Proximity to Environmental Hazards and Reported Illness in Periurban Households of the Dominican Republic.

By

Glenn A. Vorhes

A REPORT

Submitted in partial fulfillment of the requirements
For the degree of

MASTER OF SCIENCE IN ENVIRONMENTAL ENGINEERING

MICHIGAN TECHNOLOGICAL UNIVERSITY

2009

This report "Proximity to Environmental Hazards and Reported Illness in Periurban Households of the Dominican Republic" is hereby approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE IN ENVIRONMENTAL ENGINEERING.

Civil and Environmental Engineering

Master's International Program

Signatures:

Report Advisor

James R. Mihelcic

Department Chair ________________________________

William Bulleit

Date ________________________________

Abstract

This research utilized data obtained from a household socioeconomic and health survey of three periurban neighborhoods in the Dominican Republic. A proxy for the frequency and severity of illness experienced by each household was generated based on survey responses. A demographic correction model was used to remove the correlation between this proxy and the number of household inhabitants. The residuals from this model represent the measure of illness in the house which is above or below that which would be expected based on the number and ages of household inhabitants. Environmental hazards analyzed included inadequate disposal of solid waste and sanitary facilities in poor condition measured through a scoring of survey responses, and uncontained wastewater in the street measured by qualitative evaluations. The dispersion of risks was modeled through the use of raster images in ArcGIS. Values extracted from these images were compared to the geographically referenced health data for correlation and Analysis of Variance or relative risks between populations grouped by dichotomous variables or ranges of values. Significant correlations were found between the metric of illness and the prevalence of uncontained wastewater in the street and sanitary facilities in poor condition. There was also a significant correlation between inadequate disposal of solid waste and the measure of illness though the modeled dispersion suggests that solid waste disposal is of more concern to welfare of the inhabitants of the house than it is for the general health of the neighborhood. Socioeconomic factors were also analyzed with no significant correlations found between illness and income or education. Those who reported drinking tap water rather than purchased filtered water were significantly more likely to report illness. Earlier studies used dichotomous variables to model risk of exposure to environmental hazards whereas this research is unique in using spatial techniques with continuous variables to model severity of and proximity to hazards and the relation to reported waterborne illness.

Preface

This report is based on the 33 months I spent living and working in The Dominican Republic while serving as a Peace Corps Volunteer from November, 2006 to September, 2009. I worked as an Environmental Sanitation Engineer/Promoter first in the small community of La Sierra in the municipality of Altamira and later in the Santiago, the second largest city in the country with population of about 500,000.

This report is submitted to complete my master's degree in Environmental Engineering from the Master's International Program in Civil and Environmental Engineering at Michigan Technological University. Data was obtained from a health and socioeconomic diagnostic of three neighborhoods on the periphery of Santiago.

Funding for the project was provided by the Rotary Club of Joliet, Illinois. The sponsors had no role in the design or implementation of the project or in interpretation of results.

Acknowledgments

Funding was provided by Rotary Club of Joliet, IL with special recognition to Howard Hamilton and David Crow of the Santiago, Dominican Republic chapter.

Special thanks to the following people:

Dr. James Mihelcic – University of South Florida

Dr. Noel Urban – Michigan Technological University

Dr. Christopher R. Webster – Michigan Technological University

Franklin Estrella – Secretaría de Estado de Salud Pública y Asistencia Social

Eddy Hernandez Muñoz

Lissette Anyelina Frias Toribio

Table of Contents

List of Figures

List of Tables

Introduction

This research attempts to answer the scientific question, 'Is there a statistically significant relationship between the proximity to environmental hazards and the prevalence of waterborne illness which can be observed at the household scale within a neighborhood?' While it may be shown that those living closer to an environmental hazard are more likely to become ill, it may also be that they are among the less well off economically or less well educated. An identified relationship between the proximity to an environmental hazard and illness should not be assumed to be a causal relationship without considering other factors. This research does an in depth analysis of health, environmental conditions, and socioeconomic factors. By showing both the presence and absence of significant relationships between various factors and illness, the reader can make inferences of causal factors.

Environmental hazards from inadequate disposal of solid waste, sanitary facilities in poor condition, and uncontained residual water in the street were addressed. Methods included spatial analysis techniques to measure for relationships between the prevalence of waterborne illness, deduced from household survey responses, and environmental hazards. The use of continuous variables to model the distance to and severity of environmental hazards is unique to this investigation. Hazard severity is quantified by means of household survey responses as well as with qualitative observations distinct from the survey.

Periurban Sectors in the Developing World

There is no consensus definition for the *periurban sector* but the term generally refers to urban areas on the margins of the physical and regulatory city boundaries that have uncertain legal tenure, low household incomes, and lack urban services (Hogrewe et al., 1993). Services in poorer and often newer neighborhoods of many Latin American cities lag behind those in central districts (Komives, 1999). Periurban poor receive substandard or no urban services while wealthier areas are covered at subsidized rates (Paterson et al., 2007). Solid waste collection is less profitable in periurban areas due to the lower reclaimable value of discarded items and difficult collection in steep terrain or narrow streets (Coura Cuentro and Gadji, 1990). Poor neighborhoods are left unserved by sewer networks because residents cannot afford high connection charges (Wright, 1997). Levels of fecal contamination and incidence of childhood diarrhea are at least as high in periurban settings as they are in rural areas (Lopez de Romana et al. 1989, Schorling et al. 1990). Children from disadvantaged slums are malnourished from repeated episodes of diarrhea during their most important formative years (Guerrant, 1994). Accidents and environmental hazards are the major causes of illness, injury, and premature death in most urban areas of the developing world (Hancock, 1996).

It is commonly asserted that periurban communities are ignored by municipal authorities as their size and population outpaces the capacity of local planning and government (Hogrewe et al., 1993). However, Mehta (2006) found problems regardless of population size or growth. Boston sanitary surveyors reported in 1850 that, “Cities are not necessarily unhealthy, but circumstances are permitted to exist, which make them so” (Schultz and McShane, 1978). Especially in periurban settings where salient needs are basic subsistence and housing, elected officials and residents do not tend to perceive environmental health as a vital concern (Hubbard et al., 2005).

Despite their large numbers and unique situation, data are rarely collected that illuminate the plight of periurban residents. Health indicators such as infant mortality and diarrhea are much worse in crowded squatter settlements though statistics are generally not disaggregated from those of the city as a whole (Wright, 1997). While data are gathered for areas designated as rural or urban, statistics for periurban residents are not collected in Latin America or Africa (Hogrewe et al., 1993). The level of health is rarely deduced from measured variables distinctly for periurban areas such that improvements are difficult to quantify in those areas where health is generally worse than in other parts of the city (Moore et al., 2003).

Solid Waste Disposal

Improper disposal of household solid waste is a source of air, land, and water pollution and creates hazards to human health and the environment (Medina, 1999). Accumulated solid waste in urban areas can affect not only those from which it came, but can also be an environmental health hazard for nearby residents. As with concerns relating to sanitation and residual waters, inadequate solid waste disposal can put a population more at risk of exposure to disease causing agents. Uncollected garbage provides a breeding ground for disease vectors such as flies and rats (Coura Cuentro and Gadji, 1990) that contaminate food (McGranahan, et al., 1997). This is important to an analysis of waterborne illness that uses diarrhea as a metric because food contamination potentially accounts for 15 to 70% of diarrhea cases (Esrey and Feachem, 1989).

Community Level Sanitation

Sanitation at the community level is a more important measure for health benefit than is individual access to improved sanitation (Bateman et al., 1993). The percentage of residents with sewer connections is not a reliable indicator of community health (Nance, 2005) as was found by Heller (1999) in Brazil. It is suggested that at least 75% of the community should have access to improved sanitation as lower coverage puts all residents at risk from poor environmental conditions (Bateman et al., 1993). A study in Guatemala found that children living in a community with a generally high level of sanitation had low rates of growth stunting regardless of in home access to a flush toilet (Bateman and Smith, 1991). When compared in multivariate analyses, the method of wastewater disposal showed no significant relationship with health though those living near streets with uncontained wastewater had a relative risk of diarrheal morbidity of 2.38 (95% confidence: 1.87 - 3.03) over those in the wastewater absent settings (Heller, 1999). The most important intervention goal with regard to sanitation is proper disposal of all wastewater in the drainage basin in order to avoid overflows to the street (Heller, 1999).

Residual Water

Pit latrines and pour flush latrines are considered as 'improved sanitation' (WHO and UNICEF, 2004), however any sanitation technology requires safe disposal of water from washing and bathing (Kalbermatten et al., 1982), hereafter referred to as gray water. Off site sanitation such as conventional sewerage protects the user from exposure to excreta but may contaminate groundwater or increase hazards to downstream populations (Hogrewe et al., 1993). The quality of upstream solutions determines the hazards for those downstream (Heller, 1999, McGranaham, 1997).

Many residents route wastewater to storm drains or to pit latrines that can contaminate groundwater (Watson, 1995). Gray water in storm drains may come from houses that are connected to sewer systems but have routed wash water to the street instead of to the sewer (Nance, 2005). Water from bathroom facilities contains higher concentrations of contaminants but gray water is not benign. Gray water contains excreted pathogens, organic compounds, and twenty to thirty grams of biological oxygen demand per capita per day (Kalbermatten et al., 1982). Untreated wastewater should always be assumed to contain high levels of pathogens (Esrey et al., 1998).

Health improvements have not been realized solely by improved water services with no attention to drainage or sanitation. Water supply improvements without adequate means for disposal can exacerbate problems (Katakura and Bakalian, 1998, UN Habitat, 1987). Nawab et al. (2006) found that incidence of waterborne illnesses increased after the installation of a piped water system and that villagers were able to differentiate between the diseases prevalent before and after the infrastructure improvement. Environmental problems in the United States in the 19th century worsened due to larger quantities of water brought in by new supply systems (Schultz and McShane, 1978) and continued to take a toll on urban populations until the early 20th century (Weber, 1899, Hancock, 1996). That health has improved in developed world cities shows that infrastructure improvements can protect the population (Satterthwaite, 1993).

Compounding Factors

This research did not assume that solid waste collection and/or sanitation infrastructure are the only or even the most important factors with respect to general health or prevalence of illness. Though the transmission of fecal-oral disease is well understood, it may be difficult to identify the most important routes even within a given neighborhood (McGranahan, 1997). Factors pertaining to fecal-oral disease tend to be related to other environmental health problems (McGranaham, 1997). Illness associated with environmental hazards may also be linked to individual or group behavior (Moore and Carpenter, 1999). Hygiene behavior may be as important as sanitation infrastructure with respect to health. Even without latrines, diarrheal morbidity is reduced with improved hygiene practices (WHO, 1993). Simple burial of excreta breaks the transmission route (Waterkeyn & Cairncross, 2005). If maintained and used by all, any sanitation scheme should lead to health benefits (Feachem et al., 1980).

An investigation of sanitation and health should also take into account socioeconomic factors before suggesting causal relationships. Previous studies have found that metrics of education or wealth have significant relationships with measures of health. Checkley et al. (2004) found that children living in households where water was stored in small containers were more likely to have growth stunting than were those where water was stored in large containers. Mothers' education levels and articles owned were shown to be related with caloric intake of children in some settings (Bairagi, 1980).

Epidemiological Methods

A major concern of epidemiology has been to find the risk factors associated with various diseases (Haug et al., 1997). The idea of using geography in health research comes from an appreciation of non-uniformity in the distribution of illness (Mayer, 1983). The first geographically referenced epidemiological data are attributed to Dr. John Snow who mapped illness reports to suggest that a certain water pump was the source of the London cholera outbreak in 1854 (Haug et al., 1997, Moore and Carpenter, 1999).

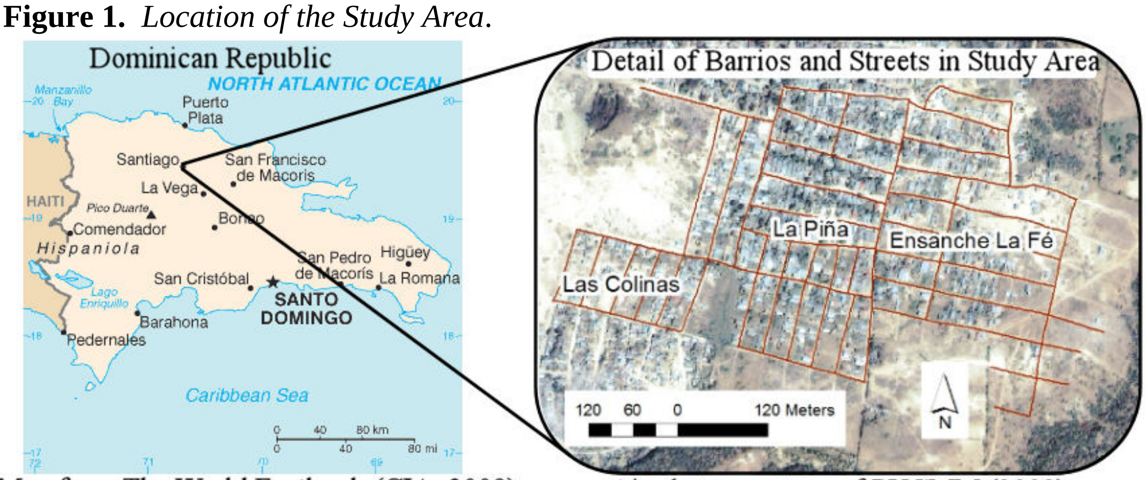
Numerical-spatial investigations such as identifying health factors within a community can be facilitated by the use of Geographic Information Systems (GIS) (Scotch et al., 2006). Moore and Carpenter (1999) provided a review of several epidemiological studies that made use of GIS technology. Spatial analysis is normally performed by plotting geographically referenced illness events and looking for hazards which have the same spatial distribution (Haug et al., 1997). Patterns in noisy data can be detected visually but maps are not good at representing complex relationships between response and explanatory variables (Westlake, 1995). A literature review found that statistical software is usually the driving force in community health assessments with GIS used primarily to display results (Scotch et al., 2006).

Methods

Study Site

The study included three neighborhoods on the periphery of Santiago (pop. 500,000), the second largest city of the Dominican Republic. Figure 1 shows a map of the Dominican Republic and detail of the study area.

Figure 1. *Location of the Study Area.*



Map from The World Factbook (CIA, 2009). Air photo courtesy of PUCMM (2009).

The neighborhoods were founded twenty years ago and are characterized by low incomes and low access to urban services. The study site is considered to be homogenous in terms of income and education levels. Electricity is available though service is not continuous as is typical in the Dominican Republic. Most residents are served with solid waste collection by the city government or private services while some, primarily those on the periphery of settlements, do not have services or choose not to use them. At the time of the study, tap water was available every second day through a network that is in some cases installed by the residents, with or without approval from the city water authority. The most common types of sanitary facilities are pit latrines and a variant of a pour flush latrine referred to as a 'septic well'. An important clarification is that the septic wells are not synonymous with septic tanks. Wastewater from an indoor bathroom is routed to a rock filled soak pit. The pit includes an overflow tube to street drains in some cases. Gray water is routed by pipes to street drains or less commonly to the residents' yards or to the latrine or septic well. The topography is such that the quantity of wastewater observed in the street varied greatly within the study area.

