MATHEMATICAL MODELING FOR PUBLIC HEALTH DECISION SUPPORT DURING ACUTE GASTROENTERITIS OUTBREAKS

by

Carl James Grafe

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biomedical Informatics

The University of Utah

May 2018

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Carl James Grafe**

has been approved by the following supervisory committee members:

R. Scott Evans	, Chair	**2/26/2018** Date Approved
Catherine Staes	, Member	**2/26/2018** Date Approved
Kensaku Kawamoto	, Member	**2/26/2018** Date Approved
Matthew Samore	, Member	**2/26/2018** Date Approved
Frederick Adler	, Member	**2/26/2018** Date Approved

and by **Wendy Chapman** , Chair/Dean of

the Department/College/School of **Biomedical Informatics**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Acute gastroenteritis (AGE) outbreaks present a significant challenge to investigating public health officials, who need to know whether point source transmission — such as contaminated food, fomites, or highly infectious individuals — has occurred to respond effectively. However, information on the mode of transmission is frequently unavailable, especially during the early stages of an outbreak when control measures have the greatest impact. Clinical decision support systems (CDS) may be used to assist outbreak investigators when only limited data are available. This dissertation research investigated a) how the guidelines for norovirus outbreaks in healthcare settings vary between state public health agencies across the U.S., b) how mathematical modeling can be used to help outbreak investigators identify potential point source outbreaks, and c) how availability of outbreak information impacts public health decision-making.

After introductory material in Chapters 1 and 2, Chapter 3 describes variation in norovirus outbreak guidelines and outcomes between states. Chapter 4 describes the development of a stochastic individual-level mathematical model for predicting whether an outbreak was likely caused by point source transmission. The model's internal and external validity were assessed, and the model was used to estimate potential misclassification in outbreaks reported to have been transmitted person-to-person. Chapter 5 describes semi-structured interviews with AGE epidemiologists about decision-making in hypothetical outbreak scenarios based on different levels of data

availability, including results from a CDS based on the mathematical model.

There was substantial variation in state healthcare-associated norovirus outbreak response guidelines, and there were differences between states with and without guidelines consistent with national guidance. The model performed well on measures of internal and external validity, and 73% of person-to-person norovirus outbreaks had at least some evidence of point source transmission. AGE epidemiologists drew different conclusions when presented with different levels of information, and there was evidence that CDS could help improve decision-making when only minimal data are available. These results demonstrate the need for CDS for transmission mode classification, the effectiveness of mathematical modeling for outbreak response decision support in some circumstances, and the potential for CDS to improve decision-making when data are sparse and public health action is required.

To Chelsea

CONTENTS

LIST OF TABLES

Tables

LIST OF FIGURES

Figures

LIST OF ACRONYMS

AGE	Acute gastroenteritis
AIC	Akaike Information Criterion
CDC	Centers for Disease Control and Prevention
CDS	Clinical Decision Support
CI	Confidence Interval
CIFOR	Council to Improve Foodborne Outbreak Response
EH	Environmental Health
EHA	Environmental Health Assessment
GLMM	Generalized Linear Mixed Effects Model
HICPAC	Healthcare Infection Control Practices Advisory Committee
LHD	Local Health Department
NACCHO	National Association of City and County Health Officials
NORS	National Outbreak Reporting System
OR	Odds Ratio
SARS	Severe Acute Respiratory Syndrome
SEIR	Susceptible-Exposed-Infected-Recovered
SIR	Susceptible-Infected-Recovered
SIS	Susceptible-Infected-Susceptible

ACKNOWLEDGEMENTS

I would like to thank my committee for their ongoing support and persistent commitment of time and energy: Scott Evans, MS, PhD; Catherine Staes, BSN, MPH, PhD; Kensaku Kawamoto, MD, PhD, MHS; Matt Samore, MD; and Fred Adler, PhD.

I'd like to thank my employers at Salt Lake County Health Department, Utah Department of Health, and the Centers for Disease Control and Prevention (CDC), whose understanding and support regarding my education have been essential and are immensely appreciated.

I am grateful for the applied epidemiologists from four local health departments and 25 state health departments whose willing participation and contributions made each phase of this research possible. Support from the NORS team at CDC, including manuscript review by Dr. Aron Hall and Mary Wikswo, was also helpful and appreciated.

Above all, I would like to thank my family, especially my wife, Chelsea, and my children, Geryn, Elsie, Gavin, Dagan, and Beau, for their encouragement and patience throughout this endeavor.

CHAPTER 1

INTRODUCTION

During acute gastroenteritis (AGE) outbreaks, information about the pathogen and route of transmission is essential to implement effective public health control measures to end an outbreak as soon as possible.[1] However, this information is frequently unavailable, especially during the early stages of an outbreak when appropriate control measures have the greatest impact.[2] Most acute gastroenteritis outbreaks are attributable to person-to-person transmission of norovirus,[3] which may require very different control measures than other pathogens or modes of transmission.[4] Mathematical modeling may be used to determine whether the data that are usually available during an outbreak are consistent with person-to-person transmission of norovirus and thence to determine the most appropriate response.

For such a mathematical model to be effective, several issues needed to be addressed. First, the processes whereby acute gastroenteritis outbreaks are investigated needed to be understood, with possible variations between jurisdictions identified and accounted for to assure the applicability and portability of a proposed model. Second, a model needed to be developed and tested on real-world data to determine its reliability and validity. Third, a decision support system incorporating the model needed to be

developed and tested in a real-world setting to evaluate its effectiveness in practice.

To these ends, the aims for this PhD research were as follows:

Aim 1. Determine how guidelines for norovirus outbreaks in healthcare facilities vary between states and how differences between states are associated with outbreak outcomes;

Aim 2. Develop and validate a mathematical model approach that effectively determines whether outbreaks of acute gastroenteritis are consistent with person-to-person transmission of norovirus based on data available during the early stages of outbreaks;

Aim 3. Investigate how public health epidemiologists make decisions during AGE outbreak investigations and have them review hypothetical outbreak scenarios to assess how different levels of data availability would affect their decision-making.

The completion of these aims is intended to provide public health practitioners in varied locations and levels of government with a means of more effectively deploying resources during acute gastroenteritis outbreaks, increasing efficiency and potentially reducing illness in the community.

References

1. Gregg M. Field Epidemiology. 3rd ed. New York [NY]: Oxford University Press, 2008.

2. Hall A, Wikswo M, Manikonda K, Roberts V, Yoder J, Gould L. Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States. *Emerg Infect Diseases* 2013;19(8):1305-9.

3. Council to Improve Foodborne Outbreak Response (CIFOR). Guidelines for foodborne disease outbreak response. 2nd ed. Atlanta [GA]: Council of State and Territorial Epidemiologists; 2009.

4. Hall A, Vinje J, Lopman B, Park G, Yen C. Updated norovirus outbreak management and disease prevention guidelines. *MMWR Morb Mort Wkly Rep* 2011;60(RR03):1-15.

CHAPTER 2

BACKGROUND

Infectious Disease Outbreak Investigations

During an infectious disease outbreak, investigation is usually undertaken by the local health department (LHD) in the community where the outbreak has occurred. During these investigations, obtaining information about the pathogen and route of transmission (Figure 2.1) is essential to implement effective public health control measures to end the outbreak as soon as possible.[1] For example, an onsite environmental health assessment of a long-term care facility kitchen might be an appropriate response during an outbreak suspected to be foodborne, but may not be necessary for an outbreak transmitted person-to-person.[2] When the specific outbreak etiology is unknown, investigation and intervention decisions can be made using information derived from the reported cases (e.g., incubation periods and incidence trends) to predict the etiology.[3] However, this information is frequently unavailable during infectious disease outbreaks, especially during the early stages when appropriate control measures have the greatest impact.[4] Over half of all outbreaks reported to public health officials are never confirmed to have been caused by a specific agent.[5] While guidance on outbreak response has been provided by the CDC[1] and other organizations,[2] there is no clear guidance on how to

make decisions about investigation and control measures in the absence of adequate data. It is common practice for public health epidemiologists to rely on intuition and/or previous experience to make decisions in such circumstances. Table 2.1 lists typical investigation activities for outbreaks that are suspected to be foodborne as described by the Council to Improve Foodborne Outbreak Response (CIFOR).[2] Recommended control measures and the involvement of environmental health staff may differ for outbreaks that are not suspected to be foodborne, but other aspects of the investigation would be similar.

Acute Gastroenteritis Outbreaks

Norovirus is the most common cause of acute gastroenteritis in the U.S.[5] Each year, norovirus causes an average of 19-21 million cases of acute gastroenteritis, 1.7-1.9 million outpatient visits, 400,000 emergency department visits, 56,000-71,000 hospitalizations, and 570-800 deaths, and is responsible for nearly 1 million pediatric medical care visits and about 50% of foodborne disease outbreaks.[6] Each year, it costs about $2 billion in the U.S. for healthcare and lost productivity from foodborne illness caused by norovirus,[7] and it follows that the total burden when nonfoodborne illness is included would be substantially greater.

Most acute gastroenteritis (AGE) outbreaks are caused by person-to-person transmission of norovirus[5] which is also the most common cause of gastroenteritis outbreaks with laboratory confirmed etiologies.[5] Different control measures are recommended depending on the pathogen and mode of transmission (Figure 2.2),[2,5,8] but the pathogen and mode of transmission are frequently never confirmed.[5] Current guidelines recommend using Kaplan's Criteria for outbreaks in which this information

remains unknown to determine whether the causal pathogen may be norovirus.[8] Several studies have shown that norovirus can be distinguished from other enteric pathogens by clinical attributes such as the incubation period and duration of illness.[4,9-11]

According to the CDC's most recent National Outbreak Reporting System (NORS) report,[5] only half (48.8%) of reported gastroenteritis outbreaks in 2009-2010 were confirmed to have been caused by specific agents. Of the outbreaks with a confirmed agent, 64.2% were caused by norovirus. Two-thirds (66.1%) of these norovirus outbreaks had a mode of transmission that was person-to-person. Among confirmed norovirus outbreaks that occurred in a single setting, 63.7% were in healthcare facilities. The health department's role in investigating outbreaks in healthcare settings may be limited to collecting information and providing education,[8] activities which can typically be completed by telephone. Foodborne outbreaks of norovirus or other pathogens usually require a visit to the affected facility to verify proper food handling practices are in place and to enable the determination of the source and distribution (traceback) of any implicated food items.[2] Since the majority of gastrointestinal outbreaks are caused by person-to-person transmission of norovirus where onsite visits by public health authorities may not be needed, it would likely not be cost-effective for a health department to do an onsite visit during each outbreak of unknown etiology. However, failing to complete an onsite visit for outbreaks that are potentially foodborne allows lapses in food handling practices to persist and contaminated food items to continue to be distributed, potentially infecting many more people. Therefore, it is important to have a strategy for identifying and implementing the more intensive investigation and control measures during outbreaks that are more likely to be associated with point source

transmission of norovirus.

