Distribution Agreement

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

Signature:

_____ _____

Alexander Emmitt Date

Pathogenic contamination of tomatoes, cantaloupe, and jalapenos on farms in Coahuila and Nuevo León, Mexico

By

Alexander Emmitt
Master of Public Health

Epidemiology

_____

Juan Leon, PhD, MPH
Committee Chair

Pathogenic contamination of tomatoes, cantaloupe, and jalapenos on farms in Coahuila and Nuevo León, Mexico

By

Alexander Emmitt
B.S., University of Michigan – Ann Arbor, 2010

Faculty Thesis Advisor: Juan Leon, PhD, MPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
Epidemiology

2013

Abstract

Pathogenic contamination of tomatoes, cantaloupe, and jalapenos on farms in Coahuila and Nuevo León, Mexico

By Alexander Emmitt

Produce-associated outbreaks contribute to healthcare costs and societal well-being. Risk for produce-related outbreaks is increasing due to globalization. The goal of this study was to assess the presence and association of indicators on produce with environmental samples on farms by aiming to quantify the extent of indicators on produce and their environmental samples and to discover associations between the environmental exposure variables (soil, irrigation water, and hands) and the outcome variable (produce) using indicators to determine presence of fecal material. Samples were tested for fecal coliforms, *E. coli*, Enterococcus, and Coliphages. Samples were collected from farms in the Coahuila and Nuevo Leon regions of Mexico. SAS 9.3 was utilized to describe concentrations and prevalence of indicators on samples and model linear and logistic regressions. When quantifying the prevalence of indicators in environmental samples, we found that many of the exposure samples of soil, irrigation water, source water, and farm worker hands contained fecal indicators. When quantifying the concentrations of indicators on environmental samples, we found that hands and produce typically had higher mean concentrations of indicators compared to soil, irrigation water, and source water. We discovered that soil and irrigation water that contained fecal indicators before harvest were significantly associated with produce containing fecal indicators before harvest ($\widehat{\beta_{Soil}}$ = -1.887) ($\widehat{\beta_{Water}}$ = 0.862) (controlling for produce type). We also found that hands that contained fecal indicators after harvest, during distribution, and at the packing shed were associated with produce contaminated with fecal indicators ($\widehat{OR_{Col}}$ = 9.67) ($\widehat{OR_{E.coli}}$ = 6.24) (controlling for produce type). An implication of this study was that remediation practices could potentially focus on hand washing behavioral changes at all stages of harvest to improve the cleanliness of the produce.

Pathogenic contamination of tomatoes, cantaloupe, and jalapenos on farms in Coahuila and Nuevo León, Mexico

By

Alexander Emmitt

B.S., University of Michigan – Ann Arbor, 2010

Faculty Thesis Advisor: Juan Leon, PhD, MPH

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
Epidemiology

2013

Acknowledgements:

I would like to personally thank Juan Leon for his guidance, his wisdom, and especially his patience.

I would also like to thank Faith Bartz and Jacquelyn Lickness for taking time to meet with me and answer specific questions in person or quickly through email. Also, thank you to the entire Clean Greens Research Team for collecting and cleaning the data which made this and other studies possible.

Table of Contents

Literature Review

Background:

Foodborne illnesses come from a variety of sources and are a major healthcare burden on the United States. Sources of foodborne illness include seafood, poultry, beef, pork, eggs, produce, dairy, wild game, lunchmeat, breads, and beverages. "Unsafe foods cause an estimated 76 million illnesses and 5,000 deaths each year in the United States" (1). Produce contamination is one of the top reasons for outbreak in the United States. The FDA and USDA regulate food production and processing, but outbreaks and illnesses are continually increasing and lack of funding still only allows inspection of companies dealing with many high-risk foods by the FDA once every five to ten years.(reviewed in (2))

In the United States, seafood consumption is the leading cause of foodborne illness outbreaks (reviewed in (2)). Between 1990 and 2006, seafood accounted for 1,140 outbreaks involving 11,809 cases of illness. Most outbreaks from seafood occurred from the naturally occurring chemical toxins scombrotoxin and ciguatoxin. Major vectors of disease were tuna, grouper, shellfish, crab cakes, tuna burgers, shrimp and lobster. Norovirus and *Vibrio* spp. were the most common causes of outbreak and illness. Harvesting beds for shellfish provided opportunities for bacterial and viral contamination.

Poultry and poultry dishes led to 620 outbreaks with 18,906 illnesses between 1990 and 2006 (reviewed in (2)). Chicken was the largest individual culprit of infection for poultry, contributing 229 outbreaks and 5,301 illnesses. Turkey contributed the second highest contamination among poultry with 103 recognized outbreaks and 5,616

illnesses. Other contaminated types of poultry included duck, game hen, and goose. Poultry dishes added 281 outbreaks and 7,875 illnesses. The most common pathogens associated with poultry included *Salmonella* spp., *Staphylococcus aureus, Clostridium perfringens*, and norovirus.

Beef accounted for a total of 518 outbreaks and 14,191 illnesses between 1990 and 2006 (reviewed in (2)). Ground beef was determined to cause 183 outbreaks and 3,370 illnesses. Other types of beef including roast beef, veal, and beef jerky accounted for 187 outbreaks and 6,543 illnesses. The pathogens most commonly associated with beef contamination included *E. coli* O157:H7, *Clostridium perfringens*, and *Salmonella* spp. Beef is continually recalled due to *E. coli* contamination. Between 1990 and 2006, 50% of *E. coli* O157:H7 contamination in food occurred in beef. For example, between February and August 2008, eight companies recalled nearly 1.5 billion pounds of ground beef.

Similarly, pork accounted for 233 foodborne illness outbreaks and 6,954 illnesses between 1990 and 2006 (reviewed in (2)). Ham was the main offender accounting for 54 outbreaks and 2,205 illnesses. The pathogen most common in pork was *Staphylococcus aureus* but norovirus, *Salmonella* spp., and *Clostridium perfringens* all contributed to outbreaks and illnesses (reviewed in (2)).

Eggs were linked to 351 outbreaks with 11,143 illnesses over the 17-year period of 1990-2006 (2). *Salmonella enteriditis* accounted for 96% of the egg-associated outbreaks. In 1998, the USDA estimated that the yearly approximation for eggs contaminated with *S. enteriditis* is around 2.3 million (reviewed in (2)).

Produce accounted for 768 outbreaks and 35,060 illnesses between 1990 and 2006 (reviewed in (2)). 279 outbreaks and 14,743 illnesses were directly associated with vegetables. Produce ranked second after seafood in total number of outbreaks. 121 outbreaks and 7,802 illnesses were directly associated with fruits. Other dishes including salads accounted for the remainder of the outbreaks and illnesses. Norovirus accounted for 41% of outbreaks in produce. *Salmonella* accounted for 18% of all produce outbreaks of disease and *E. coli* O157:H7 contributed 8% of all produce outbreaks of disease. Twenty percent of all foodborne *E. coli* O157:H7 outbreaks between 1990 and 2006 occurred due to produce contamination. Outbreaks from produce averaged 46 illnesses per outbreak, which is significantly higher compared to outbreaks from other foods such as dairy, meats, and seafood. (reviewed in (2))

Although contamination of seafood, produce, poultry, beef, and eggs are the main causes of foodborne illness and outbreaks, there are other foods that cause foodborne disease in humans. Dairy, wild game, lunchmeat, breads, and beverages round out the remainder of the transmission routes of foodborne illnesses in the United States. (reviewed in (2))

Government spending is an important aspect when viewing disease and various studies and organizations offer a level of quantifiable cost estimates. However, the costs associated with foodborne disease cannot be exactly quantified, as there are many variables involved other than monetary cost for local and national healthcare, including burden on the population. The FoodNet division of the Centers for Disease Control and Prevention offered quantifiable estimates using the Economic Research Service of the U.S. Department of Agriculture. According to the Government Accountability Office

(GAO), in 1999, the federal government programs of FSIS and the FDA spent $1 billion on food safety (reviewed in (3)). State governments contributed an additional $300.9 million.

Cost of illnesses and outbreaks aggregates personal cost with financial cost. In 1999, of the 76 million annual illnesses caused by unsafe foods, there are estimated to be about 325,000 yearly hospitalizations and 5,000 deaths (1). More recent estimates for the total cost of foodborne illness in the United States were as high as $152 billion a year (reviewed in (4)). These measurements included long-term health-related costs, medical costs, and quality of life losses. This sum also included the costs associated with insurance companies. Human illness, due to various enteric pathogens, contributes to medical costs. Costs include estimates of medical costs and productivity losses due to foodborne infections. For example, in 1998, the economic costs of human illness caused by foodborne Salmonella infections were $2.3 billion annually (5).

Pathogens Involved in Foodborne Illness and the Focus of the Clean Greens Study:

Salmonella enterica is responsible for causing as many as 1.3 billion cases of illness annually (6). There are 6 subspecies and over 2,000 serotypes of *S. enterica* (7). The six identified subspecies are enterica, salamae, arizonae, diarizonae, houtenae, and indica. Nearly all salmonella-related disease in humans and domestic animals are caused by the subspecies enterica (7). Most serovars cause diarrhea in humans (6). Human typhoid is a commonly seen occurrence from ingestion of salmonella. Following the ingestion of *S. enterica* serovar Typhi bacteria, human typhoid can occur. This occurrence is conjectured to be typically observed from close contact with the individual

or carrier, contaminated water, or animal products (8). Typhoid disease in humans typically manifests one to two weeks following inoculation by the bacterium (6).