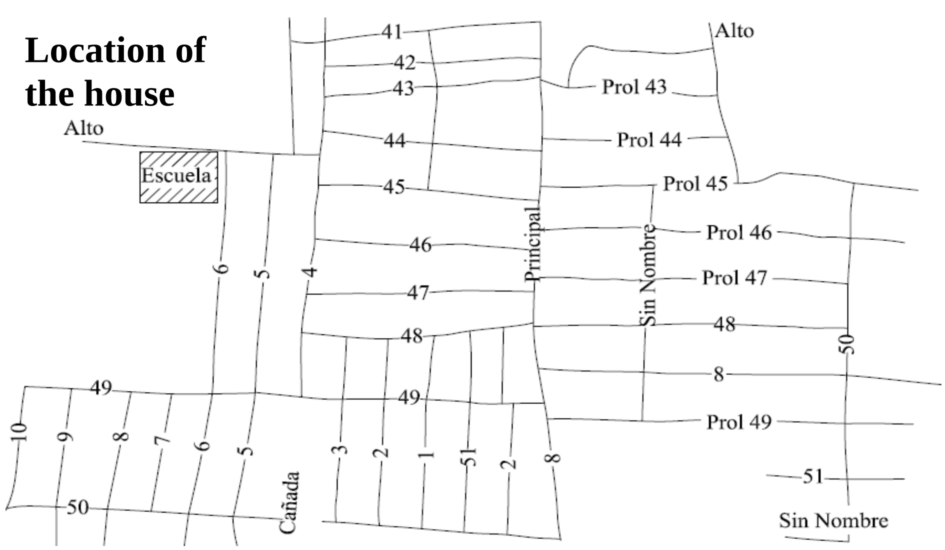
Survey Formulation

While household surveys are now performed regularly in almost all countries, surveys focusing on a broad spectrum of environmental and related health conditions are relatively rare (McGranaham, 1997). The survey for this research was formulated to extract a mix of demographic, economic, and health indicators in order to evaluate the interrelated factors pertaining to waterborne illness. Questions and the structure of the survey form were developed with the help of Dominican nationals from a variety of disciplines including personnel from the government public health agency. The survey form included the following:

1. General Neighborhood (7 questions)
2. Demographic (5 questions)
3. Economic (10 questions)
4. Hygiene (17 questions)
5. Contingent Valuation (2 questions)
6. Education (7 questions)
7. Health (6 questions)

The initial portion of form was completed by the interviewer before any questioning of the respondent took place. Figure 2 shows the most important part of this section which is a map of the survey area. Interviewers were instructed to indicate the approximate location of the house surveyed with a dot. This step allowed for a spatial reference to the survey data such that proximities to identified hazards could be included in the analyses.

Figure 2. *Study Area Street Name/Number Map Included on Each Survey Form.*



The complete survey form translated into English is included in Appendix A and the original Spanish version is provided in Appendix B.

Interviewers and Preparation

Research methodology must be carefully developed in order to establish rapport with a community and investigate what would otherwise not be available (Nawab et. al., 2006). To this end, host country nationals were trained to conduct the interviews. Ten pairs of interviewers were coordinated. Each pair included one guide from the community and one nursing student. The guide, being a member of the community, could better relate to respondents and was tasked with asking the questions. The nursing students had a background in epidemiology and were assumed to be better able to complete the survey forms based on the respondents' answers.

Interviewers were invited to a workshop on the weekend before the first day of the survey. The purpose of the workshop was threefold. It allowed the two groups, the guides from the study site and the nursing students, to become acquainted in a neutral setting before beginning work. Secondly, the participants received a detailed explanation of survey methodology and additional information for those questions that were not self explanatory. Interviewers were also trained in privacy protection for respondents. Lastly, the workshop provided two-way communication between the researcher and the interviewers. Participants offered valuable feedback on the wording of questions and the structure of the survey form such that improvements could be made.

Conducting the Survey

The survey was conducted on two consecutive Sundays. Before beginning at each house, the respondent was read a prepared introduction to the survey explaining its purpose and assuring privacy. Interviewers conducted the surveys and returned completed forms which were then kept in a secure location.

Privacy

Statistical and epidemiological methods need to be developed to protect individual and household confidentiality (Armstrong et al., 1999). Research methodology was developed in order to maintain the privacy and anonymity of respondents with regard to sensitive information. Data collection methods and proposed presentation of results were approved by the Office of Integrity and Compliance of Michigan Technological University according to protocol M0471. A copy of the Internal Review Board approval is included in Appendix E.

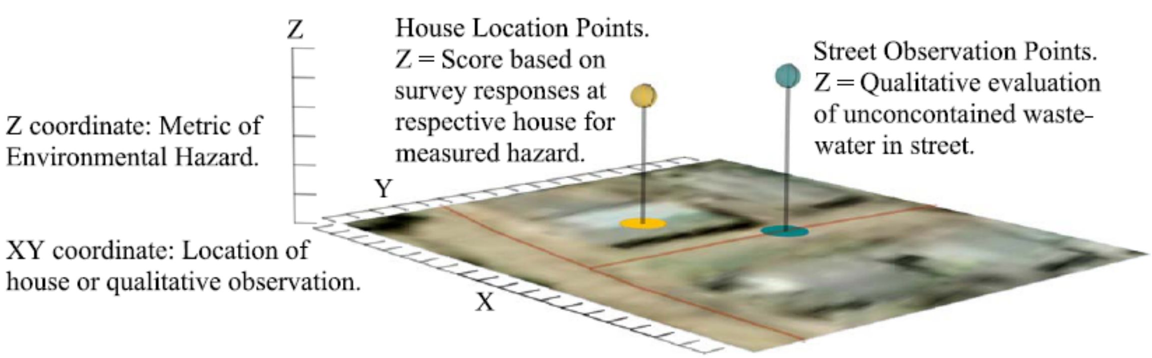
Data Entry

Digitization of the data was performed using software developed specifically for this investigation. As an improvement to entering responses into spreadsheet software as columns and rows, a graphical interface was used as an intermediary. By this method, the researcher entered the location of the house as indicated on the survey form map. Question responses, and check and option indicators were recorded as they appeared on the survey form in order to facilitate the data entry process and to minimize errors. The software made the transformation to save the responses in a database which was then linked to data points in ArcGIS 9.2 (ESRI, 2006).

Modeling Environmental Hazards

Poor environmental sanitation affects the users of the facilities as well as others in the area (McGranahan, 1997, Hogrewe et al., 1993). Spatial analysis techniques were used in this research to model the severity of environmental hazards. The environmental hazard models were generated from points in three dimensional coordinate spaces. The X and Y coordinates corresponded to a location in study area. The Z coordinate at each point was based on either a qualitative observation or on a score developed from survey responses. For the hazards deduced from respondents' answers, each of the 520 household location points had a Z coordinate generated by scoring responses of pertinent questions from the respective survey. The 287 street sanitation observation points had Z coordinates based on a qualitative evaluation of the quantity of uncontained wastewater at the location. Figure 3 shows one representative house location point and one qualitative observation point.

Figure 3. *XYZ Points used to Model Severity of Environmental Hazards.*



The set of XYZ points for each environmental hazard was used by the 3D Analyst tool in ArcGIS to develop raster images (rasters) that model the severity of the hazard at all locations in the study area. Rasters can be thought of as two dimensional grids of squares for which each square (unit of resolution, pixel, or cell) has a numerical value that corresponds to the severity of the hazard in the area represented by said square. This value is hereafter referred to as the pixel value. A note to the reader is that the terms 'pixel' and 'cell' are used interchangeably and both refer to one unit of resolution in the raster. The pixel value can be used to represent a variety of measurements or classifications. For example, GIS applications can use rasters with pixel values that represent the type of vegetation, land use classification, or the distance to feature(s). In this research, the pixel values represent the relative severity of an identified environmental hazard at the corresponding location in the study area. The rasters can be visualized as three dimensional surfaces draped over the study area where peaks and valleys represent higher or lower severities of the hazard. The use of rasters in this research is analogous to the methods used by Ali et al. (2002). In their research, a raster image with pixel values corresponding to the distance to identified areas of stagnant water was used to model the cholera hazard.

Raster images were generated from each set of environmental hazard XYZ points with the Inverse Distance Weighting algorithm in ArcGIS. Spatial dispersion of hazards was modeled by using a variety of raster cell sizes. The measure of cell size refers to the length of one side of the square cell. Small cells generally include only one XYZ point such that the pixel value in the corresponding raster cell depends almost entirely on the Z coordinate of the included point. The algorithm is such that the pixel values are not exactly equal to the average of included Z coordinates. However, the differences are small and for the sake of explanation, the pixel value can be considered to be the average of the Z coordinates of included points. Pixel values for cells without a point are generated by the algorithm according to the Z coordinates of nearby points. Larger cells include more points in a unit of resolution such that the influence of any one of the points' Z coordinates on the resultant cell's pixel value is reduced. Rasters were generated with cell sizes ranging from two to twenty meters.

Dispersion of hazards was further modeled by using the Spatial Analyst Filter Tool in ArcGIS which buffers high and low pixel values based on the values from adjacent cells. The tool creates an output raster from an input raster. The pixel value for each cell in the output raster is the un-weighted average of the pixel values from the corresponding cell in the input raster and those of its eight adjoining cells. Figure 4 shows a visual representation of the filter tool process.

With increasing cell sizes and application of the filter, the pixel value at a house's location for a given hazard becomes less associated with the Z coordinate for the house (the score generated from the house's survey responses for pertinent questions). The pixel values become increasingly reflective of the environmental hazards for the general vicinity, deduced from relevant scores from houses in the area. By varying the modeled dispersion in this way, it is possible to infer whether the household or the neighborhood /house cluster is the more adequate unit of analysis for a given hazard with respect to health. Figure 5 summarizes the relationship between the household Z coordinate and the corresponding pixel value for rasters with various cell sizes and filter settings.

Figure 4. *Filter Tool Generates Output Raster from Pixel Values in Input Raster.*

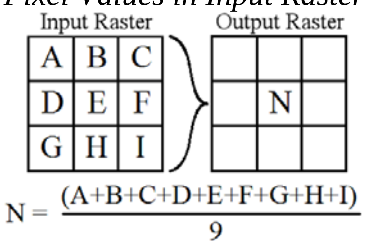
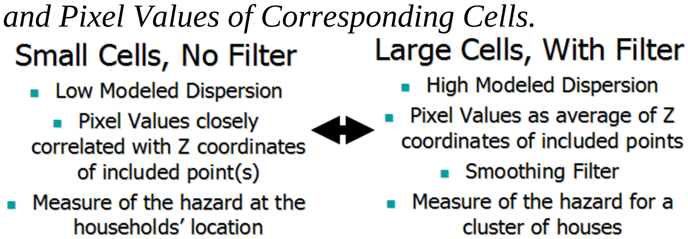


Figure 5. *Raster Cell Size and Filter Application Effects on Association Between Z Coordinate Values and Pixel Values of Corresponding Cells.*



Solid Waste Disposal

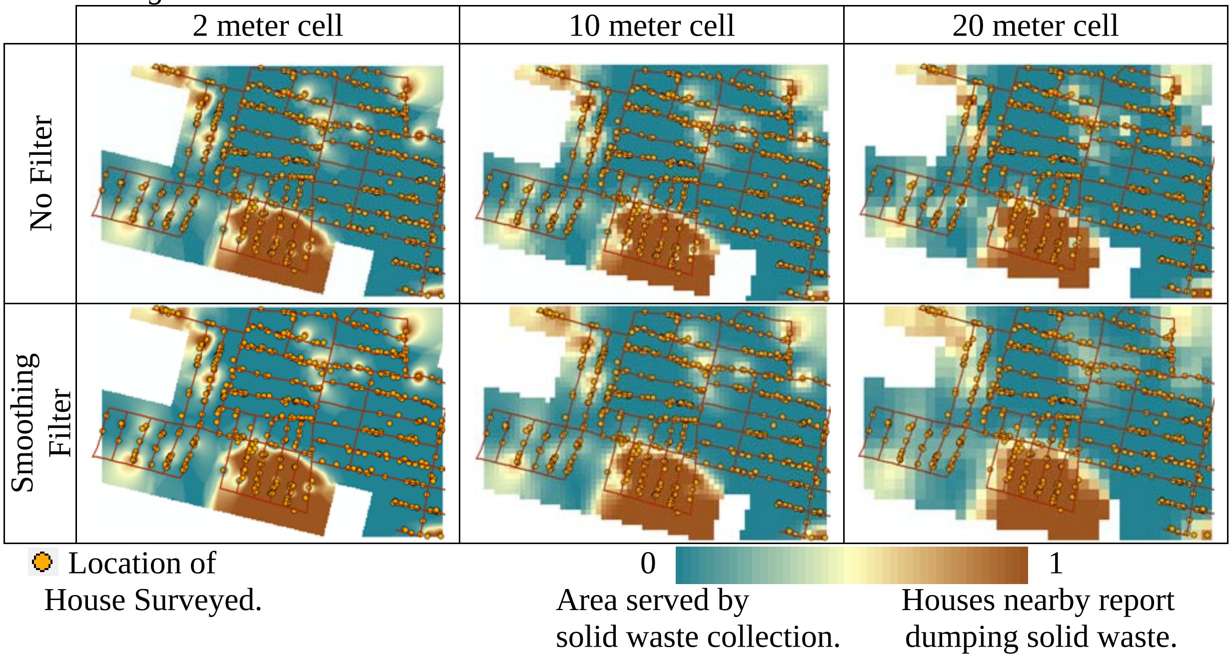
The question pertaining to solid waste disposal is shown in Table 1. Refer to Appendix A for the complete survey form.

Table 1. Survey Question Regarding Disposal of Solid Waste.

Is garbage collected?	*Yes*	*No → How is it disposed?*		
By whom?		*Thrown to:*		*Burned*
City	*Private Service*	*vacant areas*	*the street*	

The question was scored as a zero if the respondent reported solid waste collection by the city, a private service, or that it was burned. As this research concerns waterborne illness, air contamination from burning was not addressed. Disposal was considered to present a hygienic risk if it was reported as ‘Thrown to vacant areas’ or to ‘the street’ and scored as a one. Raster images were created using the location of the house as the X and Y coordinates and the score from the survey response (zero or one) as the Z coordinate as explained in the previous section. Rasters generated with cell sizes of two, ten, and twenty meters, with and without the filter applied, are shown in Figure 6. Pixel values near zero indicate the areas where residents and nearby neighbors were served by solid waste collection while values near one indicate those areas where solid waste was dumped in vacant areas or the street. Visual inspection and familiarity with the study site revealed that those living on the periphery of settlement or near vacant areas were more likely to report dumping of solid wastes.

