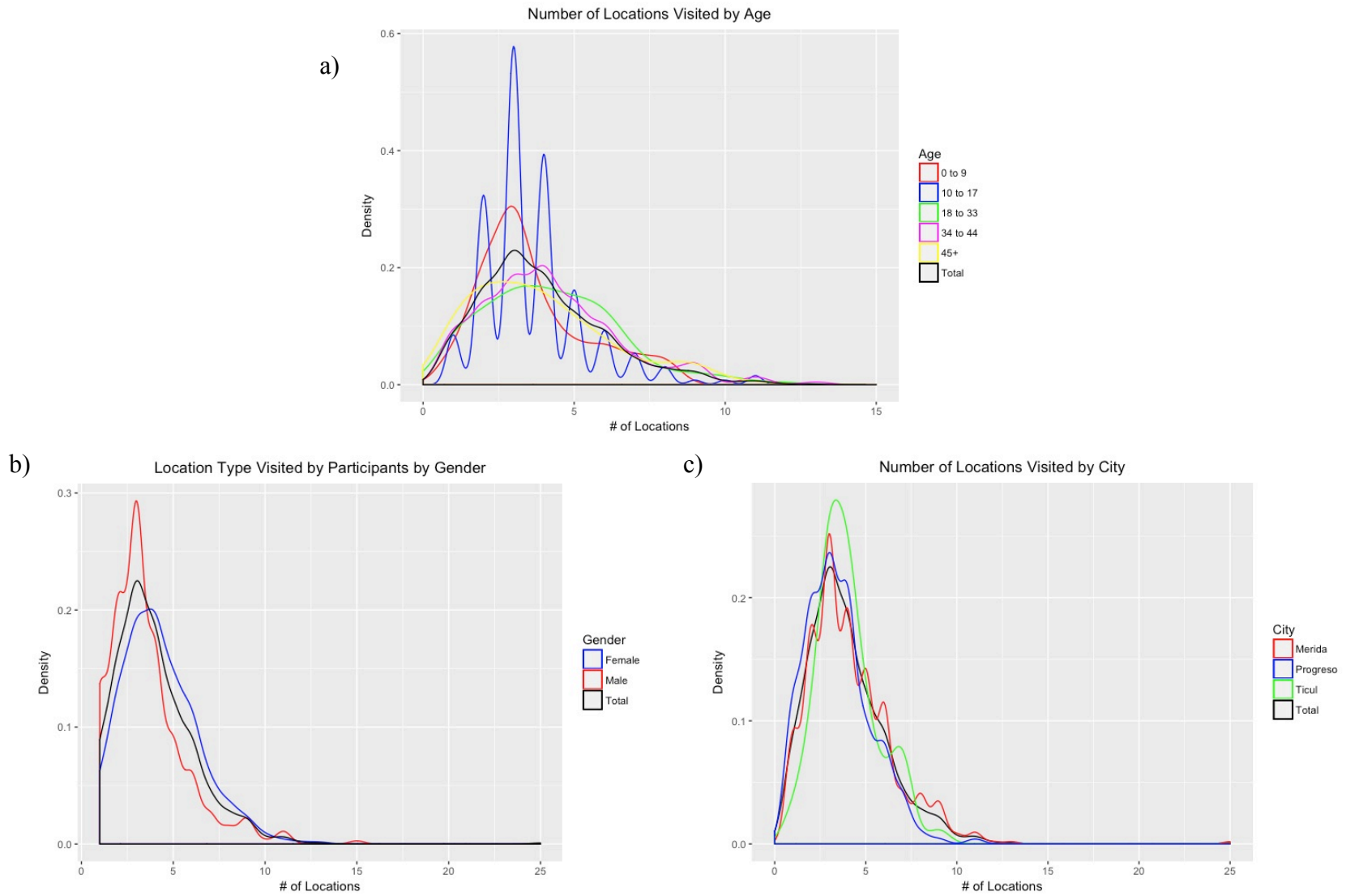**Figure 7: Types of locations visited by age.**

Figure 8: Relative density curve of number of locations visited by a) age, b) gender, and c) city of residence.