Following the ingestion of greater than 50,000 bacteria in contaminated food or water, disease in humans typically results (6). Symptoms typically occur between 6 and 72 hours after bacterial consumption (6). Without treatment for gut infections, symptoms may occur between five to seven days and resolve spontaneously (6). Salmonella infections are more likely to recur, be severe, or persist in the population when compared to other infections (9). Those who are HIV positive have higher rates of infection compared to those who are HIV negative (9). Salmonella occurs at a high rate in animals which can be transmitted to humans through rodents and manure (10). Through knowledge of Salmonella's impact on the human population, it is important that our study includes Salmonella as it is one of the most prevalent foodborne infectious diseases in the world.

E. coli O157:H7 presents another pathogenic hazard for consumption of pathogens from produce in farms. There is an urgent need to reduce the hazard to human health from this pathogen (reviewed in (11)). *E. coli* O157:H7 is associated with outbreaks of hemorrhagic colitis (12). Typically this gastrointestinal illness is characterized by sudden severe abdominal cramps combined with bloody diarrhea. *E. coli* O157:H7 is also linked to hemolytic uremic syndrome (reviewed in (13). Hemolytic uremic syndrome is a leading cause of childhood acute renal failure onset several days after the beginning of diarrheal illness.

E. coli O157:H7 thrives in many environments on farms, including intestinal tracts in livestock as well as water supplies. Most strains of *E. coli* are benign and some

inhabit the gastrointestinal tract of humans and other animals, but there also exist pathogenic *E. coli* such as strain O157:H7 (14). *Escherichia coli* O157:H7 and O157:H7 the most prevalent EHEC serotype associated with North American foodborne illness (15). Cattle have been implicated in many outbreaks of *E. coli* O157:H7 and the ability of the organism to survive in cattle feed, water, soil, and manure imply persistence within cattle populations as well as water supplies and crops (16). *E. coli* O157:H7 has been shown to survive for up to seven weeks after distribution of manure to the crops (16). *E. coli* O157:H7 has been known to survive in bovine feces for longer than 42 days (17). Also, sheep can be contaminated at time of slaughter with *E. coli* O157:H7 at rates of up to 31% (18). *E. coli* O157:H7 can be transported by rain, wind, removal and spreading of manure, as well as by animals and humans (reviewed in (11)). Numerous foods have been linked to *E. coli* O157:H7 outbreaks (reviewed in (11)). Cross contamination with bovine or ovine feces, among other animals, is suspected in the majority of the outbreaks (19, 20). There is good reason to believe that *E. coli* O157:H7 is present when fecal matter is present. Crops and soil can directly contaminate humans after crops and soil are contaminated by feed, birds, flies, cattle, farm animals and wild animals. (reviewed in (11))

Another environment in which *E. coli* O157:H7 flourishes is water. The range of *E. coli* O157:H7 concentrations are based on conditions of the water and the surrounding sediment (reviewed in (21)). Possible habitats for *E. coli* O157:H7 are manure heaps, ponds, dams, barns, straw, feed, feed troughs, water, water troughs, farm equipment, water-courses, ground and pasture. Standing water with high sediment has the most potential to contain forms of *E. coli* O157:H7. Specifically, water of great risk includes

naturally occurring lake water and water with the potential to come in contact with human feces. (reviewed in (11))

Norovirus infection weighs heavily on the individuals who contract the virus and the U.S. healthcare system. Common effects that norovirus can have on one's body are pain in the stomach, nausea, and diarrhea. Children and older adults can experience much more serious norovirus illness. Norovirus is the most common determinant of acute gastroenteritis in the U.S. contributing around 21 million illnesses, approximately 70,000 hospitalizations and 800 deaths. It is also the most common cause of foodborne-disease outbreaks in the U.S. (reviewed in (22)) Anyone can be infected by norovirus. There are also many types of norovirus, so there is a possibility that infection from norovirus leads to immunity afterward.

Norovirus is a pathogenic virus that can easily contaminate food from various environments due to its small size and high infectivity. Environmental contamination of produce by the norovirus pathogen is similar to the contamination by the other pathogens. Norovirus can be contracted from an infected person, contaminated food or water, or by touching contaminated surfaces. Norovirus can be found in a person's feces even before they experience sickness (reviewed in (22)). Even after recovery, the virus can remain in the stool for two weeks or more. Typically norovirus is transmitted from a person accidentally ingesting an infected person's stool or vomit. This can occur by eating contaminated food or drinking contaminated liquids, touching contaminated surfaces or object then putting fingers in one's mouth, or experiencing other forms of contact with someone who is infected with the norovirus. Roughly 50% of all outbreaks of foodborne illness are caused by norovirus. Often sick food handlers were involved in spreading the

virus. Contamination can occur through the hands of people handling the food who have vomit or stool on their hands when they touch the food. Also, fruit and vegetables can be contaminated in the field. (reviewed in (22))

Listeriosis outbreaks occur worldwide and are a common food-borne disease (reviewed in (23)). The bacterium involved is *Listeria monocytogenes* and it primarily affects the elderly, pregnant women, newborns, and those with a weakened immune system. Annually, about 1,600 cases occur in the United States. Pregnant women are around 13 times more likely than the general public to get listeriosis. Pregnant women may experience a premature birth or a miscarriage, and death is a possibility as well (reviewed in (24)).

The environment where *Listeria monocytogenes* exists is similar to the other pathogens mentioned and primarily exists in soil and water (reviewed in (23)). It is accepted that the consumption of contaminated food is the principal route of infection in humans (reviewed in (24)). *Listeria* bacteria can live in facilities, such as a food processing factory, for years (reviewed in (23)). Raw and uncooked foods are often contaminated because they might come in contact with the soil, water, or unwashed hands of workers that have come in contact with already-contaminated water, soil or other produce. *Listeria* is usually transmitted by herd animals and is often found in the feed and soil of grazing herd animals on a farm (reviewed in (24)).

Bacterial and Fecal Indicators:

The high cost, difficulty, and rare prevalence associated with detecting pathogens were all reasons to utilize indicators. Ideal traits of indicators include easy and rapid detection and enumeration, readily distinguishable from commensal microflora,

consistant association with the pathogen indicated, similar growth rates and immobilization rates compared to the pathogen indicated. The optimal indicator microorganism should have a well-known and stable taxonomy, exist at the same time as the target pathogen, it should be at least as abundant as the target pathogen, it should survive slightly better than the target pathogen in a range of habitats, and it should be easily quantifiable using sensitive and specific methods considering economic feasibility (reviewed in (25)). Indicators serve as more of a marker that fecal contamination has occurred, and represent a potential for pathogen presence (reviewed in (26)).

Common indicators include somatic coliphages, fecal coliforms, *Enterococcus* and generic *E. coli*. Coliphages are DNA viruses that infect host cells through the outer cell membrane and may be an indicator for pathogenic viruses (reviewed in (27)). Coliphages are at least as abundant, and are easy to enumerate by methods specifically and sensitively (28)(reviewed in (26)). Coliphages are used because fecal indicator bacteria do not adequately reflect the presence or absence of human viruses (29). The Payment and Franco study in 1993 analyzed the suitability of indicators for the contamination of water, and somatic coliphages were the only explanatory variable for the human enteric virus counts in settled water. Within coliphages, somatic coliphages are useful to sensitively indicate presence or absence of human enteric viruses (30, 31). Fecal coliforms are aerobic and facultatively anaerobic, gram negative rods that ferment lactose and serve as an indicator of the extent of fecal contamination in the environment (32, 33). For this study, fecal coliforms are preferred over total coliforms because total coliforms are generally in high numbers in the environment and do not necessarily indicate presence of fecal material (reviewed in (26)). Fecal coliforms are associated

with human and warm-blooded animals' intestinal tracts (26, 33). The test for fecal coliforms can recover strains of *E. coli, Enterobacter species,* and *Klebsiella pneumoniae* among others (34). Identifying fecal coliforms in the laboratory is not without error so near-perfect conditions must be met to ensure correct identification. Necessity of specific conditions is not limited to the laboratory. Excessive chlorine can be added to the environment, changing the true fecal number. Other viruses may be present which are not accounted for by the fecal coliforms (35).

Enterococcus and generic *E. coli* are indicator organisms because they are also almost exclusively found in the intestinal tracts of humans and animals (26). According to the U.S. Environmental Protection Agency, generic *E. coli* and *Enterococcus* are the two organisms that should be used to indicate fecal pathogens in freshwater and salt-water (36). They are not perfect fecal indicators because they have different rates of reproduction than the organisms they indicate (*E. coli* O157:H7, Salmonella) and may persist longer than pathogens. *E. coli* is also not perfect because environmental reservoirs of *E. coli* are known to exist, indicating that *E. coli* identified in this study may not be due to human contamination (37). Although presence of generic *E. coli* does not necessarily correlate with a greater probability of enteric pathogens being present, the presence or absence of *E. coli* will assist in assessing the overall quality of food and hygienic conditions during the food processing stages (33).

Enterococci are considered to be all streptococci of fecal origin that produce group D antigen (33). Enterococci are located in the intestinal tracts of warm-blooded animals, cold-blooded animals, and even some insects. *E. faecalis* and *E. faecium* are the two most common enterococci encountered. Enterococci presence and absence is

important to analyze because the different levels and counts vary with product, holding conditions, time of storage, and other factors. If these variables are held constant, the predictability of Enterococcus as a quantifiable indicator improves.