One solution to the problem is to increase laboratory testing, an approach that has been supported by the CDC in its CaliciNet program. CaliciNet provides laboratory support to states without in-house capacity for norovirus strain typing.[12] Currently, however, only about half of state public health laboratories are certified to participate,[13] so it can take several days for many health departments to get results. Since the average norovirus outbreak lasts about one week,[8] there may be little incentive for local health departments to expend time and resources to collect and submit additional specimens for testing. This situation may change with improved laboratory tests and techniques,[8] but there are also ethical concerns. The CDC promotes increased laboratory testing in their outbreak response guidelines, recommending that five stool specimens should be collected to confirm a diagnosis of norovirus,[8,14] but most persons with norovirus infection recover within three days[8] without needing to receive medical care. Requiring the collection of stool specimens from ill persons who may not derive any clinical benefit from the testing may be unduly burdensome, especially when it is unlikely the test results will be returned in time for public health officials to change their response.

Publication of Guidance Relative to Public Health Jurisdictional Relationships

The availability of guidelines for responding to acute gastroenteritis outbreaks is an important factor that influences how outbreak investigators make decisions on investigation and control measures. Though the CDC has published its own guidelines for responding to AGE outbreaks in healthcare settings,[15] many state and local health departments have prepared and distributed their own guidelines that differ from the

CDC's recommendations. In addition to the 50 state departments of health, the National Association of County & City Health Officials (NACCHO) identified 2,532 LHDs in the U.S., with 1,943 being locally governed, 402 being subunits of the state health agency, and 187 having shared governance between local and state jurisdictions.[16] If one assumes the shared or substate jurisdictions use the same guidelines as those developed by their respective states, then there may be as many as 1,993 different guidelines in place for responding to AGE outbreaks.

However, since the majority of LHDs (71%) serve populations of fewer than 50,000[16] (who would expect to have at most one AGE outbreak reported annually[2]), it is likely that most of these independent jurisdictions have not developed their own guidelines, as evidenced by their low participation in the development of local policy in general (28%).[16] Furthermore, those that do develop their own policies tend to have closer relationships with state agencies (47% of LHDs in jurisdictions with shared governance reported having local policy development).[16] Much of the variation between guidelines nationally is thus likely reflected in the state-level guidelines.

Variation in AGE guidelines between states may be due to attempts to simplify or add detail to CDC's guidelines (for example, identifying an outbreak based on "two or more cases" in Oregon[17] or "at least 3 patients/residents" in New Mexico[18] instead of using Kaplan's Criteria). Local laws and practices may also influence state guidelines (for example, norovirus is a reportable disease in Massachusetts[19] but not in Texas[20]). These state-specific guidelines have the potential to impact outbreak detection rates and outcomes and should be thoroughly evaluated. They are likely to have a significant impact on the outbreak investigation and control measures implemented in local

jurisdictions, which will impact healthcare facilities in those states. Since the majority of AGE outbreaks are caused by norovirus, and most norovirus outbreaks occur in healthcare facilities, it is important to understand what these state-specific guidelines are, and how they differ between states.

National Outbreak Reporting System

The National Outbreak Reporting System has served as a national repository of AGE outbreak data since February 2009. All local, state, and territorial health departments are encouraged to report all AGE outbreaks, which have been defined by NORS as 2 or more cases of a similar illness epidemiologically linked to a common exposure, to the system through an online portal.[5] There are four required fields for reporting an outbreak to NORS: primary mode of transmission, date first case became ill, reporting state, and estimated total number of primary cases. However, health departments are also encouraged to report detailed information on other elements, including the setting, contributing factors, implicated food vehicles, and traceback investigations.[21] Though mode of transmission is required, the determination is based largely on the findings of the investigation and may be subject to error. For example, the guidelines specify that transmission should be considered person-to-person "if most of the cases had known direct contact or likely had the opportunity for direct contact with one another." This definition allows room for interpretation, and if a common food or environmental source is not clearly identified, outbreaks may be misclassified. Before NORS was initiated in 2009, only foodborne outbreaks were reported to the CDC. A rapid decline in foodborne outbreaks was observed when the option of reporting other

modes of transmission became available, suggesting that some outbreaks had previously been incorrectly classified as foodborne. This past misclassification shows the potential for continuing misclassification, but the inclusion of additional modes of transmission in NORS is thought to have improved the accuracy of transmission mode categorization over time.[22]

Etiology determinations in NORS are based on gold standard laboratory testing, and other data fields are entered directly from reports prepared by the investigating health departments. However, characteristics of the NORS system may impact the quality of data collected about AGE outbreaks. While it is true that some data may be missing from NORS simply due to health departments' failure to report fields that are not required by the CDC, many health departments do enter complete outbreak data, and there are enough data available for determinations to be made about which elements can feasibly be collected in practice. The restriction of NORS data to those fields selected for inclusion by the CDC is an important limitation. Identifying other potentially important fields that may need to be included in NORS is a subject for future research. Despite the limitations of the data, NORS remains a rich source of information on AGE outbreaks in the U.S. A summary of data fields and availability in NORS from 2009-2011 is included in Table 2.2.

Mathematical Modeling and Clinical Profiles in Applied Public Health

Mathematical modeling is an analytical process whereby "population parameters are described by symbols and linked by algebraic formulae" to "provide a realistic representation of the real world."[23] These models have been widely used in the study of

human infectious diseases.[23]

Mathematical models have been widely used for determining optimal control measures during outbreaks of influenza,[24-25] measles,[26] Severe Acute Respiratory Syndrome (SARS),[27] and many other agents.[28-30] However, such studies appear to primarily address control measures that are to be implemented for known infectious agents during large-scale pandemics. Control measures for smaller, local outbreaks, which health departments are likely to face much more frequently than pandemics and where the infectious agent may not be known, are not well represented in the literature.

The structure of a model is dependent on the nature of the pathogen and population being modeled. For example, a Susceptible-Exposed-Infective-Recovered (SEIR) model might be appropriate for diseases that are only infectious following an incubation period. Other traditional epidemic models include the Susceptible-Infective-Recovered (SIR) and Susceptible-Infective-Susceptible (SIS) models. The results of such a model may be used to distinguish between the different types of outbreaks when the data necessary for more sophisticated analyses are unavailable. This approach is applicable to all outbreaks of infectious disease where the etiology is unknown. However, this dissertation is focused on outbreaks caused by enteric pathogens.

Several researchers have successfully used various techniques to categorize gastroenteritis outbreaks based on clinical presentation.[4,9-11] Kaplan's Criteria, which has been found to be relatively sensitive (68%) with a remarkably high specificity (99%),[4] has been recommended for use by the CDC in investigating outbreaks of unknown etiology.[3] Other techniques have also performed well, with attribution of outbreaks of unknown etiology ranging from 73%[9] to 93%[10] or 94%,[11] though no record of these

techniques being used in applied practice was readily found in the literature. Furthermore, each of these methods was applied to subsets of data containing information that a health department may not often have readily available at any point during many outbreaks. The reasons for missing data are unknown and likely vary widely due to the diversity of jurisdictions represented, but may include problems such as insufficient resources to collect data or the difficulty in estimating the incubation period in person-to-person outbreaks. Kaplan's Criteria, for example, which requires information about the percentage of cases that vomited, the mean incubation period, and the mean duration of illness, has been found to be useable in only 37% of outbreaks of unknown etiology.[4]

The success of techniques that use clinical profiles for norovirus attribution suggests that the parameters for a mathematical model for person-to-person transmission of norovirus (such as incubation period and duration of illness) can be expected to be relatively consistent between populations, allowing the potential for them to be assumed beforehand instead of collected as part of an investigation. Such assumptions would enable a model to be run using only such basic preliminary information as the current number of cases, size of the at-risk population, and the number of days since the first onset of illness – typically among the first items of information obtained during an outbreak investigation. Such a mathematical model should be able to predict with reasonable confidence whether the number of outbreak cases over time is consistent with person-to-person transmission of norovirus, with sensitivity and specificity approaching that of clinical profile-based approaches, but only requiring data that are available most or all of the time. However, even with high measures of validity, the willingness of public health practitioners to utilize such a tool is unknown, and an acceptable user interface

with adequate measures of usability must also be developed and evaluated for the mathematical model to be an effective public health resource in practice.

Table 2.1. Investigation activities for epidemiology and environmental health staff during foodborne outbreaks associated with events or establishments (as recommended by the Council to Improve Foodborne Outbreak Response[2]).

Objective	Epidemiology Staff	Environmental Health Staff
Identify etiologic agent	Contact healthcare providers of cases; interview cases to characterize symptoms, incubation period, and duration of illness; obtain stool from cases; establish case definition based on confirmed diagnosis or clinical profile	Interview management/food workers; obtain stool; obtain samples of implicated and suspected food; determine whether setting or food suggests a likely pathogen
Identify persons at risk	Obtain from event organizer a list of persons attending event/patronizing establishment; Interview persons to determine attack rates, by time	Obtain lists of reservations for establishment, receipts, inventory of foods ordered, or guest lists for events
Identify mode of transmission and vehicle	Interview cases and controls or well meal companions about all common exposure sources and calculate odds ratios; interview persons with identified exposures and determine attack rates and relative risks	Obtain menu; interview food workers to determine food-preparation responsibilities; reconstruct food flow for implicated food; identify contributing factors; obtain samples of food; obtain environmental samples
Identify source of contamination	Combine descriptive and analytical epidemiology results	Evaluate food flow for implicated food; if no contamination event identified, trace source ingredients back through distribution
Identify contributing factors	Summarize information to identify confirmed or suspected agent/confirmed or suspected food vehicle	Evaluate results of environmental investigation, given identification of agent and results of epidemiologic investigation to identify factors
Determine potential for ongoing transmission and need for abatement procedures	Create epidemic curve and evaluate to determine whether additional cases may still be occurring; review potential abatement procedures	Implement control measures to prevent further exposures: verify food workers excluded, verify contaminated foods disposed, very surfaces cleaned and sanitized, train staff; if measures cannot be verified, alert public or close premises

Table 2.2. National Outbreak Reporting System data fields, completeness, and distribution, 2009-2011 (*N* = 6,877).

Field Name	% Complete	Low Value	High Value	Description
CDCID	100.0	15	16753	Unique identifier
PrimaryMode	100.0	Animal contact	Person-to-person	Primary mode of transmission
DateFirstIll	100.0	1/1/2009	12/31/2011	Date of first onset
DateLastIll	87.6	1/1/2009	9/4/2013	Date of last onset
ReportingState	100.0	Alabama	Wyoming	Reporting state
ConfirmedPrimary	96.6	0	272	Number of confirmed primary cases
ProbablePrimary	94.7	0	1858	Number of probable primary cases
EstimatedPrimary	100.0	2	1939	Estimated number of primary cases
IncMedian	28.8	3 minutes	34 days	Median incubation period
DurMedian	51.6	0 minutes	59 days	Median duration of illness
Symptom	79.9	Abdominal cramps	Weight loss	Name of symptom
SymptomCases	73.3	0	1043	Number of cases with symptom

Table 2.2 continued.