Figure 9: Cumulative probability distribution of distance traveled by a) age, b) gender, and c) city of residence. P(ΔD) represents the probability of movement outside an individual's home.
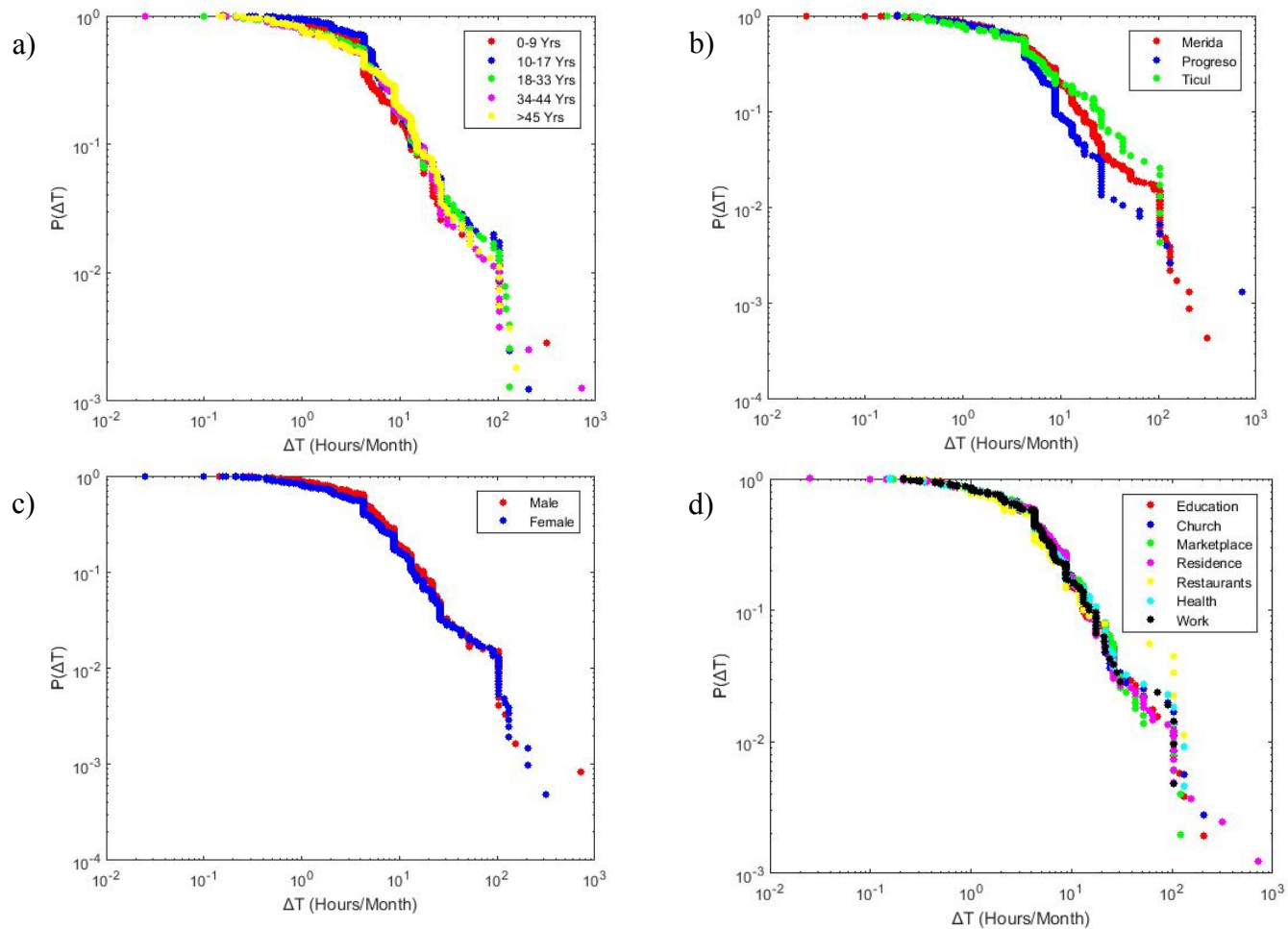
Figure 10: Cumulative probability distribution of time spent at a