Figure 6. *Solid Waste Disposal Rasters with Various Cell Sizes, With and Without Smoothing Filters.*



Condition of Sanitary Facilities

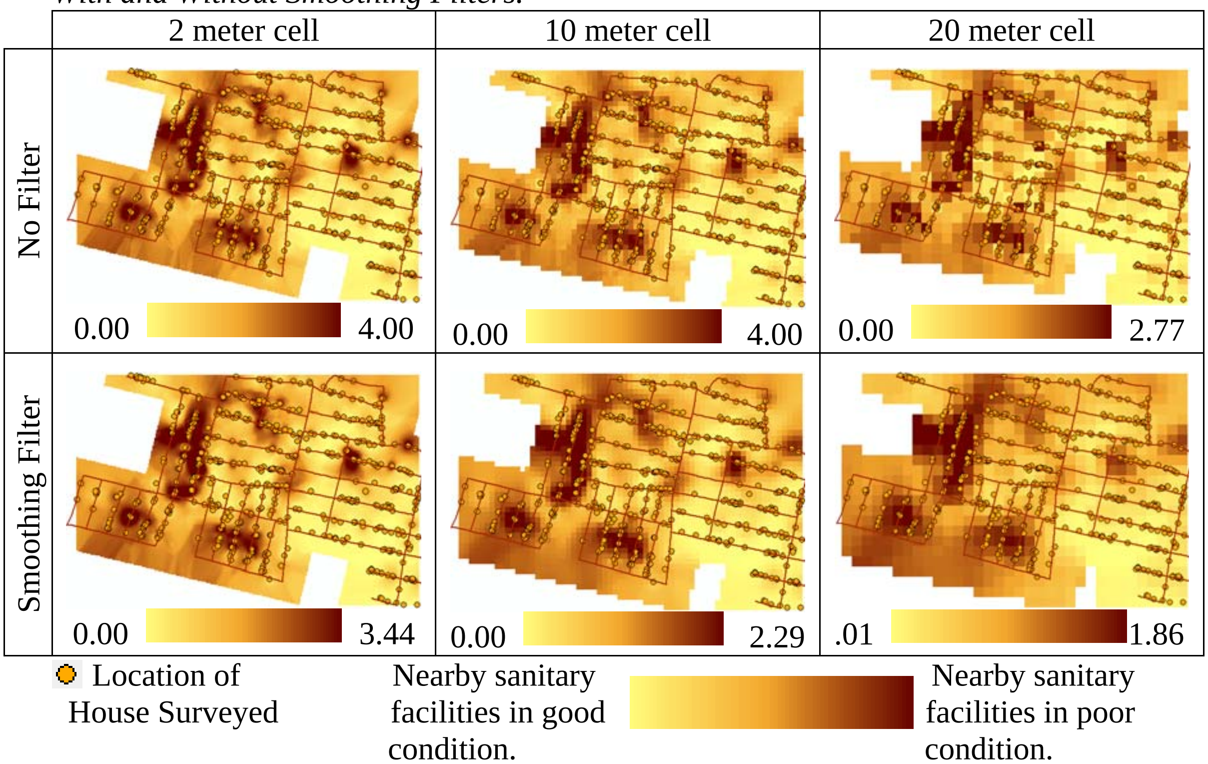
Table 2 shows the questions and scoring of responses used to quantify the general condition of sanitary facilities. The complete survey form is included in Appendix A.

Table 2. Questions and Scoring of Responses Regarding Sanitary Facility Condition.

Question	Scoring
Does it serve your needs?	Yes:0 No:1
Is the latrine or septic pit full?	Yes:1 No:0
Does the latrine or septic pit have an overflow to the street?	Yes:1 No:0
Do the latrines or septic pits fill with water when it rains?	Yes:1 No:0
	Sum for house: 0 to 4

The score for the sanitary facilities in each household was used as the Z coordinate at the XY household location to generate the sanitary facility condition rasters. High values correspond to areas with sanitary facilities scored as in poor condition. For the few surveys for which no response was given (11 of 520), conditions were assumed to be adequate. Raster images were generated with cell sizes of two, five, ten, fifteen, and twenty meters, with and without smoothing filters for each cell size. Representative images with various cell sizes and filter settings are shown in Figure 7. Larger cell sizes and the application of the filter truncate high and low values such that maximum and minimum values are different for some images.

Figure 7. *Representative Sanitary Facility Condition Rasters with Various Cell Sizes, With and Without Smoothing Filters.*



Street Sanitation Condition

Qualitative observations of the prevalence of uncontained residual water in the street were made at all intersections and at approximately thirty meter intervals according to the rubric shown in Table 3. Appendix C shows representative photos of each of the ratings.

Table 3. Rubric for Qualitative Evaluation of Wastewater Presence in Streets of Study Area.

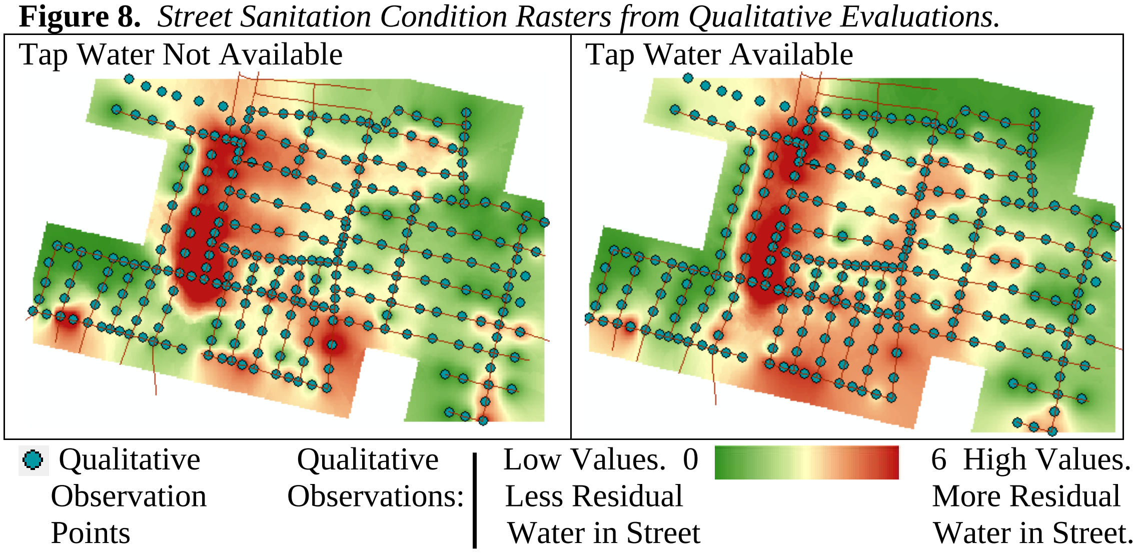
Rating	Description
0	Residual Water Absent, Streets Dry.
1	Least Detectable Levels of Residual Waters, Damp Street or Curb.

2	Marginally Observable Flow in Street Gutters, No Standing Water.
3	Plainly Observable Flow in Street Gutters and Ground Channels With or Without Very Minimal Observable Standing Water.
4	Considerable Flow and Some Standing Water Observed.
5	Considerable Flow and Standing Water Present, Odors Detected.
6	Continuous Flow and Standing Water Always Present, Strong Odors, Most Severe Conditions.

At the time of the study, water was available through the city water supply every second day. For this reason, the same evaluation was performed on two consecutive days in order to evaluate both the tap water unavailable and tap water available conditions.

The qualitative rating was entered as the Z coordinate at the respective XY coordinate location of the observation ArcGIS. Figure 8 shows the street sanitation condition rasters generated using the observations for both the tap water not available and tap water available days. Rasters for this hazard were only generated with the two meter cell size. The qualitative observations had no connection to the household survey data so it was not necessary to model dispersion of the hazard between houses.

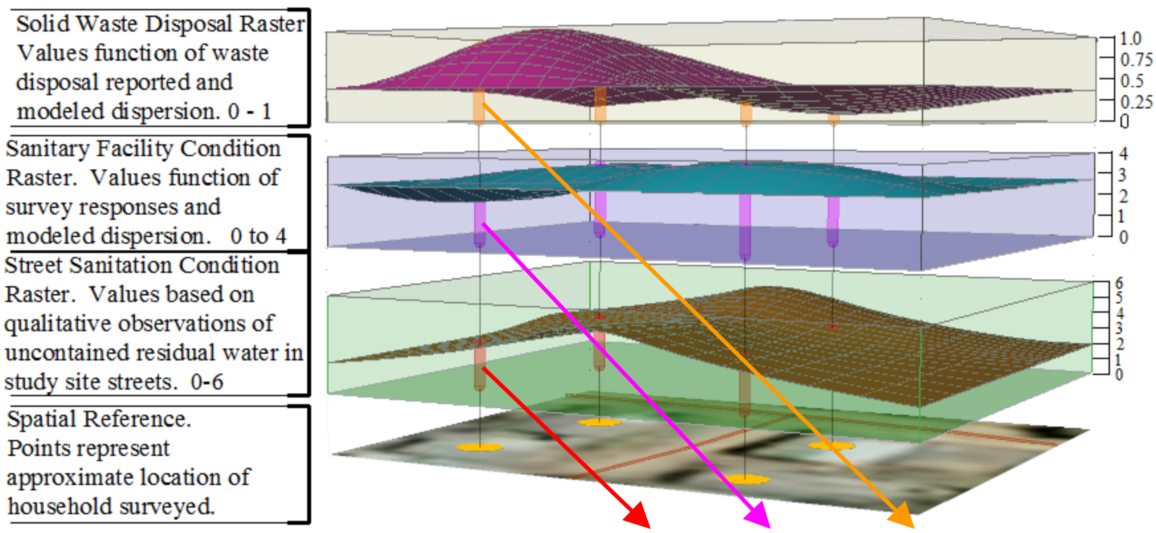
Figure 8. *Street Sanitation Condition Rasters from Qualitative Evaluations.*



Extracting Pixel Values to Survey Data

The Spatial Analyst Extraction Tool in ArcGIS allowed the pixel values from the rasters at each household's location to be extracted and appended as columns of values to the survey data. The process is visualized in Figure 9.

Figure 9. *Visual of Extraction Process: Pixel Values Extracted from Rasters and Appended to Survey Data.*



Survey-ID	Demographics	Economics	Health Data	Etc.	Street Sanitation Condition (2 day average)	Sanitary Facility Condition (10 columns)	Solid Waste Collection (6 columns)
0	...	...	...	...	3.64	0.50	0.170
1	...	...	...	...	3.13	0.86	0.005
2	...	...	...	...	3.05	0.62	0.003
...	...	...	...	...	...	...	...

The graphic represents pixel values as the height of a three dimensional surface above a datum though the rasters are two dimensional images for which the pixel values represent hazard severity. The process visualized here was repeated for each of the environmental hazard rasters at each cell size and filter setting in order to later test for relationships with health and other factors at several levels of modeled hazard dispersion.

The terms ‘Street Sanitation Condition’, ‘Sanitary Facility Condition’, and ‘Solid Waste Disposal’ are used in the remainder of this document to refer to the pixel values extracted to the survey data. The spreadsheet generated from the extraction process includes all survey data along with six solid waste disposal columns (three cell sizes each with and without filter), ten sanitary facility condition columns (five cell sizes each with and without filter), and one column pertaining to the street sanitation condition (the average of pixel values extracted from the tap water available/not available rasters).

Illness Index – Reported Illness from Survey Responses

The dependent variable for this research, the relative measure of waterborne illness for the household, was deduced from four relevant survey questions. Diarrhea was considered to be an important indicator and is mentioned in three of the four questions. Diarrhea is an acute disorder that reflects current environmental risks and there is a direct relationship between water supply and sanitation and diarrhea prevention (Bateman et al., 1993). Parasites and typhoid fever were also included in the illness scoring. Typhoid fever was not defined or verified by the interviewer. A positive response was assumed to indicate if not typhoid, a high fever.

The dependent variable is an integer value that is generated for each of the surveys. It is a function of the responses to the pertinent health questions. The survey questions were intended to extract quantifiable episodes of illness but variability in the manner in which responses were given and recorded made this impractical. Questions and responses were instead scored according to the metric shown in Table 4. In order to avoid giving excessive weight to those surveys performed with individuals who were inclined to mention every episode of illness, the scoring metric was simplified such that any positive response was recorded as a positive one for a given question and a zero for a negative or non response. Refer to Appendix A for the complete survey form.

Table 4. Survey Health Questions and Scoring of Responses.

Survey Question	Scoring
In the last three months, anyone has become ill from diarrhea, fever, or parasites? Yes Give the illness and age of the sufferer. No	Response: 1 No Response: 0
Anyone has suffered from typhoid fever? Yes No	Yes: 1 No: 0
In the last three months, anyone has missed a day of work or school due to a stomach ache or diarrhea? How Many? ☐☐☐☐☐ (Each episode was to be indicated by the number of days lost in one of the five boxes provided)	1 or more days or episodes: 1 0 days or episodes: 0
If you have needed to take someone to the hospital for diarrhea, typhoid fever, or parasites, how much did you spend on the consultation and medicine? *Consultation: ____________ Medicine: ________________*	Response: 1 No Response: 0
	Sum: Illness Index 0-4

The sum of the scores for all four questions is hereafter referred to as the ***Illness Index*** for the household. A key assumption is that respondents from those houses wherein waterborne illness was more frequent and/or more severe would be more likely to give positive responses for one or more of the questions. The Illness Index is not to be mistaken as the number of episodes of waterborne illness. This investigation ***does not quantify episodes of illness.*** The values for the Illness Indices serve as a relative measure of illness and can be compared between houses in this study.

Illness Residual – Data Transformation

There was a highly significant relationship between the number of household inhabitants and the values for the previously described Illness Index. The coefficient of linear regression was 0.172 (95% confidence: 0.087 – 0.257) with a R^2 value of 0.029. It could be logically assumed that households with more inhabitants would report more illness than would those that house fewer people. Checkley et al. (2004) used models that included age specific effects when doing regressions of the individual effects of water source, water storage, and sanitation. A simple per capita contribution model was not preferred in this investigation. To transform the data as such would have given no consideration to age specific effects. It would similarly not incorporate that children are more at risk than are adults for diarrheal disease associated with poor sanitation (Wright, 1997). A demographic correction was utilized in order to remove the association between the Illness Indices and the number of inhabitants in the household. A multiple linear regression was performed using the number of people in various age categories as the independent variables and the Illness Index as the dependent variable. The coefficients of linear regression for the various age groups are shown in Figure 10. The regression model incorporates both per capita contributions and age specific effects as they relate to the Illness Index for the household. The resultant model offers a generalized expectation of the Illness Index based on the number and ages of household inhabitants. Figure 11 compares the expected Illness Indices based on the age specific (demographic) model with those from a simple per capita contribution model.

Figure 10. *Coefficients of Linear Regression: Number of People in Various Age Categories as Explanatory Variables of the Illness Index.*

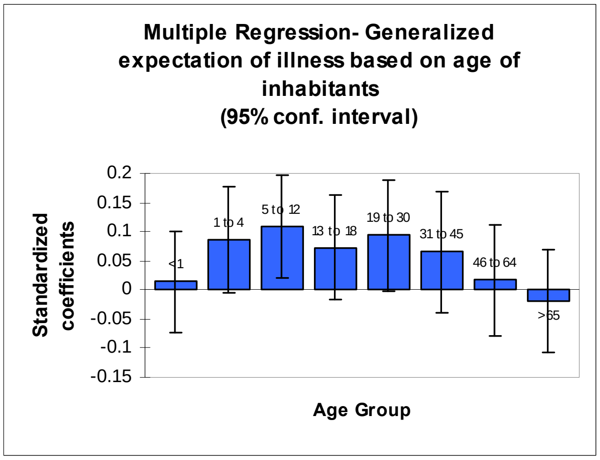
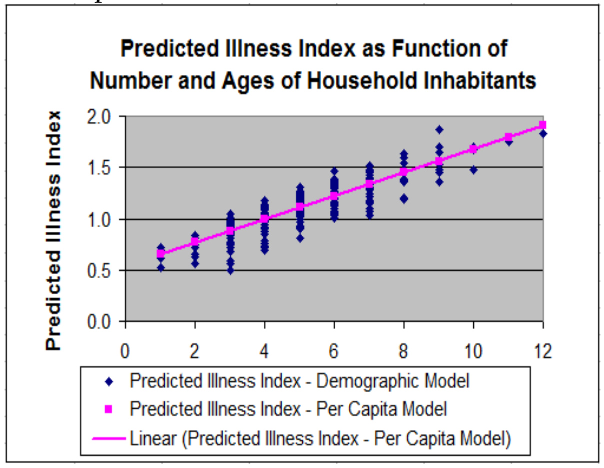


Figure 11. *Comparison of Predicted Illness Index from Age Specific (Demographic) and Per Capita Contribution Models.*



The model produced residuals; the Illness Index minus the predicted Illness Index based on the number of and ages of household inhabitants for each house. Houses with positive residuals are those with higher Illness Indices than would be expected based on the demographic prediction model and conversely, houses with negative residuals are those with lower than expected Illness Indices based on the model. The equation used to calculate the residuals used the generalized coefficients from the preceding multivariate regression and is shown in Figure 12. All subsequent analyses are conducted with the model residuals which are hereafter referred to as the ***Illness Residuals***.