Field Name	% Complete	Low Value	High Value	Description
SymptomInfo	73.1	0	528	Number of cases for whom symptom info is known
SecondaryMode	12.0	Animal Contact	Water	Secondary mode of transmission
ConfirmedSecondary	22.3	0	17	Number of confirmed secondary cases
ProbableSecondary	22.0	0	292	Number of probable secondary cases
TotalSecondary	26.1	0	299	Total number of secondary cases
TotalCases	100.0	2	1939	Total number of cases
EtiologyKnown	100.0	No (38.7%)	Yes (61.3%)	Etiology is confirmed or suspected
SpecimensTaken	94.9	No (21.1%)	Yes (11.3%)	If etiology is unknown, specimens were collected
NumberSpecimens	32.3	0	410	Number of specimens collected
TestedBacteria	100.0	No (79.2%)	Yes (20.8%)	Specimens were tested for bacteria
TestedChemicals	100.0	No (97.6%)	Yes (2.4%)	Specimens were tested for chemicals
TestedViruses	100.0	No (75.5%)	Yes (24.5%)	Specimens were tested for viruses
TestedParasites	100.0	No (98.0%)	Yes (2.0%)	Specimens were tested for parasites
GenusName	63.5	Astrovirus	Yersinia	Genus of identified etiology

Table 2.2 continued.

Field Name	% Complete	Low Value	High Value	Description
SpeciesName	50.5	A	Unknown	Species of identified etiology
SerotypeName	13.9	Agona	Unknown	Serotype of identified etiology
Confirmed	100.0	No (16.0%)	Yes (49.1%)	Etiology is confirmed
OtherCharacteristics	6.5	Norovirus	XbaI: JEGX01.0023	Other characteristics of identified etiology
PatientSpecimen	94.7	No (6.1%)	Yes (53.6%)	Etiology was detected in a patient specimen
FoodSpecimen	94.7	No (56.6%)	Yes (3.2%)	Etiology was detected in a food specimen
EnvironmentSpecimen	94.7	No (59.0%)	Yes (0.8%)	Etiology was detected in an environmental specimen
FoodWorkerSpecimen	94.7	No (57.0%)	Yes (2.8%)	Etiology was detected in a food worker specimen
NumberLabConfirmed	60.0	0	272	Number of laboratory confirmed cases
MajorSetting	44.5	Adult day care	Youth center	Where person-to-person exposure occurred
GuestExposed	32.1	0	266,944	Number of guests exposed
StaffExposed	21.9	0	4,000	Number of staff exposed
GuestIll	38.6	0	742	Number of guests ill
StaffIll	31.0	0	181	Number of staff ill

Table 2.2 continued.

Field Name	% Complete	Low Value	High Value	Description
GuestAR	32.0	0	100	Guest attack rate
StaffAR	21.6	0	100	Staff attack rate
OtherSetting	2.1	Adult day care	Workplace	Additional person-to-person exposure setting
AnimalSetting	1.7	Animal Exhibit	Zoo	Setting of animal exposure
WherePrepName	30.2	Banquet Facility	Workplace, not cafeteria	Setting of food preparation
WhereEatenName	29.2	Banquet Facility	Workplace, not cafeteria	Setting of food exposure

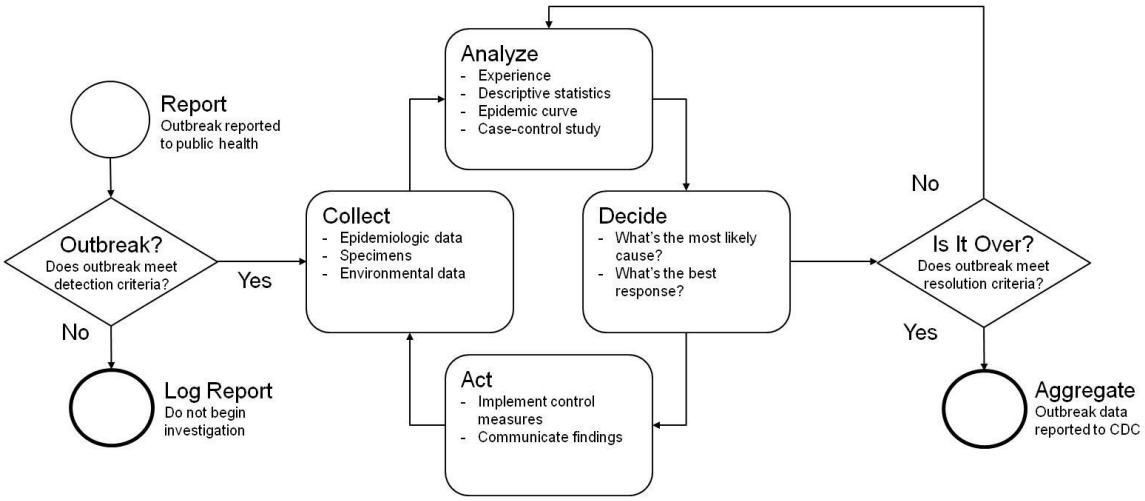


Figure 2.1. High-level processes associated with investigating an acute gastroenteritis outbreak.

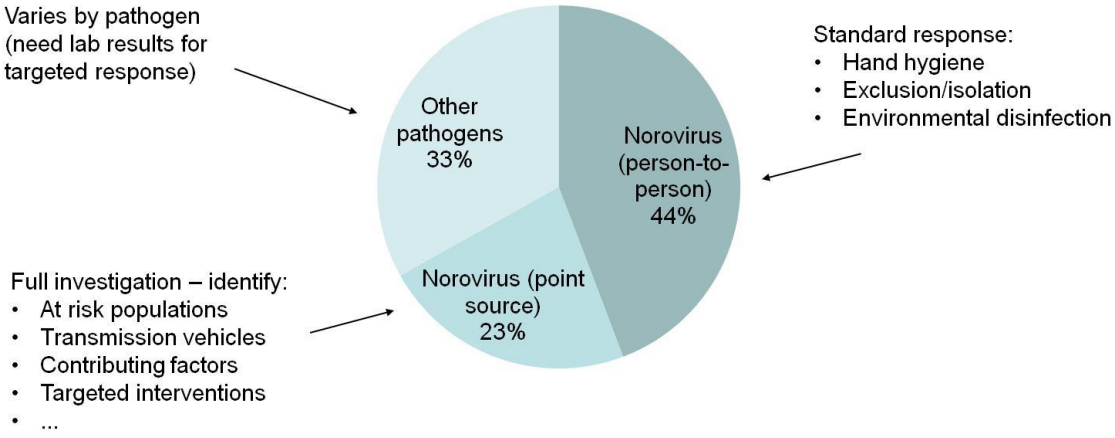


Figure 2.2. Types of acute gastroenteritis outbreaks and typical public health response (based on national guidelines and surveillance reports[2,5,8]).

References

1. Gregg M. Field Epidemiology. 3rd ed. New York [NY]: Oxford University Press, 2008.

2. Council to Improve Foodborne Outbreak Response (CIFOR). Guidelines for Foodborne Disease Outbreak Response. Atlanta [GA]: Council of State and Territorial Epidemiologists, 2009.

3. 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.

4. Turcios RM, Widdowson MA, Sulka AC, Mead PS, Glass RI. Reevaluation of epidemiological criteria for identifying outbreaks of acute gastroenteritis due to norovirus: United States, 1998-2000. *Clin Infect Dis* 2006;42(7):964–9.

5. Hall AJ, Wikswo ME, Manikonda KR, Virginia A, Yoder JS, and Gould LH. Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States. *Emerg Infect Diseases* 2013;19(8):1305-9.

6. Burden of Norovirus Illness and Outbreaks [Internet]. Atlanta [GA]: Centers for Disease Control and Prevention [updated 2017 Jul 10; cited 2017 Jul]. Available from http://www.cdc.gov/norovirus/php/illness-outbreaks.html.

7. Hoffmann S, Batz MB, Morris JG. Annual cost of illness and quality-adjusted life year losses in the United States due to 14 foodborne pathogens. *J Food Prot* 2012;75:1292-302.

8. Hall AJ, Vinjé J, Lopman B, Park GW, Yen C, Gregoricus NP, et al. Updated norovirus outbreak management and disease prevention guidelines. *MMWR Morb Mortal Wkly Rep* 2011;60(RR03):1-15.

9. Domínguez A, Broner S, Torner N, Martínez A, Jansà JM, Alvarez J, et al. Utility of clinical-epidemiological profiles in outbreaks of foodborne disease, Catalonia, 2002 through 2006. *J Food Protect* 2010;73(1):125–31.

10. Hedberg CW, Palazzi-Churas KL, Radke VJ, Selman CA, Tauxe RV. The use of clinical profiles in the investigation of foodborne outbreaks in restaurants: United States, 1982-1997. *Epidemiol Infect* 2008;136(1):65-72.

11. Verhoef LPB, Kroneman A, van Duynhoven Y, Boshuizen H, van Pelt W, Koopmans M. Selection tool for foodborne norovirus outbreaks. *Emerg Infect Diseases* 2009;15(1):31–8.

12. Vega E, Barclay L, Gregoricus N, Williams K, Lee D, Vinje J. Novel surveillance network for norovirus gastroenteritis outbreaks, United States. *Emerg Infect Diseases* 2011;17(8):1389-95.

13. Responding to Norovirus Outbreaks [Internet]. Atlanta [GA]: Centers for Disease Control and Prevention [updated 2012 Jun 24; cited 2017 Jul]. Available from: http://www.cdc.gov/norovirus/php/responding.html.

14. Duizer E, Pielaat A, Vennema H, Kroneman A, Koopmans M. Probabilities in norovirus outbreak diagnosis. *J Clin Virol* 2007;40(1):38–42.

15. Guideline for the Prevention and Control of Norovirus Gastroenteritis Outbreaks in Healthcare Settings, 2011 [Internet]. Healthcare Infection Control Practices Advisory Committee (HICPAC) [updated 2017 Feb 15]. Available from: https://www.cdc.gov/infectioncontrol/pdf/guidelines/norovirus-guidelines.pdf.

16. 2013 National Profile of Local Health Departments [Internet]. National Association of County & City Health Officials; 2014. Available from http://nacchoprofilestudy.org/reports-publications/.

17. Public Health Division [Internet]. Norovirus Outbreak Detection and Management. Oregon Health Authority; 2014. Available from: https://public.health.oregon.gov/DiseasesConditions/CommunicableDisease/Outbreaks/Gastroenteritis/Documents/Noro-Outbreak-Detection-Mgmt-Tool.pdf.

18. Norovirus in Long Term Care Facilities Outbreak Checklist. New Mexico Department of Health; 2015.

19. Norovirus & Other Caliciviruses [Internet]. Massachussetts Department of Public Health [released 2016 Sep; cited 2017 Jul]. Available from http://www.mass.gov/eohhs/docs/dph/disease-reporting/guide/norovirus.pdf.

20. Norovirus Outbreaks [Internet]. Texas Department of State Health Services [revised 2015 Apr]. Available from https://www.dshs.state.tx.us/region2-3/programs/commprep/NOROVIRUS-OUTBREAK-FORM.pdf.

21. National Outbreak Reporting System (NORS) Guidance. Atlanta [GA]: Centers for Disease Control and Prevention; 2017.

22. Imanishi, M, Manikonda, K, Murthy, BP, Gould, LH. Factors contributing to decline in foodborne disease outbreak reports, United States. *Emerg Infect Dis* 2014;20(9):1551-3.

23. Vynnycky E, White RG. An Introduction to Infectious Disease Modelling. New York [NY]: Oxford University Press; 2010.

24. Merl D, Johnson LR, Gramacy RB, Mangel M. A statistical framework for the adaptive management of epidemiological interventions. *PloS One* 2009;4(6):e5807.