visited location by a) age, b) city of residence, c) gender, and d) location type. P(ΔT) represents the probability of time spent at a given location outside the participant's home.
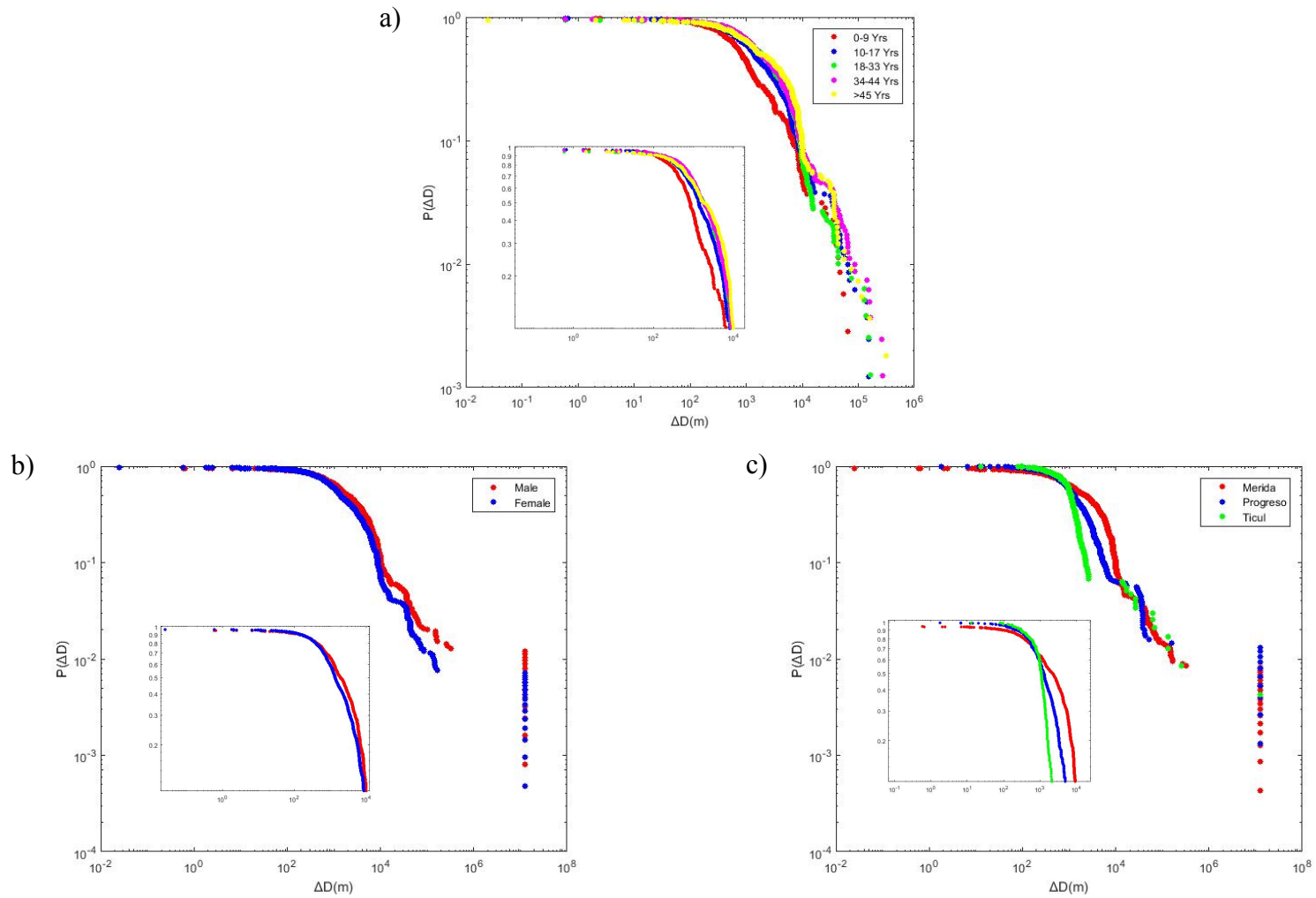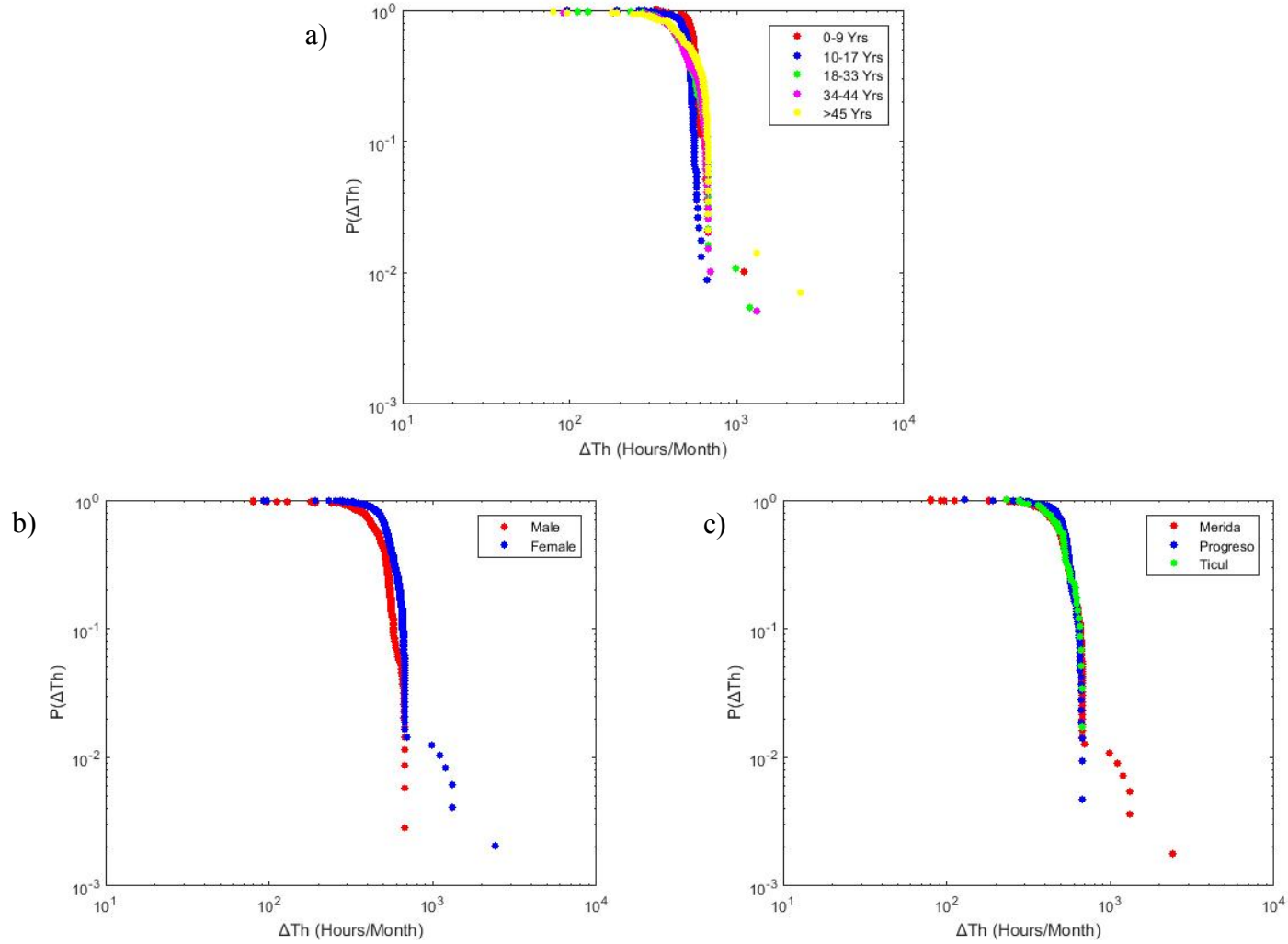