Transmission Routes:

Transmission routes hypothesized to contribute to produce contamination are soil contamination, water contamination (both irrigation and source water), and human hand contamination. Most produce contamination occurs on the surface of the fruit or vegetable, but there is evidence that pathogens may enter depressions or crevices by capillary action (38).

The soil is a potential source of contamination of crops. Before even planting the crop, growing location in soil can influence the safety of the produce (39). Produce planted in fields with a history of wild animals or livestock contaminating the soil are more likely to be contaminated with pathogens (40). Wildlife can contaminate the soil with their fecal droppings (reviewed in (26, 41)). Healthy cattle can host *E. coli* O157:H7, which can be found in their feces (42). Some pathogens can survive for months in soil fertilized with manure (reviewed in (39)). Manure used as fertilizer may also be particularly hazardous, as it is abundant with bacteria and viruses (reviewed in (26)). Proper composting of manure can improve chances of non-contamination because proper composting would mean heating the compost to a high enough temperature to kill the pathogens. A history of flooding and proximity of farms to water should also be considered in conjunction with use of manure in soil because standing water is a potential breeding ground for pathogens. Knowing the composition of soil is an important step in assisting the determination of the source of contamination in the soil.

Various factors directly influence the physical, biological, and chemical composition of soil (43, 44). One factor is the microenvironment of soil that influences the composition of soil. Other factors that physically, biologically, and chemically affect movement and number of microorganisms in soil include soil texture (42), particle size and distribution, clay type and content, organic matter type and content, pH, pore size distribution, sunlight, temperature, presence of antibiotics or toxic substances, soil water content, soil water flux, and time of year (45). Soil does not need to remain present on the produce to maintain pathogen survival. Once soil particles are washed away, enteric pathogens are *E. coli* and Salmonella can survive, but are considered to be in survival mode instead of actively reproducing (44). If produce was harvested soon after irrigation, contaminated irrigation water is more likely to remain on the produce, thus contaminating it (reviewed in (26)).

Different methods of irrigation water usage influence produce contamination to different degrees (46). For example, there is a high probability of produce contamination associated with overhead irrigation with unclean water (reviewed in (39)). Using wastewater to irrigate crops also has great potential to contaminate produce with pathogens. The processes of irrigation are also very important where wetting of the entire plant may cause contamination (reviewed in (26)). Splashing debris can also be different route of transfer to the fruit.

Although some irrigation methods may increase produce's risk for contamination, some irrigation methods may prevent contamination. Utilizing drip irrigation or hydroponics may prevent contamination. In a study of drip irrigation for tomato and cucumber plants, contamination occurs highest at the stem, followed by the roots, leaves,

and fruit (47). The fruit had a lower level of contamination compared to all other parts of the plant. However, further harvesting methods can expose the fruit to a potential transfer of viruses from the leaves, roots, or stem. Subsurface drip irrigation is vital to lowering the contamination on the plant by the water. Because wetting or splashing debris on the plant could cause produce contamination, hydroponic systems are ideal to minimize contamination but are often too expensive (reviewed in (26)).

Handling of the produce is a major potential source of contamination. Unfortunately, there are no studies examining the direct transfer of pathogens from hands onto produce. Instead there is evidence that improper personal hygiene of food service workers, especially regarding hand washing, contributes considerably to the risk of foodborne diseases (48) (49). For example, most human norovirus infection outbreaks are attributed to contamination of food by improperly washed hands by food handlers (50). Norovirus was a useful organism to focus on contamination from the Liu, Chien et al 2009 study, but Salmonella, *E. coli* O157:H7 and Listeria also contaminate produce after improper hand hygiene (51). These other pathogens were not included in this Liu, Chien et al 2009 example but were still of interest in the Clean Greens Study. In the United States between 1991 and 2000, greater than 56% of all foodborne norovirus outbreaks were contaminated by foods that were handled at one point without subsequent heating (52). Most reported norovirus outbreaks have been due to food contamination from hands of contagious workers close to the point of service. Proposed mechanisms for viral outbreaks included handling of produce by water or infected individuals at the farm or post-harvesting level (reviewed in (53)). Similarly, bacteria can also spread on hands (54). Although hands are considered to be a major source of contamination, there

is evidence directly proving hands can contaminate produce at the farm level. Although consumer and retail handling can contribute to contamination of produce, neither of these routes will be included in this study.

Goals and Aims:

The goal of this project is to assess the presence and association between indicators on produce and environmental samples, such as soil, water, and farmworker hands in Northern Mexico in the summer of 2010.

The two aims of this project attempt to perform exploratory analyses and analytical analyses on the Clean Greens project data. One aim of this project is to quantify the prevalence and concentrations of indicators on the environmental samples (hands, source water, soil, and irrigation water) as well as produce (tomatoes, cantaloupe, and jalapeno peppers) from farms in Coahuila and Nuevo Leon, Mexico. Another aim of this project is to establish if there are relationships between indicators on produce (tomatoes, cantaloupe, and jalapeno peppers) and indicators from the environment (hands, irrigation water, soil) using logistic and linear regression models based on prevalence and concentrations of those bacteria and viruses from farms in Coahuila and Nuevo Leon, Mexico.

Significance:

Establishing long-term solutions to these problems with uncleanliness of produce on farms is the most important aspect of the Clean Greens III study. The data from this analysis will assist in this goal of establishing solutions by providing information useful at the government level and farm level. At the government level, regulations can be

implemented encouraging full compliance of hygienic behavior by workers at the farm before harvest, at harvest, during distribution, and at the packing shed. Farms seeking to improve the cleanliness of their produce could benefit from this study by using the information of what environments could be improved upon. For example, if the soil measurements turn up extremely contaminated and not the hands of the workers, farms would know to clean the produce before distribution after the harvest stage. The different stages of growing and harvesting produce offer different opportunities to reduce risk of contamination. For example, further analysis can determine which stages present the highest risk for contamination and can be focused upon through remediation techniques. This study is not solely for farms in Mexico and the United States of America. It is also for the rest of the world. Improved, cost-effective sanitation methods will be useful to remediate uncleanliness at the farm level and this research could be used to come up with solutions.

Materials & Methods

Study farms and subjects:

Farms were enrolled by Universidad Autonoma de Nuevo Leon (UANL) in Mexico in the administrative divisions of Nuevo Leon and Coahuila. The IRB number for this study was IRB00035460. La Universidad Autonoma de Nuevo Leon sampled from the 16 fields of these selected nine farms, and was responsible for testing for indicators (generic fecal coliforms, *Escherichia coli*, coliphages, and Enterococcus). The subjects involved in the study were Mexican farm workers who contacted the produce manually.

Sample collection:

Samples were collected from farms spread across the Mexico-U.S. regions of Coahuila and Nuevo Leon. Three universities were involved with this study. Universidad Autonoma de Nuevo Leon (UANL) recruited these farms, sampled from the field, and tested for indicators. North Carolina State University was responsible for generic *E. coli* strain-typing. Emory University was responsible for general project management, design of interviews and surveys, data entry, and statistical analysis.

The produce involved in this study was cantaloupe, jalapeno peppers, and tomatoes. The environmental samples included in this study were farmworker hands, soil, source water, and irrigation water. There were different stages in time when samples were taken. The first stage was before harvest in which produce items, source water, irrigation water, and soil were sampled. The next stages were during harvest, during the distribution stage, and at the packing shed where the harvest was temporarily stored. Produce items and hands of the farm workers were sampled at these stages. Three randomly-selected locations were sampled and composited for each point in each chain. Not all farms had every stage of the chain. The same field could have been

sampled twice, but it was considered to be a different sample chain if it was not sampled on the same day.

To collect the indicator information on hands, groups of 6 hands were rinsed with a 2250 milliliter 0.1% peptone water. These hand-wash data were represented in the units: cfu/hand (colony forming units/hand). Groups of 54 tomatoes were rinsed with a 1500 milliliter 0.1% peptone water. These tomato-solution data were represented by the units: cfu/tomatoes (colony forming units/tomatoes). Groups of 42 jalapenos were rinsed with a 1500 milliliter 0.1% peptone water. These jalapeno-solution data were represented by the units: cfu/jalapenos (colony forming units/jalapenos). 6 cantaloupes were rinsed with a 1500 milliliter 0.1% peptone water. These melon-solution data were represented by the units: cfu/melons (colony forming units/melons). In conjunction with produce, indicators in irrigation and source water were measured from a 100 milliliters water sample. The units for water were: cfu/100 ml water (colony forming units per 100 milliliters of water). Indicators in soil were measured by taking 25 grams and rinsing with either 75 or 225 milliliters of 0.1% peptone water. Microbial analyses began within 24 hours of sample collection. Membrane filtration was utilized to concentrate the indicators. Fecal coliforms and generic *E. coli* were grown on RAPID'E.coli 2 (Bio-Rad Laboratories, Inc., Hercules, CA). Enterococcus was grown on KF Streptococcus agar (Oxoid Limited, Basingstoke, Hampshire, UK). Coliphages were detected using FastPhage MPN Quanti-tray (Charm Sciences, Inc., Laurence, MA). *E. coli* colonies were differentiated from other fecal coliforms by the color of the colonies.

Experimental procedures:

Fecal coliform, Enterococcus, *E. coli*, and coliphage samples were analyzed either through direct plating or membrane filtration. The number of samples varied by farm ID

(A-I), chain time (before harvest, after harvest, during distribution, at the packing shed), produce type (tomatoes, jalapenos, melons/cantaloupe) and sample type (soil, irrigation water, source water, hands).

