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:

_____ _____

Elizabeth Bitler Date

The Associations between Norovirus Outbreak Transmission Mechanisms and Vehicles with Attack Rate, Genogroup Distribution, and GII.4 Strain Distribution: An Outbreak Meta-Analysis

By

Elizabeth Bitler

Master of Public Health

Epidemiology

_____

Juan Leon
Committee Chair

The Associations between Norovirus Outbreak Transmission Mechanisms and Vehicles with Attack Rate, Genogroup Distribution, and GII.4 Strain Distribution: An Outbreak Meta-Analysis

By

Elizabeth Bitler

B.A., Haverford College, 2008

Thesis Committee Chair: 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
in Epidemiology
2012

Abstract:

The Associations between Norovirus Outbreak Transmission Mechanisms and Vehicles with Attack Rate, Genogroup Distribution, and GII.4 Strain Distribution: An Outbreak Meta-Analysis

By Elizabeth Bitler

Norovirus outbreaks are responsible for significant morbidity and mortality worldwide. Norovirus outbreaks can result from foodborne, waterborne, and environmental transmission, with commonly implicated food vehicles including shellfish, produce, and ready-to-eat (RTE) prepared foods, and commonly implicated water vehicles including tap, ground, surface, and recreational water. Attack rate, genogroup distribution, and GII.4 strain are important outbreak outcomes, and may assist in implicating a particular transmission mechanism or vehicle. The goal of this study was to assess the association between outbreak transmission mechanisms and vehicles with attack rate, genogroup distribution, and GII.4 strain distribution. We used bivariate and multivariate techniques to control for other outbreak characteristics. We observed that attack rate did not vary by transmission or food vehicles, upon controlling for other outbreak characteristics, but it did vary by water vehicle. In contrast, genogroup distribution did significantly vary by transmission and food vehicles upon controlling for other outbreak characteristics, but it did not vary by water vehicle. GII.4 strain did not vary by transmission, food vehicles, or water vehicles. We also observed other significant associations between outbreak characteristics (e.g. setting, season, and hemisphere) and outbreak outcomes. Taken together, these results suggest that attack rate may be useful for implicating water vehicles, and genogroup may be useful for implicating transmission mechanisms or food vehicles, however GII.4 strain distribution may not be useful for implicating transmission mechanisms or vehicles during an outbreak investigation. Knowledge of these relationships may help public health workers to more rapidly identify transmission mechanisms or vehicles during norovirus outbreak investigations to reduce morbidity and mortality.

The Associations between Norovirus Outbreak Transmission Mechanisms and Vehicles with Attack Rate, Genogroup Distribution, and GII.4 Strain Distribution: An Outbreak Meta-Analysis

By

Elizabeth Bitler

B.A., Haverford College, 2008

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
in Epidemiology
2012

Acknowledgements:

I would like to thank everyone in my life that supported me prior to and during my time at Emory. This thesis would not have been possible without the support of my family, friends, and colleagues. Thank you all for your encouragement. To my mother, thank you for always supporting my endeavors, academic and otherwise, over the years. I owe my deepest gratitude to my faculty advisor, Juan Leon. Thank you for your patience as you continually offered advice to shape my thesis and my career in public health. You provided encouragement every step of the way, and inspired me to challenge myself.

Table of Contents

Chapter I: Literature Review

Norovirus Incidence and Prevalence

Norovirus is responsible for outbreaks of significant morbidity and mortality that sicken millions of humans in both developed and developing countries. These outbreaks alone account for as many as 60% to 90% of all nonbacterial acute infectious diarrhea outbreaks, making norovirus the most common cause of acute non-bacterial gastroenteritis worldwide.(1)(reviewed in 2-5) Global estimates of norovirus exposure, based on antibody prevalence, indicate that greater than 90% of adults have been exposed to norovirus.(6, 7)(reviewed in 8) In addition to a high incidence and widespread exposure, the number of outbreaks rises as new variants emerge. The characteristics of outbreaks (e.g. seasonality and frequency of outbreaks) also change over time due to rising epidemic prevalence of specific norovirus strains.(9) By better understanding the characteristics of norovirus outbreaks, effective interventions may be employed to minimize associated morbidity and mortality. Due to the extensive occurrence of norovirus outbreaks, and changing outbreak epidemiology, describing outbreak characteristics is critical to reducing norovirus outbreak burdens.

Due to the high prevalence of norovirus infection, and the typically self-resolving and short-term symptoms, individuals often disregard norovirus outbreaks as an important public health issue. However, morbidity and mortality of norovirus are significant.(reviewed in 10) Among foodborne illnesses in the United States, norovirus caused 58% of all illnesses, with 5.5 million cases annually.(11) In addition to debilitating symptoms, the disease may cause complications in high-risk groups. Those with increased risk of complication include infants, the elderly, and immuno-compromised individuals.(reviewed in 12) As a result, norovirus is the second leading cause of hospitalization and the third leading cause of death among US foodborne illnesses.(11)

Historical Perspective

Norovirus disease outbreaks were first observed as part of a common seasonal influx of acute gastrointestinal illness outbreaks termed "winter vomiting disease."(13) In 1969, Adler and Zickl described the epidemiology of one such outbreak in Norwalk, Ohio, and were unable to implicate a bacterial source.(14) At that time, they hypothesized that a viral agent was responsible for the illness. Samples from this outbreak were later studied by immune electron microscopy, and the a virus particle was visualized and named after the location of the outbreak – Norwalk virus.(15) Subsequent outbreak viral agents with epidemiological characteristics similar to the Norwalk outbreak were named Norwalk-like viruses (NLV), caliciviruses, or small round-structured viruses. In the 1990s, reverse transcription polymerase chain reaction (RT-PCR) techniques enabled the amplification, detection, and classification of distinct norovirus strains which belong to the genus *Norovirus* and family *Caliciviridae*.(16)

Clinical Presentation

Norovirus infectivity is high, although not all infected individuals become symptomatic.(17) Among those that develop symptoms, the incubation period is 2-61 hours, with a median of 34 hours for outbreaks.(reviewed in 18) Symptom presentation also varies by individual.(19) The disease typically manifests with symptoms of diarrhea, vomiting, abdominal pain, abdominal cramps, nausea, fever, chills, and mylagia.(14)(reviewed in 2, 20, 21) Symptoms cease in approximately 1-8 days for outbreaks.(reviewed in 18)

While infected, individuals shed the virus in stools and vomitus.(reviewed in 4) Norovirus is shed not only by those with clinical disease, but also by those recovering from

disease and by infected individuals who do not manifest symptoms.(15, 17)(reviewed in 22) The viral particles are stable in the environment, persisting for extended periods of time and in extreme conditions (viral particles are stable across a pH range of 3-7, and in temperatures up to 55 degrees Celsius).(23) Only a few viral particles are required for infection, with a 50% infectious dose (ID_{50}) of 18 virions.(24, 25)(reviewed in 12)Additionally, individuals may have multiple infections in their lifetime.(17) As a result, infected individuals readily propagate norovirus.

Epidemiology

Understanding factors associated with norovirus outbreaks is important for understanding how to best prevent future outbreaks and successfully intervene to mitigate and halt ongoing outbreaks. To develop effective interventions, it is necessary to understand the manner in which norovirus particles are spread and how people become exposed. In addition to transmission, it is also important to understand other outbreak characteristics that are meaningful for outbreak prevention, investigation, and intervention. Important outbreak characteristics include attack rate, genotype distribution, season, setting, size, and duration, which will be considered here.

Transmission Routes and Common Vehicles

Norovirus is spread by oral contact with infected feces or emesis (vomitus). Some of the most common modes of transmission include foodborne, waterborne, and environmental transmission. Although person-to-person transmission is an important mode of transmission in norovirus outbreaks, and occurs relatively frequently in healthcare settings, it poses unique considerations and challenges warranting separate investigation, and as such will not be

considered in the present analysis.(reviewed in 26) Foodborne, waterborne, and environmental disease transmission each result from contamination of a common source.(reviewed in 8) They are also similar in that these transmission modes are likely to occur in both developed and developing countries, and among individuals of all ages.

Foodborne transmission occurs by the ingestion of viral particles with food products. Contaminated food products may result in a focal outbreak if the foods are distributed only in an immediate area, such as at a picnic, or they may result in a dispersed outbreak if the foods are distributed regionally or internationally. Foods can become contaminated at any point prior to consumption, and specific food items are more likely to become contaminated. Shellfish, produce, and ready-to-eat (RTE) foods are common culprits as norovirus transmission vehicles. These vehicles are responsible for a large number of foodborne associated norovirus outbreaks, and each represent unique means of food contamination.(reviewed in 8) Because these vehicles become contaminated through different routes, intervention methods must vary accordingly to prevent norovirus transmission.

Filter-feeding bivalve molluscan shellfish are a common transmission vehicle.(reviewed in 8) Among seafood, the filter-feeding shellfish are uniquely capable of transmitting norovirus. Their physiology results in selective concentration of viruses, including norovirus, in a way that other seafood items (e.g. fish or arthropods such as lobsters and crabs) do not. When shellfish intake contaminated water (e.g. water contaminated by sewage, polluted runoff, or dumping of infected boaters' vomitus or stools), norovirus particles present in the water adhere to the shellfish digestive tissues.(27-30) Subsequent preparation and cooking of shellfish may not be sufficient to inactivate the norovirus prior to consumption.(31) Thus prevention of shellfish contamination may occur by monitoring harvesting waters for sewage pollution, treating contaminated shellfish, and relocating growing beds.(reviewed in 8, 26)

Produce also represents an important vehicle for foodborne norovirus transmission.(reviewed in 4) There are several means by which fresh and fresh-frozen produce may become contaminated. For example, produce may be contaminated in the field if contaminated water is used for irrigation (reviewed in 32) or if farm workers defecate in the open. In addition to the fields, several other opportunities exist for produce contamination during the production stages. Produce may be contaminated during picking, packing, or any other manual handling.(reviewed in 8) It may also become contaminated if washed with contaminated water.(33)(reviewed in 26)Contaminated produce that is consumed without sufficient cooking exposes individuals to norovirus. Because fresh produce may become contaminated by a variety of means, control measures first require accurate identification of the contamination route. Efforts may then target the implicated source, such as worker behavior or water treatment.(reviewed in 8)