Figure 12. *Generation of Illness Residuals based on Multiple Linear Regression Coefficients.*

$$\mathbf{IR_i = II_i - \left(\begin{array}{l} 0.015{*}N_i^{<1} + 0.086{*}N_i^{1\text{-}4} + .109{*}N_i^{5\text{-}12} + 0.073{*}N_i^{13\text{-}18} + \\ 0.094{*}N_i^{19\text{-}30} + 0.066{*}N_i^{31\text{-}45} + 0.017\ N_i^{46\text{-}64} - 0.020{*}N_i^{>65} \end{array} \right)}$$

IR_i = Illness Residual for House i.
II_i = Illness Index for House i.
N_i = Number of inhabitants in house i in given age group.
Age Groups: <1, 1-4, 5-12, 13-18, 19-30, 31-45, 46-64, and >65

Coefficients were generated from a multivariate linear regression using the number of people in each age group as the independent variables and the Illness Index as the dependent variable. Coefficients are shown with three significant digits.

Analysis of the residuals brings the investigation one step closer to inferring causality as the association with the number of household inhabitants was removed by the transformation. The Illness Residual is the dependent variable to which other variables are compared by linear regression or for variance between groups in cases of dichotomous or divisive variables. Table 5 provides further explanation of the relationship between the demographic data, the Illness Indices, demographic model prediction, and the Illness Residuals.

Table 5. Relationship Between Demographic Data, *Illness Index*, Demographic Prediction Model, and *Illness Residual*.

Survey-ID	Inhabitants in Age Category <1	1 to 4	5 to 12	13 to 18	19 to 30	31 to 45	46 to 64	65+	Illness Index (Relative Measure of Illness based on survey responses)	Expected Illness Index (based on number of and ages of inhabitants)	Illness Residual (Illness Index - Model Prediction)	All Other Data
0	0	1	2	0	1	1	0	0	2	1.225	0.775	...
1	0	0	2	1	1	1	1	0	4	1.191	2.809	...
2	0	3	0	0	1	1	0	0	0	1.308	-1.308	...
...	...	...	...	...	...	...	...	...	...	...	...	...

Data Analysis

Analyses of the data included a combination of Analysis of Variance (ANOVA), simple linear regressions, and relative risk assessments. Statistical analyses were performed using the following software:

- ANOVA – Data Analysis in Microsoft Excel (Microsoft Corporation, 2003)
- Regression – XLStat 2009, Statistics add in for Microsoft Excel (Addinsoft, 2009)
- Relative Risk – ClinTools Odds Ratio Generator (Devilly, 2007)

Seven survey forms were excluded from all analyses because there was no geographic reference point indicated on the survey form. Analyses were performed with the available data. In cases for which no response was recorded for a given question, the survey was excluded from that particular analysis. Appendix D provides a summary of the variables used in the analyses and the number of responses given for each question.

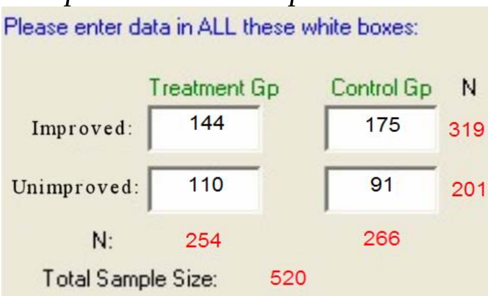
ANOVA and Relative Risk

An ANOVA was used to test for statistically significant differences between populations in cases for which it was practical to divide the surveys into distinct groups. The population was divided either by a dichotomous variable, categories, or by a cut off value selected in a range of a continuous variable as with the values from the environmental hazard rasters. Populations were considered to be significantly different if the ANOVA returned a P-value of 0.0500 or less.

Relative Risk (RR) and the closely related odds ratio provide a convenient way to present results for dichotomous variables (McGranahan, 1997) as with present/absent conditions or in other cases for which the population can be divided into treatment and control groups. The RR analyses used the Illness Residual as follows. Recall that the Illness Residual is the Illness Index for a given household minus the predicted Illness Index based on the number and ages of inhabitants as per the demographic prediction model. Households for which the Illness Residual was less than zero were designated as improved; greater than zero as unimproved.

Households were divided into treatment and control groups according to either a dichotomous variable or by selecting a cut off value in analyses involving either multiple choice responses or continuous variables as with the values from the environmental hazard rasters. For analyses involving the latter, the specific value chosen as a cut off was varied within the variable's range for a series of analyses. The calculated relative risks were compared for different distinctions of the control and treatment groups. A screenshot showing the format of the program used in RR analyses is shown in Figure 13.

Figure 13. *Screenshot from Odds Ratio Generator Showing Format of Entry for Control and Treatment Groups with Improved and Unimproved Cases in Each.*



Adapted from Devilly (2007).

The relative risk values should not be taken as the increase in risk of illness episodes or morbidity but rather as a relative measure to be interpreted with respect to other RR analyses in this research. A relative risk was considered to be significant if the 95% Confidence Interval (CI) did not include unity.

Linear Regression

The goal of this research was to analyze the relationships between waterborne illness and a number of parameters including environmental hazards and socioeconomic factors. Dependent and independent variables are suggested in the form of hypothetical questions though this should not be taken to imply of causality. In some cases, the distinction of dependent and independent variables may not be intuitive. For example, if a significant relationship were found between education and income, it would not be immediately obvious whether higher levels of education lead to higher incomes, vice versa, or if there is some other causal factor involved. Linear regression results are presented in a manner that allows the reader to make inferences of causality.

As was done by Ali et al. (2002) in their analyses of illness with respect to sanitation and socioeconomic factors, a linear regression was performed in cases for which it was appropriate to test for a relationship between two continuous variables. All variables were normalized such that the minimum and maximum values ranged from zero to 100 in order to allow for a direct comparison between the coefficients from the various regression models. The coefficients with 95% confidence intervals and R^2 values are presented.

Results and Discussion

The parameters measured, the hypothetical questions to be tested, and findings for the ANOVA and Relative Risk analyses are presented in Table 6; those of simple linear regression in Table 7. The narrative refers to specific results in the tables by referencing the row number.

Number of Household Inhabitants

A positive relationship between the number of people in the house and the Illness Residuals could suggest that crowding is a factor with regard to illness. Recall that the Illness Residuals incorporate a correction for per capita and age specific effects in the measure of illness for the household. As shown in Table 7, Row 1, no significant relationship was found between the Illness Residuals and the number of people in the house. From this result, crowding was considered not to be a factor. The goal of the demographic correction was to remove the association between the metric of illness and the number of people in the house. The finding of no significance in the linear regression

Table 6. Analysis of Variance and Relative Risk – Analyses and Results.

Row	Grouping	Test Parameter	Hypothesis to be Tested	Result: Significant / Not Significant
1	Drinking Water Source	Illness Residual	*Do those who report drinking tap water report higher frequency of illness than those who buy filtered water?*	P=0.0057 Tap Water RR 1.27 (CI 1.02 - 1.59) (Significant)
2		Reported Weekly Income	Do bottled water users generally have higher incomes?	P=0.0002 (Significant)
3		Formal Education	*Do bottled water users gernerally have higher levels of formal education?* Highest grade achieved by heads of household (male/female), average, and maximum.	Signficant for all four measures of education. P<=0.0016 (Significant)
4	Solid Waste Disposal, Dumped or Collected	Illness Residual	*Are those who reported dumping trash in vacant areas or the street more likely to report illness?*	P=.0283 Thrown RR 1.48 (CI 1.16 - 1.90) (Significant)
5		Reported Weekly Income	*Do those who reported dumping trash in vacant areas or the street have lower reported weekly incomes?*	P=0.0076 (Significant)
6		Formal Education	*Do those who reportdumping trash in vacant areas or street have lower levels of formal education?* Highest grade by heads of household (male/female), average, and maximum.	Not Significant for any of the four measures. P>=0.2788 (Not Significant)
7	Latrine or In House Toilet with Septic Pit	Illness Residual	*Are those living in houses with latrines more at risk than those having indoor facilities?*	P=0.8911 RR includes 1 in confidence interval. (Not Significant)
8		Reported Sanitary Facility Condition	*Are residents with latrines more likely to report poor sanitary facilities?*	P=0.1036 (Not Significant)
9		Reported Weekly Income	*Do residents with latrines report lower weekly incomes than those with toilets?*	P=0.0009 With Indoor Toilet RD$2099/week With Latrine RD$1504/week (Significant)
10		Formal Education	*Do indoor toilet owners/users report higher level of formal education?* Highest grade achieved by heads of household (male/female), average, and maximum.	Indoor toilet users average higher formal education is 1.5 grade levels higher, $P<1.264 \times 10^{-5}$ (Significant)
11	Self Reported San. Facilities	Illness Residual	*Do those who report poor conditon of sanitary facilities also report more frequent illness?* An integer value was calcualted according to the survey responses from 0 to 4 with higher values associated with worse conditions. The control group was designated as those with a score of 0, 0-1, 0-2, and 0-3 for a series of analysis.	Difference between groups, P=0.0001 RR by control group distinction. **0**: 1.31 (1.05 -1.62) **0-1:** 1.54 (1.22 -1.95) **0-2**: 1.93 (1.51 - 2.48) **0-3**: 2.19 (1.5 - 3.18) (Significant)

Table 6 Continued. Analysis of Variance and Relative Risk – Analyses and Results.

Row	Grouping	Test Parameter	Hypothesis to be Tested	Result: Significant Not Significant
12	Street Sanitaton Condition Raster Classes 0-6	Illness Residual Classes	*Is there a significant difference between households divided by ranges of observed sanitation values? If so, at what level of observed street sanitation are the control and exposed groups most distinct?* Surveys were divided according the observed sanitation values in groups 0-.99, 1 -1.99, 2-2.99, up to 6. Varied designations of exposed and treatment groups.	Significant difference between groups. P=0.0229 Street Sanitation Variable: Control Group <1.99, Exposure Goup >=2.00, RR 1.27 (1.02 - 1.57)
13	Self Reported Street Sanitation	Illness Residual	*Are those who reported more frequent residual water in the street also reporting higher incidence of illness?* The control group was designated as those reporting residual waters were never present, never-sometimes, and never-sometimes-often.	Significant difference between groups, P=0.0096 RR by control group designation, water present: Never 1.32 (1.04 - 1.68) Never-Sometimes 1.31 (1.05 - 1.63) Never-Sometimes-Often 1.34(1.08-1.66)
14		Observed Street Sanitation	*Is there a statistically significant difference in the extracted observed street sanitation values for those reporting residual water is never, sometimes, often, or always in the street near their homes?*	$P=1.26\text{x}10^{-11}$
15	Children Play in Residual Water Near House	Illness Residual	*Do respondents reporting children playing in residual waters nearby also report more frequent illness within the household?* Based on the response to a yes or no survey question as to whether children could be observed playing in residucal waters in the proximity of the house.	Significance at 90% confidence level though not at 95%. P=0.0680 Households responding yes, RR 1.29 (1.03 - 1.61)
16		Observed Street Sanitation	*Are households where respondents report children playing in residual waters nearby also in areas with higher observed street sanitation risk?*	P= 0.0479
17	Reported To Know Waterborne Diseases Are Transmitted	Illness Residual	*Are those households where respondents did not know how waterborne illness is transmitted more likely to be ill?*	Opposite expected. Those reporting no knowledge of transmission have lower Illness Residuals. P=0.0383
18		Formal Education	*Are formal education levles lower where respondents did not know how waterborne illness is transmitted ?* Highest grade achieved by heads of household (male/female), average, and maximum.	Significantly lower for female head of household (P=0.0100), not for male, average, or maximum (P>=.1265).

Table 6 Continued. Analysis of Variance and Relative Risk – Analyses and Results.

Row	Grouping	Test Parameter	Hypothesis to be Tested	Result: Significant Not Significant
19	Works as Solid Waste Separator	Illness Residual	*Are households where inhabitants report working as solid waste separators more likely to report frequent illness?*	P=0.3520 RR included 1 in confidence interval.
20		Reported Weekly Income	*Do households where inhabitants report working as solid waste separators report lower incomes?*	P=0.5401
21		Formal Education	*Do households where inhabitants report working as solid waste separators report lower levels of formal education?* Highest grade achieved by heads of household (male/female), average, and maximum.	Significantly lower formal education levels for separators (P<=.0192 except for male education (P=0.0863)
22	Property Owners or Renters	Illness Residual	*Are tenants more likely to be afflicted with illness than are property owners?*	P=0.6982 RR includes 1 in confidence interval.
23		Reported Weekly Income	*Doenants or property owners have significantly lower reported weekly incomes?*	P=0.5924
24		Reported Sanitary Facility Condition	*Do rental properties have sanitation facilities that are in poorer condition or are tenants more critical of the facilities?*	P=0.6623
25		Reported Street Sanitation	*Do renters report higher sanitation risk near their homes or are tenants more critical of the conditions around their dwelling?*	P=0.9168
26		Observed Street Sanitation	*Are rental properties located in areas with higher observed sanitation risk?*	P=0.8172
27		Time living in barrio	*Do owners report longer time living in barrio than do renters?* Serves as a methods check; it can be assumed that renters have less time than property owners.	Average Time in barrio: Renters 2.99 yr, Owners, 7.30 yr. $P=3.531x10^{-14}$
28	Man or Woman Interviewed	Illness Residual	*Are interviews with men likely to report less illness than those with women?*	P=0.6076

Table 7. Simple Linear Regression Analyses and Results.

Row	Independent Varible	Dependent Variable	Hypothesis to be Tested	Result: Significant / Not Significant
1	Number in House	Illness Residual	*Does overcrowding within a household lead to increased risk of waterborne illness?*	Coefficient 0.000 (CI: -0.086 - 0.086) $R^2 = 0.000$
2	Number in House	Reported Weekly Income	*Does household income tend to increase with the number of occupants?*	Coefficient 0.024 (CI: -0.067 - 0.115) $R^2 = 0.001$
3	Reported Weekly Income	Illness Residual	*Do higher reported incomes correlate to lower frequency of illness?*	Coefficient -0.008 (CI: -0.083 - 0.100) $R^2 = 0.000$
4	Reported Weekly Income	Street Sanitation Condition	*Are those households with lower reported weekly incomes more likely to be in those areas with the worst observed sanitation conditions?*	Opposite expected. Positive correlation between reported weekly income and poor street sanitation condition. Coefficient 0.130 (CI: 0.007 - 0.254) $R^2 = 0.009$
5	Reported Weekly Income	Formal Education	*Do higher levels of formal education for the heads of household correlate to higher reporrted household income?* Highest grade achieved by heads of household (male/female), average, and maximum.	Significant positive association for all four measures . $R^2 >= 0.024$
6	Per Capita Weekly Income	Illness Residual	*Does per capita income have any correlation with the incidence of illness in the household?*	Coefficient 0.012 (CI: -0.103 - 0.079) $R^2 = 0.000$
7	Formal Education	Illness Residual	*Does higher level of formal education for the heads of household correlated to redusced frequency of illness?* Highest grade achieved by heads of household (male/female), average, and maximum.	No significant correlation at any of the four measures. $R^2 = 0.000$

Table 7 Continued. Simple Linear Regression Analyses and Results.