Once plated, the absence or presence of indicators was noted, as well as the number of colonies if countable. The indicators were grouped into 7 different categories regarding the average of their plated countable colony forming units adjusted by the dilution factor. The first group was that all plate counts were zero. The second group was that any plate count was less than 25 but not zero colony forming units. The third group indicates that any of the plate counts fall within 25 and 250 colony forming units. The fourth group indicates that all plate counts were greater than 250 but still countable. The fifth group indicates that the plate counts were all outside the countable range and can include counts less than 25 and others greater than 250. The sixth group was that all plates colony forming units were too numerous to count. The seventh group was that all plates contain either 0 or too numerous to count colony forming units. Only three samples were included in this seventh group and they all averaged to be too numerous to count. For linear regression analysis, this study focused only on the groups with countable plates under 25 (group 2), those between the range of 25 and 250 cfu (group 3), and those with cfu greater than 250 but not all too numerous to count (group 4). For logistic regression analysis, this study included all groups.

Data Analysis and Statistical Methods:

All statistical analyses in this cross-sectional study were performed with Statistical Analysis Systems software 9.3 (SAS Institute, Cary, North Carolina). The significance level for all odds ratios were $P < 0.05$. The significance level for the logistic and linear

regression figures were $P < 0.10$. Descriptive statistics included means, standard deviations and prevalence. The data for linear regression and means and standard deviation were transformed because the distributions of the data were not normal across groups 2, 3 and 4. Common logarithms (log_{10}) were used in normalizing this data. Because there was an observed correlation between source water and irrigation water, to prevent multicollinearity in the linear and logistic models, irrigation water was used when both source water and irrigation water were available. Irrigation water was the preferred measurement because irrigation water draws from the source water and comes into direct contact with the produce. If only source water measurements were available, source water was used as the water measurement. This only occurred twice, however.

The logistic models were also stratified based on chain time and sample type. The criteria for determining if fecal coliform, *E. coli,* or Enterococcus indicators were absent in the sample were if the average plate count type was defined as zero. Otherwise, the indicators were classified as present. Coliphages were present if the most probable number was between 1 and 2,419.6 or over.

Kleinbaum (55)'s methods for linear and logistic regression were utilized to determine the best models and most significant confounders, considering produce type, chain time, and Farm ID.

Linear Regression:

The variables in the maximum model included all basic predictor and control variables with other justifiable interaction terms (combinations of water, soil, and hands with potential confounders produce type, chain time, and Farm ID). Variable selection was done using the backward selection strategy. Potential confounders were selected based

on whether there was a significant difference of 10% of the beta coefficient of the main effect variable(s) in the model.

After the final linear model was selected, regression diagnostics were performed. These diagnostics tested for the linearity, normality, independence, and homoscedasticity assumptions. The data were checked for outliers using the Cook's Distance residual plot, and other partial plots to examine homoscedasticity and linearity. Linearity was assessed and corrected before model selection because violations of linearity were serious and these data needed to be logarithmically transformed. Independence was assessed by observing residual plots and multicollinearity. Normality was assessed by observing the distribution of the normal probability plot in the SAS (9.3) output. Variance Inflation Factors greater than 10 indicated a problem with multicollinearity.

Logistic regression:

The modeling strategy included variable specification, interaction assessment, and confounding assessment followed by precision considerations. The models before harvest included the soil and water as main exposure variables. The models after harvest included the hand main as the exposure variable. For both models, potential confounders were produce type and farm ID, along with potential interaction terms between all confounders and exposure variables. Subsequently, a test for collinearity using SAS macro was performed. If the Condition Index was greater than 30, there was likely a collinearity problem and the variable with a VDP of greater than 0.5 was likely the source of collinearity and was deleted. Interaction presence or absence was determined by using the likelihood ratio test between models with the interaction terms and those without. If interaction terms were selected, to remain hierarchically-well-formulated, all lower-order exposure and confounding variables in the interaction terms must have been included in

the model. After this step, a gold standard model was produced and confounding was assessed assessing the precision by eliminating various combinations of potential confounders. When all subsets of the gold standard model were accumulated, the ones with the odds ratios of the main exposure variables that fell within 10% of the gold standard model were assessed for increases in precision. This precision was assessed using the 95% confidence interval of the odds ratio using CI width and CI ratios.

To run diagnostics on logistic regression models, goodness of fit testing and receiver operating characteristic (ROC) curves were used. The goodness of fit test used for this data was the Hosmer-Lemeshow Statistic. If the Hosmer-Lemeshow Statistic had a p-value of less than 0.05, this represented a poorly fitted model. The receiver operating characteristic curves measured sensitivity against 1-specificity indicating how well the fitted model discriminated between cases and non-cases. The *a priori* rules were that when ROC=0.5, there was no discrimination; when ROC was between 0.7 and 0.8, this was an acceptable discrimination; when ROC was between 0.8 and 0.9, this was excellent discrimination; and when ROC was greater than 0.9, this was outstanding discrimination.

Results

The indicators fecal coliforms, *E. coli*, Enterococcus, and Coliphages were isolated from produce, soil, irrigation water, source water, and hand samples for this study. 409 total sets of samples were analyzed, isolating these indicators for analysis. (Tables 1-3) provide a description of this collected information.

Descriptive Results:

Before linear and logistic models were used to see which measurements were best associated with indicators on produce, we must first describe the data. To determine indicator concentrations and prevalence of different farming processes, we measured several types of produce and several potential sources of environmental exposure. Stratifying on produce type as the main outcome variable, the exposure variables associated with the process of tomato harvest were described in (Table 1). We found that hands, source water, irrigation water, and soil associated with tomatoes were all likely to be contaminated by at least one indicator (*E. coli*, fecal coliforms, Enterococcus, or Coliphages). When observing Enterococcus specifically, hands were contaminated in 100% of the samples with a concentration log_{10} mean value of 3.53. Source water and irrigation water were both contaminated in 93% of the enterococcus samples with concentration log_{10} mean values of -1.35 and -1.16 respectively. Tomatoes were contaminated in 67% of the enterococcus samples with a concentration log_{10} mean value of 1.96. Soil was contaminated in 35% of the enterococcus samples with a concentration log_{10} mean value of 1.16. *E. coli* was observed at highest prevalence in source water and irrigation water, at lower prevalence on hands and tomatoes, and not present in soil. The highest concentration of *E. coli* was on hands (log_{10} = -0.10), then produce (log_{10} = -

0.48). Irrigation water and source water had similar concentrations of *E. coli* with concentrations of log_{10} = -1.00 and -1.13, respectively. Soil did not have any measurements of *E. coli* concentrations due to the study parameters. Fecal coliforms did not show a trend as most samples were contaminated to some extent between 88% and 100%. Fecal coliform concentration measurements were both high and similar on tomatoes and hands, still high in soil, but much lower in source and irrigation water. In the tomato-associated samples, most Coliphage measurements had medium prevalence between 56% and 86%. Coliphage concentration measurements were low for hands, soil, and tomatoes and a little higher for source and irrigation water ranging from log_{10}= -0.70 to -0.24 for all exposures. In conclusion, the investigation demonstrated a quantitative description of tomato farming data.

Before linear and logistic models were used to see which measurements were best associated with indicators on produce, we must first describe the data. To determine indicator concentrations and prevalence of different farming processes, we measured several types of produce and several potential sources of environmental exposure. Stratifying on produce type as the main outcome variable, the exposure variables associated with the process of cantaloupe (melon) harvest were described in (Table 2). We found that hands, source water, irrigation water, and soil associated with cantaloupes were all likely to be contaminated by at least one indicator (*E. coli,* fecal coliforms, Enterococcus, or Coliphages). When observing *E. coli* specifically, hands were contaminated in 68% of the samples with a concentration log_{10} mean value of 0.91. Source water was contaminated in 36% of the *E. coli* samples with a concentration log_{10} value of -1.54. Irrigation water was contaminated in 38% of the *E. coli* samples

with a concentration log_{10} mean value of -0.82. Cantaloupe samples were contaminated in 65% of the *E. coli* samples with a concentration log_{10} mean value of 0.47. Soil was contaminated in 50% of the *E. coli* samples with a concentration log_{10} mean value of -1.07. Enterococcus was observed at highest prevalence on hands and produces at 100% prevalence, while at a lower prevalence in irrigation water, source water, and soil. Enterococcus concentrations were highest observed on melons (2.99) and hands (2.47). Soil had a medium concentration log_{10} = 0.55, but irrigation water and source water had low concentrations both around log_{10} = -1.00. Fecal coliforms did not show a trend as all samples were contaminated at 100% frequency, only differing in log_{10} concentration, highest on produce (2.87), hands (1.66), and soil (1.66). In the melon-associated samples, most Coliphage measurements had too small of a sample size to draw conclusions from prevalence or concentration information. In conclusion, the investigation demonstrated a quantitative description of cantaloupe farming data.