25. van Kerkhove MD, Ferguson NM. Epidemic and intervention modelling--a scientific rationale for policy decisions? Lessons from the 2009 influenza pandemic. *Bull World Health Organ* 2012;90(4):306–10.

26. Grais RF, Conlan AJK, Ferrari MJ, Djibo A, Le Menach A, Bjørnstad ON, et al. Time is of the essence: exploring a measles outbreak response vaccination in Niamey, Niger. *J Royal Soc Interface* 2008;5(18):67–74.

27. Hsieh Y-H, King C-C, Chen CWS, Ho M-S, Hsu S-B, Wu Y-C. Impact of quarantine on the 2003 SARS outbreak: a retrospective modeling study. *J Theor Biol* 2007;244(4):729–36.

28. Tildesley MJ, Savill NJ, Shaw DJ, Deardon R, Brooks SP, Woolhouse MEJ et al. Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. *Nature* 2006;440(7080):83–6.

29. Ball FG, Lyne OD. Optimal vaccination policies for stochastic epidemics among a population of households. *Math Biosci* 2002;177-178:333-354.

30. Elderd BD, Dukic VM, Dwyer G. Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. *Proc Natl Acad Sci USA* 2006;103(42):15693–7.

CHAPTER 3

STATE-LEVEL ADOPTION OF NATIONAL GUIDELINES FOR NOROVIRUS OUTBREAKS IN HEALTHCARE SETTINGS

This chapter is a precopyediting, author-produced copy of an article accepted for publication in the American Journal of Infection Control following peer review. It is published here in accordance with the journal's policies allowing prior publication as part of an academic thesis.

Introduction

In the U.S., there is growing interest in nationally interoperable public health informatics undertakings that work across sectors and jurisdictional boundaries, such as electronic case reporting[1] and Public Health 3.0.[2] However, public health remains largely local and fragmented, with policies and information systems often developed for a single state or local jurisdiction.[2] Understanding existing guidelines and expectations currently applied in jurisdictions across the country is an important first step toward such ambitious development projects.

Clinical decision support (CDS) systems can help investigating officials follow best practices when responding to infectious disease outbreaks,[3] but variation in local guidelines would make such systems difficult to develop and implement. It was our efforts to develop such a system for acute gastroenteritis outbreak response and the impact of variation in guidelines for norovirus in healthcare settings on our approach that led to the present study. In 2011, the Healthcare Infection Control Practices Advisory Committee (HICPAC) published guidelines concerning norovirus outbreaks in healthcare settings[4] (hereinafter referred to as "the HICPAC guidelines"). These guidelines gave detailed recommendations for prevention and control of norovirus and similar acute gastroenteritis (AGE) outbreaks in healthcare settings based on the existing literature and other evidence. Noroviruses are responsible for at least 50% of AGE outbreaks worldwide, causing approximately 20 million illnesses each year.[5] Local or state health departments investigate outbreaks to institute interventions and prevent future outbreaks. The results are reported to CDC’s National Outbreak Reporting System (NORS), an Internet-based system launched in 2009 for foodborne and waterborne outbreaks, AGE

outbreaks caused by contact with infected persons, animals, or environmental sources, and AGE outbreaks caused by other or unknown modes of transmission.[6]

Many state and local jurisdictions had released their own guidelines prior to the HICPAC guidelines, and most have since published or updated their own norovirus or general AGE outbreak guidelines with recommendations tailored to the perceived needs of their individual jurisdictions. Healthcare facilities are the most common setting of reported AGE outbreaks,[6] and most of these guidelines are specific to the management of outbreaks in such settings. This emphasis on healthcare-associated AGE outbreaks can also be explained by the vulnerable populations that reside in such settings, the potential for successful interventions, and the regulatory relationships that exist between public health entities and healthcare facilities in many states. The guidelines may also include guidance on investigation activities (e.g., when to conduct environmental health assessments), administrative details (e.g., how local health departments should report findings to the state), regulatory requirements (e.g., when to close a healthcare facility's food services), or any other topics deemed by a state to be important to outbreak response.

Beyond their use by a state health department, state guidelines may also be used by local health departments or healthcare facilities or adapted when these entities create their own guidelines. Failure to adopt national guidelines could result in the use of ineffective strategies for outbreak response and lead to disparities in outbreak data collection and reporting between jurisdictions. The objectives of this study were a) to identify the extent to which state-level guidelines adhere to national guidelines for response to norovirus outbreaks in healthcare settings and b) to explore whether the

differences in guidelines may impact the information about outbreaks reported to NORS. The impact of such differences is currently unknown. A better understanding of the potential relationships between guidance and outbreak outcomes would improve the interpretability of national outbreak data, potentially lead to improved guidelines, and encourage development of CDS.

Methods

Guidelines Collection and Verification

In February 2016, internet searches for norovirus guidelines in each of the 50 U.S. states were conducted using Google Search. Search terms included the individual names of states combined with "state norovirus guidelines." The first 20 search results were reviewed, and state health department documents addressing norovirus or AGE outbreak response were collected. Next, state epidemiology, infectious disease, and/or healthcare-associated infections pages were navigated manually from each state health department website and inspected for AGE guidelines. If search functionality was available on a state's website, then the search terms "norovirus" and "gastroenteritis" were used and results that addressed norovirus or AGE outbreak response in the first 20 search results were inspected. Since guidelines designed to address AGE outbreaks in general were not distinguished from guidelines specific to norovirus, we will refer to "AGE outbreaks" unless specifically referring to individual guidelines or outbreaks solely related to norovirus.

To verify that the guidelines identified were current and the preferred guidance documents, we attempted to contact an authority from each state. A directory of email

and web form addresses was created by manually searching the state health departments' epidemiology, infectious disease, and/or healthcare-associated infections web pages for contact information. For states where the email or web form addresses at the department level could not be found, higher levels of electronic contact information were sought using websites' navigation trees up to the main health department level. A standard letter requesting verification of online guidelines and copies of any other guidelines was sent via email or web form to the 45 states for which addresses were obtained in June 2016. States were provided four weeks to respond. Once all internet searches were completed and all state responses collected, available guidelines were reviewed and restricted to those including guidance specific to healthcare settings.[7-54]

Guidelines Review and Tabulation

The HICPAC guidelines for norovirus outbreaks in healthcare settings include 12 recommendation topics[4] and were used as a template for comparison. Additional topics were included if a new recommendation occurred in at least 10 state-based guidelines. When analyzing the states' guidelines, general recommendations that fell under one of the HICPAC guidelines' recommendation topics (e.g., *increased hand hygiene* and *patient cohorting*) were grouped broadly under the associated HICPAC-defined topic, while recommendations found only in state guidelines were analyzed in greater detail. Quantitative recommendations, such as the number of stool specimens to be collected or the time range for patient isolation, were grouped based on the specific ranges indicated. Variations between states for each of these recommendations were summarized, and implications for outbreak reporting were identified.

Impact Analysis

Data about outbreaks reported to NORS in 2015 (dates of first onset between January 1 and December 31, 2015) that were attributed at least in part to laboratory-confirmed or suspected norovirus in healthcare facilities were used to assess the potential impact of variation in guidelines between states. The year 2015 was the most recent year for which data were available. An outbreak was considered healthcare-associated if it was reported as occurring at least in part in one of the following settings (based on where the food was eaten for foodborne outbreaks): hospitals, long-term care facilities, such as nursing homes and assisted living facilities, and other healthcare facilities.[55]

Outbreak-related elements in the NORS data were stratified using the following criteria from the states' guidelines:

- number of specimens: a) median below five (five is the HICPAC recommendation[56]) (e.g., "1-3 specimens"), b) median at or above five (including ranges with no upper bound), or c) no guidance or unspecified number (e.g., "Collect specimens within 48-72 hours");
- outbreak detection: a) two or more epidemiologically-linked cases occurring within a specified time frame (NORS definition[57]), b) three or more epidemiologically-linked cases, "more than expected" case counts, or outbreaks based on unspecified suspicion; c) no guidance;
- ill patient isolation: a) minimum of two days (HICPAC recommendation[4]), b) less than two days; c) no guidance;
- ill employee exclusion: minimum of two days (HICPAC recommendation[4]), b) less than two days; c) no guidance;

- outbreak resolution: a) two incubation periods or more (a standard time frame in applied epidemiology used to identify transmission from subclinical infections or unrecognized cases[58]), b) less than two incubation periods, or c) no guidance;
- mode of transmission: a) any guidance, or b) no guidance;
- environmental health assessments: a) any guidance, or b) no guidance).

Descriptive statistics were reported and used to compare outbreak outcomes in states with different guidelines. Categories were combined or comparisons were not made if recommendation categories were in fewer than five states' guidelines. Rate ratios were used to compare rates between states. Two-proportion *Z*-tests and two-sample *t*-tests were used to compare states with reference recommendations (those consistent with the HICPAC guidelines or otherwise representing model practice) to states with other recommendations. The impact of differences in recommendations for stool specimen collection was assessed by comparing the percentages of outbreaks with confirmed etiology (as indicated in NORS) between states. For ill patient isolation and ill employee exclusion, the average outbreak duration and the average number of cases per outbreak were compared. For outbreak detection, the number of outbreaks per 100,000 people using state population estimates from the 2015 American Community Survey[59] and the average number of cases per outbreak were compared. For outbreak resolution, the average outbreak duration and the average number of cases per outbreak were compared. Average outbreak duration only included outbreaks for which the first and last onset dates were available (i.e., outbreaks with missing dates were excluded). Average number of cases was based on the estimated number of primary cases. Two incubation periods

was defined as four days for this analysis. For states with and without guidelines on determining the mode of transmission and when to conduct environmental health assessments, the percentages of outbreaks attributed to "Indeterminate/Other/Unknown" transmission were compared.

Human Participant Compliance Statement

The Institutional Review Board at the University of Utah reviewed the procedures outlined above and determined that the study did not meet the criteria for human subjects research.

Results

State-based guidelines specific to AGE outbreak response in healthcare settings were identified for 41 states. Four states reported they had no guidelines, and five states did not respond to our queries or have guidelines available on their websites. Of the 50 U.S. states, 38 (76%) were found to have online guidelines available through the attempted internet searches. Email or web form contact information were readily available for 45 states, and public health officials from 25 of these states responded to queries for guidelines, for a response rate of 56%. Altogether, 12 responding states provided additional or current documents, nine verified existing online resources, and four reported they had no available guidelines. Of the 20 states that did not respond to queries, 15 had guidelines specific to AGE outbreak response in healthcare settings available on their websites, and 5 did not. The five states for which no electronic contact addresses could be found also had healthcare-specific guidelines available online, making

a total of 41 states with guidelines available for analysis. These guidelines were dated from 2001 to 2016, with a median year of 2013. Four states had undated guidelines. Guidelines in states that responded to queries had date ranges that were slightly more recent (2006-2016, median 2014) than those of nonresponding states (2001-2016, median 2012).

The 12 prevention and control measures identified in the HICPAC guidelines, together with additional guidance concerning outbreak detection, outbreak resolution, mode of transmission, and environmental health assessments, were including in 10 or more states' guidelines and thus met the threshold for further analysis. These categories are detailed below. The following guideline elements were identified in fewer than 10 states' guidelines and thus did not meet the threshold to be included for this analysis: outbreak confirmation, events warranting additional public health notification, additional documents, and other content.