Figure 11: Cumulative probability distribution of time spent at home by a) age, b) gender, and c) city of residence. P(ΔTh) represents the probability of time spent at home.
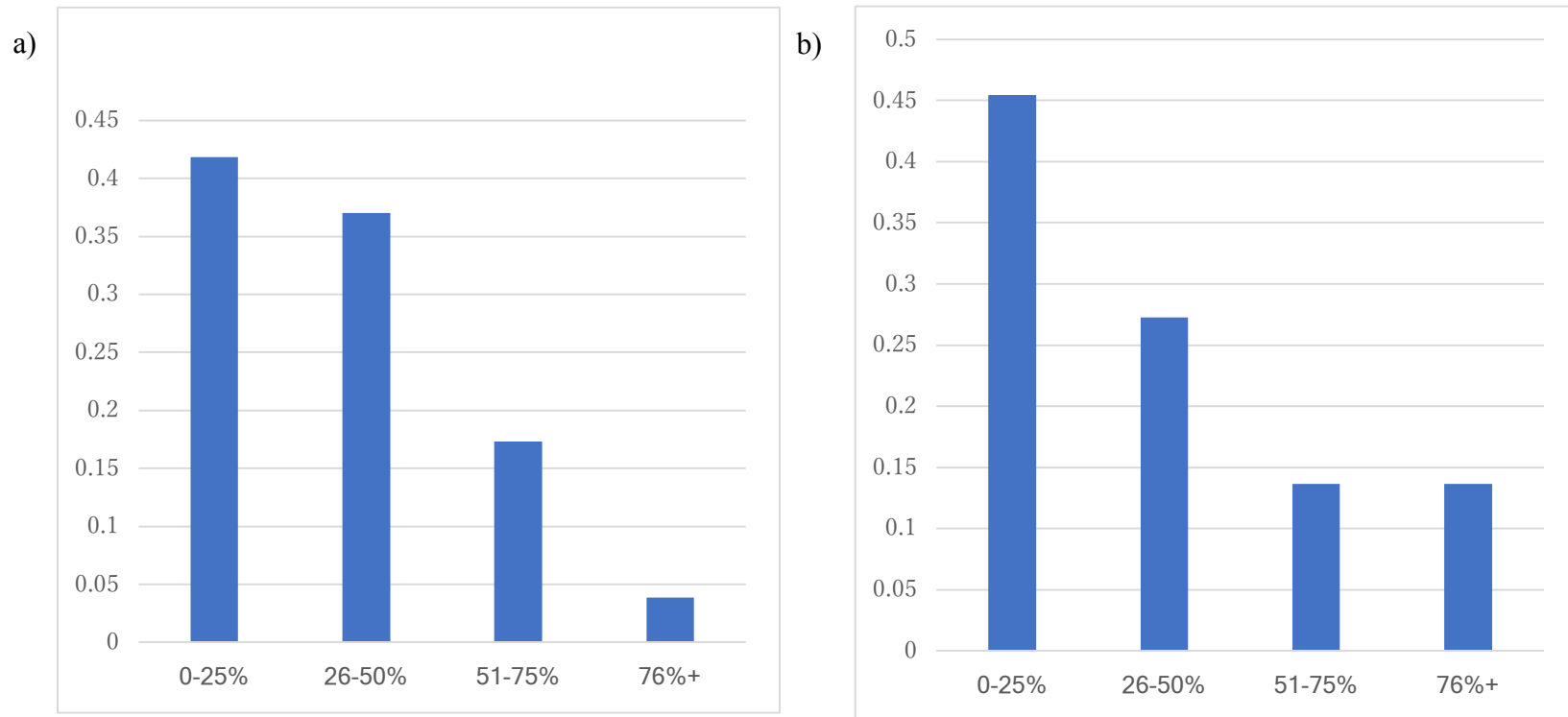
Figure 12: Bar plots of location overlap percentage for each household for a) prospective cohort and b) retrospective cohort.