Row	Independent Varible	Dependent Variable	Hypothesis to be Tested	Result: Significant / Not Significant
8	Solid Waste Disposal Rasters	Illness Residual	*Does inadequate disposal of solid waste correlate to reported illness and to what extent do the risks associated with improper disposal exted to nearby houses?*	Positive correlation for all rasters and filter settings (t-Statistic >= 0.009). Strongest correlation with 2 meter, non filtered raster. Coefficient 0.122 (CI: 0.005 - 0.207) $R^2 = 0.015$ (Significant)
9	Sanitary Facility Condition Rasters	Illness Residual	*Does the condition of sanitation facilities correlate to reported illness and to what extent do risks associated with facilities in poor condition affect nearby households?*	Positive correlation for all rasters and filter settings (R^2 >= 0.021). Strongest correlation 5 meter, non filtered raster. Coefficient 0.195 (CI: 0.111 - 0.280) $R^2 = 0.038$ (Significant)
10	Street Sanitation Condition Raster	Illness Residual	*Is there a correlation between the severity of the observed sanitation conditions in the street and frequency of illness reported in nearby households?*	Coefficient 0.097 (CI: 0.011 - 0.183) $R^2 = 0.009$ (Significant)
11	Street Sanitation Condition Raster	Sanitary Facility Condition	*Are values generated from the observed street sanitaiton condition raster and those from the sanitary facility raster corrletated?*	Significant positive correlations for all raster settings. (R^2 >= 0.065) Highest strength of correlation with the filtered, 15 m cell raster. Coefficient 0.449 (CI: 0.372 - 0.526) $R^2 = 0.202$ (Significant)
12	Formal Education	Street Sanitation Condition	*Do those reporting lower levels of formal education live in areas with higher observed sanitaiton risks?* Highest grade achieved by heads of household (male/female), average, and maximum.	Not significant for any of the four measures. Coefficients not significantly greater or less than 0. $0.001 <= R^2 <= 0.006$ (Not Significant)

between the number of inhabitants and the Illness Residuals validated the desired transformation.

It was hypothesized that more people in the household could mean more potential wage earners and higher reported weekly household income. However, no significant relationship was observed as shown in Table 7, Row 2.

Reported Weekly Income, Per capita income

Occupants in households with higher weekly incomes could be assumed to be better able to sustain health by means of more and higher quality food and medical care should it be necessary. However, a linear regression showed no significant relationship between reported weekly income (Table 7, Row 3) or per capita weekly income (Table 7, Row 6) and the Illness Residual. As there was no significant relationship between the Illness Residuals and income, it should not be assumed that other significant relationships with health must be traced back to household income to find root causality.

The hypothesis was tested that those reporting lower weekly incomes would live in areas with higher street sanitation hazards as measured by the qualitative observations of residual water in the street. The street sanitation condition values were compared by linear regression to reported weekly income to test for an inverse relationship. The opposite trend was found as is shown in Table 7, Row 4. There was a significant relationship between weekly income and the street sanitation condition values suggesting that wealthier households tend to be in those areas with less sanitary street conditions. This may be due to the likelihood of higher income individuals to have more in-house plumbing and higher water use resulting in more residual water. Higher income houses also tend to be in the older part of the study area.

Table 7, Row 5 shows the results of linear regression between income and formal education measured as the highest grade achieved by the female and male heads of household and the average and maximum of the two. As was hypothesized, there was a positive relationship between formal education levels and reported weekly income for all four measures of education. As was mentioned in earlier comments on regression, it is difficult to state with certainty whether higher education led to higher incomes, higher incomes led to higher education, or if there was another factor at play that led to the positive association.

Education Level of Heads of Household

Formal education levels were measured as the highest grade achieved by the female head of household, the male head of household, and the average and maximum of the two. A linear regression was performed between the measures of formal education and the Illness Residuals in order to evaluate if higher education was inversely related to illness. This tendency has been observed in the literature as Ali et al. (2002) found lower rates of cholera in those households with higher levels of education. However in this research, no significant relationship was found between the Illness Residuals and any of the four measures of formal education as shown in Table 7, Row 7. Analogous to the findings with respect to weekly income, that no significant relationship was found indicates that other factors with significant relationships do not need to be traced back to formal education levels in an effort to find the root causality.

`The hypothesis that those with low levels of formal education would live in areas with poor street sanitation was not supported. Linear regressions between the measures of formal education and the qualitative measure of the street sanitation condition did not attain statistical significance as shown in Table 7, Row 12.

A more direct question about knowledge of hygiene and disease transmission was posed as an additional measure of education. Respondents were asked if they knew how illnesses such as diarrhea and parasites are transmitted. A positive or negative response to this question was assumed to be a better measure of hygiene knowledge than the level of formal education. Interestingly, the Illness Residual was found to be significantly higher (P-value = 0.0383) in those households where the respondent reported a lack of knowledge about waterborne illness transmission as shown in Table 6, Row 17. For those reporting knowledge of waterborne illness, the formal education levels were only significantly higher for the measure of female education with a P-value of 0.0100 as shown in Table 6, Row 18. This suggests that there was a weak link between formal education and knowledge of illness transmission.

Drinking Water Source

The population that reported drinking tap water was found to have significantly higher Illness Residuals (P-value = 0.0057) than the population which reported purchasing filtered water as shown in Table 6, Row 1. The relative risk for those drinking tap water was 1.27 (CI 1.02 – 1.59) over those purchasing filtered water. No distinction was made for if or how tap water was treated before drinking in order to avoid what is often a suspect response.

Reported weekly income (Table 6, Row 2) and formal education levels (Table 6, Row 3) were significantly higher for those who purchased filtered water than for those drinking tap water. As mentioned earlier, linear regressions found no relationship between income, per capita income, or formal education levels and the Illness Residual. For the lack of association with these socioeconomic factors, the investigation can come one step closer to inferring a causal relationship between drinking tap water and illness.

Property Owners or Renters

Analyses by Cifuentes et al. (2002) showed a higher risk of diarrhea for children from rented homes than for those living in proprietary dwellings. However in this study, no significant difference was found by ANOVA between tenants and property owners for the Illness Residuals (Table 6, Row 22), income (Table 6, Row 23), sanitary facility condition as scored by responses (Table 6, Row 24), or the street sanitation condition evaluated by either self ratings (Table 6, Row 25) or values from the qualitative observations (Table 6, Row 26). The only test returning statistical significance between property owners and renters was the comparison of the time reported to have lived in the neighborhood (Table 6, Row 27). This question was intended more as a test of methods as one could logically assume that renters are more transient than are property owners.

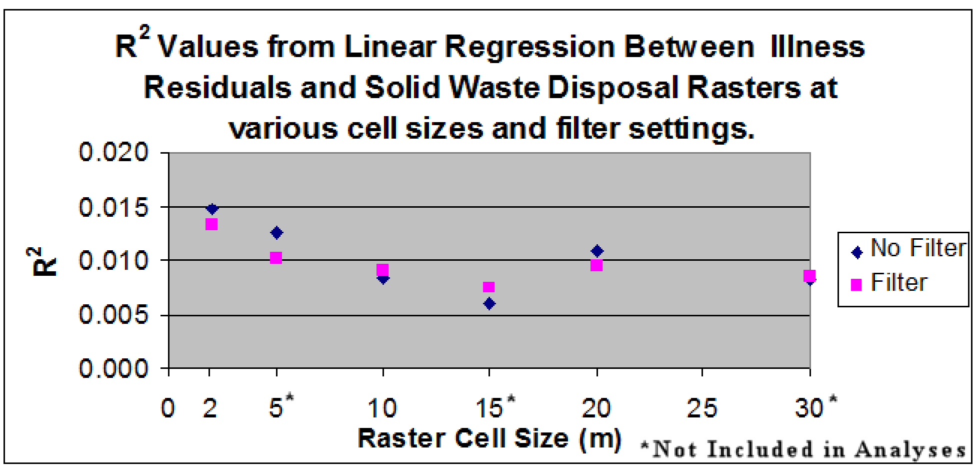
Solid Waste Separators

McGranahan et al. (1997) found that the two groups most directly exposed to solid waste are children and waste separators. The study site is near the city dump and a portion of the local economy comes from separating and selling reusable materials. ANOVA found that those reporting working as solid waste separators did not have significantly higher Illness Residuals (Table 6, Row 19). Reported weekly income of separators was not significantly different than that of the general population (Table 6, Row 20). It is a misconception to assume that solid waste separators are among the poorest as, in some cities, separators can be among the highest earners (Medina, 1999). There was no significant difference between groups in the measures of formal education except for that of the highest grade achieved by the female head of household that was, on average, lower for households with occupants working as separators as shown in Table 6, Row 21. These results should be treated only casually as a low number (15 of 482 with 38 non responses) indicated working as a solid waste separator.

Solid Waste Disposal

The results of the linear regressions between values from the solid waste disposal rasters with various cell sizes and filter settings and the Illness Indices are shown in Table 7, Row 8. There was a significant relationship between the solid waste disposal values and Illness Residuals for all raster cell sizes and filter settings. The R^2 values for linear regressions between the Illness Residuals and the solid waste disposal raster values with different cell sizes and filter settings is shown in Figure 14.

Figure 14. *Linear Regression of Solid Waste Disposal Raster Values and Illness Residuals: R^2 Values by Raster Image Cell Size and Filter Setting.*



The relationship was strongest for the two meter, non filtered raster which represents the lowest modeled dispersion. This suggests that the household is the more appropriate unit for analysis with respect to solid waste; that in-house solid waste solutions and behaviors are more important to the health of residents than are the practices in neighboring houses.

ANOVA and relative risk analyses revealed similar findings as those households which reported inadequate disposal had significantly higher Illness Residuals and a RR of 1.48 (CI: 1.16 – 1.90) over those reporting collection or burning (Table 6, Row 4). Other comparisons found that the reported weekly income for those households that reported dumping solid waste was significantly lower on average (Table 6, Row 5). No significant difference was found in formal education levels (Table 6, Row 6).

Sanitary Facilities

Latrine or Indoor Toilet

Cifuentes et al. (2002) found that illness rates were lower for children living in houses that were connected to sewers than for those with latrines though no distinction was made for the condition of the latrine. This research found no significant difference in the Illness Residuals between households with latrines and those with indoor toilets connected to septic wells (Table 6, Row 7). The scores generated from responses regarding the sanitary facility condition were not significantly different for those with latrines from those with toilets (Table 6, Row 8). It is important that there was uniformity in this evaluation metric as the sanitary facility condition raster was generated with these scores. Income (Table 6, Row 9) and formal education levels (Table 6, Row 10) were significantly higher for those with indoor toilets than for those with latrines.

Sanitary Facility Condition Measure as Score from Responses

The survey population was divided into five groups according to the zero to four integer score from the sanitary facility condition questions described in the methods section. An ANOVA revealed significantly different average Illness Residuals as the groups with incrementally higher scores had similarly higher average Illness Residuals (P-value = .0001). The average Illness Residuals for the five groups are shown in Table 8.

Table 8. Average Illness Residuals for Population Divided by Sanitation Facility Condition Score.

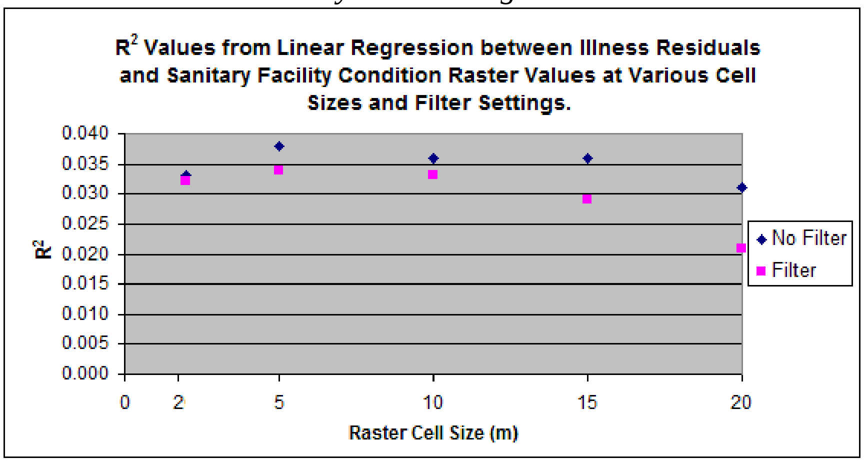
San. Fac. Condition Score	Count	Average Illness Residual	Variance
0	301	-0.09492	1.472935
1	139	-0.03778	1.10147
2	46	0.060835	1.283976
3	28	0.87745	1.815323
4	6	1.172945	2.260662

The five populations were then grouped into control and treatment groups in a series of RR analyses. Regardless of the division of control and treatment groups (example: 'control 0, treatment 1-4', 'control 0-1, treatment 2-4', etc), the relative risk was significantly higher for the group with sanitary facilities scored as in poorer condition. Results from the ANOVA and RR analyses are shown in Table 6, Row 11.

Sanitary Facility Condition Raster

The values generated from the sanitary facility condition rasters were compared to other factors by linear regression and ANOVA. The results of the linear regressions between values from rasters with various cell sizes and filter settings and the Illness Indices are shown in Table 7, Row 9. A positive relationship was found between the sanitary facility condition rasters and the Illness Residuals regardless of the raster cell size or filter setting. The R^2 values from the regressions are shown in Figure 15.

Figure 15. *Linear Regression of Sanitary Facility Condition Raster Values and Illness Residuals: R^2 Values by Raster Image Cell Size and Filter Setting.*



The strongest relationship was found with the non filtered, five meter raster. The level of dispersion modeled by this setting would primarily reflect the condition of on site sanitary facilities with a slight influence from the condition of neighbors' facilities.

Street Sanitation – Presence of Residual Water

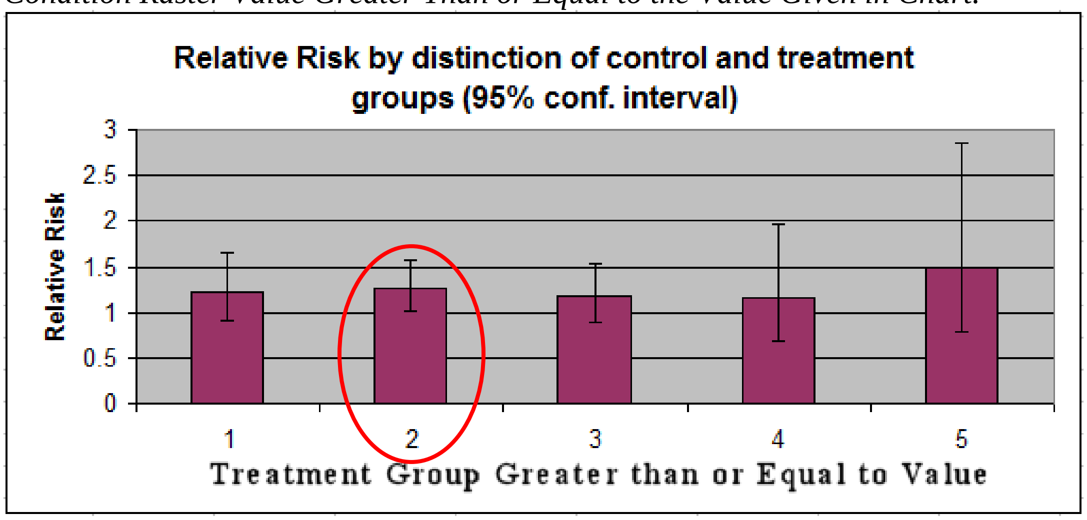
Sanitation Condition Raster

Values from the street sanitation condition raster were compared to other factors by linear regression, ANOVA, and RR. The street sanitation condition values were also subjected to natural log and exponential transformations to test for non linear relationships. The results from the linear regression analyses are shown in Table 7, Row 10. A significant relationship was found between the street sanitation condition values and the Illness Residuals which supports the hypothesis that a higher prevalence of wastewater in the street and incidence of illness are related. The linear regression between the natural log transformation of the street sanitation condition values and the

Illness Residual barely attained statistical significance while the exponentially transformed values did not.