RTE foods also present a significant risk for norovirus transmission.(reviewed in 4) RTE products are any food products that do not require additional preparation before consumption. Food preparation includes washing, cooking, or other processing.(34) Prepared (versus homemade) potato salad, deli sandwiches, and packaged fruit cups are examples of RTE foods. RTE foods pose a unique means of contamination because food workers handle ingredients. Thus, an ill or asymptomatic food worker may pass norovirus onto food items if they lack sufficient personal hygiene to prevent transfer.(35)(reviewed in 26, 32) Consumers are then exposed to any contamination that may have occurred during processing, as RTE items are not cooked any further after preparation.(reviewed in 4) Control of RTE foodborne transmission relies primarily on safe food worker practices, including staying home from work while ill or while caring for ill family members.(reviewed in 26)

Waterborne transmission may occur after ingestion of contaminated water. This may happen by coming into contact with surface water, while navigating through accumulated

floodwaters, or while swimming in lakes and rivers. Contaminated ground water may result in norovirus transmission via wells used to provide household water. Additionally, if municipal water supplies or bottled water become contaminated and are not treated appropriately, drinking water may also spread norovirus and result in an outbreak.(reviewed in 2, 26, 36) Similar to foodborne transmission outbreaks, waterborne outbreaks may result in focal or dispersed outbreaks depending on the distribution of contaminated water. For example, contaminated recreational water will result in a focal outbreak, extending only to those that are exposed at the site of contamination, while municipal water contamination may result in a dispersed outbreak that spans the water supply. Once water sources have been implicated in an outbreak, effective control measures include repairing drinking water delivery systems to prevent contamination or chemically treating recreational water.(reviewed in 26)

Environmental transmission provides another means for contact with viral particles. This typically occurs via individual contact with infectious fomites.(reviewed in 2, 26) Fomites are any environmental object and these fomites may become contaminated with fecal or vomitus. For example, an individual may be exposed to norovirus while using shared toilet facilities or pressing elevator buttons after contamination by an infected individual.(37, 38) These fomites will result in infection in those individuals that come into contact with the contaminated item, resulting in a focal outbreak. Control of environmental transmission requires disinfection of the contaminated surfaces. It may be difficult to sanitize contaminated fomites due to the environmental stability of norovirus, and disinfection recommendations vary and may not be completely effective.(reviewed in 26)

Attack Rate

Norovirus outbreaks are often described in terms of the attack rate. Attack rate is a measure of the number of cases within a group of exposed individuals. The use of "rate" is misleading, as it is not a measure of disease incidence over a period of time, and is instead a measure of incidence. Thus attack rate indicates the efficacy of the virus to infect individuals, enabling the spread of norovirus. In norovirus challenge studies, attack rates of 55% have been observed.(39) It is not uncommon for outbreaks to have attack rates greater than 50%.(40)

Attack rates may vary by outbreak for a variety of reasons. Transmission mode and particular vehicles may promote high attack rates if they encourage greater exposure to norovirus or more efficient internalization of viral particles. It may be easier to ingest enough particles to cause illness by eating contaminated foods than by coming into contact with contaminated fomites in the environment. Likewise, a greater attack rate may be associated with a RTE item prepared by a sick food worker (reviewed in 26) than with a produce item that was contaminated by a sick field worker but was subsequently processed in a manner that reduced the number of viral particles.

Genotype

Noroviruses are a member of the *Caliciviridae* family. They are categorized by genogroup, genotype or cluster, and subgenotype or strain.(41) Noroviruses are currently classified into five genogroups GI-GV, of which three cause disease in humans- GI, GII, and GIV.(reviewed in 42) The number of genotypes varies according to genotyping and nomenclature strategy.(41, 43)(reviewed in 44) New strains of norovirus emerge as the viruses undergo mutation and genetic recombination in response to population pressure.(reviewed in 42) Recombinant noroviruses can be classified as intergenogroup, intergenotype, and

intersubgenotype. These recombinant viruses represent combinations of viruses from different genogroups, different genotypes within the same genogroup, and different strains within the same genotype, respectively.(43) Because recombinant noroviruses cannot be adequately grouped using traditional nomenclature, the nomenclature strategies continue to change with better understanding of norovirus phylogeny.(43)

GII.4 strains are common in outbreaks, and have distinct epidemiological patterns. GII.4 strains have been implicated in several global outbreaks and are more commonly associated with norovirus outbreaks than strains of other clusters, accounting for 85.8% of outbreaks in one study.(45)(reviewed in 26) Although the GII.4 cluster is the most commonly implicated in outbreaks, this observation is likely confounded by transmission type and setting.(reviewed in 26) GII.4 strains are commonly implicated in institutional settings and person-to-person outbreaks, which are more likely to be reported. In contrast, GI and GII strains, other than GII.4, are common in untreated and treated sewage (reviewed in 26), which may be implicated in contamination of water, shellfish, or produce that lead to waterborne or foodborne outbreaks(46, 47). GII.4 strains have a high mutation rate relative to other norovirus strains.(48) The evolution rate of GII.4 strains may promote epidemiological fitness, resulting in the dominance of this cluster and continued circulation over the past several decades.(48-54)(reviewed in 42)

GII.4 strains may result in different clinical presentations during an outbreak than non-GII.4 strains. Among those infected by GII.4 norovirus strains, the occurrence of vomiting is greater than among those infected by other norovirus genotypes.(55) GII.4 strains are also associated with prolonged illness versus non-GII.4 strains.(56) Strains that result in increased vomiting or duration of illness also have increased potential for spreading more virions, which may further propagate outbreaks. GII.4 outbreaks also appear to have higher attack rates than

non-GII.4 outbreaks, and may have higher infectivity.(55) Virulence may also vary by strain, as differences have been observed among GII.4 norovirus strains.(57)(reviewed in 42)

In summary, various strains, clusters, or genogroups may be associated with particular outbreak transmission types, settings, vehicles, and even clinical presentation. The distribution of genotypes is an important epidemiological outcome with regard to outbreaks. Early identification of a genotype distribution during an outbreak may implicate one transmission route or vehicle over others. Thus, better understanding of these distribution patterns may lead to quicker and more efficient outbreak intervention.

Distinct genogroup or genotype patterns may also reflect how a vehicle becomes contaminated. One would expect to observe a different pattern if contamination occurs by one person (i.e. one or few strains), such as an ill food-worker preparing RTE items, than if contamination occurs by many people (i.e. multiple strains), such as sewage contamination of a water supply. Sewage contains noroviruses circulating in the population, and contamination with sewage is likely to result in outbreaks with multiple strains.(58, 59) Therefore, it is likely that foods contaminated by sewage (e.g. early in the processing chain) compared to foods contaminated by food-handler (e.g. late in the processing chain) will vary in their norovirus genotype distribution. As a result, shellfish or produce contaminated by sewage (e.g. early in the processing chain) are likely to have multiple norovirus genotypes present that differ from RTE foods that become contaminated by an infected food worker (e.g. late in the processing chain).(59, 60)(reviewed in 32)

Various means of contamination and environmental conditions may result in distinct patterns of norovirus genogroup or genotype distribution among different transmission routes or vehicles. For example, the proportion of non-GII.4 norovirus strains to GII.4 strains is greater for foodborne outbreaks than for person-to-person outbreaks.(61) Waterborne transmission is also

likely to result in outbreaks with multiple norovirus strains especially when contaminated by sewage.(58, 62) Additionally, GI strains are more often associated with waterborne outbreaks, while GII strains are more often associated with foodborne outbreaks, and this may be due to the stability of GI strains in water.(46, 63, 64) Similarly, GI strains are more often associated with shellfish outbreaks, while GII strains are more often associated with foodborne outbreaks.(60) For shellfish in particular, bioaccumulation of norovirus in shellfish tissues fluctuates by strain, further differentiating the genotype profile of shellfish-related outbreaks.(65) Shellfish contaminated with multiple virus strains also provide opportunities for genetic recombination, and the emergence of novel norovirus strains.(reviewed in 32)

Additional epidemiological characteristics

Norovirus outbreaks are marked with a distinct seasonality, with an increase in cases during the winter months. The winter increase was first noted in the original account of "winter vomiting disease."(13) The winter seasonality may be due to the increased stability of norovirus virions in winter climates (66), and indoor activities may promote the spread of norovirus.(67) Supporting the hypothesis that indoor activities and the winter climate promote norovirus spread, a study of environmental samples taken at food catering locations observed the same trends in seasonality for contaminated surfaces as seasonal infections.(68) There continues to be a wintertime peak for norovirus outbreaks, although outbreaks occur year-round with smaller spring and summer outbreaks.(53, 69-72)(reviewed in 73) Interestingly, outbreaks are most common in the spring and summer months in the Southern hemisphere.(74, 75)(reviewed in 20) Spring and summer outbreaks tend to have greater genetic diversity than outbreaks during the winter season.(70, 76) It has been proposed that this genetic seasonality promotes the selection, during the spring and summer, of the strains that will predominate the remainder of the season.

As new strains emerge, they are often associated with increased attack rate or out of season infections, which may explain the summertime outbreaks.(53) Genetic diversity of norovirus strains decreases throughout the season, as favorable strains dominate.(54)(reviewed in 42) Specific strains seem to differ in their seasonality. For example, a study of long-term care facilities observed that GII.4 strains displayed the typical seasonality with an increase during winter months, while non-GII.4 strains did not.(56) The relationship between seasonality and strain extends to contamination of shellfish, as well. GI.1 accumulates in shellfish tissues with a seasonal pattern, while GII.4 and GII.3 do not.(65) In addition to norovirus strains, seasonality also varies by outbreak setting. Outbreaks in healthcare settings follow the seasonal fluctuation with a winter-time peak, while outbreaks in non-healthcare settings do not.(77) This observation may be due to the seasonality of different genotypes, which also vary by setting.(78)

Attack rate and genotype distribution of norovirus outbreaks are associated with the setting of the outbreak.(40)(reviewed in 20) For example, Harris *et al.* observed greater attack rates for outbreaks healthcare settings, hospitals, and nursing homes than for other semi-enclosed settings, although there were few outbreaks in other settings and the observed difference is not significant. With regard to genotype distribution, Bruggink and Marshall found that GII.4 to be more common in healthcare settings while GI.2 and GIIb genotypes were more common in non-healthcare settings.(78) Because of the observation of numerous GII.4 strains circulating in hospital environments, it has been hypothesized that GII.4 strains have a competitive advantage in healthcare settings.(79)