Prevention and Control

The state guidelines generally followed the HICPAC guidelines with regard to recommendations on prevention and control. Coverage ranged from 71% having recommendations related to *patient transfer and ward closure* to 100% having recommendations regarding *hand hygiene* and *staff leave and policy* (see Table 3.1). However, even within these well-established categories there was considerable variation between states. For example, HICPAC recommends that ill residents be isolated and ill staff excluded until at least two days after their symptoms resolve. However, the 27 available state recommendations on ill patient isolation ranged from 24 hours up to 72

hours after resolution, with five different recommended time frames within that range. Similarly, the 40 state recommendations concerning ill employee exclusion ranged from allowing employees to return immediately after symptom resolution to 72 hours after resolution, with six different recommended time frames within that range (see Tables 3.2 and 3.3).

The HICPAC guidelines on norovirus outbreaks in healthcare settings refrain from giving specific guidance on the number of stool specimens to collect during an outbreak, instead directing healthcare facilities to consult with their state or local public health authorities. However, in the "Updated Norovirus Outbreak Management and Disease Prevention Guidelines," CDC recommends that stool specimens be collected from at least five ill persons during norovirus outbreak investigations.[56] In contrast, of the 36 states giving guidance on the number of stool specimens to collect, only two recommended consulting with public health officials, two mentioned only that stool specimens should be collected, and the remaining 32 states gave 21 different recommended ranges, with ranges beginning or ending on almost every integer from 1 to 12 (see Figure 3.1).

Outbreak Detection

At least 11 different outbreak definitions were recommended for outbreak detection in 40 (98%) states' guidelines (see Table 3.3). Only one state (2%) did not provide guidance on outbreak detection. The HICPAC guidelines recommend using clinical laboratory diagnostics or Kaplan's Criteria[60] to identify the likely etiology of an AGE outbreak, but they do not address how to determine that an AGE outbreak has

occurred. However, NORS defines an outbreak as "two or more cases of similar illness associated with a common exposure,"[57] and 15 states adopted a definition for AGE outbreaks similar to the NORS definition. There was still variation in how different states defined the linkage between cases — for example, whether the cases had to occur within one, two, or three days, or just be otherwise "epidemiologically linked"— but all 15 guidelines required associations in place, time, or common exposure between at least two cases. The other 25 states defined outbreaks more liberally as a) three or more cases, b) more cases than expected in a unit or facility during a given time period or time of year, or c) gave no guidance beyond directives to report any "suspected" or "apparent" outbreaks or clusters.

Outbreak Resolution

At least 14 different criteria for when to conclude an outbreak investigation were recommended in 29 (71%) states' guidelines (see Table 3.3). About one third of the states (*N* = 12, 29%) gave no definition for when to consider an outbreak to be over. Among the states that did provide guidance, most (86.2%) gave a specific amount of time to wait before declaring an outbreak over, generally ranging from two to seven days. Eleven states recommended that the outbreak be considered over after waiting two incubation periods, but there was some disagreement on this value as well, with the associated time frames for nine states ranging from four to six days, and two states not specifying how long two incubation periods should be. However, even when the number of days was the same, there was disagreement about when to start counting those days, whether after the last onset of illness, after the last new case is identified, or after the last symptoms have

resolved. Three states recommended waiting to start counting until after the outbreak "appears over," and two others only recommended that control measures continue "during periods of increased illness" or "until the outbreak has resolved" without specifying how to make those determinations.

Mode of Transmission

At least six different methods for determining mode of transmission were recommended in 15 (37%) states' guidelines (see Table 3.3). Twenty-six states (63%) provided no guidance for identifying the suspected mode of transmission of an outbreak, which is a required field for all outbreaks reported to NORS (though it can be entered as unknown). Of the states that did provide guidance, recommendations included determining whether there was a common source (*N* = 2, 13%), looking at the shape of the epidemic curve (*N* = 4, 27%), counting the number of incubation periods from the initial case to subsequent cases (*N* = 3, 20%), identifying a rapid increase in cases over a short period of time (*N* = 3, 20%), checking for either a rapid increase in the number of cases or a common exposure (*N* = 1, 7%). Two states (13%) mentioned the mode of transmission, but gave no guidance beyond directives to identify it.

Environmental Health Assessments

At least six different sets of criteria concerning when to conduct an environmental health assessment at a facility experiencing an outbreak were recommended in 13 (32%) states' guidelines (see Table 3.3).Twenty-eight states (68%) provided no guidance on this topic. Of the states that did provide guidance, recommendations included that an

assessment should be conducted if the outbreak appears to be foodborne (*N* = 1, 8%) or caused by a point source (*N* = 2, 15%), if food handlers become ill (*N* = 4, 31%), or if either the outbreak appears to be foodborne or food handlers become ill (*N* = 1, 8%). Three states (23%) mentioned that an environmental health assessment should be conducted, without any guidance on any times when it should not be conducted, and two others (15%) indicated that it should only be done if determined to be necessary using unspecified criteria.

NORS Analyses

Several elements identified in the guidelines review were found to have implications for outbreak reporting. Variation in guidelines on outbreak detection could potentially impact the number of outbreaks reported in a state, while different definitions for outbreak resolution could affect the duration of the outbreak and final case counts. Lack of guidance on determining the mode of transmission or conducting environmental health assessments could result in fewer outbreaks having a mode of transmission identified and fewer determined to be foodborne or attributed to other causes. Lower recommendations for the number of stool specimens to collect could result in smaller proportions of outbreaks having a confirmed etiology.

A total of 1,253 outbreaks attributed at least in part to laboratory-confirmed or suspected norovirus in healthcare settings were reported to NORS in 2015, 1,124 (90%) of which occurred in the 41 states with guidelines analyzed in this study. Six states did not report any outbreaks that year, and two states reported fewer than five outbreaks, but were retained in the analysis due to consistency with historical trends in those states.

Excluding the two states with fewer than five outbreaks from the analysis resulted in qualitatively identical results and similar effect sizes. None of the outbreaks were reported to have had exposures in multiple states.

Outbreak metrics for states with and without recommendations in alignment with the HICPAC guidelines are presented in Table 3.4. States whose recommended ranges for specimen collection were consistent with the five specimens recommended by the CDC (Figure 3.1) had percentages of norovirus outbreaks with confirmed etiologies 17 percentage points higher than states with lower ranges (proportion with confirmed etiology of 61.8% vs. 45.2%, respectively; $p < 0.0001$) and 13 percentage points higher than states with no guidance (61.8% vs. 48.8%, respectively; $p = 0.0056$). Meanwhile, states with definitions of outbreaks in line with the CDC definition reported 25% fewer outbreaks per 100,000 people, on average, than states with less restrictive definitions (outbreak rate of 0.37 vs. 0.49 per 100,000 people, respectively; $p < 0.0001$). States recommending at least two days for ill patient isolation had 15% lower average case counts than states recommending less than two days (average case count of 30.1 vs. 35.6 cases, respectively; $p = 0.0021$) and 29% longer average outbreak duration than states with no guidelines on patient isolation (average outbreak duration of 14.0 vs. 10.8 days, respectively; $p = 0.0082$). States recommending at least two days for ill employee exclusion had 10% lower average case counts than states recommending less than two days for employee exclusion (average case count of 30.5 vs 34.0, respectively; $p = 0.0400$). States with recommendations to continue interventions for at least two incubation periods after the last case reported 26% shorter outbreak durations than states with recommendations to continue interventions for fewer than two incubation periods

after the last case (average outbreak duration of 11.5 vs. 15.4 days, respectively; p = 0.0109). There also appeared to be fewer outbreaks categorized as "Indeterminate/Other/Unknown" in states providing guidance for identifying an outbreak's mode of transmission (1.6% vs. 2.6%, respectively) or when to conduct an environmental health assessment (1.6% vs. 2.6%, respectively) compared to states providing no guidance, though these proportions were small and the differences were not statistically significant.

Discussion

This descriptive study is the first to characterize the variability in state-level guidelines for AGE outbreak response in healthcare settings. The results provide evidence that substantial variation exists in guidelines between states despite the release of national guidelines in 2011. The HICPAC guidelines focus on prevention and control activities, for which there were higher levels of agreement between states than for other outbreak response activities, such as outbreak detection and stool specimen collection, which were not directly addressed in the HICPAC guidelines. Outbreak detection, stool specimen collection, and other outbreak response activities have been addressed by the CDC elsewhere.[56-57]

With the exception of patient isolation, this study found more favorable outbreak outcomes in states with recommendations in alignment with HICPAC and CDC guidelines. These findings support the importance of close adherence to evidence-based national guidelines when health departments develop guidelines for their jurisdictions. However, there are important limitations of the dataset used, and associations should be

interpreted with caution. As a passive reporting system, not all states submit outbreak data consistently to NORS. Among reported outbreaks, optional variables such as last illness date (for determining duration), setting, and etiology may not be consistently reported. Investigators of outbreaks reported to NORS may or may not have adhered to the state guidelines in the state where they reside. The analysis was strictly observational and associations may not be causal, and no multivariate analyses were performed or other attempts made to control for potential confounding, which may have significantly impacted the observed associations.

Despite these limitations, the presence of such apparent disparities indicates that more can be done to improve the equitability of AGE outbreak response in the U.S. Whether conditions favorable to large outbreaks with unconfirmed etiologies result in less stringent guidelines or if it is the guidelines themselves not being consistent with CDC guidance that causes the less desirable outbreak outcomes is unclear, and there may be other factors that explain or contribute to the relationship. For instance, it is possible that differential access between states to new diagnostic tests and screening systems for norovirus and other AGE pathogens may explain some of the findings. Further research is needed to determine whether improvement in state guidelines and access to testing would lead to more favorable outcomes. Other factors that were not formally tested included whether different guidelines clustered geographically and whether guidelines specific to norovirus differed from guidelines meant for AGE outbreaks in general. Neither seems likely, as there was no readily apparent clustering among states in the categories analyzed in this study, and almost all of the general AGE guidelines still put greater emphasis on norovirus than other pathogens.

Given that 50% of states did not confirm or directly provide any guidelines for their states, there is some potential for bias in the results based on nonresponse. Among states that responded, 72% had guidelines available online, which is similar to the rate of guideline availability (75%) among nonresponding states. However, since additional unposted guidelines obtained from responding states increased the percentage with guidelines to 84%, it is possible that the five nonresponding states without online guidelines, if contacted, could have provided at least one or two more sets of guidelines for review, though their inclusion would be unlikely to substantially alter the study results. It is also possible that nonresponding states had additional or updated guidelines that were different from those available on their websites. However, since the publication dates were comparable between responding and nonresponding states and only one responding state stated that their online guidelines were out of date, the impact of these missing documents is likewise expected to be small.

AGE outbreak response guidelines are also developed and published at the local and territorial levels of public health, but this study is limited to guidelines at the state level. The focus on state guidelines restricted the data collection to a manageable scope, avoided language barriers (e.g., Puerto Rico has guidelines written in Spanish), and reduced variation due to differing government roles, while still covering a diverse range of U.S. regions and utilizing the centralized source of guidelines for the majority of jurisdictions. Not all states have written guidelines, so the full variability at any given time may not be fully captured by document review. Furthermore, several states reported that their guidelines were in the process of being updated. Among states with available guidelines, element categories were grouped generously, ignoring small deviations in

favor of grouping guidelines that were generally similar in nature. Thus, this study presents a conservative analysis of the variation in state guidelines at the time the study was conducted, and additional variation, though not documented, was certainly present and may be expected to change over time. However, there was no apparent convergence of the state guidelines over time, and similar patterns of variation were observed regardless of the dates indicated in the guidelines.