The population was divided into six groups based on whole number ranges of values from the street sanitation condition value such as 0-0.99, 1-1.99, up to the highest value of 6. An ANOVA revealed that Illness Residuals were significantly different between the six populations with a P-value of 0.0229. Relative risk analyses were performed with various cut off values within the zero to six range to designate the control and treatment groups. Figure 16 shows the relative risk according to the value used to separate the control and treatment groups.

Figure 16. *Relative Risk Results. Population Divided Into Control and Treatment Groups. Treatment Group Designated as Those Houses with a Street Sanitation Condition Raster Value Greater Than or Equal to the Value Given in Chart.*



The cut off value at which a significant relative risk was found between the control and treatment groups merits further discussion. When the treatment group was designated as those houses with a street sanitation condition value of two to six and the control group as those houses with values less than 1.99, a significant relative risk was found for those in the treatment group. Referring to the scoring rubric (0. Streets Dry 1. Wet ground 2. Least detectable flow), the designation of a location being a zero, a one, or even up to a two could be considered as the wastewater absent condition. Ratings of three and above refer to areas where there was plainly observable flow. As measured in this research, the rating of two would roughly delimit the presence/absence determination. The findings suggest that the presence or absence determination for wastewater in the street used by Heller (1999) was an ideal measure of sanitation as it relates to public health. Results from ANOVA and RR analyses are shown in Table 6, Row 12.

Sanitation Condition Survey Responses

Each respondent was asked to rate the sanitation condition in the street near their home by indicating if uncontained wastewater was never, sometimes, often, or always present. Grouping respondents according to their answer, the populations reporting wastewater present never, sometimes, often, and always had progressively higher Illness Residuals (P-value = .0096) as shown in Table 9.

Table 9. Illness Residuals for Populations Grouped by Reported Frequency of Uncontained Wastewater in the Street.

Repsonse: How often is there residual water in the street near the house?	*Count*	*Average Illness Residual*	*Variance*
Never	201	-0.16207	1.1936
Sometimes	107	-0.10822	1.230545
Often	32	0.020022	1.466934
Always	170	0.239342	1.759886

Dividing the study population into control and treatment groups for RR based on the response to this question yielded similar results. Those reporting more frequent wastewater in the street had higher relative risks, again with the improved/unimproved designation based on the value of the Illness Residual. The analysis was performed three times such that the control group was designated as those who responded never, never-sometimes, and never-sometimes-often. The relative risk produced for each analysis was 1.32, 1.30, and 1.33 respectively with the lower limit of the confidence interval being no less than 1.04. This suggests that regardless of the designation of the control and treatment groups, the group reporting more prevalent wastewater would have higher rates of illness. The results for ANOVA and RR analyses are shown in Table 6, Row 13.

As a check of methods, the street sanitation condition values generated from the qualitative observations were compared to the respondents' ratings. Again dividing the population based on the never to always scale, those groups who reported uncontained wastewater present in the street never, sometimes, often, or always had progressively higher street sanitation condition raster values with a P-value of 1.263×10^{-11} as shown here in Table 10 and in Table 6, Row 14.

Table 10. Street Sanitation Condition Raster Values for Populations Grouped by Reported Frequency of Uncontained Wastewater in the Street.

Response	*Count*	*Average Street Sanitation Value*	*Variance*
Never	201	1.691037	0.918386
Sometime:	107	1.874366	0.905912
Often	32	1.649074	0.707593
Always	170	2.46927	1.499675

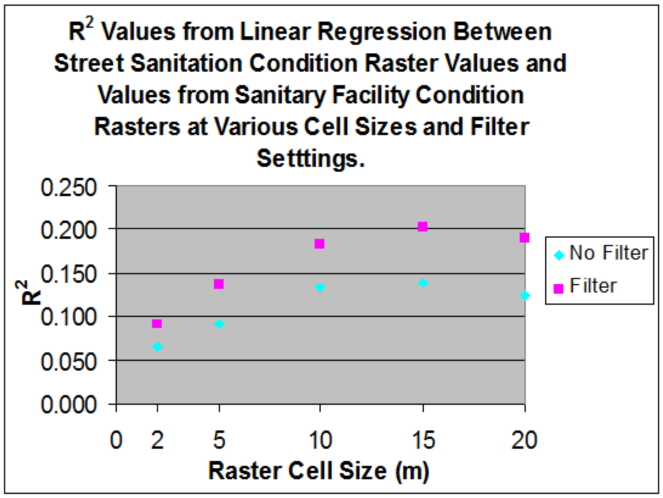
Respondents were asked if children could be observed playing in wastewater in the street near the house. The response to this question was used to group the population for an ANOVA. The population of those houses for which the response was affirmative had a higher average Illness Residual though the difference between was only significant at 90% confidence with a P-value of 0.0680 as shown in Table 6, Row 15. Houses that reported children playing in uncontained wastewater had a significantly higher average value for the street sanitation condition with a P-value of 0.0479 as shown in Table 6, Row 16.

Relationship between Street Sanitation and Sanitary Facilities Raster

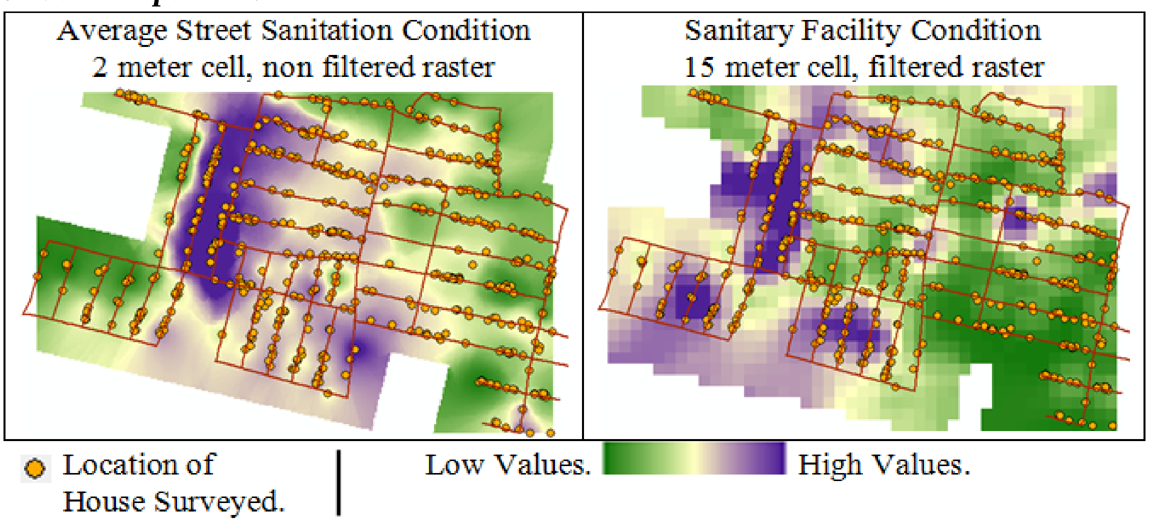
Households in those areas that reported full and overflowing latrines and septic pits may also be in the high hazard areas according to the qualitative evaluations of street. As shown in Table 7, Row 11, linear regressions between values for the street sanitation condition and those of the sanitary facility condition rasters at all cell sizes and filter settings revealed strong associations. This is understandable as the sanitary facility condition rasters are built with higher values corresponding to latrines or septic pits that are full, have overflow tubes, or fill in heavy rains. The areas that have many such latrines or septic wells could be assumed to have more wastewater draining to the street.

Poor condition of the sanitary facilities was associated with the quantity of uncontained residual water observed in the street. The strength of the relationship between the street sanitation condition values those from the sanitary facility condition rasters increases with raster cell size of the latter, peaking at the fifteen meter cell size before declining at the twenty meter cell size. The strongest association between the street sanitation and sanitary facility condition values was found with the filtered, fifteen meter cell raster with a R^2 value of 0.202. Figure 17 shows the R^2 values from linear regressions between the street sanitation condition values and those of the sanitary facility condition rasters with different cell sizes and filter settings. Also included for visual comparison is the street sanitation condition raster and the filtered, fifteen meter cell sanitary facility condition raster

Figure 17. *Association between Values from Hazard Models: Street Sanitation Condition and Sanitary Facility Condition.* R^2 *Values and Visual Comparison Shown.*



Visual Comparison.



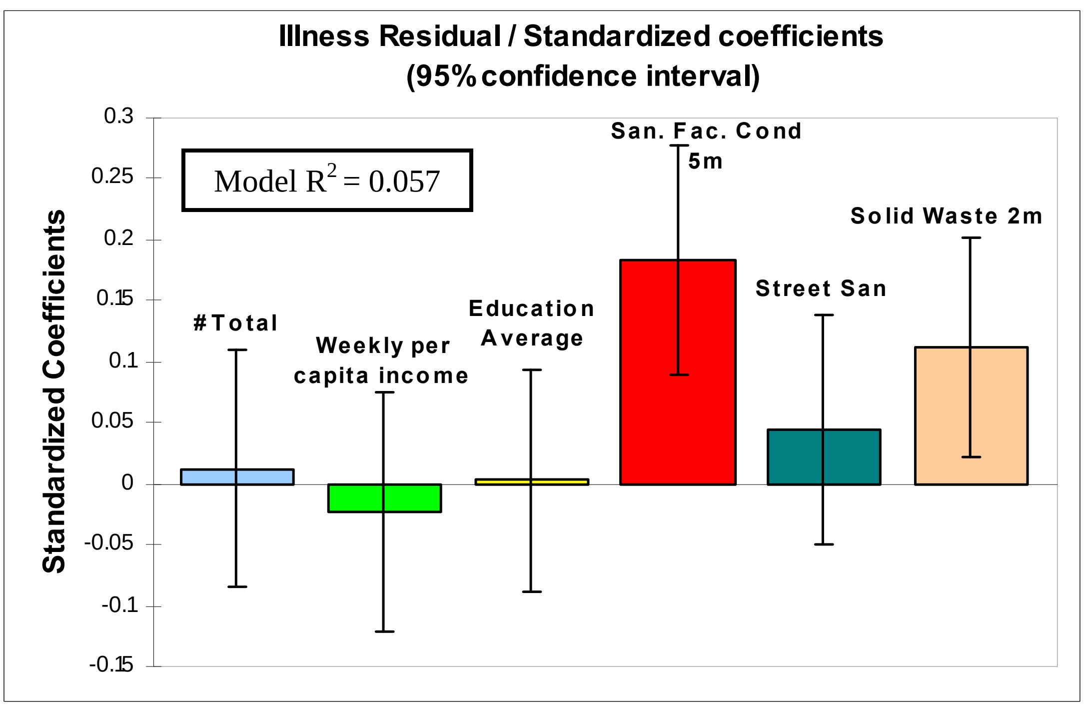
Man or Woman Interviewed

Men are less knowledgeable about the household or neighborhood environment than are women in many cultures (McGranaham, 1997). Women are the primary caretakers in the study area and it could be assumed that would be better able to report illness in the household. However, when the study population was divided according to the gender of the respondent, there was no significant difference between the male respondent and female respondent groups in the Illness Residuals with a P-value of 0.6076 as shown in Table 6, Row 28.

Multivariate Analysis – Relative Importance of Factors

All linear regressions have included only one independent and one dependent variable with the exception of the demographic prediction model that produced the Illness Residuals. The following multivariate linear regression includes the street sanitation condition values, sanitary facility condition values from the non filtered, five meter cell raster, and solid waste disposal values from the non filtered, two meter cell raster. Though earlier analyses showed no significant relationship between total number of people in the house, per capita weekly income, or average level of formal education of the heads of household, these factors were included in a the multivariate analysis in order to show their relative importance. Figure 18 shows the model coefficients for the selected factors as explanatory variables of the Illness Residual. All variables ranges were normalized to 0-100 in order to allow for direct comparison of the coefficients.

Figure 18. *Coefficients of Multiple Linear Regression with Illness Residual.*



The total number of inhabitants in the house, weekly per capita income, and average education of the heads of household had no significant relationship with the Illness Residual. Only the sanitary facility condition values and the solid waste disposal values had significant relationships with the Illness Residuals. Interestingly, while the street sanitation condition values and the Illness Residuals were significantly related in the pair wise linear regression, significance was not maintained when combined with the other factors. It appears that though the sanitary facility condition values and the street sanitation condition values are related with one another, the sanitary facility condition values explain more of the variance in the Illness Residuals.

This simple model explains more of the variation in the Illness Residuals ($R^2 = 0.057$) than did any of the pair wise linear regressions. The multivariate analysis graphic offers a comparison of the relative importance of the variables as they relate to the measure of waterborne illness in this study.

Conclusion

This research has shown that statistically significant relationships can be observed at the household scale within a neighborhood between the proximity to environmental hazards and a measure of waterborne illness frequency. Statistically significant relationships were found between the measure of illness and the hazards from inadequate disposal of solid waste, sanitary facilities in poor condition, and uncontained wastewater in the street. The results from the modeled dispersion of hazards suggest that adequate disposal of solid waste is of greater importance to the health of the residents of the individual household than for the neighborhood as a whole. No significant difference in health was observed between households that have an indoor toilet or an outdoor latrine but the condition of the sanitary facilities was as important as or more so than hazards from the prevalence of uncontained wastewater in streets. The findings presented here support the conclusions of Heller (1999); that the presence or absence of residual wastewater in the street is an important measure as it relates to the incidence of waterborne illness in nearby houses.

In this relatively homogenous study site, no significant linear regression relationships or differences between populations were found between illness and socioeconomic factors such as income, per capita income, education, or whether the residents were renters or owners. The drinking of tap water from the municipal network may be a risk factor as those who reported drinking tap water had significantly higher measures of illness.

The Illness Indices and Illness Residuals that were used as metrics did not attempt to quantify episodes of illness but were sufficient to suggest the relative importance of different factors as they relate to public health. Quantifying diarrheal episodes would require more detailed data collection methods such as the daily visits in research by Checkley et al. (2004) or at most, the two week recall period for health surveys recommended by WHO (1984) that was used by Cifuentes et al. (2002) and McGranahan (1997). Additional research could use methods similar to those in this study but substitute quantifiable episodes of illness for Illness Index and Residuals.

The residents of the periurban area which were the subjects of this study did not enjoy the same level of access to urban services as do their counterparts in most developed-world urban areas. The hazards from inadequate disposal of solid waste, sanitary facilities in poor condition, and the presence of uncontained wastewater were addressed. However, it is necessary to consider that an individual living in a contaminated environment does not necessarily make contact with the hazard agent. 'Exposure' cannot be defined by the location of a residence (Jarup, 2004). This study does not identify an 'exposed population' but rather what could be considered a population more likely to be exposed. The research implications of this warrant more rigorous methods for those studies attempting to quantify the effects of exposure. It would be more difficult but more scientifically meaningful to quantify morbidity in a population identified as having been exposed to a hazard.

This research offers a model that suggests the relative importance of a variety of factors with respect to waterborne illness. The highest explanation of variance was achieved by the multivariate analysis which included several independent variables. There are surely other factors which were not included in this research that would improve the model. Hygiene behavior was not deduced from measured variables and would be a useful addition in later research. Additionally, as the data were obtained from surveys with one visit per house, a relatively high level of randomness was likely. Further research is required to adequately measure environmental hazards, risks generated from those hazards, and to obtain data that quantify the results of exposure.

References

Addinsoft Inc. (2009). XLstat 2009. [Computer software]. Paris, France: Addinsoft Inc.

Ali, M., Emch, M., Donnay, J. P., Yunus, M., and Sack, R. B. (2002). Identifying environmental risk factors for endemic cholera: a raster GIS approach. *Health and Place, 8*, 201-210.

Armstrong, M. P., Rushton, G., and Zimmerman, D. L. (1999). Geographically masking health data to preserve confidentiality. *Statistics in Medicine, 18,* 497-525.

Bairagi, R. (1980). Is income the only constraint on child nutrition in rural Bangladesh? *Bulletin of the World Health Organization, 58,* 767-772.