Most families have activity space overlap less than 50% of the time for both cohorts.
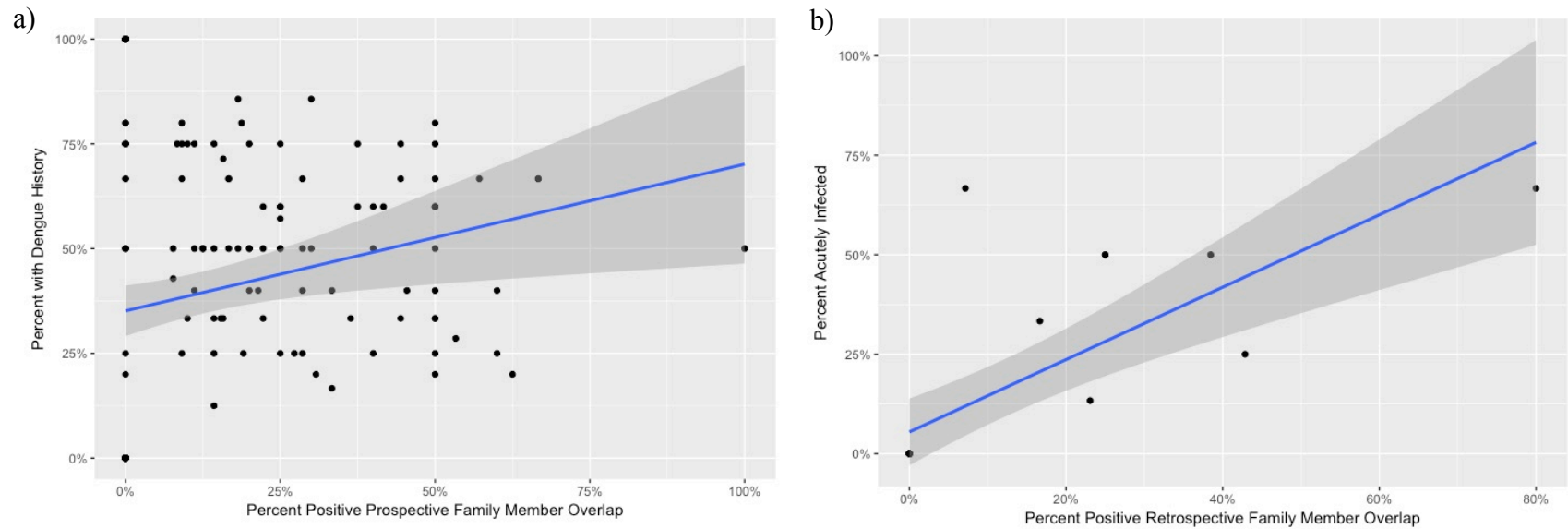
Figure 13: Linear regression between percent family overlap with positive members and percent positive member count in household for a) prospective cohort and b) retrospective cohort. There is a significant positive correlation between percent positive family member overlap and percent positive member count within household despite the medium to low correlation coefficient (prospective: $r^2 = 0.032$, $p < 0.001$; retrospective: $r^2 = 0.532$, $p < 0.001$).