Implications for Policy and Practice

This study has strong implications for public health practice. Despite clear, evidence-based national guidance regarding norovirus and related AGE outbreaks in healthcare settings, most state health departments have developed their own guidelines which often deviate from those issued by HICPAC. Substantial variation also exists in other areas that are not directly addressed in the HICPAC guidelines. Variations in state guidelines are potentially associated with important outbreak characteristics that may influence the size and duration of an outbreak and whether similar outbreaks can be prevented in the future. The causes of the observed variation are unknown, but all states would likely benefit from reviewing their guidelines and addressing deviations from national recommendations. It is important for the public health community to have a broader conversation about how local guidelines are developed and the importance of adhering to national guidelines, or updating national guidelines if local practices are shown to be superior. Other potential next steps may include further investigation into topics with especially wide variation (such as stool specimen collection) and efforts to bridge gaps in the current national guidelines.

Conclusion

This study identified substantial variation in state healthcare-associated AGE outbreak response guidelines beyond that which would reasonably be expected based on national guidelines and the scientific evidence. Such variation must be considered when planning national CDS systems for outbreak response and other public health use cases. This variation appeared to be associated with differences in outbreak characteristics reported to NORS, including number of outbreaks, outbreak duration, case counts, modes of transmission, and etiology confirmation. These findings may be used to encourage public health officials to carefully consider any divergence from national guidelines as they develop and revise jurisdictional guidelines. More research is needed to understand why this variation exists, how it impacts outbreak outcomes, and where improvements in evidence-based recommendations and communication of national guidance are needed.

Acknowledgements

Outbreak investigation documents were verified or directly provided by public health officials from the Alabama Department of Public Health, Colorado Department of Public Health & Environment, Connecticut Department of Public Health, Florida Department of Health, Idaho Department of Health and Welfare, Indiana State Department of Health, Kentucky Department for Public Health, Louisiana Department of Health & Hospitals, Massachusetts Department of Public Health, Maine Center for Disease Control and Prevention, Minnesota Department of Health, Mississippi State Department of Health, Montana Department of Public Health and Human Services, New Hampshire Department of Health and Human Services, Nevada State Health Division,

New York State Department of Health, Oregon Health Authority, Rhode Island Department of Health, South Carolina Department of Health and Environmental Control, South Dakota Department of Health, Tennessee Department of Health, Utah Department of Health, Virginia Department of Health, Washington State Department of Health, and Wyoming Department of Health. All other state guidelines were obtained from state health department websites. Data for other analyses were provided by the National Outbreak Reporting System Team at the U.S. Centers for Disease Control and Prevention.

Table 3.1. Numbers of states with public health recommendations aligned with CDC categories for prevention and control during acute gastroenteritis outbreaks in healthcare settings (*N* = 41 states with available guidelines).

Recommendations for prevention and control	Count (%)
(1) Patient cohorting & isolation	40 (98)
(2) Hand hygiene	41 (100)
(3) Patient transfers & ward closure	29 (71)
(4) Food handlers	37 (90)
(5) Diagnostics	38 (93)
(6) Personal protective equipment	37 (90)
(7) Environmental cleaning	39 (95)
(8) Staff leave & policy	41 (100)
(9) Visitors	34 (83)
(10) Education	35 (85)
(11) Active case-finding	32 (78)
(12) Communication & notification	39 (95)

Table 3.2. Numbers of states with recommendations for duration of patient isolation and staff exclusion (*N* = 41 states with available guidelines).

Recommendation for duration of ill patient isolation and ill staff exclusion	Patients Count (%)	Staff Count (%)
Recommendation provided	27 (66)	40 (98)
Until symptoms resolve	0 (0)	4 (10)
1 day after resolution of symptoms	2 (7)	4 (10)
1-3 days after resolution of symptoms	3 (11)	1 (3)
2 days after resolution of symptoms	13 (48)	17 (43)
2-3 days after resolution of symptoms	1 (4)	6 (15)
3 days after resolution of symptoms	8 (30)	8 (20)
No recommendation provided	14 (34)	1 (2)

Table 3.3. Recommendation categories found in at least 10 states' guidelines for healthcare-associated norovirus outbreak response (*N* = 41 states with available guidelines).

Recommendation category	Number of states with recommendations (%)	Total number of distinct recommendations
Ill patient isolation duration	27 (66)	5
Ill employee exclusion duration	40 (98)	6
Stool specimen collection ranges	36 (88)	21
Outbreak detection definition	40 (98)	14
Outbreak resolution definition	29 (71)	11
Mode of transmission determination	15 (37)	6
Environmental health assessment	13 (32)	6

Table 3.4. Description of norovirus outbreaks in healthcare settings reported to the National Outbreak Reporting System (NORS) in 2015, stratified by features of norovirus outbreak response guidance collected from state health departments in 2016.

Outbreak measure/recommendation (number of states)	Number of outbreaks	Value (%, rate, or mean)
Percentage with confirmed etiology (*N* = 35)	1,124	51.2%
No. of stool specimens to be collected: median of range at or above 5 (*N* = 13)[§]	369	61.8%
No. of stool specimens to be collected: median of range less than 5 (*N* = 13)	595	45.2%**
No guidance or unspecified number of specimens (*N* = 9)	160	48.8%*
Number of outbreaks per 100,000 pop. (*N* = 35)	1,124	0.43
Outbreak defined as 2+ epidemiologically-linked cases (*N* = 14)[§]	502	0.37
Outbreak defined as 3+ epidemiologically-linked cases, "more than expected", or suspicion (*N* = 21)	622	0.49**
Average number of cases per outbreak (*N* = 35)	1,124	31.8
Patient isolation for 2+ days after resolution (*N* = 19)[§]	631	30.1

Table 3.4 continued.

Outbreak measure/recommendation (number of states)	Number of outbreaks	Value (%, rate, or mean)
Patient isolation for < 2 days after resolution (*N* = 5)	191	31.6
No guidance on patient isolation (*N* = 11)	302	35.6*
Employee exclusion for 2+ days after resolution (*N* = 26)[§]	766	30.5
Employee exclusion for < 2 days after resolution (*N* = 7)	271	34.0*
No guidance on employee exclusion (*N* = 2)	—	—
Defined as 2+ epidemiologically-linked cases (*N* = 14)[§]	502	31.7
3+, variable, or unspecified definitions (*N* = 21)	622	32.0
Wait ≥ 2 incubation periods to declare an outbreak is resolved (*N* = 14)[§]	555	31.0
Wait < 2 incubation periods to declare an outbreak is resolved (*N* = 12)	445	32.5
No guidance on outbreak resolution (*N* = 9)	124	33.3

Table 3.4 continued.

Outbreak measure/recommendation (number of states)	Number of outbreaks	Value (%, rate, or mean)
Average outbreak duration in days (*N* = 35)	1,124	13.2
Patient isolation for 2+ days after resolution (*N* = 19)[§]	631	14.0
Patient isolation for < 2 days after resolution (*N* = 5)	191	10.8*
No guidance on patient isolation (*N* = 11)	302	13.4
Employee exclusion for 2+ days after resolution (*N* = 26)[§]	766	13.8
Employee exclusion for < 2 days after resolution (*N* = 7)	271	11.9
No guidance on employee exclusion (*N* = 2)	—	—
Wait ≥ 2 incubation periods to declare an outbreak is resolved (*N* = 14)[§]	555	11.5
Wait < 2 incubation periods to declare an outbreak is resolved (*N* = 12)	445	15.4*
No guidance on outbreak resolution (*N* = 9)	124	12.6

[§]Reference recommendation – consistent with HICPAC and/or CDC guidance
**p*-value < 0.05 when compared to reference recommendation
***p*-value < 0.005 when compared to reference recommendation
—No comparisons made (< 5 states with recommendation)

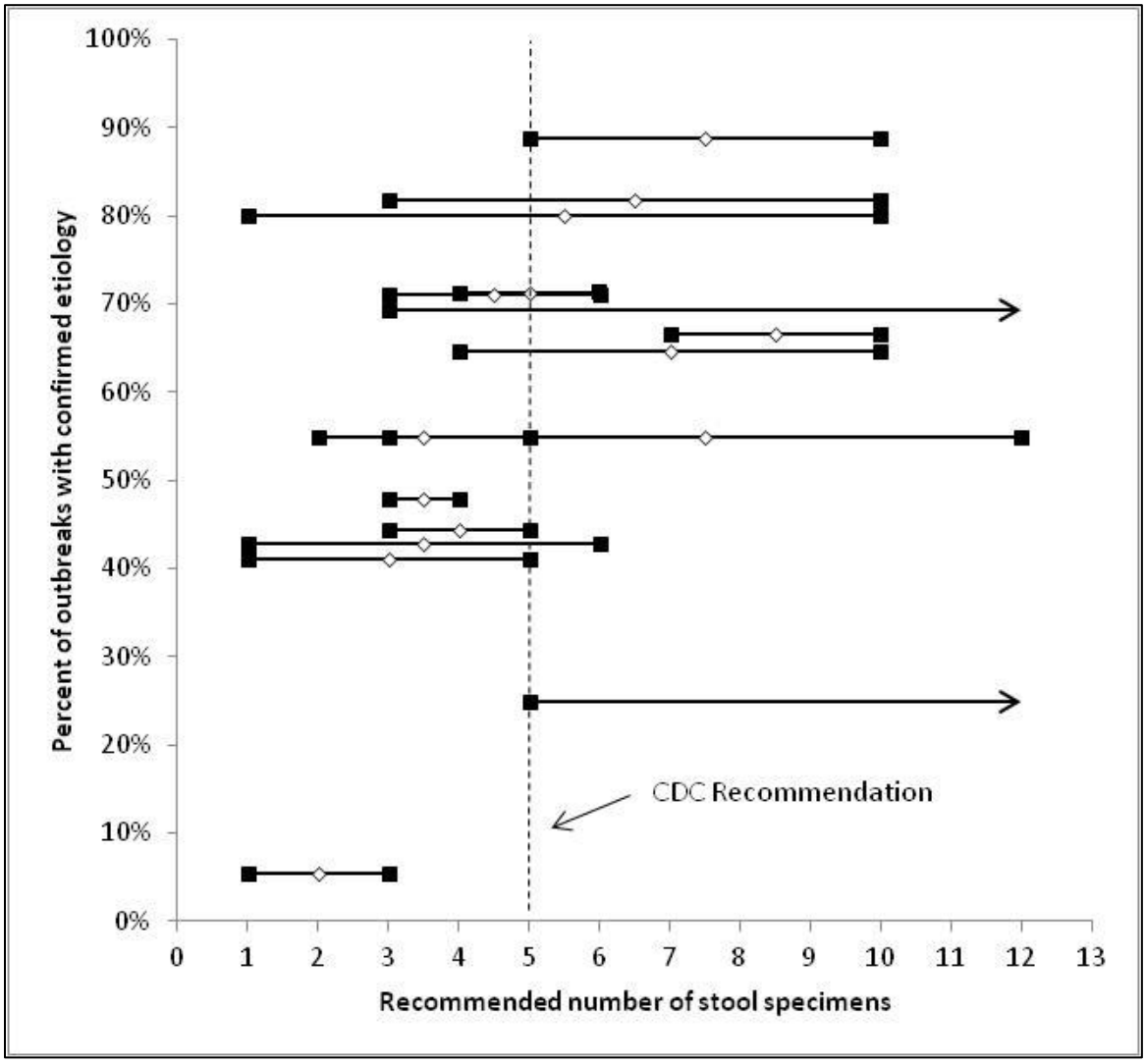


Figure 3.1. Percent of outbreaks with confirmed etiology reported to the National Outbreak Reporting System in 2015 by recommended ranges for stool specimen collection obtained from state health departments in 2016 (diamonds indicate range medians; arrows indicate ranges with no upper bounds) (*N* = 16 states with available outbreak data). Correlation coefficient = 0.71 (*p* = 0.004) for stool specimen range midpoints and percent of outbreaks with confirmed etiology (*N* = 14).