Outbreak size and duration are important markers of the impact of an outbreak. Outbreaks may range in size from a group of fewer than ten people to tens of thousands of people.(80, 81) The total number of people affected in an outbreak varies by setting. For example more people were affected in outbreaks on ships than in healthcare, education, or other

recreational settings.(40) Size also varies by transmission. Outbreaks involving a common source, such as contaminated food or water, result in more cases than those involving person-to-person transmission.(reviewed in 73) Duration is also an important outbreak characteristic. Outbreak duration may vary from one day to several months.(reviewed in 73) Longer outbreaks are associated with continued exposure to a contaminated common source or continual introduction of susceptible people to those infected.(reviewed in 73) Rosenthal *et al.* observed that duration was longer in larger, long-term care facilities with increased potential for continual introduction than in smaller facilities.(56) Outbreak duration is significantly longer for healthcare settings than for non-healthcare settings, displaying variety by setting, as well.(40)

Analysis of multi-outbreak data

Previous research has attempted to characterize risk determinants for particular norovirus outbreak outcomes. For example, the Centers for Disease Control and Prevention compiled data from norovirus outbreaks reported in the United States between July 1997 and June 2000 and analyzed it to describe epidemiologic and molecular trends. Bivariate analyses were employed to assess the relationship between setting, transmission, and severity with norovirus strain type; however only a significant association between setting and genotype was observed.(1) Similar methods were used in conjunction with the Foodborne Viruses in Europe (FBVE) surveillance network to observe associations between genotype profiles and contamination sources.(60) Multivariate modeling was used with data from the FBVE network to assess the relationship between genotype and setting, transmission mode, and seasonality for reported norovirus outbreaks. Transmission and seasonality significantly predicted genotypes among reported outbreaks in the FBVE network. For example, person-to-person transmission was associated with GII.4 outbreaks but not other GII outbreaks. Additionally, winter seasonality was clearest for

outbreaks caused by GII.4 strains, with greater peaks in outbreak incidence during the winter months. A weaker seasonality pattern was observed for GII strains that are not GII.4, and no seasonality pattern was observed for non-GII strains. Although the data did describe genotype trends, the researchers did not observe previously described associations between genotype and setting.(82) Verhoef *et al.* also used FBVE network data and multivariate strategies to calculate the potential of an outbreak to be related to food contaminated early in the processing chain as opposed to a food-handler or person-to-person transmission.(61) Verfhoef *et al.*'s multivariate modeling is important for considering numerous outbreak characteristics; however these data are limited to passive surveillance efforts in Europe and contain information only for suspected foodborne outbreaks, preventing the extension of findings to other transmission types, such as waterborne or environmental transmission outbreaks.(60, 61, 82) Recently, Harris *et al.* conducted a meta-analysis of outbreaks published in peer-reviewed articles. They analyzed data from 72 outbreaks described in 47 papers related to semi-enclosed settings, such as nursing homes, cruise ships, and schools. Attack rate, duration, and number of cases by setting were compared for outbreaks in which infection control measures were used and those for which they were not. Attack rate and duration of outbreak did vary by setting, however the authors did not employ multivariate techniques to control for confounding.(40) Previous studies highlighted the importance of multivariate modeling to assess the relationships between outbreak characteristics and outcomes such as attack rate and genotype. Recently, our group employed multivariate techniques to describe the relationships for transmission and setting outcomes with attack rates and genogroup distribution for all published norovirus outbreaks since 1992, but did not assess GII.4 strain distribution or commonly implicated vehicles.(83) At present, the norovirus outbreak literature lacks a comprehensive analysis of the relationships for attack rate, genogroup distribution, and GII.4 strain distribution with foodborne, waterborne, and environmental transmission, or with commonly implicated food and water vehicles, while controlling for other

outbreak characteristics. Thus, an analysis of all published norovirus outbreaks would support the existing literature and would enable multivariate modeling for several transmission types and commonly implicated vehicles.

Goal and aims

To address this need, the goal of this thesis is to assess the association between outbreak transmission mechanisms and vehicles with attack rates, genogroup distribution, and GII.4 strain distribution, while controlling for other outbreak characteristics (e.g. setting, season, and hemisphere) for published worldwide foodborne, waterborne, and environmental norovirus outbreaks published between December, 1993 and May, 2011, that meet inclusion criteria for analysis.

To address this goal the following aims are proposed:

Aim 1. To assess the relationships between risk determinants and attack rate and genotype distribution between foodborne, waterborne, and environmental transmission norovirus outbreaks.

Aim 2. To assess the relationships between risk determinants and attack rate and genogroup distribution for foodborne norovirus outbreaks overall, and those associated with particular food products- specifically, shellfish, produce, and ready-to-eat prepared foods.

Aim 3. To assess relationships between risk determinants and attack rate and genogroup distribution for waterborne norovirus outbreaks overall, and those associated with particular transmission vehicles- specifically, contaminated groundwater, surface water, and drinking water.

Public health significance

The large dataset provides a unique opportunity for data analysis not previously possible. This present study will provide novel insights into the overarching trends of norovirus outbreaks, with a large volume of data allowing detailed analysis. This detailed analysis will help to inform the predictors of attack rate and genogroup for specific norovirus transmission modes and vehicles. By analyzing a large, worldwide database of norovirus outbreaks according to transmission type, we can inform public health practitioners of significant relationships between outbreak characteristics and norovirus outcomes. This analysis will provide a better understanding of the different predictors for foodborne, waterborne, and environmental transmission norovirus outbreaks, with an emphasis on the predictors that remain significant while controlling for other associated variables. Findings regarding risk determinants associated with increased norovirus outbreak attack rates or genogroup patterns will be communicated in a manner that is useful for prevention efforts or informing outbreak investigations by public health field workers. The findings may enable more rapid identification of outbreak sources, and indicate the most appropriate intervention techniques given a set of outbreak characteristics distinctive of foodborne, waterborne, or environmental norovirus outbreaks.

Chapter II: Manuscript

Title

The Associations between Norovirus Outbreak Transmission Mechanisms and Vehicles with Attack Rate, Genogroup Distribution, and GII.4 Strain Distribution: An Outbreak Meta-Analysis

Authors

E. J. Bitler, J. E. Matthews, B. W. Dickey, J. N. S. Eisenberg, C. L. Moe, J. S. Leon

(subject to change)

Abstract

Norovirus outbreaks are responsible for significant morbidity and mortality worldwide. Norovirus outbreaks can result from foodborne, waterborne, and environmental transmission, with commonly implicated food vehicles including shellfish, produce, and ready-to-eat (RTE) prepared foods, and commonly implicated water vehicles including tap, ground, surface, and recreational water. Attack rate, genogroup distribution, and GII.4 strain are important outbreak outcomes, and may assist in implicating a particular transmission mechanism or vehicle. The goal of this study was to assess the association between outbreak transmission mechanisms and vehicles with attack rate, genogroup distribution, and GII.4 strain distribution. We used bivariate and multivariate techniques to control for other outbreak characteristics. We observed that attack rate did not vary by transmission or food vehicles, upon controlling for other outbreak characteristics, but it did vary by water vehicle. In contrast, genogroup distribution did

significantly vary by transmission and food vehicles upon controlling for other outbreak characteristics, but it did not vary by water vehicle. GII.4 strain did not vary by transmission, food vehicles, or water vehicles. We also observed other significant associations between outbreak characteristics (e.g. setting, season, and hemisphere) and outbreak outcomes. Taken together, these results suggest that attack rate may be useful for implicating water vehicles, and genogroup may be useful for implicating transmission mechanisms or food vehicles, however GII.4 strain distribution may not be useful for implicating transmission mechanisms or vehicles during an outbreak investigation. Knowledge of these relationships may help public health workers to more rapidly identify transmission mechanisms or vehicles during norovirus outbreak investigations to reduce morbidity and mortality.

Introduction

Norovirus is responsible for outbreaks of significant morbidity and mortality that sicken millions of humans, and is the most common cause of acute non-bacterial gastroenteritis worldwide.(1)(reviewed in 2-5, 10) The disease typically manifests with symptoms of diarrhea, vomiting, abdominal pain, abdominal cramps, nausea, fever, chills, and mylagia.(14)(reviewed in 2, 20, 21) Symptoms cease in approximately 1-8 days for outbreaks.(reviewed in 18) Norovirus infection may cause severe complications in high-risk groups, and morbidity and mortality of norovirus are significant.(reviewed in 10, 12) To develop effective interventions and reduce morbidity and mortality of norovirus outbreaks, it is necessary to understand the manner in which norovirus particles are spread and how people become exposed.

Norovirus is spread by oral contact with infected feces or vomitus. Oral-fecal contact can occur by ingesting contaminated food or water, or by oral contact with a contaminated object in the environment. Thus foodborne, waterborne, and environmental transmission outbreaks each can result from contamination of a common source.(reviewed in 8) Although person-to-person transmission is important (reviewed in 26), it poses unique considerations and challenges warranting separate investigation, and will not be considered in the present analysis. Foodborne transmission occurs by the ingestion of viral particles with food products. Foods can become contaminated at any point prior to consumption, and specific food items are more likely to become contaminated. Shellfish, produce, and ready-to-eat (RTE) prepared foods are common culprits as norovirus transmission vehicles.(reviewed in 4, 8) When shellfish intake norovirus-contaminated water, norovirus particles present in the water adhere to the shellfish digestive tissues.(27-30) Produce may be contaminated if irrigated or washed by contaminated water, exposed to feces by workers defecating in the field, or manually handled by an infected worker.(33)(reviewed in 8, 26, 32) RTE food items (foods that do not require additional

preparation before consumption) may become contaminated when handled by an infected food worker.(35)(reviewed in 26, 32) Waterborne transmission may occur upon ingestion of contaminated water. This may happen by ingesting contaminated municipal tap water, contaminated ground and well water, surface water (e.g. floodwaters or lakes), or recreational water (e.g. pool water).(reviewed in 2, 26) Environmental transmission occurs upon ingestion of viral particles after contact with infectious fomites.(reviewed in 2, 26) Fomites are any environmental object that becomes contaminated with fecal or vomitus (e.g. shared toilet facilities or elevator buttons.)(37, 38) Taken together, contaminated food vehicles, water vehicles, and environmental fomites result in numerous norovirus outbreaks.