Bateman, O. M. and Smith, S. (1991). *A Comparison of the Health Effects of Water Supply and Sanitation in Urban and Rural Guatemala.* (Rep. No. Field Report No. 352). Water and Sanitation for Health.

Bateman, O. M., Smith, S., and Roark, P. (1993). *A comparison of the health effects of water supply and sanitation in urban and rural areas of five African countries.* (Rep. No. 398). Washington, DC: U.S. Agency for International Development.

United States Central Intelligence Agency (CIA). (2009). Dominican Republic. In The World Factbook. Retrieved October 15, 2009, from https://www.cia.gov/library/publications/the-world-factbook/geos/dr.html.

Checkley, W., Gilman, R. H., Black, R. E., Epstein, L. D., Cabrera, L., and Sterling, C. R. (2004). Effect of water and sanitation on childhood health in a poor Peruvian peri-urban community. *The Lancet, 363,* 112-118.

Cifuentes, E., Mazari-Hiriart, M., Carneiro, F., Bianchi, F., and Gonzales, D. (2002). The risk of enteric diseases in young children and environmental indicators in sentinel areas of Mexico City. *International Journal of Health Research, 12,* 53-62.

Coura Cuentro, S. and Gadji, D. M. (1990). The Collection and Management of Household Garbage. In J.E.Hardoy, S. Cairncross, and D. Satterthwaite (Eds.), *The Poor Die Young.* (pp. 169-188). London: Earthscan Publications LTD.

Devilly, G. J. (2007). ClinTools Software for Windows. (Version 4.1) [Computer software]. Melbourne, Austrailia.

Esrey, S., Gough, J., Rapaport, D., Sawyer, R., Simpson-Hèbert, M., Vargas, J., and Winblad, U. (1998). *Ecological Sanitation.* (Rep. No. 9158676120). Stockholm: Swedish International Development Cooperation Agency.

Esrey, S. A. and Feachem, R. G. (1989). *Interventions for the control of diarrhoeal diseases among young children: Promotion of food hygiene.* (Rep. No. 89.30). Geneva: World Health Organization.

Environmental Systems Research Institute, Inc. (ESRI). (2006). ArcGIS. (Version 9.2) [Computer software]. Redlands, CA: ESRI.

Feachem, R., Garelick, H., and Mara, D. (1980). *Health Aspects of Excreta and Wastewater Management.* Baltimore: Johns Hopkins University Press / World Bank.

Guerrant, R. (1994). Twelve messages from enteric infections for science and society. *American Journal of Tropical Medicine and Hygiene, 51,* 26-35.

Hancock, T. (1996). Health and sustainability in the urban environment. *Environmental Impact Assessment Review, 16,* 159-277.

Haug, D. B., MacEachren, A. M., Boscoe, F. P., Brown, D., Marra, M., and Beedasy, J. (1997). Implementing exploratory spatial data analysis methods for multivariate health statistics. In (pp. 205-213). Cincinnati, OH: *Proceedings of GIS/LIS '97 Annual Conference and Exposition*.

Heller, L. (1999). Who really benefits from environmental sanitation services in the cities? An intra-urban analysis in Betim, Brazil. *Environment and Urbanization, 11,* 133-144.

Hogrewe, W., Joyce, S. D., and Perez, E. A. (1993). *The Unique Challenges of Improving Peri-Urban Sanitation.* (Rep. No. 86). Washington, DC: USAID.

Hubbard, B., Gelting, R., Baffigo, V., and Sarisky, J. (2005). Community environmental health assessment strengthens environmental public health services in the Peruvian Amazon. *International Journal of Hygiene and Environmental Health, 208,* 101-107.

Jarup, L. (2004). Health and environment information systems for exposure and disease mapping, and risk assessment. *Environmental Health Perspectives, 112,* 995-997.

Kalbermatten, J. M., Julius, D. S., and Gunnerson, C. G. (1982). *Appropriate Sanitation Alternatives: A Technical and Economic Appraisal.* Baltimore, Maryland: Johns Hopkins University Press.

Katakura, Y. and Bakalian, A. (1998). *People, Poverty And Pipes: A Program Of Community Participation and Low-Cost Technology Bringing Water and Sanitation to Brazil's Urban Poor.* Washington, DC: World Bank.

Komives, K. (1999). *Designing pro-poor water and sewer concessions: Early lessons from Bolivia.* The World Bank.

Lopez de Romana, G., Brown, K. H., and Black, R. E. (1989). Longitudinal studies of infecctious disease and physical growth of infacts in Huascar, and underprivileged peri-urban community in Lima, Peru. *American Journal of Epidemiology, 129,* 769-784.

Mayer, J. D. (1983). The role of spatial analysis and geographic data in the detection of disease causation. *Social Science and Medicine, 17,* 1213-1221.

McGranaham, G., Leitmann, J., and Surjadi, C. (1997). *Understanding environmental problems in disadvantaged neighborhoods: Broad spectrum surveys, participatory appraisal and contingent valuation.* (Rep. No. Working Paper No. 16). Washington, D.C. UNDP/UNCHS (Habitat) World Bank.

Medina, M. (1999). *Globalization, Development, and Municipal Solid Waste Management in Third World Cities.* Working paper from El Colegio de la Frontera Norte, Tijuana, Mexico.

Mehta, N. M. (2006). The preference bias in sanitation: Explaining failures in public provision. In New Orleans, LA: 2006 Public Choice Meetings.

Microsoft Corporation. (2003). Microsoft Office Exel 2003. [Computer software]. Redmond, WA: Microsoft Corporation.

Moore, D. A. and Carpenter, T. E. (1999). Spatial Analytical Methods and Geographic Information Systems: Use in health Research and Epidemiology. *Epidemiologic Reviews, 21,* 143-161.

Moore, M., Gould, P., and Keary, B. (2003). Global urbanization and impact on health. *International Journal of Hygiene and Environmental Health, 206,* 269-278.

Nance, E. B. (2005). Multistakeholder evaluation of condominial sewer services. *American Journal of Evaluation, 26,* 480-500.

Nawab, B., Nyborg, I. L. P., Esser, K. B., and Jenssen, P. D. (2006). Cultural preferences in designing ecological sanitation systems in North West Frontier Province, Pakistan. *Journal of Environmental Psychology, 26,* 236-246.

Paterson, C., Mara, D., and Curtis, T. (2007). Pro-poor sanitation technologies. *Geoforum, 38*, 901-907.

Pontificia Universidad Católica Madre y Maestra (PUCMM). (2009). Satellite image. Santiago, Dominican Republic, Laboratorio de Cartografia, Centro de Estudios Urbanos y Regionales.

Satterthwaite, D. (1993). The impact on health of urban environments. *Environment and Urbanization, 5,* 87-111.

Schorling, J. B., Wanke, C. A., and Schorling, S. K. (1989). A prospective study of persistent diarrhea among children in an urban Brazilian slum: patterns of occurrence and etiologic agents. *American Journal of Epidemiology, 132*, 144-156.

Schultz, S. K. and McShane, C. (1978). To engineer the metropolis: sewers, sanitation, and city planning in late-nineteenth-century America. *The Journal of American History,* 389-411.

Scotch, M., Parmanto, B., Gadd, C. S., and Sharma, R. K. (2006). Exploring the role of GIS during community health assessment problem solving: experiences of public health professionals. *International Journal of Health Geographics, 5*.

UN Habitat (1987). *Water-Supply and Waste-Disposal Managment: Impact-Evaluation Guidelines.* Nairobi: United Nations Centre for Human Settlements.

Waterkeyn, J. and Cairncross, S. (2005). Creating demand for sanitation and hygiene through community health clubs: A cost-effective intervention in two districts in Zimbabwe. *Social Science and Medicine, 61*, 1958-1970.

Watson, G. (1995). *Good Sewers Cheap? Agency-Customer Interactions in Low-Cost Urban Sanitation in Brazil.* Washington, DC: The World Bank.

Weber, A. F. (1899). *The growth of cities in the nineteenth century: A study in statistics.* New York: Macmillan.

Westlake, A. (1995). Strategies for the use of geography in epidemiologic analysis. In M.de Lepper, H. J. Scholten, and R. M. Stern (Eds.), *The added value of geographical information systems in public and environmental health.* (pp. 135-144). Boston, MA: Academic Press.

World Health Organization (WHO). (1984). *Guidelines for a sample survey of diarrhoeal disease morbidity, mortality and treatment rates.* Geneva: WHO.

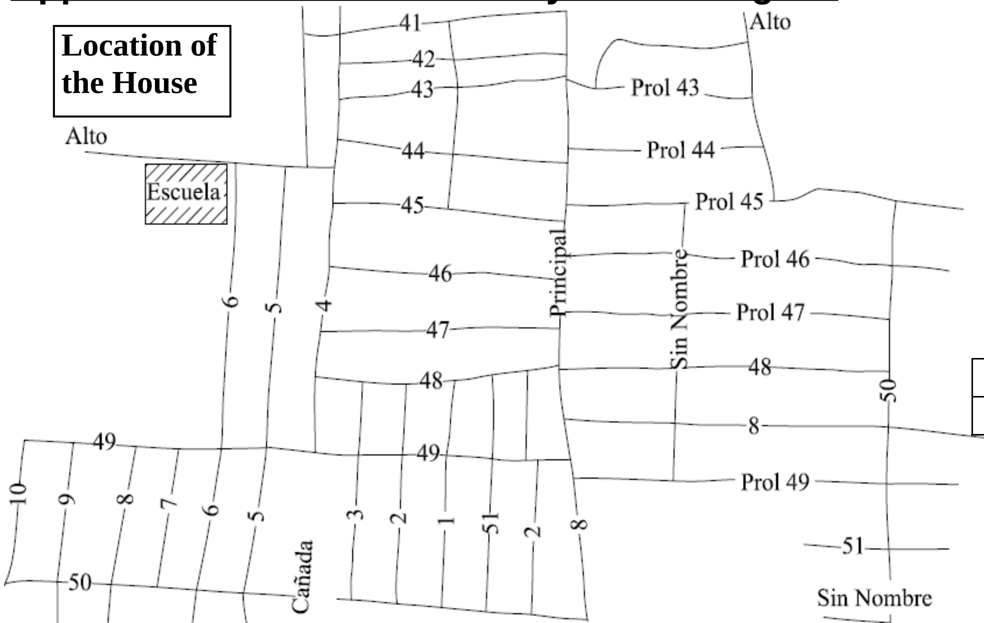
World Health Organization (WHO). (1993). *Improving water and sanitation hygiene behaviours for the reduction of diarrhoeal disease. The report of an informal consultation.* (Rep. No. CWS/93.10). Geneva: World Health Organization.

World Heatlh Organization (WHO) and United Nations Children's Fund (UNICEF). (2004). *Meeting the MDG Drinking Water and Sanitation Target: A mid-term assessment of progress.* Geneva and New York: World Health Organization and United Nations Children's Fund.

Wright, A. M. (1997). *Toward a strategic sanitation approach: improving the sustainability of urban sanitation in developing countries.* Washington, DC: International Bank for Reconstruction and Development / World Bank.

Appendix A. Household Survey Form - English

Location of the House



Interviewers:________________________________				
Type of Dwelling	House	Apartment	Pension	With Store
¿What is the house made of?				With Workshop

Floor		Wall			Roof		
Cement	Dirt	Block	Wood	Zinc	Cement	Zinc	Cans
Observations:							

Who was interviewed?	Man	Woman	Child	Other
Notes:				

What is the name of the barrio here? ____________________

Do you participate in community groups?

Yes Which ones?____

No, Why not? ____

What are the most important community organiations?

____

If the community needs something, who makes it happen? Name individuals or organizations.

____

What is something that the community has achieved together?

____

Where do youth meet and what to they do?

____

What does the barrio need?

____

____

Demographics

What are the ages and genders of those who live here?

	<1	1-4	5-12	13-18	19-30	31-45	46-64	65+
F								
M								

How many adults are legally declared? | __ of ___ | None | All |

How many children are legally declared? | __ of ___ | None | All |

How many years have you lived here? ______

Where is your family from?		Region of Cibao		Santo Domingo
Santiago	North	South	East	Other:_________

Do you own or rent the dwelling? Owners Renters

Owners	Renters
Do you have the legal title of the house? Sí No	How much do you pay in rent?______

What are the principie occupations of those who live here?

Free Trade Zone	Separator	Public Employee	Handyman
Pensioned	Military/Police	Private Employee	Housewife
Unemployed	Private Business	Traveling Sales	__________

At times, anyone from the house works as a separator in the landfill?

Yes How many times a week? ______	No

What is the weekly household income?

Thousands	<1	1	2	3	4	5	6	7	8	9	10	10+

Which of the following do you have?

Car/Pickup	Stereo Equipment	Television	Refrigerator	Washer
Moped/Motorcycle	Computer	Power Inverter	Stove	Fan

Do you relieve a remittance from another country?	Yes How much? _________ From what country?__________	No

Do you Cook at home or eat out?	At Home	Eat Out	Both

Hygiene

How many times a week do you have water? _____   For how many hours? _____

Where does the water come from?		
In house tap	Yard tap	Trucks

Where do you store water?	55 gal	Bucket	Jugs	Tank

If the water doesn't come, how long does stored water last?__	Is water covered? Yes No

Where do you get drinking water?	Tap	Bottled	Rain

What treatment do you do before drinking (if not bottled water)?	Filter	Boil	Bleach	None

Bathroom Questions What kind of bathroom do you have?

Latrine	Toilet	Other________________	None

How much did it cost? ___________ How many years ago? ____

It serves your needs? Yes No	Is it full? Yes No

The hole/pit has an overflow to the street? Yes No

The hole/pit fills when it rains? Yes No

Comments:

Where do you dispose of wash and bath water?

Drain to the street	Latrine hole/pit	Yard

How frequently are there residual waters in the street or gutter near your house?

Never	Sometimes	Often	Always

Children play in residual water near your house? Yes No

How much do you pay in these services monthly?
Elec. ______ Water______ Tel./Cel. ______ Cable ______

If there were the opportunity to connect to municipal sewer, there would be the costs of tubes on your property plus an initial quota. Would you connect if the quota were 1500 pesos?

1500 Yes → 2500 Yes No
1500 No → 1000 Yes No

What is the maximum that you would pay to connect? _____

Do you consider 150 pesos monthly a fair charge for municipal water and sewer?

150 Yes → 250 Yes No
150 No → 100 Yes No

What is the maximum that you consider fair? ____

Trash is collected? __ Sí → By whom? / No → What do you do?

By whom?		Dumped to:		Burned
Municipal	Others	Patio/ Other	Street	

Education What is the highest level of education achieved?

	None	Primary	High School	University
Woman	0	1 2 3 4 5 6 7 8	1 2 3 4	Strt. Fin.
Man	0	1 2 3 4 5 6 7 8	1 2 3 4	Strt. Fin.

How many children go to school?____ How many adults? ___
To the university? ____ Do you attend technical classes? Yes No

What are the obstacles to going to school?

____

Health What are the most common illnesses in the community?

____

Do you know what causes illnesses such as diarrhea and parasites?
Sí No

In the last three months, anyone has become ill of the following illneses: Diarrhea, Fever, Parasites

Yes Give the age and illness of the affected	No

Anyone has died of one of the given illnesses?