References

1. Mac Kenzie WR, Davidson AJ, Wiesenthal A, Engel JP, Turner K, Conn L, et al. The promise of electronic case reporting. *Public Health Rep* 2016;131(6):742-746.

2. Office of the Assistant Secretary for Health. Public Health 3.0: A Call to Action to Create a 21st Century Public Health Infrastructure. U.S. Department of Health and Human Services; 2016.

3. Li Y, Fang L, Gao S, Wang Z, Gao H, Liu P, et al. Decision support system for the response to infectious disease emergencies based on WebGIS and mobile services in China. *PLoS ONE* 2013;8(1):e54842.

4. MacCannel T, Umscheid CA, Agarwal, RK, Lee I, Kuntz G, Stevenson KB. Guideline for the prevention and control of norovirus gastroenteritis outbreaks in healthcare settings. *Infect Control Hosp Epidemiol* 2011;32(10):939-69.

5. Hall AJ, Lopman BA, Payne DC, Patel MM, Gastanaduy PA, Vinje J, et al. Norovirus disease in the United States. *Emerg Infect Diseases* 2013;19(8):1198-1205.

6. Hall AJ, Wikswo ME, Manikonda K, Roberts VA, Yoder JS, Gould LH. Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States. *Emerg Infect Diseases* 2013;19(8):1305-6.

7. Healthcare-Associated Infection Program. Norovirus Outbreak Detection and Management: Guidance for Long-Term Care Facilities. Arizona Department of Health Services; 2012.

8. Division of Communicable Disease Control. Recommendations for the Prevention and Control of Viral Gastroenteritis Outbreaks in California Long-Term Care Facilities. Richmond [CA]: California Department of Health Services; 2006.

9. Investigation and Management of Outbreaks of Known or Suspected Norovirus in Health Care and Residential Facilities. Colorado Department of Public Health & Environment; 2015.

10. Applewhite C, Weeks T. Foodborne Disease Outbreak Investigations: A Practical Guide for Local Health Departments. Version 3. Connecticut Department of Public Health; 2015.

11. Division of Public Health. Guidelines for Norovirus Prevention and Control in Long Term Care Facilities. Delaware Health and Social Services; 2011.

12. Recommendations for the Control of Viral Gastroenteritis Outbreaks in Long-

Term Care Facilities. Georgia Department of Public Health; Undated.

13. Center for Acute Disease Epidemiology. Guide to Surveillance, Investigations, and Reporting. Iowa Department of Public Health; 2014.

14. Division of Public Health. Idaho Investigative Guidelines: Norovirus. Idaho Department of Health and Welfare; 2015.

15. Conway JA. Guideline for the Prevention and Control of Viral Acute Gastroenteritis Outbreaks in Illinois Long-Term Care Facilities, 2012. Illinois Department of Public Health; 2012.

16. Stauffer J. Indiana Epidemiology Newsletter: Norovirus Outbreak Prevention and Control in Long-Term Care Settings. Indiana State Department of Health; 2014.

17. Gastrointestinal Virus Infection Control Measures: Healthcare Facilities. Indianapolis [IN]: Indiana State Department of Health; 2014.

18. Norovirus Outbreak Response Guide. Revision 3. Kentucky Department for Public Health; 2014.

19. Infectious Disease Epidemiology Section. Norovirus Infections. Louisiana Department of Health & Hospitals; 2013.

20. Infectious Disease Epidemiology Section. Guidelines for the Control of a Suspected or Confirmed Outbreak of Viral Gastroenteritis in a Long Term Care Facility. Louisiana Department of Health & Hospitals; 2010.

21. Division of Epidemiology and Immunization. Gastrointestinal Disease: Infection Control Guidelines for Long-Term Care Facilities. Massachusetts Department of Public Health; 2009.

22. Bureau of Communicable Disease Control. Norovirus & Other Caliciviruses. Massachusetts Department of Public Health; 2006.

23. Guidelines for the Epidemiological Investigation of Gastroenteritis Outbreaks in Long Term Care Facilities. Maryland Department of Health and Mental Hygiene; 2001.

24. Norovirus Outbreak Guidance for Long Term Care. Augusta [ME]: Maine Center for Disease Control and Prevention; 2015.

25. Communicable Disease Division. Guidelines for the Control of a Suspected or Confirmed Outbreak of Viral Gastroenteritis in a Nursing Home. Lansing [MI]: Michigan Department of Community Health; 2015.

26. Communicable Disease Division. Norovirus Investigations in Long Term Care Facilities: Things to Consider. Lansing [MI]: Michigan Department of Community Health; 2015.

27. Checklist for Nursing Homes/Long Term Care Facilities When Outbreaks of Gastroenteritis Are Suspected. Minnesota Department of Health; Undated.

28. Division of Community and Public Health. Communicable Disease Investigation Reference Manual: Outbreak Investigation: Acute Gastroenteritis. Missouri Department of Health and Senior Services; 2012.

29. Office of Communicable Diseases. Epidemiology Manual: Guidelines for Surveillance Reporting and Control. Jackson [MS]: Mississippi State Department of Health; 2016.

30. Norovirus Outbreak Management and Disease Prevention Guidelines: Health Care Settings. Montana Department of Public Health and Human Services; 2016.

31. North Carolina Statewide Program for Infection Control and Epidemiology. Control of Suspected Norovirus Outbreaks in Long Term Care Facilities: N.C. Public Health Recommendations. North Carolina Division of Public Health; 2015.

32. Division of Disease Control. What Do I Need To Know? Norovirus in Health-Care Facilities. North Dakota Department of Health; 2006.

33. Schaefer J, Safranek TJ. Health Alert Network Update: Preventing and Containing Norovirus Outbreaks. Nebraska Department of Health and Human Services; 2013.

34. NH Division of Public Health Services. NH DHHS Health Alert: Increase in Norovirus Activity in New Hampshire. Concord [NH]: New Hampshire Department of Health and Human Services; 2016.

35. Communicable Disease Service. Guidelines for the Control of Gastroenteritis Outbreaks in Long-Term Care and Other Institutional Settings. New Jersey Department of Health and Senior Services; 2011.

36. Norovirus in Long Term Care Facilities: Outbreak Checklist. New Mexico Department of Health; Undated.

37. Bureau of Health Statistics, Planning, Epidemiology & Response. Norovirus Management Toolkit: Outbreak Preparedness & Response. Version 1.1. Nevada State Health Division; 2012.

38. Enos L, Peek J, Graber P, Gomez M, Sowadsky R, Urrutia M. Norovirus in

Skilled Nursing Facilities in Nevada: 2012 Update. Nevada State Health Division; 2012.

39. Bureau of Health Care Quality and Compliance. Recommendations for Preventing and Controlling Viral Gastroenteritis Outbreaks in Medical and Residential Care Facilities. Nevada State Health Division; Undated.

40. Bureau of Communicable Disease Control. Health Advisory: Prevention and Control of Nosocomial Gastrointestinal Illness Outbreaks. Albany [NY]: New York State Department of Health; 2006.

41. General Infection Prevention and Control Recommendations: Residential and Long-term Care Facilities. Ohio Department of Health; 2012.

42. Acute Disease Service. Public Health Fact Sheet: Norovirus Infection Control Recommendations for Healthcare, Long-Term Care, and Residential Care Settings. Oklahoma State Department of Health; 2011.

43. Oregon State Public Health Division. Norovirus Outbreak Detection and Management: Instruction for Long-Term Care Facilities. Portland [OR]: Oregon Health Authority; 2014.

44. Center for Acute Infectious Disease Epidemiology. Gastrointestinal Illness Cluster/Outbreak Investigation in a Long-Term Care Facility. Version 1. Rhode Island Department of Health; 2016.

45. Guidance: Suspected Norovirus Outbreaks in Institutional Settings. South Carolina Department of Health and Environmental Control; 2014.

46. Communicable and Environmental Diseases and Emergency Preparedness. Recommendations for the Prevention and Control of Viral Gastroenteritis Outbreaks in Long-Term Care Facilities. Nashville [TN]: Tennessee Department of Health; 2013.

47. Infectious Disease Control Unit. The Emerging and Acute Infectious Disease Guidelines: Norovirus Outbreaks. Texas Department of State Health Services; 2015.

48. Utah Public Health - Disease Investigation Plans: Norovirus (Norwalk-Like Virus). Utah Department of Health; 2010.

49. Norovirus Prevention Guidance for Institutions/Facilities. Virginia Department of Health; 2013.

50. Office of Epidemiology. Norovirus Outbreaks. Virginia Department of Health;

2014.

51. Kelso P. Health Advisory: Norovirus Update. Vermont Department of Health; 2012.

52. Bureau of Communicable Diseases and Emergency Response. Recommendations for the Prevention and Control of Viral Gastroenteritis Outbreaks in Wisconsin Long-Term Care Facilities. Wisconsin Division of Public Health; 2009.

53. West Virginia Bureau for Public Health. Guidelines for Suspected Norovirus Outbreaks in Healthcare Facilities. West Virginia Department of Health & Human Services; Undated.

54. Epidemiology Unit. Guidelines for the Control of a Suspected or Confirmed Outbreak of Viral Gastroenteritis in a Nursing Home or Institution. Wyoming Department of Health; Undated.

55. Prevention Strategies for Seasonal Influenza in Healthcare Settings [Internet]. Atlanta [GA]: Centers for Disease Control and Prevention [updated 2016 May 26; cited July 3, 2016]. Available from: http://www.cdc.gov/flu/professionals/infectioncontrol/healthcaresettings.htm.

56. Hall AJ, Vinjé J, Lopman BA, Park GW, Yen C, Gregoricus N, Parashar U. Updated norovirus outbreak management and disease prevention guidelines. *MMWR Morb Mortal Wkly Rep* 2011;60(RR03):1-15.

57. NORS team. National Outbreak Reporting System (NORS) Guidance. Atlanta [GA]: Centers for Disease Control and Prevention; 2017, p. 4.

58. Faulkner A, Skoff T, Cassiday P, Tondella ML, Liang J. Measles. In: Rousch SW, Baldy LM, editors. Manual for the Surveillance of Vaccine-Preventable Diseases. Atlanta [GA]: Centers for Disease Control and Prevention; 2008, p. 7.11.