Attack rate and genotype distribution are commonly described outbreak outcomes, with implications for outbreak investigations.(55, 56) Attack rate is a measure of the number of cases within a group of exposed individuals. The attack rate indicates the efficacy of the virus to infect individuals, and may be higher for transmission modes or vehicles that encourage greater exposure to norovirus or more efficient internalization of viral particles.(reviewed in 26) For example, it may be easier to ingest enough particles to cause illness by eating contaminated foods than by coming into contact with contaminated fomites in the environment. With regards to genotype distribution, norovirus outbreaks are often characterized by genogroups or strains present. Noroviruses are a member of the *Caliciviridae* family, and are categorized by genogroup, genotype or cluster, and strain.(41) Noroviruses are classified into five genogroups GI-GV, of which three cause disease in humans- GI, GII, and GIV.(reviewed in 42) GII.4 cluster strains are the most common outbreak strains, accounting for 85.8% of outbreaks in one study.(45)(reviewed in 26) Both genogroup and GII.4 strain distribution may be associated with different outbreak exposure routes. For example, GI strains are more often associated with waterborne outbreaks,(63, 64) while GII strains are more often associated with foodborne outbreaks,(46) and this may be due to the stability of GI strains in water.(46, 63) The presence of both GI and GII

strains in fecal or vomit samples from an outbreak victim may indicate food or water contamination by sewage, as sewage contains noroviruses circulating in the population, and is likely to result in outbreaks with multiple strains.(58, 59) Additionally, Verhoef *et al.* suggested that the proportion of non-GII.4 norovirus strains to GII.4 strains is greater for foodborne outbreaks than for person-to-person outbreaks.(61) As identification of attack rate or genotype distribution during an outbreak may implicate one transmission route or vehicle over others, a better understanding of the relationships between these outcomes with transmission modes and vehicles may facilitate implication of one transmission mechanism or vehicle over another.

Other outbreak characteristics may confound the relationships between exposure routes and outcomes (attack rate and genotype distribution). For example, both attack rate and genotype distribution of norovirus outbreaks were associated with the setting of the outbreak.(40, 78)(reviewed in 20) Genotype distribution is also associated with season, as spring and summer outbreaks tend to have greater genetic diversity than outbreaks during the winter season.(70, 76) Outbreak patterns also varied by hemisphere.(74, 75)(reviewed in 20) In order to effectively characterize the relationships between outbreak outcomes and particular modes of transmission or vehicles, it is important to control for other outbreak characteristics that may confound the relationships (e.g. setting, season, hemisphere). In one study, multivariate methods were employed to distinguish between outbreaks associated with food contaminated early in the processing chain, as opposed to a food-handler or person-to-person transmission, using genotype profiles.(60) In another study, multivariate methods distinguished between foodborne and person-to-person outbreaks using GII.4 strain, number of cases, and setting.(61) The two studies described were performed with data from the Foodborne Viruses in Europe (FBVE) surveillance network, limiting the data to surveillance efforts in Europe.(60, 61) Recently, our group employed multivariate techniques to describe the relationships for transmission and setting outcomes with attack rates and genogroup distribution for all published norovirus outbreaks since

1992, but did not assess GII.4 strain distribution or commonly implicated vehicles.(83) At present, the norovirus outbreak literature lacks a comprehensive analysis of the relationships for attack rate, genogroup distribution, and GII.4 strain distribution with foodborne, waterborne, and environmental transmission, or with commonly implicated food and water vehicles, while controlling for other outbreak characteristics.

The goal of this study was to assess the association between outbreak transmission mechanisms and vehicles with attack rates, genogroup distribution, and GII.4 strain distribution, while controlling for other outbreak characteristics (e.g. setting, season, and hemisphere). The current study uses a large collection of worldwide-published data to enable both bivariate and multivariate analysis. Identifying the relationships between transmission mechanisms and vehicles with outbreak outcomes (attack rates, genogroup distribution, and GII.4 strain distribution) will enable public health workers to use outcomes as evidence for associated transmission modes or vehicles during outbreak investigations.

Methods

IRB

This research did not require Institutional Review Board (IRB) review because it did not meet the definition of research involving "human subjects" or the definition of "clinical investigation" as set forth in Emory policies and procedures and federal rules.

Outbreak Data

Norovirus outbreak data were collected from peer-reviewed articles published between December 1993 and May 2011. Data abstraction methods were discussed in depth in Matthews *et al.*(83) Of 902 outbreaks included in the database, 435 contained information about norovirus transmissions for vehicles of interest. Dichotomous variables were constructed for primary transmission or vehicle identified in each outbreak. The transmission variables indicated foodborne, waterborne, or environmental outbreaks. Food vehicles included were produce, shellfish, and RTE-associated outbreaks. Water vehicles included were tap and municipal water, ground water, surface water, and recreational water. In outbreaks where multiple vehicles were implicated, the vehicle identified as most likely associated with the outbreak was used for analysis. For example, a specific vehicle was implicated if the authors explicitly mentioned that there was stronger circumstantial evidence in favor of that vehicle, or if stronger epidemiological evidence (e.g. a higher significant odds ratio) was presented in favor of that vehicle. The outcome variables of interest were attack rate, genogroup distribution, and strain. Attack rate was determined as the number of cases out of all persons at risk for each outbreak. Genogroup was categorized for each outbreak according to the presence of GII strains only, GI strains only, or both GII and GI strains. Strain was categorized for each outbreak as either the presence of any GII.4 strain or the presence of only non-GII.4 strains.

Data Analysis

Analyses were performed using Statistical Analysis Software 9.3 (SAS Institute, Carey, N.C.). The relationships between predictors and attack rate (via ANOVA tests followed by Tukey's post hoc tests) or genogroup and strain (via chi-square tests followed by multiple comparisons tests for proportions (84)) were assessed. Chi Square analysis was also performed for subsets of the outbreak characteristic variables and outcome variables if the analysis on the full dataset could not be performed (e.g. zero observations for any cell or expected counts less than five). Multivariate analyses were performed using linear regression for continuous variable attack rate, polytomous regression for the nominal, three-level variable for genogroup (GII only, GI only, and both GII and GI), and logistic regression for strain (GII.4 and non-GII.4 strains). An interaction term for season and setting was included to assess effect modification, however data were too sparse for this assessment. Data were analyzed to ensure that linear, polytomous, and logistic regression model assumptions were met.(85, 86) Backward elimination was performed with Partial F tests to determine which variables did not significantly improve prediction models and did not confound the relationship between the predictors of interest and outbreak outcome variables. Transmission, food vehicle, and water vehicle variables were not eligible for backward elimination, because they must remain in the models to describe the relationships between these predictors and outbreak outcome variables. In the instance of semi-complete separation of the predictor of interest and the outcome, logistic regression modeling was supplemented with the Firth option to obtain estimates.(87-90) An alpha level of 0.05 was used for all tests of significance.

Results

Bivariate Analysis

To determine which relationships between outbreak characteristics and the outcomes of attack rate, genogroup distribution, and strain were significant, bivariate methods were employed for each pair of predictors and outcomes (Table 1). 435 outbreaks with transmission or vehicle information were included in the analysis. For attack rate, there was a significant overall effect of transmission on attack rate ($p=0.0214$), however Tukey post hoc tests did not reveal any significant pairwise mean differences. Shellfish outbreaks (mean=59.9, SD=26.8) had a significantly higher attack rate than RTE outbreaks (mean=40.9, SD=26.8). Surface water outbreaks (mean=69.8, SD=11.4) had a significantly higher attack rate than tap water outbreaks (mean=26.9, SD=15.8), on average. Foodservice outbreaks (mean=55.1, SD=25.7) had a significantly higher attack rate than either leisure setting outbreaks (mean=40.6, SD=22.1) or hospital/nursing setting outbreaks (mean=30.4, SD=16.2). Neither season nor hemisphere was significantly associated with attack rate. In summary, attack rate varied by transmission, food vehicle, water vehicle, and setting, but not by season or hemisphere. For genogroup distribution, there were associations between genogroup distribution with foodborne and waterborne transmission, and between genogroup distribution with shellfish and RTE foods ($p < 0.0001$) (data not shown). There was also a significant association between genogroup and season ($p=0.0114$). However, post hoc analysis could not be performed as both genogroup and season have more than two categories, and the multiple comparisons test necessitates that one variable have only two categories. There was no significant association between GII.4 outbreak strain and any of the outbreak characteristics. In conclusion, we observed significant associations for attack rate with transmission, vehicles, and setting, and for genogroup with transmission, food vehicle, and season.

Multivariate Analysis

Transmission

To determine what relationships between transmission and the outcomes of attack rate, genogroup distribution, and strain were significant upon controlling for other outbreak characteristics, multivariate methods were employed for each outcome (Table 2). 432 outbreaks with foodborne (n=352), waterborne (n=69), or environmental (n=11) transmission were eligible for inclusion in models of attack rate, genogroup, and strain. Transmission was not associated with attack rate or GII.4 strain, but was associated with genogroup distribution. Season and GII.4 strain could be eliminated without confounding the relationship between transmission and attack rate. Fall outbreaks had 9.44 fewer cases per persons at risk than winter season outbreaks (SE=4.55). Southern hemisphere outbreaks had 14.89 fewer cases per persons at risk than Northern hemisphere outbreaks (SE=6.43). Fewer cases per persons at risk were observed with leisure (18.88, SE=4.40) and hospital/nursing (25.63, SE=10.18) than foodservice setting outbreaks. Genogroup distribution did significantly differ between waterborne and foodborne outbreaks. Specifically, waterborne outbreaks, compared to foodborne outbreaks, were more likely associated with GI strains over GII strains, and both GI and GII strains over GII strains only. Hemisphere could be eliminated without confounding the relationship between transmission and genogroup distribution. Spring and fall outbreaks, compared to winter outbreaks, were more likely associated with GI strains over GII strains. GII.4 strain did not significantly vary by transmission. Season and hemisphere could be eliminated without confounding the relationship between transmission and GII.4 strain. Leisure and hospital/nursing setting outbreaks, compared to foodservice setting outbreaks, were more likely associated with GII.4 strains over non-GII.4

strains. In summary, among all of the outcomes, transmission was only associated with genogroup distribution.

Food Vehicles

To determine what relationships between food vehicles and the outcomes of attack rate, genogroup distribution, and strain were significant upon controlling for other outbreak characteristics, multivariate methods were employed for each outcome (Table 3). 206 outbreaks with produce (n=28), shellfish (n=133), or RTE (n=45) food vehicles were eligible for inclusion in models of attack rate and genogroup. Food vehicle was not associated with attack rate, but was associated with genogroup distribution. Season could be eliminated without confounding the relationship between food vehicles and attack rate. Southern hemisphere outbreaks had 39.8 fewer cases per persons at risk than Northern hemisphere outbreaks (SE=9.40). Genogroup distribution did significantly differ between shellfish outbreaks and produce outbreaks. Specifically, shellfish outbreaks, compared to produce outbreaks, were more likely associated with both GI and GII strains over GII strains only and GII strains only over GI strains. Southern, compared to Northern, hemisphere outbreaks were more likely due to GII over multiple strains. Outbreak characteristics did not significantly predict strain (data not shown.) In summary, among attack rate and genogroup outcomes, food vehicle was only associated with genogroup distribution.