Before linear and logistic models were used to see which measurements were best associated with indicators on produce, we must first describe the data. To determine indicator concentrations and prevalence of different farming processes, we measured several types of produce and several potential sources of environmental exposure. Stratifying on produce type as the main outcome variable, the exposure variables associated with the process of jalapeno harvest were described in (Table 3). We found that hands, source water, irrigation water, and soil associated with jalapenos were all likely to be contaminated by at least one indicator (*E. coli*, fecal coliforms, Enterococcus, or Coliphages). When observing Enterococcus specifically, hands were contaminated in

100% of the samples with a concentration log_{10} mean value of 3.12. Source water and irrigation water were contaminated in 92% and 86% of the enterococcus samples with concentration log_{10} mean values of -0.97 and -1.48 respectively. Jalapenos were contaminated in 56% of the enterococcus samples with a concentration log_{10} mean value of 2.01. Soil was contaminated in 29% of the enterococcus samples with a concentration log_{10} mean value of 0.92. *E. coli* was observed at highest prevalence on hands (55%), at lower prevalence in irrigation water (25%), source water (36%), and on produce (23%), but not present in soil. *E. coli* concentrations were highest on hands, log_{10} = 0.52, but low on produce (-0.67), irrigation water (-0.80), and source water (-0.91). Fecal coliforms did not show a trend in prevalence as most samples were contaminated to some extent between 84% and 100%. Fecal coliforms were of high concentrations in soil, on hands, and on produce ranging from log_{10} = 1.38 to 2.57 and low in irrigation water (-0.87) and source water (-1.05). In the jalapeno-associated samples, most Coliphage measurements had medium to high prevalence on produce, hands, and in soil between 59% and 78%, but low prevalence in irrigation water and source water (20%-25%). Coliphage mean concentrations are all within a similar range for soil, hands, and produce between log_{10} = -0.87 and -0.73 with insufficient data for irrigation water and source water. In conclusion, the investigation demonstrated a quantitative description of jalapeno farming data.

Some discernible differences in this data were that the fecal-coliform measurements were highly prevalent. This was observed in tomatoes (Table 1) with prevalence between 88% and 100%. This was also observed in melons (Table 2) with

prevalence of 100% on produce, hand-rinse, source water, irrigation water, and soil. This was also observed in jalapenos (Table 3) with prevalence between 84% and 100% for all environmental samples. Regardless of the produce type, the fecal coliform measurements tend to be at least as prevalent as the other indicators.

There were several means with negative log_{10} values. This indicated that quite a few of the mean values for colony forming units were below 1 on average for their particular strata.

Regression Results:

Before models were run, a simple t-test was done to see if the frequency of source water had a similar distribution to the frequency of irrigation water. The t-statistic was 0.16 with a p-value of 0.87, indicating that these variables were not close to being statistically different and could be used interchangeably. (Tables 1-3) still included source water as a measurement to be thorough. Also, to keep analysis constant, produce type was maintained as a control variable throughout the models at all levels including all linear and logistic models.

The indicators fecal coliforms, *E. coli*, Enterococcus, and Coliphages were separately analyzed before harvest and after harvest. Before harvest (Tables 4, 6), the main exposure variables were log_{10} indicators on soil or water (Table 4), or presence and absence of indicators on soil or water (Table 6). After harvest, during distribution, and at the packing shed were the stages grouped together to provide information to show association of potential contamination after harvest (Table 5, 7). After harvest, the main

exposure variables were log_{10} indicators on hands (Table 5), or presence and absence of indicators on hands (Table 7).

Before the harvest of produce, it was important to determine the extent to which of the main exposure variables influenced the outcome variable. To determine the possible effect on the outcome variable of indicators on $produce$, we measured several variables including the main exposure variables of indicators in $soil$ and indicators in $water$ and included potential confounding variables of Farm ID and produce type, but only controlled for produce type because most models recognized it as a significant predictor (Table 4). We found that models measuring Enterococcus and *E. coli* were significant for at least one of the main exposure variables when controlling for produce type, but *E. coli* was not normally distributed. The main effect variables of $soil$ and $water$ were both significant for Enterococcus, controlling for produce type. One main effect variable for $soil$ had a positive effect on the outcome variable: log_{10} Enterococcus indicators on produce. As Enterococcus indicators in soil increased ($soil$ = -1.89), there was an observed increase in Enterococcus indicators on produce. As Enterococcus indicators in water increased ($water$ = 0.86), there was an observed increase in Enterococcus indicators on produce. Enterococcus information provided the best model for this data, statistically with the significant main predictors and an R-square value of .9552. In conclusion, this study demonstrated that the exposure variables of log_{10} indicators in soil and water could have correlative effects on determining the number of log_{10} indicators on produce.

The *E. coli* model was not valid with a soil value because there were only 3 matching *E. coli* measurements between soil and produce. The Coliphage model was

also invalid because it only had an MPN value for 7 water samples and 16 soil samples and fewer MPN values after controlling for produce type.

After the harvest of produce, it was important to determine the extent to which the main exposure variables influenced the outcome variable. To determine the possible effect on the outcome variable of log_{10} of indicators on produce, we measured several variables including the main exposure variable of log_{10} of indicators on hands and potential confounding variables of Farm ID, chain time, and produce type, but only controlled for produce type due to cumulative significance of models (Table 5). We found that model measuring Enterococcus was significant for the main exposure variable of log_{10} indicators on hands when controlling for produce type. The parameter estimates for hands in these models were all positive, controlling for produce type. For example, as indicators on $hands$ increased, so did the indicators on $produce$ (Hands = 0.375). . In conclusion, this study demonstrated that the exposure variables of log_{10} indicators on hands could have positive correlative effects on determining the number of log_{10} indicators on produce.

Before the harvest of produce, it was important to determine the extent to which of the main exposure variables influenced the outcome variable. To determine the possible effect on the outcome variable of presence of fecal indicators on produce, we measured several variables including the main exposure variables of presence of indicators in soil and presence of indicators in water and potential confounding variables of Farm ID and produce type, but only controlled for produce type due to cumulative

significance (Table 6). We found that none of these models were significant for either of the main exposure variables when controlling for produce type. In conclusion, the investigators demonstrated that the exposure variables of indicators in soil and water likely do not have correlative effects on determining presence of indicators on produce.

The fecal coliforms model did not result in a valid model because there were 47 positive produce samples for the fecal coliform indicator and only 3 samples that tested negative. All water samples tested positive for fecal coliforms and only two out of 53 soil samples tested negative.

After the harvest of produce, it was important to determine the extent to which the main exposure variables influenced the outcome variable. To determine the possible effect on the outcome variable of presence of fecal indicators on produce, we measured several variables including the main exposure variable of presence of indicators on hands and potential confounding variables of Farm ID, chain time, and produce type, but only controlled for produce type due to cumulative significance (Table 7). We found that two of these models were significant for the main exposure variable when controlling for produce type. Both *E. coli* and Coliphage turned up a significant odds ratio. Enterococcus on hand samples were all positive for every value of Enterococcus measured as positive on produce samples, so there was no differentiation and no model as a result. Hands that contain *E. coli* indicators were estimated to be approximately 6.2 times more likely to be associated with produce that contains *E. coli* than hands without *E. coli* indicators. In addition to *E. coli* on hands, hands with Coliphages were estimated to be approximately 9.7 times more likely to be associated with produce that contained Coliphages than hands without Coliphage presence. In conclusion, the investigators

demonstrated that the main exposure variable of indicators on hands likely does have correlative effects on determining presence of indicators on produce.

Discussion

The goal of this study was to assess the presence and association between indicators on produce and environmental samples, such as soil, water, and farmworker hands on farms in Northern Mexico in the summer of 2010. An important finding of this study was the high frequency and concentrations of indicators on the hands of farm workers. Another result of this study was that exposure to contaminated soil, water, and hands could have resulted in contamination of produce, controlling for produce type.

Description Analyses:

Farm workers had high frequencies and concentrations of indicators on their hands, potentially contributing to the risk of foodborne diseases. As a result of these high frequencies, hands were a major source of potential contamination (Tables 1, 2, 3). There could be many reasons as to why farm worker hands contained such a high frequency of indicators in this study. Improper hygiene of food service workers, especially regarding hand washing, contributes considerably to the risk of foodborne diseases (48, 49). Available hand-washing facilities could potentially prevent this uncleanliness. However, only the packing shed contained hand-washing facilities, while workers did not have access to any form of hand-washing before harvest, after harvest, or during distribution. Most farm workers did not have access to washing their hands after defecating, which is not uncommon for farmers (reviewed in (56)). This would likely be the reason for unsanitary hands. A speculative reason for unsanitary hands could have been the hands of farm workers in the packing shed being cross-contaminated by the already contaminated produce from earlier stages.

Regression Analyses:

Before harvest, contaminated soil was associated with contaminated produce. Our data indicated that the number of enterococci in soil was a positive association ($\widehat{\beta_1}$= -1.89) for the outcome number of enterococci on produce (Table 4). This positive association could have resulted from the physical, biological, or chemical composition of soil potentially contaminating the produce by coming into contact with it (43, 44). Some soil compositions make it optimal for pathogenic growth depending on particle size, clay type, organic matter type, pH, sunlight, temperature, water content, water flux, and time of year (45). This association could have also occurred due to humans contaminating the soil via contamination by defecation in the fields and not sanitizing hands or animal contamination by manure spread onto the produce (40)(reviewed in (26, 39)). The use of manure and direct fecal contamination of workers in the field due to defecation then not sanitizing their hands were likely the mechanisms to produce contamination.

These data also indicated that there was also a positive association between enterococci in produce and irrigation water ($\widehat{\beta_2}$ = 0.86) (Table 4). This positive association could have resulted from irrigation methods and farming practices of perforated hose drip irrigation from an unclean water source or an unclean hose (46)(reviewed in (39)). The fruit is typically the least contaminated part of the plant due to irrigation, so this water could potentially be highly contaminated (47). Using wastewater or unclean water as the irrigation water on the farm could be the reason for contamination (reviewed in (39)). Quick harvest after wetting the plant or wetting the plant in specific areas can lead to direct transmission, as well (reviewed in (26)).