	0-9	10-17	18-33	34-44	45+
0-9	-	D=0.187 $p <0.001$	D = 0.245 $p < 0.001$	D = 0.250 $p < 0.001$	D = 0.241 $p < 0.001$
10-17	-	-	D = 0.086 $p < 0.001$	D = 0.081 $p = 0.010$	D = 0.116 $p < 0.001$
18-33	-	-	-	D = 0.047 $p = 0.330$	D = 0.067 $p = 0.109$
34-44	-	-	-	-	D = 0.066 $p = 0.113$
45+	-	-	-	-	-

Table 1a: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of distance traveled by age group. The p-value associated with each test is the probability of the two sample groups not coming from a common distribution.

	Female
Male	D=0.081 $p < 0.001$

Table 1b: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of distance traveled by gender. The p-value associated with each test is the probability of the two sample groups not coming from a common distribution.

	Merida	Ticul	Progreso
Merida	-	D = 0.409 $p < 0.001$	D = 0.226 $p < 0.001$
Ticul	-	-	D = 0.235 $p < 0.001$
Progreso	-	-	-

Table 1c: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of distance traveled by city of residence. The *p*-value associated with each test is the probability of the two sample groups not coming from a common distribution.

	0-9	10-17	18-33	34-44	45+
0-9	-	D=0.227 $p <0.001$	D=0.107 $p < 0.001$	D=0.126 $p < 0.001$	D = 0.133 $p < 0.001$
10-17	-	-	D = 0.170 $p < 0.001$	D = 0.194 $p < 0.001$	D = 0.195 $p < 0.001$
18-33	-	-	-	D = 0.045 $p = 0.287$	D = 0.052 $p = 0.346$
34-44	-	-	-	-	D = 0.037 $p = 0.754$
45+	-	-	-	-	-

Table 2a: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of time spent at a location outside the household by age group. The *p*-value associated with each test is the probability of the two sample groups not coming from a common distribution.

	Female
Male	D=0.096 $p < 0.001$

Table 2b: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of time spent at a location outside the household by gender. The *p*-value associated with each test is the probability of the two sample groups not coming from a common distribution.

	Merida	Ticul	Progreso
Merida	-	D = 0.09	D = 0.125
		$p = 0.070$	$p < 0.001$
Ticul	-	-	D = 0.113
			$p = 0.022$
Progreso	-	-	-

Table 2c: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of time spent at a location outside the household by city of residence. The *p*-value associated with each test is the probability of the two sample groups not coming from a common distribution.

	Education	Church	Marketplace	Residence	Restaurants	Health	Work
Education	-	D = 0.061 p = 0.4157	D = 0.062 p = 0.2832	D = 0.053 p = 0.346	D = 0.089 p = 0.586	D = 0.062 p = 0.602	D = 0.045 p = 0.926
Church	-	-	D = 0.043 p = 0.835	D = 0.049 p = 0.580	D = 0.124 p = 0.227	D = 0.040 p = 0.982	D = 0.036 p = 0.995
Marketplace	-	-	-	D = 0.027 p = 0.983	D = 0.128 p = 0.167	D = 0.048 p = 0.877	D = 0.061 p = 0.651
Residence	-	-	-	-	D = 0.114 p = 0.245	D = 0.056 p = 0.652	D = 0.068 p = 0.421
Restaurants	-	-	-	-	-	D = 0.097 p = 0.590	D = 0.102 p = 0.535
Health	-	-	-	-	-	-	D = 0.027 p = 1
Work	-	-	-	-	-	-	-

Table 2d: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of time spent at a location outside the household by location type. The *p*-value associated with each test is the probability of the two sample groups not coming from a common distribution.

	0-9	10-17	18-33	34-44	45+
0-9	-	D=0.459 p <0.001	D=0.408 p < 0.001	D=0.432 p < 0.001	D = 0.342 p < 0.001
10-17	-	-	D = 0.271 p < 0.001	D = 0.307 p < 0.001	D = 0.403 p < 0.001
18-33	-	-	-	D = 0.069 p = 0.761	D = 0.154 p = 0.004
34-44	-	-	-	-	D = 0.150 p = 0.050
45+	-	-	-	-	-

Table 3a: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of time spent at home by age group. The *p*-value associated with each test is the probability of the two sample groups not coming from a common distribution.

	Female
Male	D=0.283 p < 0.001

Table 3b: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of time spent at home by gender. The *p*-value associated with each test is the probability of the two sample groups not coming from a common distribution.