Water Vehicles

To determine what relationships between water vehicles and the outcomes of attack rate, genogroup, and strain were significant upon controlling for other outbreak characteristics,

multivariate methods were employed for each outcome (Table 4). 60 outbreaks with tap (n=24), ground (n=26), surface (n=6), or recreation (n=4) water vehicles were eligible for inclusion in models of attack rate and strain (Table 4). Water vehicle was associated with attack rate, but not with GII.4 strain. Surface water outbreaks had 41.97 more cases per persons at risk than tap water outbreaks (SE=12.13). Outbreak characteristics did not significantly predict genogroup (data not shown.) Surface and recreation water outbreaks could not be included in the model for GII.4 strain due to sparse data that prevented model stability. GII.4 outbreak strain did not significantly differ by water vehicle or by other outbreak characteristics for ground water and tap water outbreaks. In summary, among attack rate and strain outcomes, water vehicle was only associated with attack rate.

Discussion

The goal of this study was to assess the association between outbreak transmission and vehicles with attack rates, genogroup distribution, and GII.4 strain distribution. We observed that attack rate did not vary by transmission or food vehicles, upon controlling for other outbreak characteristics, but it did vary by water vehicle. In contrast, genogroup distribution did significantly vary by transmission and food vehicles upon controlling for other outbreak characteristics, but it did not vary by water vehicle. GII.4 strain distribution did not vary by transmission, food vehicles, or water vehicles. We also observed other significant associations between outbreak characteristics (e.g. setting, season, and hemisphere) and outbreak outcomes (attack rates, genogroup distribution, and GII.4 strain distribution).

In general, after controlling for confounding variables, attack rate was not associated with transmission and food vehicle, but was associated with water vehicle. These results seem to contradict the findings in the literature that indicate associations between attack rate and transmission or food vehicles.(reviewed in 26) Because we did observe associations between attack rate and transmission or food vehicles in the bivariate analyses (Table 2), the apparent contradiction may be explained by the role of confounding variables. Upon controlling for confounding variables, the associations for attack rate with transmission and food vehicle did not hold. This observation indicated that the apparent associations result from relationships between attack rate, transmission, and food vehicle with other variables, such as setting, rather than with each other. The observation that controlling for other outbreak characteristics eliminated these significant relationships is a novel finding. An additional hypothesis to explain this finding is that stratifying by several variables resulted in insufficient power to detect differences in attack rate for transmission mechanisms and food vehicles. However, in water vehicles, which had fewer outbreaks to model than transmission mechanisms or food vehicles, we did detect differences in

attack rate for surface water outbreaks and tap water outbreaks. This indicated that the models for transmission and food vehicles should have sufficient power to detect between group differences. Another hypothesis that may explain the lack of relationships between attack rate and transmission or food vehicles, once controlling for confounding variables, is the pathogenicity of norovirus. The infectious dose of norovirus is low, and the low infectious dose may enable a large opportunity for infection through any transmission mechanisms or upon contact with nearly any food vehicle that is contaminated. We also observed that attack rate was positively associated with surface water vehicles compared to tap water vehicles, a finding that was not in the existing published literature. One would expect a greater attack rate for intentional ingestion of tap water by nearly all exposed than accidental ingestion of surface waters by a subset of those exposed. This difference in attack rate may result from an actual epidemiological difference in the number of cases or persons at risk, or from a methodological bias in reporting. One epidemiological hypothesis is that surface waters may have higher levels of norovirus contamination, which could contribute to an increased attack rate by providing increased opportunities for exposed individuals to ingest viral particles. In support of this hypothesis, researchers have previously observed a greater prevalence of norovirus contamination for surface waters than for ground water in Slovenia.(91) One methodological hypothesis may be that the observed difference may result from the challenge of identifying persons exposed to contaminated surface water, leading to a decreased attack rate denominator. In instances where it was difficult to identify all individuals exposed to a contaminated source, we would expect the reported number of persons at risk to be lower than the actual number of those exposed. We did observe a significantly lower number of persons at risk for surface water outbreaks than for tap water outbreaks, supporting this hypothesis (data not shown). Because the persons at risk varied between surface and tap water outbreaks for our data, we would lend more weight to the hypothesis that attack rate was higher for surface water due to difficulties identifying all persons at risk. However, additional research

should address the hypothesis that higher levels of contamination in surface water may lead to an increased attack rate.

Genogroup distribution was significantly associated with transmission routes and food vehicles, but not for water vehicles. The different genogroup profiles likely represent the varying stability of strains in different media and different contamination methods. The increased likelihood of GI only strains and both GI and GII strains over GII only strains among waterborne compared to foodborne transmission may be due to an increased stability of GI strains in water than GII strains.(46, 63) In contrast, the genogroup distribution for shellfish and produce vehicles may be due to contamination methods. There was an increased likelihood of GII strains only over GI strains only, and an increased likelihood of both GI and GII strains over GII strains only, for shellfish compared to produce. The increased likelihood of GII strains over GI strains among shellfish compared to produce may be due to the high prevalence of GII strains circulating among individuals,(45)(reviewed in 26) as shellfish harvest waters are more likely to become contaminated by sewage, while produce is more likely to become contaminated by an individual (e.g. ill farm worker). The observation of increased likelihood of both GI and GII strains over GII strains for shellfish compared to produce may be due to the increased likelihood of shellfish to become contaminated by sewage with several strains of norovirus than by a single ill individual that spreads only a single strain or a few norovirus strains.(58, 62)

GII.4 strain distribution does not appear to vary by transmission or vehicle. Both bivariate and multivariate analysis indicate that GII.4 strain distribution did not vary according to transmission, water vehicle, or food vehicle, although multivariate analysis was not possible for food vehicles. It is interesting that, while genogroup distribution varied to some extent by transmission or vehicle, the presence or absence of GII.4 strains in particular did not. GII.4 strains have been widely implicated in person-to-person outbreaks,(92) but may not represent an

important outcome for foodborne, waterborne, and environmental outbreaks. This hypothesis is supported by Zheng *et al.*'s observation that GII.4 strains tend to predominate in settings with person-to-person transmission, while non-GII.4 strains were associated with outbreaks in settings with foodborne and environmental transmission.(92)

In addition to transmission and vehicles, these data indicate that other outbreak characteristics have important relationships with outbreak outcomes. The role of setting and season appears to be important for the relationship between transmission or vehicle and outbreak outcome. When generating reduced, efficient models, in all instances where setting could have been included in the model, it remained in the model and season dropped out. In all instances where setting was not included in the model, season remained in the model. Inclusion of both season and setting in the model did not result in collinearity assumption violations. Season is likely associated with setting. For example, we observed a significant overall association between season and setting, with a significantly higher percentage of leisure setting outbreaks in the summer months than foodservice setting outbreaks (data not shown.) Unfortunately, data were too sparse to formally evaluate the interaction between season and setting. Hemisphere was also important for some of the relationships between transmission or vehicles and outbreak outcomes (attack rates, genogroup distribution, and GII.4 strain distribution). Attack rate varied by hemisphere for transmission and food vehicle models, and genogroup distribution varied by hemisphere for the food vehicle model. There is evidence in the literature that outbreaks may vary by hemisphere, as outbreaks occur more frequently in the cooler months in the Northern hemisphere, and more frequently in the warmer months in the Southern hemisphere.(75)(reviewed in 20) The observed role of hemisphere could also reflect differences in reporting between hemispheres, as more than 90% of reported outbreaks occur in the Northern hemisphere.(83) The potential for differences by hemisphere further demonstrates the need to control for covariates when analyzing norovirus outbreak trends.

Limitations and Strengths

These data provide a comprehensive approach to characterizing published foodborne, waterborne, and environmental outbreaks. However, as these data are from published norovirus outbreak reports, they are subject to reporting bias and publication bias. Our data are dependent on passive and surveillance efforts, and may under represent regions with limited surveillance capacity. This restricts the potential for extrapolation to norovirus outbreaks in areas with less reporting or decreased genotypic capabilities, such as developing countries. These data represent only foodborne, waterborne, or environmental outbreaks, and do not include person-to-person outbreaks, which comprise the majority of norovirus outbreaks.(92) Attack rate, genogroup, or strain may have different relationships with other outbreak characteristics than for person-to-person outbreaks, and exclusion of person-to-person outbreaks limits the applicability of our findings to a large portion of norovirus outbreaks. However, we feel that vehicle-associated outbreaks remain understudied, and these data provide novel insight into the complex interactions of multiple outbreak characteristics for vehicle-associated outbreaks.

This data set and the analysis techniques provided a thorough approach to characterization of norovirus outbreaks by transmission and commonly implicated vehicles. This is the largest meta-analysis of worldwide norovirus outbreaks, and is a clear strength of our study. The data included in our meta-analysis represented a greater geographic region than previous meta-analysis efforts with norovirus outbreaks.(82) The size of this data set provided the ability to examine vehicle-specific data with multivariate techniques. The multivariate approach to characterize norovirus outbreak trends is another strength of this study. Our bivariate findings were consistent with the existing literature, supporting the validity of our study.(92)(reviewed in 26) However, upon adjustment for confounding variables, such as setting and season, some relationships between outbreak outcomes and transmission or vehicle changed. Because of the interrelatedness of outbreak characteristics, we hypothesize that the adjusted relationships were

more valid than the unadjusted relationships. The adjusted results offer both novel findings and support for multivariate approaches. Additionally, the presence of significant predictors in our multivariate models suggests sufficient sensitivity to capture large differences between groups, despite stratification across several variables.

Implications

These data can be employed to better understand the underlying relationships for transmission mode and vehicles with outbreak outcomes that are confounded by other outbreak characteristics. Although attack rate has important clinical implications for outbreaks, public health practitioners may not want to factor attack rate into arguments for or against a particular transmission mechanism or vehicle during an outbreak investigation, as it did not vary by transmission or food vehicle, and appeared to be driven instead by setting, season, hemisphere, and genotype. On the other hand, genogroup distribution may strengthen the case against a particular transmission type or food vehicle because it did vary by transmission and food vehicles. GII.4 norovirus strains predominate in norovirus outbreaks; however these data indicate that they may not be an important outcome for foodborne, waterborne, and environmental outbreaks, as they did not vary by transmission, food vehicle, or water vehicle. As a result, GII.4 strain distribution may not be a useful outcome for implicating transmission mechanisms or vehicles during an outbreak investigation. Based on study findings, it is recommended that future investigative efforts consider the interrelationships of outbreak characteristics and utilize multivariate techniques when possible.