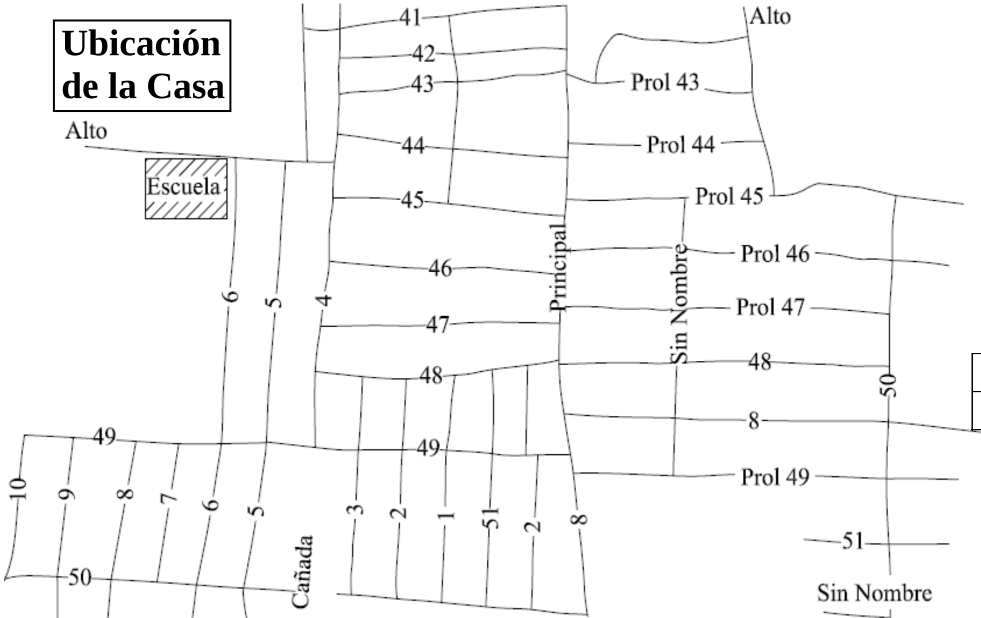
Yes Give the age and illness.	No

Has anyone suffered from typhoid fever? Yes No

In the last three months, anyone has missed a day of work or school due to a stomach pain or diarrhea? How many days? [| | | |]

If you needed to take someone to the hospital for diarrhea, fever or parasites, how much did you spend? Yes No
Consultation: _____________ Medicine: ___________________

Appendix B. Household Survey Form - Spanish



Entrevistadores/as:____________________				
Tipo de Vivienda	Casa	Apartamento	Pensión	Tiene Colmado
¿De qué está hecha la vivienda?				Tiene Taller

Piso		Pared			Techo		
Cemento	Tierra	Block	Madera	Zinc	Cemento	Zinc	Hojas
Observaciones:							

¿A quién se le hará la entrevista?	Hombre	Mujer	Hijo	Familiar
Notas:				

¿Como se llama el barrio donde Usted vive? ____________________

¿Usted participa en reuniones de un grupo comunitario?

Sí, ¿Cuáles? ______________________________________

No, ¿Por qué no?______________________________________

¿Cuáles son las organizaciones comunitarias más importantes?

____

¿Si la comunidad necesita algo, cuáles son los que se mueven para lograrlo? Nombre individuales u organizaciones.

____

¿Cuente un trabajo que la comunidad ha realizado junto?

____

¿Dónde se reúnen los jóvenes y qué hacen?

____

¿Que le falta al barrio?

____

____

Demográficos

¿Cuáles son las edades y los géneros de los que viven en esta casa?

	<1	1-4	5-12	13-18	19-30	31-45	46-64	65+
H								
V								

¿Cuántos de los adultos están declarados? | __ de __ | Ninguno | Todos |

¿Cuántos de los niños están declarados? | __ de __ | Ninguno | Todos |

¿Cuantos años tienen viviendo aquí? ______

¿De donde es la familia suya?		Región del Cibao		Santo Domingo
Santiago	Del Norte	Del sur	Del este	Otro:__________

¿Ustedes son dueños de la casa o es alquilada? Dueños Alquilada

¿Ustedes tienen los papeles legales de la casa? Sí No	¿Cuánto pagan de alquiler? ________

¿Cuáles son los empleos principales de los que viven aquí?

Zona Franca	Buzo/a, separador	Empleado publico	Chiripero/a
Pensionado/a	Militar/Policía	Empleado privado	Ama de casa
Desempleado/a	Negocio privado	Vendedor ambulante	__________

¿Algunas veces, alguien de la casa trabaja como separador o buzo?

Sí ¿Cuántas veces por la semana? ______	No

¿Cuánto es el ingreso semanal de la casa?

Miles	<1	1	2	3	4	5	6	7	8	9	10	10+

¿Cuales de los siguientes artículos Ustedes tienen?

Carro/guagua	Equipo de música	Televisor	Nevera	Lavadora
Motor/pasola	Computadora	Inversor	Estufa	Abanico

¿Ustedes reciben una remesa de afuera del país?	Sí ¿Cuánto por mes? _________ ¿De cuál país?______________	No

¿Comen en la casa o salen a comer? | En casa | Afuera | Los Dos |

Higiene

¿Cuántas veces por semana viene el agua? _____
¿Cuántas horas dura? ______

¿De dónde proviene el agua?		
Llave en la casa	Llave en el patio	Camiones

¿Donde almacenan el agua?	Tanques	Cubos	Galones	Tinaco

¿Si no viene el agua, se les termina en cuántos días?____

¿El agua está tapada? Sí No

¿De donde proviene el agua de tomar?	Llave	Botellón	Lluvia

¿Qué le hace antes de bebérsela? (si no es de botellón)	Filtrar	Hervir	Cloro	Nada

Preguntas del baño ¿Que tipo de baño tienen?

Sanitario	Inodoro	Otro________________	Ninguno

¿Cuánto costó para hacer? ___________ ¿Hace cuántos años? ____

¿Les sirve bien? Sí No

¿Esta lleno? Sí No

¿El hoyo o pozo tiene un tubo o desagüe a la calle? Sí No

¿Los hoyos o los pozos de sanitario se llenan cuando llueve? Sí No

Comentarios del baño:

¿Para dónde salen las aguas de bañarse, lavar, y fregar?

Por tubo hacia la calle o el contén	Hoyo o Pozo Séptico	El Patio

¿Con qué frecuencia hay aguas negras en la calle o el contén en el frente de su casa?

Nunca	A Veces	Muchas Veces	Siempre

¿Niños juegan en las aguas sucias de la calle cerca de la casa? Sí No

¿Cuánto pagan de estos servicios mensualmente?
La luz ______ El agua______ Tel./Cel. ______ Cable ______

Habiendo la oportunidad conectar a una cloaca, habría una cuota de conexión más los gastos de los tubos dentro de la propiedad suya. ¿Usted conectaría si la tarifa inicial es 1500 pesos?

1500 Sí → 2500 Sí No
1500 No → 1000 Sí No

¿Cuál es el máximo que Usted pagaría para conectar? _____

¿Usted considera 150 pesos mensual una tarifa justa para el servicio del agua y de la cloaca?

150 Sí → 250 Sí No
150 No → 100 Sí No

¿Cuál es el máximo que Usted considera justo? _____

¿Se recoge la basura? Sí → ¿Por quién?

Ayuntamiento	Otros pagados

No → ¿Qué hace con la basura?

Se tira:		Se quema
al patio/ al monte	a la calle	

Educación ¿Hasta que grado de la escuela lograron?

	Ninguno	Primaria	Bachiller	Universidad
Mujer	0	1 2 3 4 5 6 7 8	1 2 3 4	Epz. Term.
Hombre	0	1 2 3 4 5 6 7 8	1 2 3 4	Epz. Term.

¿Cuántos de los menores van a la escuela?____De los mayores de 18 ___

¿Cuántos van a la universidad? ____ ¿Asisten clases técnicas? Sí No

¿Cuáles son los obstáculos de ir a la escuela o centro educativo?

____

Salud ¿Cuales son las enfermedades que más afectan la comunidad?

____

¿Usted sabe de donde provienen las enfermedades tales como diarrea y parásitos? Sí No

¿En los últimos tres meses, alguien se le ha enfermado por algunas de estas enfermedades: Diarrea, fiebre, o parásitos?

Si Diga la enfermedad y edad del afectado.	No

¿Alguien se le ha muerto de una de tales enfermedades?

Si Diga la enfermedad y edad.	No

¿Alguien de la casa ha sufrido alguna vez de fiebre tifo? Sí No

¿En los últimos tres meses, alguien ha perdido un día de trabajo o de la escuela por un dolor de barriga o por diarrea? ¿Cuántos?

¿Han tenido que llevar a alguien al hospital por diarrea, fiebre tifo, o parásitos, cuanto gastó en la consulta y en los medicamentos? Sí No

Consulta: ________________ Medicamentos: ____________________

Appendix C. Street Sanitation Condition: Qualitative Evaluation Criteria for Presence of Residual Water in Streets of Study Area.

Rating 0. Residual Water Absent, Streets Dry.

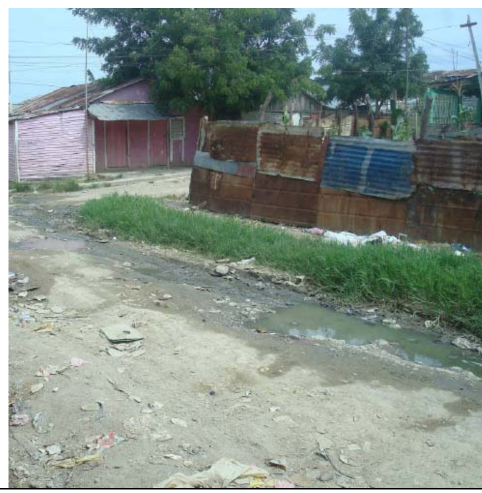
Rating 1. Least Detectable Levels of Residual Waters, Damp Street or Curb.

2. Marginally Observable Flow in Street Gutters, No Standing Water.

Rating 3. Plainly Observable Flow in Street Gutters and Ground Channels With or Without Very Minimal Observable Standing Water.

Rating 4. Considerable Flow and Some Standing Water Observed.

Rating 5. Considerable Flow and Standing Water Present, Odors Detected.



Rating 6. Continuous Flow and Standing Water Always Present, Strong Odors, Most Severe Conditions.

Appendix D. Summary of Variables and Numbers of Responses

	Variable	Description	Number of Values or Responses
Continuous Variables	Solid Waste Disposal Raster Values	Value extracted from solid waste disposal raster at the household's location. Rasters were generated based on reponses to question about waste disposal method.	520
	Sanitary Facility Condition Raster Valules	Value extracted from sanitary facility condition raster at the household's location. Rasters were generated based on scoring responses about the condition of sanitary facilities.	520
	Street Sanitation Condition Raster Values	Value generated from sanitary facility condition raster at the household's location. Rasters were generated from qualitative observations of the presence of uncontained wastewater in streets.	520
Integer Values Or Values Derived From Integer Values	Total Number in House	The sum of the number of individuals reported in the various age groups.	520
	Illness Index	Value based on scoring of responses regarding waterborne illness in the household. Number of positive repsonses for individual questions as follows: Has been sick in last three months: 168 Has had typhoid fever (typhoid not verified): 119 Has lost a day of school or work for illness: 119 Has been to hospital for diahrrea: 132	Illness Index / # 0 / 256 1 / 92 2 / 90 3 / 62 4 / 20 Total: 520
	Illness Residual	The Illness Index minus the predicted Illness Index based on the number and ages of people in the house (demographic correction model).	520
	Formal Education	Highest grade of formal education achieved by the female head of household, the male head of household, and the average and maximum of the two.	Female Level: 482 Male Level: 393 Average: 518 Maximum: 518
	Reported Weekly Income	Reported average weekly household income.	466
	Per Capita Weekly Income	Reported weekly income divided by the number of household inhabitants.	466
	Years Living in Area	Reported time having lived in the neighborhood.	509
	Self Reported Sanitary Facility Condition	Value based on scoring of responses regarding aspects of the sanitary facilities in the household. Number of positive repsonses for individual questions as follows: Facilities do not serve needs: 48 The latrine of septic pit is full: 107 The latrine or septic pit as a drain to the street: 89 The latrine or septic pit fills with water when it rains: 95	Score / # 0 / 301 1 / 139 2 / 46 3 / 28 4 / 6 Total: 520

	Variable	Description	Number of Values or Responses
Multiple Choice Response	Self Reported Street Sanitation Condition	Respondents were asked if there is never, sometimes, often, or always uncontained wastewater in the street near their house.	Never: 201 Sometimes: 107 Often: 32 Always: 170
Dichotomous Variables	Drinking Water Source	Reported drinking water source. Primary water sources were purchased filtered water and tap water.	Filtered: 260 Tap: 248 Total: 508
	Solid Waste Disposal Method	Respondents were asked if trash was collected, burned, or thrown to the street or vacant areas.	Collected: 385 Burned: 34 Thrown to: Street: 4 Vacant Areas: 71 Total: 494
	Sanitary Facility Type	If the sanitary facility is a latrine or a septic pit with an indoor toilet.	Latrine: 144 Septic Pit: 363 Total: 507
	Works as Solid Waste Separator	Some residents separate and sell usable materials from the nearby city dump.	Separator: 15 Not Separator: 469 Total: 484
	Property Owner or Renter	If the property is owned by the inhabitants or is rented.	Rent: 102 Own: 401 Total: 503
	Gender of Respondent	Female or male respondent.	Female: 344 Own: 134 Total: 478
	Know How Waterborne Illness is Transmitted	Repsondents were asked if they knew how waterborne illness is transmitted.	Do Not Know: 127 Know: 377 Total: 504
	Children Play in Residual Water	Respondents were asked if children can be regularly observed playing in residual waters in the street in the area around the house.	Play in Water: 147 Do not: 328 Total: 475

Appendix E. Human Subjects in Research Approval.

MichiganTech
Michigan Technological University

Office of Research Integrity
and Compliance

317 Admin. Building
1400 Townsend Drive
Houghton, MI 49931
906.487.2902

MEMO

TO: Dr. William Bulleit, Civil and Environmental Engineering
CC: Glenn Vorhes, Civil and Environmental Engineering
Jim Mihelcic, Civil and Environmental Engineering

FROM: Joanne Polzien, Director Research Integrity and Compliance Joanne Polzien

DATE: August 6, 2009

SUBJECT: Approval M0471

Your application to use human subjects in research or classroom situations has been reviewed with the following determination:

Protocol #: M0471
Protocol Title: "Environmental hazards and incidence of disease in households of periurban Dominican Republic: An analysis using GIS"
Approved Dates: August 10, 2009 through August 9, 2010

Approvals are granted for up to a one year period. You will need to request a continuation for each year of the project six weeks prior to the end date indicated above for each year of the project. The Office of Research Integrity and Compliance will make every effort to send the Principal Investigator annual reminders. However, the Principal Investigator is responsible for submitting annual Continuation Forms in advance of the expiration date for the project. It is very important that these expiration dates are not missed. Failure to submit annual review materials on time will result in the termination of this protocol.

This approval applies only for this project, and only under the conditions and procedures described in the application; if any changes are made in the protocol or conditions set forth in the application, the principal investigator must obtain a separate approval before these changes take place. The approved project will be subject to surveillance procedures requiring periodic review. This review will consist of consulting with the principal investigator and examining the appropriate project records.

Individual identification of human subjects in any publication is an invasion of privacy. Before beginning a project involving human subjects, and only if required, the principal investigator must obtain a properly executed informed consent from each subject and/or the person legally responsible for the subject. **If a consent form has been reviewed and approved it has been attached with an official date stamp on it. Only copies of the official date stamped informed consent is to be distributed to participants relating to this project. If any changes or modifications are needed regarding this form, you must first submit the revised document for review and approval prior to use.** The principal investigator must retain informed consent forms on file for at least three years after the end of the project. If a project involves a high level of risk, copies of the signed informed consent forms must be filed with the Human Subjects Committee; if this is the case, you will be notified.

This document is on file in the Office of Research Integrity and Compliance. If you have any questions, please contact me at 487-2902 or jpolzien@mtu.edu.