After harvest, and during the distribution and packing shed stages, there was a significant association between contamination of hands and contamination of produce.

The presence of fecal indicators on hands was significantly associated with the presence of *E. coli* and Coliphage indicators on produce (Table 7). A hypothesis as to why these values were significant for hands was that hands were highly contaminated on the farms, and hands came in direct contact with produce on the farms at multiple stages in the process (48-50). Most stations did not have options to hand wash, which could have led to contamination of the produce (reviewed in (56)). Only the packing shed had any type of hand washing option. However, the produce were either contaminated by then or the workers did not utilize this hand washing option. There also could have been cross-contamination from produce to hands and back to other produce.

Strengths and Limitations:

This study was not without its strengths and weaknesses. One strength of this study was that the samples were retrieved from 9 farms and 3 different produce types, reducing chance of selection bias. Another strength of this study was that the cross-sectional nature of the study could be reproduced in other locations with similar predictors over a reasonable amount of time. A weakness of this study could have been unmeasured confounders, which could have led to different associations. Another weakness of this study was that the *E. coli* was not normally distributed, so the results gleaned from its models were biased. A way to improve upon this study would be to have more samples, increasing the power and decreasing the chance of bias.

Conclusion and Implications:

In conclusion, contamination of soil, water, and hands could have led to contamination of produce, controlling for produce type. Fecal indicators were useful to determine potential problems with these processes within harvesting practices in farms. The information from this study could potentially be used to summarize where most of

the problems in harvesting produce exist. Soil and water were significant predictors of contamination in produce before harvest. Hands were significant predictors of contamination on produce after harvest. One implication of this study was that hands need to be sanitized before handling of produce if one wanted to lower the chance of exposure to organisms indicated on produce. Another implication is that while soil, irrigation water, and hands might be contaminated, remediation techniques could be further analyzed for each level.

Reference List

1. Mead PS, Slutsker L, Dietz V, et al. Food-related illness and death in the United States. *Emerging infectious diseases* 1999;5(5):607-25.
2. DeWaal CS, Tian XA, Bhuiya F. Outbreak Alert! 2008. *Center for Science in the Public Interest* 2008(10):1-22.
3. GAO US. Food Safety: Overview of Federal and State Expenditures. GAO, 2001.
4. Scharff RL. Health Related Costs from Foodborne Illness in the United States. *Produce Safety Project*: Georgetown University, 2010.
5. Frenzen PD, Buzby JC, Roberts T. An Updated Estimate of the Economic Costs of Human Illness Due to Foodborne *Salmonella* in the United States. Presented at 3rd International Symposium on the Epidemiology and Control of Salmonella in Pork, Washington D.C., August 5-7 1999.
6. Coburn B, Grassl GA, Finlay BB. Salmonella, the host and disease: a brief review. *Immunology and cell biology* 2007;85(2):112-8.
7. Fierer J, Guiney DG. Diverse virulence traits underlying different clinical outcomes of Salmonella infection. *The Journal of clinical investigation* 2001;107(7):775-80.
8. Hornick RB, Greisman SE, Woodward TE, et al. Typhoid fever: pathogenesis and immunologic control. *The New England journal of medicine* 1970;283(13):686-91.
9. Altekruse S, Hyman F, Klontz K, et al. Foodborne bacterial infections in individuals with the human immunodeficiency virus. *Southern medical journal* 1994;87(2):169-73.

10. Altekruse SF, Cohen ML, Swerdlow DL. Emerging foodborne diseases. *Emerging infectious diseases* 1997;3(3):285-93.

11. Bach SJ, McAllister TA, Veira DM, et al. Transmission and control of Escherichia coli O157:H7 - A review. *Canadian Journal of Animal Science* 2002:475-90.

12. Riley LW, Remis RS, Helgerson SD, et al. Hemorrhagic colitis associated with a rare Escherichia coli serotype. *The New England journal of medicine* 1983;308(12):681-5.

13. Rowe PC, Orrbine E, Lior H, et al. Risk of hemolytic uremic syndrome after sporadic Escherichia coli O157:H7 infection: results of a Canadian collaborative study. Investigators of the Canadian Pediatric Kidney Disease Research Center. *The Journal of pediatrics* 1998;132(5):777-82.

14. Levine MM. Escherichia coli that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. *The Journal of infectious diseases* 1987;155(3):377-89.

15. Bettelheim KA. A spectrum of verocytotoxigenic Escherichia coli. *Journal of medical microbiology* 1998;47(12):1037-8.

16. Lynn TV, Hancock DD, Besser TE, et al. The occurrence and replication of Escherichia coli in cattle feeds. *Journal of dairy science* 1998;81(4):1102-8.

17. Wang G, Zhao T, Doyle MP. Fate of enterohemorrhagic Escherichia coli O157:H7 in bovine feces. *Applied and environmental microbiology* 1996;62(7):2567-70.

18. Kudva IT, Blanch K, Hovde CJ. Analysis of Escherichia coli O157:H7 survival in ovine or bovine manure and manure slurry. *Applied and environmental microbiology* 1998;64(9):3166-74.

19. Armstrong GL, Hollingsworth J, Morris JG, Jr. Emerging foodborne pathogens: Escherichia coli O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. *Epidemiologic reviews* 1996;18(1):29-51.

20. Mead PS, Griffin PM. Escherichia coli O157:H7. *Lancet* 1998;352(9135):1207-12.

21. Savageau MA. Escherichia coli Habitats, Cell Types, and Molecular Mechanisms of Gene Control. 1983.

22. CDC. Norovirus. 2012, (Centers for Disease Control and Prevention)

23. CDC. Listeria (Listeriosis). 2012, (Centers for Disease Control and Prevention)

24. Ramaswamy V, Cresence VM, Rejitha JS, et al. Listeria--review of epidemiology and pathogenesis. *Journal of microbiology, immunology, and infection = Wei mian yu gan ran za zhi* 2007;40(1):4-13.

25. Pearson DL. Selecting indicator taxa for the quantitative assessment of biodiversity. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 1994;345(1311):75-9.

26. Leon JS, Jaykus L, Moe CL. Food Safety Issues and the Microbiology of Fruits and Vegetables. In: Heredia NL, Wesley IV, Garcia S, eds. *Microbiologically Safe Foods*: Wiley, 2009:255-90.

27. Telliard WA. Method 1601: Male-specific (F+) and Somatic Coliphage in Water by Two-step Enrichment Procedure. Office of Water, Washington, D.C., 2001, (Agency USEP

28. Payment P, Franco E. Clostridium perfringens and somatic coliphages as indicators of the efficiency of drinking water treatment for viruses and protozoan cysts. *Applied and environmental microbiology* 1993;59(8):2418-24.

29. Jiang SC, Chu W, He JW. Seasonal detection of human viruses and coliphage in Newport Bay, California. *Applied and environmental microbiology* 2007;73(20):6468-74.

30. Griffin DW, Gibson CJ, 3rd, Lipp EK, et al. Detection of viral pathogens by reverse transcriptase PCR and of microbial indicators by standard methods in the canals of the Florida Keys. *Applied and environmental microbiology* 1999;65(9):4118-25.

31. Havelaar AH, van Olphen M, Drost YC. F-specific RNA bacteriophages are adequate model organisms for enteric viruses in fresh water. *Applied and environmental microbiology* 1993;59(9):2956-62.

32. Elmund GK, Allen MJ, Rice EW. Comparison of Escherichia coli, Total Coliform, and Fecal Coliform Populations as Indicators of Wastewater Treatment Efficiency. *Water Environment Research* 1999;71(3):332-9.

33. Kornacki JL, Johnson JL. *Enterobacteriaceae, Coliforms, and Escherichia coli as Quality and Safety Indicators. Compendium of Methods for The Microbiological Examination of Foods.*: American Public Health Association; 2001.

34. Splittstoesser DF, Queale DT, Bowers JL, et al. Coliform Content of Frozen Blanched Vegetables Packed in the United States. *Journal of Food Safety* 1980;2(1):1-11.

35. Berg G, Dahling DR, Brown GA, et al. Validity of fecal coliforms, total coliforms, and fecal streptococci as indicators of viruses in chlorinated primary sewage effluents. *Applied and environmental microbiology* 1978;36(6):880-4.

36. Anderson KL, Whitlock JE, Harwood VJ. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. *Applied and environmental microbiology* 2005;71(6):3041-8.

37. Chordash RA, Insalata NF. Incidence and pathological significance of Escherichia coli in human fecal samples. *Food Technology* 1978;32(54).

38. Petterson SR, Teunis PF, Ashbolt NJ. Modeling virus inactivation on salad crops using microbial count data. *Risk analysis : an official publication of the Society for Risk Analysis* 2001;21(6):1097-108.

39. Brackett RE. Incidence, contributing factors, and control of bacterial pathogens in produce. *Postharvest Biology and Technology* 1998;15:305-31.

40. Tauxe RV. Emerging foodborne diseases: an evolving public health challenge. *Emerging infectious diseases* 1997;3(4):425-34.

41. USFDA (U.S. Food and Drug Administration). Guidance for Industry: Guide to Minimize Hazards of Fresh-cut Fruits and Vegetables. 2008.

42. Kearney TE, Larkin MJ, Levett PN. The effect of slurry storage and anaerobic digestion on survival of pathogenic bacteria. *The Journal of applied bacteriology* 1993;74(1):86-93.