Conclusions

Meta-analysis of published norovirus outbreaks with multivariate techniques enabled a unique opportunity to assess the relationships between transmission or vehicle and outbreak outcomes such as attack rate, genogroup, or strain. Attack rate did not vary by transmission or food vehicles, upon controlling for other outbreak characteristics, but did vary by water vehicle. In contrast, genogroup distribution did significantly vary by transmission and food vehicles upon controlling for other outbreak characteristics, and it did vary by water vehicle. GII.4 strain distribution did not vary by transmission, food vehicles, or water vehicles. Based on our results, attack rate may be useful for implicating water vehicles, and genogroup distribution may be useful for implicating transmission mechanisms or food vehicles. However, GII.4 strain distribution may not be useful for implicating transmission mechanisms or vehicles during an outbreak investigation. As many variables can impact norovirus outbreak outcomes, these data highlight the importance of controlling for potential confounders when examining the relationships between outbreak characteristics and outcomes.

References

1. Fankhauser RL, Monroe SS, Noel JS, Humphrey CD, Bresee JS, Parashar UD, et al. Epidemiologic and molecular trends of "Norwalk-like viruses" associated with outbreaks of gastroenteritis in the United States. J Infect Dis 2002;186(1):1-7.

2. Goodgame R. Norovirus gastroenteritis. Curr Gastroenterol Rep 2006;8(5):401-8.

3. Atmar RL, Estes MK. The epidemiologic and clinical importance of norovirus infection. Gastroenterol Clin North Am 2006;35(2):275-90, viii.

4. Koopmans M, Duizer E. Foodborne viruses: an emerging problem. Int J Food Microbiol 2004;90(1):23-41.

5. Koo HL, Ajami N, Atmar RL, DuPont HL. Noroviruses: The leading cause of gastroenteritis worldwide. Discov Med 2010;10(50):61-70.

6. Black RE, Greenberg HB, Kapikian AZ, Brown KH, Becker S. Acquisition of serum antibody to Norwalk Virus and rotavirus and relation to diarrhea in a longitudinal study of young children in rural Bangladesh. J Infect Dis 1982;145(4):483-9.

7. Greenberg HB, Valdesuso J, Kapikian AZ, Chanock RM, Wyatt RG, Szmuness W, et al. Prevalence of antibody to the Norwalk virus in various countries. Infect Immun 1979;26(1):270-3.

8. Leon J, Moe C, Potter M. Role of viruses in foodborne disease. In: Food consumption and disease risk: consumer-pathogen interactions (Potter M. ed.). Baltimore: Woodhead Publishing in Food Science, Technology and Nutrition; 2006. p. 309-342.

9. Lopman B, Vennema H, Kohli E, Pothier P, Sanchez A, Negredo A, et al. Increase in viral gastroenteritis outbreaks in Europe and epidemic spread of new norovirus variant. Lancet 2004;363(9410):682-8.

10. Koopmans M. Progress in understanding norovirus epidemiology. Curr Opin Infect Dis 2008;21(5):544-52.

11. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, et al. Foodborne illness acquired in the United States--major pathogens. Emerg Infect Dis 2011;17(1):7-15.

12. Hutson AM, Atmar RL, Estes MK. Norovirus disease: changing epidemiology and host susceptibility factors. Trends Microbiol 2004;12(6):279-87.

13. Zahorsky J. Hyperemesis hiemis or the winter vomiting disease. Arch Pediatr 1929;46:391–5.

14. Adler JL, Zickl R. Winter vomiting disease. J Infect Dis 1969;119(6):668-73.

15. Kapikian AZ, Wyatt RG, Dolin R, Thornhill TS, Kalica AR, Chanock RM. Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. J Virol 1972;10(5):1075-81.

16. Ando T, Monroe SS, Gentsch JR, Jin Q, Lewis DC, Glass RI. Detection and differentiation of antigenically distinct small round-structured viruses (Norwalk-like viruses) by reverse transcription-PCR and southern hybridization. J Clin Microbiol 1995;33(1):64-71.

17. Lindesmith L, Moe C, Marionneau S, Ruvoen N, Jiang X, Lindblad L, et al. Human susceptibility and resistance to Norwalk virus infection. Nat Med 2003;9(5):548-53.

18. Todd EC, Greig JD, Bartleson CA, Michaels BS. Outbreaks where food workers have been implicated in the spread of foodborne disease. Part 4. Infective doses and pathogen carriage. J Food Prot 2008;71(11):2339-73.

19. Dolin R, Blacklow NR, DuPont H, Formal S, Buscho RF, Kasel JA, et al. Transmission of acute infectious nonbacterial gastroenteritis to volunteers by oral administration of stool filtrates. J Infect Dis 1971;123(3):307-12.

20. Thornton AC, Jennings-Conklin KS, McCormick MI. Noroviruses: agents in outbreaks of acute gastroenteritis. Disaster Manag Response 2004;2(1):4-9.

21. Nelson KE, Williams C. Infectious Disease Epidemiology: Theory and Practice. Second Edition ed: Jones & Bartlett Learning; 2007.

22. Leon JS, Souza M, Wang Q, Smith ER, Saif LJ, Moe CL. Immunology of norovirus infection (Vajdy M. ed.). In: Immunity Against Mucosal Pathogens. Boston: Springer Science; 2008. p. 219-262.

23. Ausar SF, Foubert TR, Hudson MH, Vedvick TS, Middaugh CR. Conformational stability and disassembly of Norwalk virus-like particles. Effect of pH and temperature. J Biol Chem 2006;281(28):19478-88.

24. Teunis PF, Moe CL, Liu P, Miller SE, Lindesmith L, Baric RS, et al. Norwalk virus: how infectious is it? J Med Virol 2008;80(8):1468-76.

25. Moe CL, et al. Abstract. In: International Workshop on Human Caliciviruses. Atlanta, GA; 1999. p. 4-6.

26. Mattison K. Norovirus as a foodborne disease hazard. Adv Food Nutr Res 2011;62:1-39.

27. Le Guyader F, Loisy F, Atmar RL, Hutson AM, Estes MK, Ruvoen-Clouet N, et al. Norwalk virus-specific binding to oyster digestive tissues. Emerg Infect Dis 2006;12(6):931-6.

28. Huppatz C, Munnoch SA, Worgan T, Merritt TD, Dalton C, Kelly PM, et al. A norovirus outbreak associated with consumption of NSW oysters: implications for quality assurance systems. Commun Dis Intell 2008;32(1):88-91.

29. Berg DE, Kohn MA, Farley TA, McFarland LM. Multi-state outbreaks of acute gastroenteritis traced to fecal-contaminated oysters harvested in Louisiana. J Infect Dis 2000;181 Suppl 2:S381-6.

30. Kohn MA, Farley TA, Ando T, Curtis M, Wilson SA, Jin Q, et al. An outbreak of Norwalk virus gastroenteritis associated with eating raw oysters. Implications for maintaining safe oyster beds. JAMA 1995;273(6):466-71.

31. Kirkland KB, Meriwether RA, Leiss JK, Mac Kenzie WR. Steaming oysters does not prevent Norwalk-like gastroenteritis. Public Health Rep 1996;111(6):527-30.

32. Newell DG, Koopmans M, Verhoef L, Duizer E, Aidara-Kane A, Sprong H, et al. Food-borne diseases - the challenges of 20 years ago still persist while new ones continue to emerge. Int J Food Microbiol 2010;139 Suppl 1:S3-15.

33. Ailes EC, Leon JS, Jaykus LA, Johnston LM, Clayton HA, Blanding S, et al. Microbial concentrations on fresh produce are affected by postharvest processing, importation, and season. J Food Prot 2008;71(12):2389-97.

34. U.S. Public Health Service F. Food Code. In: Services UsDoHaH, editor. College Park, MD; 2009. p. 19.

35. Daniels NA, Bergmire-Sweat DA, Schwab KJ, Hendricks KA, Reddy S, Rowe SM, et al. A foodborne outbreak of gastroenteritis associated with Norwalk-like viruses: first molecular traceback to deli sandwiches contaminated during preparation. J Infect Dis 2000;181(4):1467-70.

36. Beuret C, Kohler D, Luthi T. Norwalk-like virus sequences detected by reverse transcription-polymerase chain reaction in mineral waters imported into or bottled in Switzerland. J Food Prot 2000;63(11):1576-82.

37. Ho MS, Glass RI, Monroe SS, Madore HP, Stine S, Pinsky PF, et al. Viral gastroenteritis aboard a cruise ship. Lancet 1989;2(8669):961-5.

38. Wu HM, Fornek M, Schwab KJ, Chapin AR, Gibson K, Schwab E, et al. A norovirus outbreak at a long-term-care facility: the role of environmental surface contamination. Infect Control Hosp Epidemiol 2005;26(10):802-10.

39. Graham DY, Jiang X, Tanaka T, Opekun AR, Madore HP, Estes MK. Norwalk virus infection of volunteers: new insights based on improved assays. J Infect Dis 1994;170(1):34-43.

40. Harris JP, Lopman BA, O'Brien SJ. Infection control measures for norovirus: a systematic review of outbreaks in semi-enclosed settings. J Hosp Infect 2010;74(1):1-9.

41. Zheng DP, Ando T, Fankhauser RL, Beard RS, Glass RI, Monroe SS. Norovirus classification and proposed strain nomenclature. Virology 2006;346(2):312-23.

42. Marshall JA, Bruggink LD. The dynamics of norovirus outbreak epidemics: recent insights. Int J Environ Res Public Health 2011;8(4):1141-9.

43. Phan TG, Kaneshi K, Ueda Y, Nakaya S, Nishimura S, Yamamoto A, et al. Genetic heterogeneity, evolution, and recombination in noroviruses. J Med Virol 2007;79(9):1388-400.

44. Patel MM, Hall AJ, Vinje J, Parashar UD. Noroviruses: a comprehensive review. J Clin Virol 2009;44(1):1-8.

45. Bull RA, Tu ET, McIver CJ, Rawlinson WD, White PA. Emergence of a new norovirus genotype II.4 variant associated with global outbreaks of gastroenteritis. J Clin Microbiol 2006;44(2):327-33.

46. Lysen M, Thorhagen M, Brytting M, Hjertqvist M, Andersson Y, Hedlund KO. Genetic diversity among food-borne and waterborne norovirus strains causing outbreaks in Sweden. J Clin Microbiol 2009;47(8):2411-8.