	Merida	Ticul	Progreso
Merida	-	D = 0.069 p = 0.988	D = 0.178 p < 0.001
Ticul	-	-	D = 0.185 p = 0.088
Progreso	-	-	-

Table 3b: Kolmogrov-Smirnov Test results of cumulative probability distribution comparisons of time spent at home by city of residence. The *p*-value associated with each test is the probability of the two sample groups not coming from a common distribution.

a)

Predictor	b	b 95% CI [LL, UL]	Fit
(Intercept)	0.32**	[0.25, 0.39]	
Overlap Percentage	0.33*	[0.19, 0.47]	
Family Size	0.01	[-0.01, 0.03]	
			R^2 = .023**

b)

Predictor	b	b 95% CI [LL, UL]	Fit
(Intercept)	0.10	[0.05, 0.15]	
Overlap Percentage	0.89**	[0.71, 1.08]	
Family Size	-0.01	[-0.02, 0.00]	
			R^2 = .504**

Table 4a: Linear regression analysis results accounting for family size using a) percent dengue history (prospective cohort) as the criterion and b) percent acutely infected (retrospective cohort) as the criterion. * indicates *p* < .05; ** indicates *p* < .01.

a)

Predictor	b	b 95% CI [LL, UL]	Fit
(Intercept)	0.32**	[0.24, 0.40]	
Overlap Percentage	0.35	[-0.08, 0.78]	

Predictor	b	b 95% CI [LL, UL]	Fit
Family Size	0.01	[-0.01, 0.03]	
Overlap Percentage:Family Size	-0.01	[-0.11, 0.09]	
			R^2 = .019

b)

Predictor	b	b 95% CI [LL, UL]	Fit
(Intercept)	-0.17	[-0.34, 0.00]	
Overlap Percentage	2.22*	[1.42, 3.02]	
Family Size	0.09	[0.03, 0.15]	
Overlap Percentage:Family Size	-0.46	[-0.73, -0.19]	
			R^2 = .550**

Table 4b: Linear regression analysis results accounting for family size and the interaction term using a) percent dengue history (prospective cohort) as the criterion and b) percent acutely infected (retrospective cohort) as the criterion. * indicates $p < .05$; ** indicates $p < .01$.

Appendix:

EMORY UNIVERSITY | Institutional Review Board

TO: Gonzalo Martin Vazquez Prokopec, PhD
Principal Investigator

DATE: March 13, 2018

RE: **Clarification of Exempt Status**

IRB00082848

Indoor residual spraying to support integrated dengue management

Thank you for requesting clarification of whether the above-referenced study still qualifies as exempt in view of a change in plans. We have reviewed the information you submitted to the IRB on 1/22/2018 and find that the project, as changed, is still exempt from further IRB review. This amendment includes the following: The addition of Yoshihara Masato to the study team to use already collected anonymous data to retrospectively analyze patterns of exposure related to the current protocol in measuring type of location and distance from the treatment area. This change does not impact the exemption status of the current protocol and may continue.

As you know, a determination of exempt status is good indefinitely unless something changes substantively in the project that affects our analysis. The PI is responsible for contacting the IRB for clarification about any *substantive* changes in the project. Therefore, please *do* notify us if you plan to:

- Add a cohort of *children* to a survey or interview project, or to a study involving the observation of public behavior in which the investigators are participating.
- Change the study design so that the project no longer meets the exempt categories (e.g., adding a medical intervention or accessing identifiable and potentially damaging data)
- Make any other kind of change that does not appear in the list below.

Please *do not* notify us of the following kinds of changes:

- Change in personnel, except for the PI
- Change in location
- Change in number of subjects to be enrolled or age range for *adults*
- Changes in wording or formatting of data collection instruments that have no *substantive* impact on the study design
- For more information about the exemption categories, please see our Policies & Procedures at www.irb.emory.edu. In future correspondence about this study, please refer to the IRB file number, the name of the Principal Investigator, and the study title. Thank you.

Sincerely,

Regina Drake, M.Div, CIP
Senior Research Protocol Analyst

Appendix I: IRB Approval Letter