43. Mawdsley J, Bardgett R, Merry R, et al. Pathogens in livestock waste, their potential for movement through soil and environmental pollution. *Applied Soil Ecology* 1995;2:15.

44. Warriner K, Huber A, Namvar A, et al. Recent advances in the microbial safety of fresh fruits and vegetables. *Advances in food and nutrition research* 2009;57:155-208.

45. Abu-Ashour J, Joy DM, Lee H, et al. Transport of Microorganisms Through Soil. 1992.

46. Stine SW, Song I, Choi CY, et al. Application of microbial risk assessment to the development of standards for enteric pathogens in water used to irrigate fresh produce. *Journal of food protection* 2005;68(5):913-8.

47. Alum A, Enriquez C, Gerba C. Impact of Drip Irrigation Method, Soil, and Virus Type on Tomato and Cucumber Contamination. *Food Environ Virol* 2011;3:78-85.

48. Green LR, Radke V, Mason R, et al. Factors related to food worker hand hygiene practices. *Journal of food protection* 2007;70(3):661-6.

49. Todd EC, Greig JD, Bartleson CA, et al. Outbreaks where food workers have been implicated in the spread of foodborne disease. Part 5. Sources of contamination and pathogen excretion from infected persons. *Journal of food protection* 2008;71(12):2582-95.

50. Liu P, Chien YW, Papafragkou E, et al. Persistence of human noroviruses on food preparation surfaces and human hands. *Food Environ Virol* 2009;1:141-7.

51. Medeiros LC, Hillers VN, Kendall PA, et al. Food safety education: what should we be teaching to consumers? *Journal of nutrition education* 2001;33(2):108-13.

52. Widdowson MA, Sulka A, Bulens SN, et al. Norovirus and foodborne disease, United States, 1991-2000. *Emerging infectious diseases* 2005;11(1):95-102.

53. Berger CN, Sodha SV, Shaw RK, et al. Fresh fruit and vegetables as vehicles for the transmission of human pathogens. *Environmental microbiology* 2010;12(9):2385-97.

54. Fuerst R. *Frobisher and Fuerst's Microbiology in health and disease: foods as vectors of microbial disease*. W.B. Saunders Company; 1983.

55. Kleinbaum D, Kupper L, Nizam A, et al. *Applied Regression Analysis and Other Multivariable Methods*. 4 ed. Belmont, CA: Thomson Brooks/Cole; 2008.

56. Hansen E, Donohoe M. Health issues of migrant and seasonal farmworkers. *Journal of health care for the poor and underserved* 2003;14(2):153-64.

Tables

Table 1. Characteristics of microbial indicators on tomatoes, hands, source water, irrigation water, and soil associated with growing tomatoes on farms in Coahuila and Nuevo León, México, 2010.

Sample Type	Indicator	Prevalence N (%)	log_{10} Mean (SD)*	n
Produce	*E. coli*	16 (28%)	-0.48 (1.10)	16
	Enterococcus	38 (67%)	1.96 (1.15)	36
	Fecal Coliforms	52 (96%)	2.31 (1.53)	39
	Coliphage	44 (77%)	-0.53 (0.47)	38
Hand-Rinse	*E. coli*	10 (25%)	-0.10 (0.89)	8
	Enterococcus	40 (100%)	3.53 (1.21)	39
	Fecal Coliforms	36 (90%)	2.41 (1.32)	20
	Coliphage	28 (70%)	-0.70 (0.47)	24
Source Water	*E. coli*	14 (88%)	-1.13 (0.46)	14
	Enterococcus	13 (93%)	-1.35 (0.52)	13
	Fecal Coliforms	12 (100%)	-0.43 (0.62)	12
	Coliphage	4 (57%)	-0.26 (0.55)	3
Irrigation Water	*E. coli*	12 (86%)	-1.00 (0.29)	12
	Enterococcus	13 (93%)	-1.16 (0.53)	13
	Fecal Coliforms	12 (100%)	-0.41 (0.83)	12
	Coliphage	6 (86%)	-0.24 (0.35)	5
Soil	*E. coli*	0 (0%)	N/A	0
	Enterococcus	6 (35%)	1.16 (0.38)	6
	Fecal Coliforms	14 (88%)	1.87 (0.83)	14
	Coliphage	9 (56%)	-0.66 (0.75)	9

*****The data for means and standard deviation were transformed because the distributions of the data were not normal. Common logarithms (log_{10}) were used in normalizing this data. *E. coli* was not normally distributed, even after transformation.

Table 2. Characteristics of microbial indicators on cantaloupe, hands, source water, irrigation water, and soil associated with growing cantaloupe on farms in Coahuila and Nuevo León, México, 2010.

Sample Type	Indicator	Prevalence		
		N (%)	log_{10} Mean (SD)*	n
Produce	*E. coli*	39 (65%)	0.47 (1.47)	24
	Enterococcus	60 (100%)	2.99 (1.24)	17
	Fecal Coliforms	60 (100%)	2.87 (0.90)	11
	Coliphage	56 (93%)	N/A	0
Hand-Rinse	*E. coli*	25 (68%)	0.91 (1.91)	19
	Enterococcus	38 (100%)	2.47 (2.44)	9
	Fecal Coliforms	38 (100%)	1.66 (1.35)	7
	Coliphage	32 (84%)	N/A	0
Source Water	*E. coli*	4 (36%)	-1.54 (0.87)	4
	Enterococcus	8 (73%)	-1.14 (0.96)	7
	Fecal Coliforms	11 (100%)	-0.73 (0.40)	9
	Coliphage	4 (40%)	-0.10 (0.00)	1
Irrigation Water	*E. coli*	9 (38%)	-0.82 (0.36)	7
	Enterococcus	19 (79%)	-0.78 (1.42)	17
	Fecal Coliforms	24 (100%)	-0.97 (1.00)	16
	Coliphage	5 (21%)	-1.15 (1.21)	2
Soil	*E. coli*	12 (50%)	-1.07 (0.72)	11
	Enterococcus	18 (75%)	0.55 (1.28)	18
	Fecal Coliforms	24 (100%)	1.66 (1.19)	19
	Coliphage	1 (4%)	-0.10 (0.00)	1

*****The data for means and standard deviation were transformed because the distributions of the data were not normal. Common logarithms (log_{10}) were used in normalizing this data. *E. coli* was not normally distributed, even after transformation.

Table 3. Characteristics of microbial indicators on jalapenos, hands, source water, irrigation water, and soil associated with growing jalapenos on farms in Coahuila and Nuevo León, México, 2010.

Sample Type	Indicator	Prevalence		
		N (%)	log_{10} Mean (SD)*	n
Produce	*E. coli*	10 (23%)	-0.67 (0.86)	7
	Enterococcus	24 (56%)	2.01 (1.64)	24
	Fecal Coliforms	36 (84%)	1.38 (1.61)	25
	Coliphage	29 (78%)	-0.73 (0.53)	19
Hand-Rinse	*E. coli*	16 (55%)	0.52 (1.24)	9
	Enterococcus	29 (100%)	3.12 (1.24)	28
	Fecal Coliforms	27 (93%)	1.71 (1.46)	10
	Coliphage	17 (59%)	-0.77 (0.60)	12
Source Water	*E. coli*	5 (36%)	-0.91 (0.93)	5
	Enterococcus	12 (92%)	-0.97 (0.81)	12
	Fecal Coliforms	11 (100%)	-1.05 (0.56)	7
	Coliphage	1 (20%)	N/A	0
Irrigation Water	*E. coli*	2 (25%)	-0.80 (0.00)	1
	Enterococcus	6 (86%)	-1.48 (0.77)	5
	Fecal Coliforms	7 (100%)	-0.87 (0.55)	5
	Coliphage	1 (25%)	N/A	0
Soil	*E. coli*	0 (0%)	N/A	0
	Enterococcus	4 (29%)	0.92 (0.96)	4
	Fecal Coliforms	13 (93%)	2.57 (0.64)	13
	Coliphage	8 (62%)	-0.87 (0.33)	8

*****The data for means and standard deviation were transformed because the distributions of the data were not normal. Common logarithms (log_{10}) were used in normalizing this data.
E. coli was not normally distributed, even after transformation.

Table 4. Linear Regression of log_{10} indicators on produce measuring the main effect predictors of log_{10} number indicators in soil and log_{10} number of indicators in water controlling for produce type (Coahuila and Nuevo León, México, 2010).

	$Soil$ (SE)	p-value	$Water$ (SE)	p-value	r^2
Fecal Coliforms	0.273 (0.295)	0.377	-0.410 (0.321)	0.230	0.7677
*E. coli**	----------	-------	-3.507 (0.916)	0.012	0.8245
Enterococcus	-1.887 (0.430)	0.022	0.862 (0.242)	0.038	0.9552

**E. coli* was not normally distributed even after log_{10} transformation

Table 5. Linear Regression of log_{10} indicators on produce measuring the main effect predictor of log_{10} number of indicators on hands controlling for produce type (Coahuila and Nuevo León, México, 2010).

	$Hands$ (SE)	p-value	R-Square
Fecal Coliforms	0.386 (0.197)	0.059	0.202
E. coli	0.369 (0.230)	0.141	0.521
Enterococcus	0.375 (0.121)	0.003	0.201
Coliphage	0.232 (0.135)	0.135	0.360

Table 6. Logistic Regression of presence of indicators on produce measuring the main effect predictors of presence of indicators in soil and presence of indicators in water controlling for produce type (Coahuila and Nuevo León, México, 2010).