47. Kageyama T, Shinohara M, Uchida K, Fukushi S, Hoshino FB, Kojima S, et al. Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to Norovirus in Japan. J Clin Microbiol 2004;42(7):2988-95.

48. Bok K, Abente EJ, Realpe-Quintero M, Mitra T, Sosnovtsev SV, Kapikian AZ, et al. Evolutionary dynamics of GII.4 noroviruses over a 34-year period. J Virol 2009;83(22):11890-901.

49. Buesa J, Montava R, Abu-Mallouh R, Fos M, Ribes JM, Bartolome R, et al. Sequential evolution of genotype GII.4 norovirus variants causing gastroenteritis outbreaks from 2001 to 2006 in Eastern Spain. J Med Virol 2008;80(7):1288-95.

50. Bull RA, Eden JS, Rawlinson WD, White PA. Rapid evolution of pandemic noroviruses of the GII.4 lineage. PLoS Pathog 2010;6(3):e1000831.

51. Bull RA, White PA. Mechanisms of GII.4 norovirus evolution. Trends Microbiol 2011;19(5):233-40.

52. Reuter G, Pankovics P, Szucs G. Genetic drift of norovirus genotype GII-4 in seven consecutive epidemic seasons in Hungary. J Clin Virol 2008;42(2):135-40.

53. Lopman BA, Reacher M, Gallimore C, Adak GK, Gray JJ, Brown DW. A summertime peak of "winter vomiting disease": surveillance of noroviruses in England and Wales, 1995 to 2002. BMC Public Health 2003;3:13.

54. Gallimore CI, Iturriza-Gomara M, Xerry J, Adigwe J, Gray JJ. Inter-seasonal diversity of norovirus genotypes: emergence and selection of virus variants. Arch Virol 2007;152(7):1295-303.

55. Friesema IH, Vennema H, Heijne JC, de Jager CM, Teunis PF, van der Linde R, et al. Differences in clinical presentation between norovirus genotypes in nursing homes. J Clin Virol 2009;46(4):341-4.

56. Rosenthal NA, Lee LE, Vermeulen BA, Hedberg K, Keene WE, Widdowson MA, et al. Epidemiological and genetic characteristics of norovirus outbreaks in long-term care facilities, 2003-2006. Epidemiol Infect 2011;139(2):286-94.

57. Bailey D, Karakasiliotis I, Vashist S, Chung LM, Rees J, McFadden N, et al. Functional analysis of RNA structures present at the 3' extremity of the murine norovirus genome: the variable polypyrimidine tract plays a role in viral virulence. J Virol 2010;84(6):2859-70.

58. ter Waarbeek HL, Dukers-Muijrers NH, Vennema H, Hoebe CJ. Waterborne gastroenteritis outbreak at a scouting camp caused by two norovirus genogroups: GI and GII. J Clin Virol 2010;47(3):268-72.

59. Gallimore CI, Pipkin C, Shrimpton H, Green AD, Pickford Y, McCartney C, et al. Detection of multiple enteric virus strains within a foodborne outbreak of gastroenteritis: an indication of the source of contamination. Epidemiol Infect 2005;133(1):41-7.

60. Verhoef L, Vennema H, van Pelt W, Lees D, Boshuizen H, Henshilwood K, et al. Use of norovirus genotype profiles to differentiate origins of foodborne outbreaks. Emerg Infect Dis 2010;16(4):617-24.

61. Verhoef LP, Kroneman A, van Duynhoven Y, Boshuizen H, van Pelt W, Koopmans M. Selection tool for foodborne norovirus outbreaks. Emerg Infect Dis 2009;15(1):31-8.

62. Gallimore CI, Cheesbrough JS, Lamden K, Bingham C, Gray JJ. Multiple norovirus genotypes characterised from an oyster-associated outbreak of gastroenteritis. Int J Food Microbiol 2005;103(3):323-30.

63. Maunula L, Miettinen IT, Von Bonsdorff CH. Norovirus outbreaks from drinking water. Emerging infectious diseases 2005;11(11):1716-1721.

64. Riera-Montes M, Brus Sjolander K, Allestam G, Hallin E, Hedlund KO, Lofdahl M. Waterborne norovirus outbreak in a municipal drinking-water supply in Sweden. Epidemiol Infect 2011:1-8.

65. Maalouf H, Schaeffer J, Parnaudeau S, Le Pendu J, Atmar RL, Crawford SE, et al. Strain-dependent norovirus bioaccumulation in oysters. Appl Environ Microbiol 2011;77(10):3189-96.

66. Greer AL, Drews SJ, Fisman DN. Why "winter" vomiting disease? Seasonality, hydrology, and Norovirus epidemiology in Toronto, Canada. Ecohealth 2009;6(2):192-9.

67. Kvitsand HM, Fiksdal L. Waterborne disease in Norway: emphasizing outbreaks in groundwater systems. Water Sci Technol 2010;61(3):563-71.

68. Boxman IL, Verhoef L, Dijkman R, Hagele G, Te Loeke NA, Koopmans M. Year-round prevalence of norovirus in the environment of catering companies without a recently reported outbreak of gastroenteritis. Appl Environ Microbiol 2011;77(9):2968-74.

69. Mounts AW, Ando T, Koopmans M, Bresee JS, Noel J, Glass RI. Cold weather seasonality of gastroenteritis associated with Norwalk-like viruses. J Infect Dis 2000;181 Suppl 2:S284-7.

70. Iritani N, Kaida A, Kubo H, Abe N, Goto K, Ogura H, et al. Molecular epidemiology of noroviruses detected in seasonal outbreaks of acute nonbacterial gastroenteritis in Osaka City, Japan, from 1996-1997 to 2008-2009. J Med Virol 2010;82(12):2097-105.

71. Miyoshi T, Uchino K, Matsuo M, Ikeda Y, Yoshida H, Sibata H, et al. Characteristics of Norovirus outbreaks during a non-epidemic season. Jpn J Infect Dis 2006;59(2):140-1.

72. Morillo SG, Luchs A, Cilli A, Ribeiro CD, Calux SJ, Carmona Rde C, et al. Large gastroenteritis outbreak due to norovirus GII in Sao Paulo, Brazil, summer 2010. Rev Inst Med Trop Sao Paulo 2011;53(2):119-20.

73. Kaplan JE, Gary GW, Baron RC, Singh N, Schonberger LB, Feldman R, et al. Epidemiology of Norwalk gastroenteritis and the role of Norwalk virus in outbreaks of acute nonbacterial gastroenteritis. Ann Intern Med 1982;96(6 Pt 1):756-61.

74. Marshall JA, Dimitriadis A, Wright PJ. Molecular and epidemiological features of norovirus-associated gastroenteritis outbreaks in Victoria, Australia in 2001. J Med Virol 2005;75(2):321-31.

75. Marshall JA, Hellard ME, Sinclair MI, Fairley CK, Cox BJ, Catton MG, et al. Incidence and characteristics of endemic Norwalk-like virus-associated gastroenteritis. J Med Virol 2003;69(4):568-78.

76. Verhoef L, Depoortere E, Boxman I, Duizer E, van Duynhoven Y, Harris J, et al. Emergence of new norovirus variants on spring cruise ships and prediction of winter epidemics. Emerg Infect Dis 2008;14(2):238-43.

77. Lopman BA, Adak GK, Reacher MH, Brown DW. Two epidemiologic patterns of norovirus outbreaks: surveillance in England and wales, 1992-2000. Emerg Infect Dis 2003;9(1):71-7.

78. Bruggink L, Marshall J. The relationship between health care and nonhealth care norovirus outbreak settings and norovirus genotype in Victoria, Australia, 2002-2005. J Microbiol Immunol Infect 2011;44(4):241-6.

79. Morter S, Bennet G, Fish J, Richards J, Allen DJ, Nawaz S, et al. Norovirus in the hospital setting: virus introduction and spread within the hospital environment. J Hosp Infect 2011;77(2):106-12.

80. Boxman IL, Dijkman R, te Loeke NA, Hagele G, Tilburg JJ, Vennema H, et al. Environmental swabs as a tool in norovirus outbreak investigation, including outbreaks on cruise ships. J Food Prot 2009;72(1):111-9.

81. Chatterjee NK, Moore DW, Monroe SS, Glass RI, Cambridge MJ, Kondracki SF, et al. Molecular epidemiology of outbreaks of viral gastroenteritis in New York State, 1998-1999. Clin Infect Dis 2004;38 Suppl 3:S303-10.

82. Kroneman A, Verhoef L, Harris J, Vennema H, Duizer E, van Duynhoven Y, et al. Analysis of integrated virological and epidemiological reports of norovirus outbreaks collected within the Foodborne Viruses in Europe network from 1 July 2001 to 30 June 2006. J Clin Microbiol 2008;46(9):2959-65.

83. Matthews J, Dickey B, Miller R, Felzer J, Dawson B, Lee A, et al. The epidemiology of published norovirus outbreaks: a review of risk factors associated with attack rate and genogroup. Epidemiology and Infection 2012;1(1):1-12.

84. Elliott AC, Reisch JS. Implementing a multiple comparison test for proportions in a 2xC crosstabulation in SAS¬Æ. Proceedings of the SAS User's Group International 31, San Francisco, CA, March 26‚Äì29, 2006. Paper 204 2006;31.

85. Kleinbaum K, Nizam M, Education TH. Applied Regression Analysis and Other Multivariable Methods, 4e. 2008.

86. Kleinbaum DG, Klein M. Logistic regression: A self learning text, 3e. In: New York: Springer‚ÄìVerlag Inc; 1994.

87. Firth D. Bias reduction of maximum likelihood estimates. Biometrika 1993;80(1):27-38.

88. Heinze G. The application of Firth‚Äôs procedure to Cox and logistic regression: Technical Report 10/1999, update in January 2001, Section of Clinical Biometrics, Department of Medical Computer Sciences, University of Vienna; 1999.

89. Heinze G, Schemper M. A solution to the problem of separation in logistic regression. Statistics in medicine 2002;21(16):2409-2419.

90. Allison PD. Convergence failures in logistic regression. In; 2008: Citeseer; 2008. p. 2008.

91. Steyer A, Torkar KG, Gutierrez-Aguirre I, Poljsak-Prijatelj M. High prevalence of enteric viruses in untreated individual drinking water sources and surface water in Slovenia. Int J Hyg Environ Health 2011;214(5):392-8.

92. Zheng DP, Widdowson MA, Glass RI, Vinje J. Molecular epidemiology of genogroup II-genotype 4 noroviruses in the United States between 1994 and 2006. J Clin Microbiol 2010;48(1):168-77.

Tables

Table 1: Bivariate Relationships between Outbreak Characteristics and Outcomes of Attack Rate, Genogroup, and GII.4 Strain

	Attack Rate			Genogroup						Strain				
				GII		GI		GII and GI		GII.4		non GII.4		
Variable	Mean	SD	p value**	n	% †	n	% †	n	% †	p value**	n	% †	n	% †	p value**
Transmission			0.0214*							◦					0.2535
Foodborne	52.7	27.4		206	85.1	41	69.5	63	77.8		84	78.5	268	82.5	
Waterborne	44.1	23.5		28	11.6	18	30.5	18	22.2		18	16.8	51	15.7	
Environmental	32.2	24.0		8	3.3	0	0.0	0	0.0		5	4.7	6	1.8	
Food Vehicles			0.0007*							◦					0.5827
Produce	48.0	21.2		17	18.1	6	26.1	1	1.7		3	8.8	25	14.5	
Shellfish	59.9	26.8	φ	48	51.1	9	39.1	58	98.3		22	64.7	111	64.5	
Ready to Eat	40.9	25.5	φ	29	30.9	8	34.8	0	0.0		9	26.5	36	20.9	
Water Vehicles			0.0090*							◦					◦
Tap	26.9	15.8	φ	6	31.6	6	42.9	8	47.1		8	50.0	16	36.4	
Ground	45.6	23.4		10	52.6	7	50.0	7	41.2		5	31.3	21	47.7	
Surface	69.8	11.4	φ	1	5.3	0	0.0	2	11.8		2	12.5	4	9.1	
Recreation	34.3	11.0		2	10.5	1	7.1	0	0.0		1	6.3	3	6.8	
Setting			<.0001*							◦					◦
Foodservice	55.1	25.7	φ	135	68.9	31	67.4	42	79.2		48	58.5	189	72.4	
Leisure	40.6	22.1	φa	40	20.4	13	28.3	10	18.9		23	28.0	56	21.5	
School/Daycare	39.9	21.6		9	4.6	2	4.3	1	1.9		3	3.7	12	4.6	
Hospital/Nursing	30.4	16.2	φb	12	6.1	0	0.0	0	0.0		8	9.8	4	1.5	
Season			0.2030							0.0114					0.6068
Winter	54.1	28.4		102	44.5	16	28.1	44	55.0		45	44.1	132	41.9	
Spring	51.2	26.3		49	21.4	12	21.1	19	23.8		19	18.6	77	24.4	
Summer	47.3	27.1		42	18.3	12	21.1	10	12.5		17	16.7	53	16.8	
Fall	45.2	25.0		36	15.7	17	29.8	7	8.8		21	20.6	53	16.8	
Hemisphere			0.1732							0.8811					◦
Northern	51.4	27.0		225	95.3	57	96.6	77	96.3		105	99.1	308	94.2	
Southern	42.2	26.8		11	4.7	2	3.4	3	3.8		1	0.9	19	5.8	

† Column percentages reported for genogroup and strain

** ANOVA tests for normally distributed continuous variables and Chi Square test for categorical variable assessment

* Significant at α=0.05. Significantly different means and frequencies indicated by φ. φa and φb indicate frequencies different from φ, but not each other.

◦ Chi Square test not valid due to low expected cell values

Table 2: Relationships between Transmission and Outcomes of Attack Rate, Genogroup, and GII.4 Strain, Controlling for Other Characteristics

	Attack Rate (n=232)			Genogroup (n=237) **				GII.4 Strain (n=232) **	
				GI		Both GI and GII			
Variable †	Beta	SE	p value	OR	95% CI	OR	95% CI	OR	95% CI
Transmission									
Foodborne	Reference Category			Reference Category		Reference Category		Reference Category	
Waterborne	9.15	5.90	0.1226	4.25*	1.56-11.58	2.80*	1.10-7.17	0.30	0.09-1.06
Environmental	-5.20	9.45	0.5825	φ	φ	φ	φ	0.64	0.09-4.44
Season									
Winter	Reference Category			Reference Category		Reference Category		Reference Category	
Spring	*1.07*	*4.16*	*0.7981*	3.29*	1.09-9.92	1.34	0.61-2.92	*0.86*	*0.36-2.05*
Summer	*-3.75*	*4.84*	*0.4395*	1.67	0.47-5.86	0.38	0.13-1.17	*1.08*	*0.40-2.91*
Fall	*-9.44**	*4.55*	*0.0394*	6.31*	2.12-18.78	0.77	0.28-2.17	*1.27*	*0.52-3.13*
Hemisphere									
Northern	Reference Category			Reference Category		Reference Category		Reference Category	
Southern	-14.89*	6.43	0.0213	*0.98*	*0.18-5.47*	*0.75*	*0.15-3.84*	*0.23*	*0.03-1.90*
Attack Rate	°	°	°	1.01	0.99-1.02	1.01	1.00-1.02	1.00	0.98-1.01
Setting									
Foodservice	Reference Category			φ	φ	φ	φ	Reference Category	
Leisure	-18.88*	4.40	<0.0001	φ	φ	φ	φ	2.43*	1.06-5.57
School/Daycare	-13.84	7.86	0.0797	φ	φ	φ	φ	0.36	0.04-3.03
Hospital/Nursing	-25.63*	10.18	0.0126	φ	φ	φ	φ	25.45*	2.55-254.03
Strain									
non-GII.4	Reference Category			°	°	°	°	°	°
GII.4	*-3.09*	*4.00*	*0.4406*	°	°	°	°	°	°
Intercept	59.49*	2.91	<0.0001	n/a		n/a		n/a	

† Italic covariates are eliminated by backward elimination (alpha= 0.05) without confounding covariate estimates.

** Reference category for polytomous model is Genogroup GII. Reference category for Strain is non-GII.4.

* Significant Beta, Intercept or OR estimate

° Variables not eligible for inclusion due to relationship with outcome variable

φ Variables or categories not included due to sparse data and model instability

Table 3: Relationships between Food Vehicles and Outcomes of Attack Rate and Genogroup, Controlling for Other Outbreak Characteristics

	Attack Rate (n=91)			Genogroup (n=120) **			
				GI		Both GI and GII	
Variable †	Beta	SE	p value	OR	95% CI	OR	95% CI
Food Vehicles							
Produce	Reference Category			Reference Category		Reference Category	
Shellfish	16.1	9.41	0.0909	0.14*	0.02-0.96	15.14*	1.64-139.86
Ready to Eat	-6.69	8.25	0.4198	0.50	0.12-2.15	0.14§	0.01-3.86
Season							
Winter	Reference Category			Reference Category[δ]		Reference Category[δ]	
Spring	*-0.60*	*5.61*	*0.9158*				
Summer	*3.96*	*8.43*	*0.6402*	1.46	0.39-5.44	1.28	0.47-3.51
Fall	*-6.12*	*9.53*	*0.5232*				
Hemisphere							
Northern	Reference Category			Reference Category		Reference Category	
Southern	-39.8*	9.40	<0.0001	2.30	0.13-39.33	0.16*	0.02-0.997
Attack Rate	°	°	°	1.01	0.99-1.04	0.99	0.97-1.01
Setting							
Foodservice	Reference Category			φ	φ	φ	φ
Leisure	-17.76	9.76	0.0726	φ	φ	φ	φ
School/Daycare	28.45	17.05	0.0990	φ	φ	φ	φ
Hospital/Nursing	-26.67	13.73	0.0556	φ	φ	φ	φ
Genogroup							
GII	Reference Category			°	°	°	°
GI	-0.71	8.72	0.9351	°	°	°	°
Both GI and GII	-4.03	6.53	0.5388	°	°	°	°
Intercept	50.83*	8.48	<0.0001	n/a		n/a	

† Italic covariates are eliminated by backward elimination (alpha= 0.05) without confounding estimates.

** Reference category for polytomous model is Genogroup GII

* Significant Beta, Intercept or OR estimate

° Variables not eligible for inclusion due to relationship with outcome variable

φ Variables or categories not included due to sparse data and model instability

δ Spring, Summer, and Fall are collapsed to obtain valid model estimates.

§ Estimates obtained using logistic model and Firth correction

Table 4: Relationships between Water Vehicles and Outcomes of Attack Rate and GII.4 Strain, Controlling for Other Outbreak Characteristics

	Attack Rate (n=39)			GII.4 Strain (n=31) ɣ	
Variable †	Beta	SE	p value	OR	95% CI
Water Vehicles					
Tap	Reference Category			Reference Category	
Ground	12.88	8.60	0.1442	1.00	0.11-8.79
Surface	41.97**	12.13	0.0016	°	°
Recreational	2.29	13.69	0.8685	°	°
Season					
Winter	Reference Category			Reference Category φ	
Spring	7.84	8.73	0.3764		
Summer	16.67	8.49	0.0588	0.43	0.06-3.15
Fall	8.88	12.96	0.4983		
Attack Rate	*	*	*	0.93	0.87-1.01
Strain					
non-GII.4	Reference Category			*	*
GII.4	-12.74	7.85	0.1149	*	*
Intercept	26.49**	9.06	0.0064	n/a	

† Reference category for strain is non-GII.4
** Significant Beta or Intercept estimate
* Variables not eligible for inclusion due to relationship with outcome variable
° Variables or categories not included due to sparse data and model instability
φ Spring, Summer, and Fall collapsed together to obtain valid model estimates.

Appendix A: IRB Letter of Exemption

EMORY UNIVERSITY | Institutional Review Board

May 23, 2011

Elizabeth Bitler
Department of Epidemiology
Rollins School of Public Health
Emory University

RE: Determination: No IRB Review Required
Norovirus Thesis Data
PI: Elizabeth Bitler

Dear Ms. Bitler:

Thank you for requesting a determination from our office about the above-referenced project. Based on our review of the materials you provided, we have determined that it does not require IRB review because it does not meet the definition of research involving "human subjects" or the definition of "clinical investigation" as set forth in Emory policies and procedures and federal rules.

Specifically, in this project, you will be performing secondary data analysis on a dataset of Norovirus outbreaks previously collected by a Rollins School of Public Health student. The dataset will contain no personally identifiable information or links to identifiable data. As the data you will receive will not have any attached identifiers or links to any identifiable data, there is no interaction with "human subjects" as defined in 45 CFR 46.102(f). As such, this project does not fall under the purview of the IRB.

This determination could be affected by substantive changes in the study design or identifiability of data. If the project changes in any substantive way, please contact our office for clarification.

Thank you for consulting the IRB.

Sincerely,

Sean Kiskel
Research Protocol Analyst
Emory University IRB