	OR $Soil$	(95% CI)	p-value	OR $Water$	(95% CI)	p-value
E. coli	0.93	(0.15, 5.63)	0.9364	1.16	(0.30, 4.58)	0.8293
Enterococcus	3.87	(0.60, 24.99)	0.1556	1.07	(0.07, 16.65)	0.9605
Coliphage	6.53	(0.26, 161.64)	0.2518	0.57	(0.05, 6.24)	0.6462

Table 7. Logistic Regression of presence of indicators on produce measuring the main effect predictor of presence of indicators on hands controlling for produce type (Coahuila and Nuevo León, México, 2010).

	OR Hands	(95% CI)	p-value
Fecal Coliforms	6.94	(0.48, 99.44)	0.1538
E. coli	6.24	(2.37, 16.42)	0.0002
Coliphage	9.67	(2.90, 32.32)	0.0002

Appendix A: IRB CLEARANCE

EMORY UNIVERSITY | Institutional Review Board

TO: Juan Leon, PhD
Principal Investigator
Global Health

DATE: June 8, 2012

RE: **Continuing Review Expedited Approval**
CR2_IRB00035460
IRB00035460
Identification and Control of Microbiological Hazards in Imported Fresh Fruits and Vegetables: A Field Epidemiological and Intervention Study in Northern Mexico

Thank you for submitting a renewal application for this protocol. The Emory IRB reviewed it by the expedited process on 06/05/2012, per 45 CFR 46.110, the Federal Register expeditable categories F(7) and Subpart D 46.404. This reapproval is effective from **06/29/2012** through **06/28/2013**. Thereafter, continuation of human subjects research activities requires the submission of another renewal application, which must be reviewed and approved by the IRB prior to the expiration date noted above. Please note carefully the following items with respect to this reapproval:

- A waiver of documentation of written/signed informed consent has been renewed.
- A waiver of parental consent has also been renewed.

Documents reviewed with this application:

- Clean Greens scientific protocolCLEAN6-16-10
- consentimiento_enjuaguemanos_07.14.2011
- Informacion-Encuesta Manipulador 23 MAR 2011
- Informacion-Encuesta-Productor-Manager 23 MAR 2011
- Oral Script for Written Consent_FarmManagerSurvey_Spanish_4.26.2011
- Oral Script for Written Consent_FarmManagerSurvey_ver4.26.2011_CLEAN
- OralScript_Hand Rinsing_ver7.14.2011_CLEAN

Any reportable events (e.g., unanticipated problems involving risk to subjects or others, noncompliance, breaches of confidentiality, HIPAA violations, protocol deviations) must be reported to the IRB according to our Policies & Procedures at www.irb.emory.edu, immediately, promptly, or periodically. Be sure to check the reporting guidance and contact us if you have questions. Terms and conditions of sponsors, if any, also apply to reporting.

Before implementing any change to this protocol (including but not limited to sample size, informed consent, and study design), you must submit an amendment request and secure IRB approval.

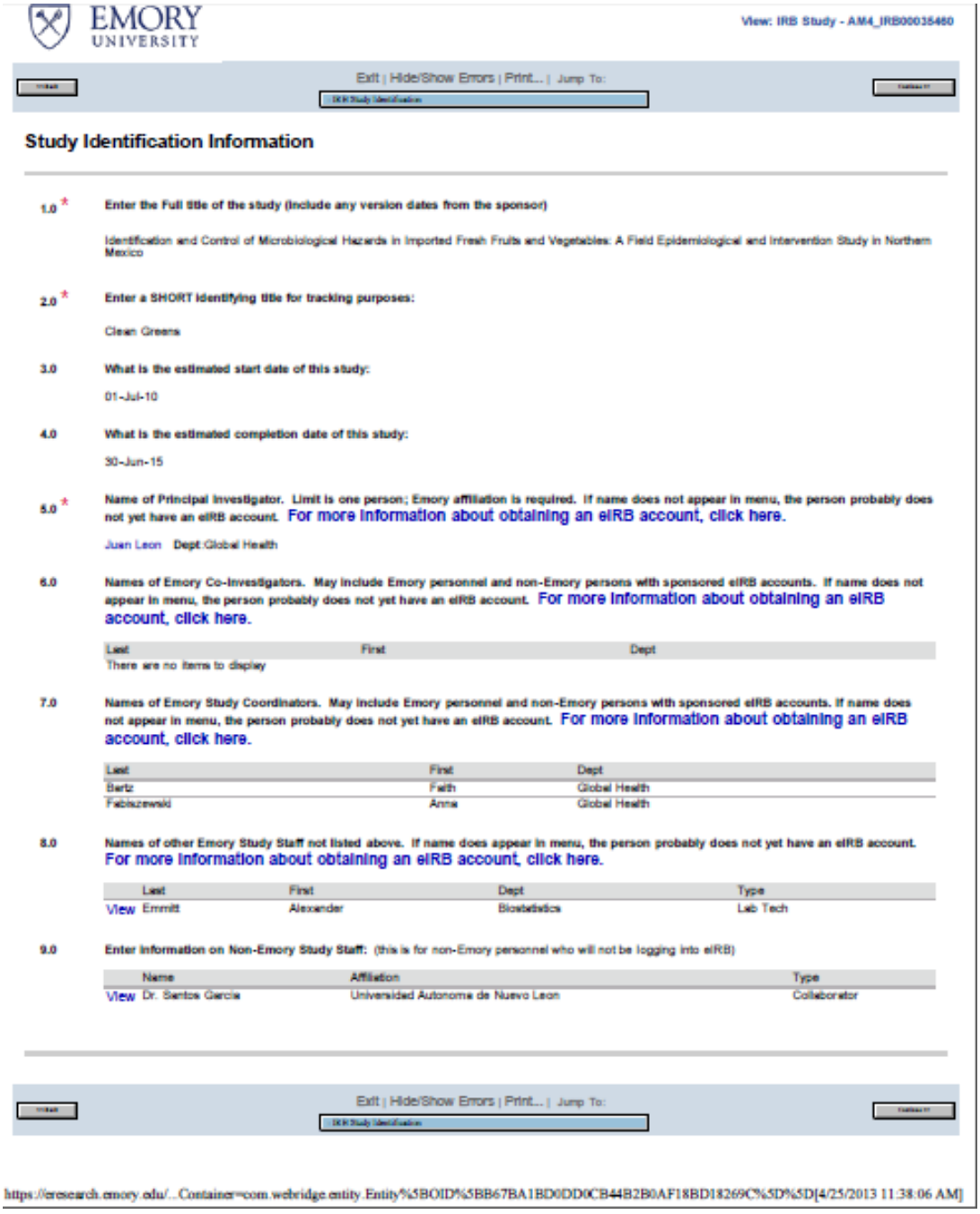
In future correspondence about this matter, please refer to the IRB file ID, name of the Principal Investigator, and study title. Thank you.

Sincerely,

Carol Corkran, MPH, CIP
Senior Research Protocol Analyst
This letter has been digitally signed

CC:	Bartz	Faith	Global Health
	Fabiszewski	Anna	Global Health

Emory University
1599 Clifton Road, 5th Floor - Atlanta, Georgia 30322
Tel: 404.712.0720 - Fax: 404.727.1358 - Email: irb@emory.edu - Web: http://www.irb.emory.edu/
An equal opportunity, affirmative action university



EMORY UNIVERSITY

View: IRB Study - AM4_IRB00035460

Exit | Hide/Show Errors | Print... | Jump To:

Study Identification Information

1.0 * **Enter the Full title of the study (include any version dates from the sponsor)**

Identification and Control of Microbiological Hazards in Imported Fresh Fruits and Vegetables: A Field Epidemiological and Intervention Study in Northern Mexico

2.0 * **Enter a SHORT identifying title for tracking purposes:**

Clean Greens

3.0 **What is the estimated start date of this study:**

01-Jul-10

4.0 **What is the estimated completion date of this study:**

30-Jun-15

5.0 * **Name of Principal Investigator. Limit is one person; Emory affiliation is required. If name does not appear in menu, the person probably does not yet have an eIRB account.** For more information about obtaining an eIRB account, click here.

Juan Leon **Dept** Global Health

6.0 **Names of Emory Co-Investigators. May include Emory personnel and non-Emory persons with sponsored eIRB accounts. If name does not appear in menu, the person probably does not yet have an eIRB account.** For more information about obtaining an eIRB account, click here.

Last	First	Dept
There are no items to display		

7.0 **Names of Emory Study Coordinators. May include Emory personnel and non-Emory persons with sponsored eIRB accounts. If name does not appear in menu, the person probably does not yet have an eIRB account.** For more information about obtaining an eIRB account, click here.

Last	First	Dept
Bertz	Faith	Global Health
Fabiszewski	Anna	Global Health

8.0 **Names of other Emory Study Staff not listed above. If name does appear in menu, the person probably does not yet have an eIRB account.** For more information about obtaining an eIRB account, click here.

	Last	First	Dept	Type
View	Emmitt	Alexander	Biostatistics	Lab Tech

9.0 **Enter Information on Non-Emory Study Staff:** (this is for non-Emory personnel who will not be logging into eIRB)

	Name	Affiliation	Type
View	Dr. Santos Garcia	Universidad Autonoma de Nuevo Leon	Collaborator

Exit | Hide/Show Errors | Print... | Jump To:

https://eresearch.emory.edu/...Container=com.webridge.entity.Entity%5BOID%5BB67BA1BD0DD0CB44B2B0AF18BD18269C%5D%5D[4/25/2013 11:38:06 AM]