59. Population Division. American Community Survey. Annual Estimates of the Resident Population: April 1, 2010 to July 1, 2016 [Internet]. U.S. Census Bureau, 2016 [released 2017 Mar; cited 2017 Jul]. Available from: https://factfinder.census.gov/.

60. 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.

CHAPTER 4

MATHEMATICAL MODELING TO SUPPORT TRANSMISSION MODE CLASSIFICATION FOR ACUTE GASTROENTERITIS OUTBREAK RESPONSE AND REPORTING

Introduction

Acute gastroenteritis (AGE) outbreaks are a major cause of illness in the U.S., and most outbreaks are caused by person-to-person transmission of norovirus.[1] Recommended interventions to control AGEs vary based on the pathogen and mode of transmission,[2,3] but many outbreaks reported to the Centers for Disease Control and Prevention (CDC) are never confirmed to have been caused by a specific agent or mode of transmission.[1] This lack of information is challenging for public health officials making decisions about how to respond to outbreaks, especially early in outbreaks when only limited information may be available. Early identification of point source transmission events, such as a contaminated food, fomite, or highly infectious individual, allows an outbreak to be flagged for a more targeted public health response than would occur for outbreaks that are assumed to be transmitted person-to-person.[2]

The National Outbreak Reporting System (NORS) was launched in February 2009, and all local, state, and territorial health departments are encouraged to report all

acute gastroenteritis outbreaks through an online portal.[1] AGE outbreaks are defined by NORS as 2 or more cases of similar illness epidemiologically linked to a common exposure. The four required fields for reporting an outbreak to NORS are the primary mode of transmission, date the first case became ill, reporting state, and the estimated total number of primary cases.[4] The infection attack rate (*number infected*/*number exposed*) is an important statistic used to characterize an outbreak,[4] and thus the number exposed, or size of the population affected by an outbreak, is often also reported or at least estimated. Due to their importance in outbreak characterization and outbreak reporting, these data fields may often be collected at the initial report of the outbreak to public health. The inclusion of additional detailed information on the setting, contributing factors, implicated food vehicles, etc., is also strongly encouraged.[4]

The mode of transmission selected by an investigator completing the NORS web-based reporting forms determines which additional data elements are presented to the user of the input screen. The mode of transmission may be identified as "foodborne," "waterborne," "animal contact," "person-to-person," "environmental contamination," or "indeterminate/unknown." More generally, the mode of transmission may be classified as either person-to-person, point source (including, animal contact, environmental contamination, foodborne, and waterborne), or unknown. Though it is a required field in NORS, the selection of the mode of transmission is largely based on the findings of the investigation and may be subject to error. For example, the NORS guidelines specify that transmission should be considered person-to-person "if most of the patients had known direct contact or likely had the opportunity for direct contact with one another" or where there is "evidence of ongoing, propagated transmission (i.e., non-point source)."[4] This

definition allows room for interpretation, and if a common food or environmental source is not readily identified, it could result in point source outbreaks being misclassified as person-to-person. There is some historical evidence of misclassification, as the CDC found a decline in foodborne outbreaks after 2009, when the option for modes of transmission other than foodborne was first introduced in the NORS reporting system.[5] This decline suggests that some outbreaks that would currently be classified as person-to-person in NORS were previously being classified as foodborne. It is possible that the inclusion of additional modes of transmission has improved the accuracy of transmission mode classification over time, but the potential for misclassification remains high, especially in settings such as long-term care facilities where point source transmission may not be easily recognized and modes of transmission other than person-to-person transmission may be plausible.

Mathematical models have been widely used for determining optimal control measures during outbreaks.[6] However, these models appear to have primarily addressed control measures that are to be implemented for known infectious agents during large-scale pandemics. Modeling for smaller, local outbreaks, which health departments are likely to face more frequently than pandemics and where the infectious agent and mode of transmission may not be known, are not well represented in the literature. However, the information that is expected to be known early in outbreaks (such as case count, affected population size, and number of days since the first onset of illness) is sufficient for the development of simple mathematical models for transmission. The results of such models can then be used to predict whether the observed outbreak data are consistent with person-to-person transmission or may indicate point source transmission. Assisting

investigators with identifying point source transmission is useful for guiding initial outbreak response, collection of needed data, and reporting of outbreaks to NORS. Thus, such models could provide important clinical decision support (CDS) for outbreak investigators.

This study sought to a) validate the ability of a mathematical model to detect point source transmission using outbreak data from NORS and b) estimate the number of norovirus outbreaks classified as person-to-person in NORS that have some evidence of point source transmission according to the model.

Methods

Outbreaks reported to NORS during 2009-2014 were obtained for analysis, including outbreaks due to animal contact, environmental contamination, foodborne transmission, person-to-person transmission, and indeterminate/other/unknown sources (Figure 4.1). This set of outbreaks does not include waterborne outbreaks because they are reported to a separate reporting system. Outbreaks due to foodborne sources and animal contact were excluded because detailed information required to develop the model, such as the total number at risk of exposure, was not systematically collected in NORS during the study period. In addition, outbreaks with affected population sizes larger than 5,000 persons were excluded due to concerns regarding reporting accuracy and a focus on more typical, smaller scale outbreaks. The resulting subset of NORS outbreaks with sufficient data for modeling is hereinafter referred to as "the NORS dataset."

NORS fields with potential relevance for mathematical modeling (mode of

transmission, incubation period, duration of symptoms, and size of the affected population) were analyzed for missing information, and the percentage of outbreaks with complete information was reported. The model was then developed and validated. Outbreaks were categorized by the primary mode of transmission reported to NORS, and a subset were reclassified for a subanalysis as described below. Confirmed and suspect etiologies were included in subanalyses, and outbreaks with multiple etiologies were attributed to norovirus if norovirus was included among the identified etiologies.

Mathematical Model Development

Basic reproduction number

The basic reproduction number (R_0) is an important measure of how easily a pathogen is transmitted person-to-person. Becker's method[7] estimates the basic reproduction number R_0 using only three statistics from an outbreak: the final number of cases (C), the number susceptible before the outbreak (S), and the total population size (N). The estimate, as simplified by Mossong and Muller,[8] is given by

$$\hat{R}_0 = \frac{N-1}{C} \sum_{i=S-C+1}^{S} \frac{1}{i} \qquad (1)$$

with standard error

$$\mathrm{SE}(\hat{R}_0) = \frac{N-1}{C} \left(\sum_{i=S-C+1}^{S} \frac{1}{i^2} + \frac{C\hat{R}_0{}^2}{(N-1)^2} \right)^{\frac{1}{2}}. \qquad (2)$$

These formulae were used to estimate the basic reproduction number for each individual in each outbreak, with $S = N - 1$ in the equations due to the assumptions of

100% susceptibility and a single initial case. To simulate individual-level heterogeneity in transmissibility, each susceptible individual in each outbreak was assigned a unique basic reproduction number sampled from a gamma distribution with mean and standard deviation set to the population-wide value estimated for the outbreak using equations 1 and 2. This formulation prevented most individuals (99%) from having basic reproduction numbers more than double a typical basic reproduction number for norovirus,[9] consistent with a "low-transmission" setting for norovirus.[10]

Model parameters and structure

Given the limited nature of the available data during AGE outbreaks, simplifying assumptions were necessary in the development of the mathematical model. A susceptible-exposed-infectious-recovered (SEIR) model structure was selected due to the clinical course followed by AGE infections, with transitions occurring daily for simplicity. Parameters for the incubation period ($1/\sigma$, where σ is the daily transition probability between exposed and infectious) and duration of symptoms ($1/\gamma$, where γ is the daily transition probability between infectious and recovered) were approximated based on estimates from the CDC for norovirus.[3] These parameters were rounded to the nearest full day for simplicity and consistency with NORS units for outbreak duration. The incubation period was allowed to vary uniformly between 1 and 2 days, and the duration of symptoms was allowed to vary uniformly between 1 and 3 days (Table 4.1) for each infected individual. The daily transition probabilities between disease states for these individuals were then derived by taking the inverse of the selected values for incubation period and duration of symptoms. The probability of an infectious contact

between individuals (γR_0, the daily transition probability between susceptible and exposed) was determined by multiplying the inverse of the duration of symptoms by an estimate of the basic reproduction number.[11] For each day of an outbreak, each individual in the affected population was assessed to determine, based on their current status (S, E, I, or R), their individual basic reproduction number, and the randomly selected incubation period and duration of symptoms, whether they would progress to the next disease state according to the transition probabilities described above.

Parameters for diseases other than norovirus were not included due to the relative abundance of norovirus outbreaks in NORS and the distinct clinical profile of norovirus that make it more easily distinguishable from other gastrointestinal pathogens in the absence of stool testing results.[12] The proportion of the population susceptible to infection (S) was assumed to be 100% due to findings that infection rates for norovirus approach 100% among older adults[13] and the frequency of reported outbreaks in the NORS dataset where all at-risk subjects were reported to have become ill. For simplicity, a single initial case, absence of asymptomatic shedding, and closed populations were all assumed, and no distinctions were made between employees and nonemployees in relevant settings.

Model estimates

Monte Carlo simulations, with values sampled from probability distributions instead of relying on fixed parameter estimates, were used to better capture the random variation expected in small populations.[14] The model was run 1,000 times for each outbreak in the NORS dataset. The mean of these model runs served as an estimate of the

final case count expected for the outbreak under the model assumptions. An interval containing 95% of simulated final case counts was built around this estimate using the 2.5^{th} and 97.5^{th} quantiles as the lower and upper limits, respectively. If the actual final case count (*C*) fell within this interval, an outbreak was considered to be consistent with person-to-person transmission of norovirus. If the actual final case count fell outside the interval limits, the outbreak was considered to be inconsistent with person-to-person transmission of norovirus due to point source transmission by contaminated food, fomite, highly infectious individual, or other means, or to a pathogen with a clinical profile significantly different from norovirus.

Trace plots, mean plots, and autocorrelation plots for the estimated final case counts were visually inspected for a subset of outbreaks to ensure that each final case count converged to a fixed value. Simulations were run using R version 2.15.1.[15]

Model Validation Using Simulated Outbreaks

The internal validity of the modeling approach was evaluated by running additional simulated outbreaks with and without artificially-introduced point source transmission and then evaluating the performance of the model on those outbreaks. Artificial case counts were estimated using the affected population size and duration of each outbreak in the NORS dataset, the assumed incubation period and duration of symptoms for norovirus, and an assumed basic reproduction number of 1.64[9-10,16-17] for each case except the initial case. This value was considered typical of person-to-person transmission of norovirus in low-transmission settings.[9] The initial case acted as a "super-spreader" with a basic reproduction number one times (typical person-to-person

transmission), two times, four times, or six times the basic reproduction number of the other individuals in the models. Such "super-spreader" events would be expected to behave similarly to point source events with other causes, such as a vomiting incident in a common area or a contaminated food item. The resulting artificial case counts were then used to run simulations using the Becker method as described above to estimate the basic reproduction numbers, and the expected case count intervals for these simulations were compared to the artificial case counts. The specificity (for the typical person-to-person R_0) and sensitivity (for two, four, or six multiples of the typical R_0) of the model predictions were calculated, and the Wilson score method was used to estimate 95% confidence intervals for sensitivity estimates.

Model Validation Comparing Likely Point Source vs.

Person-to-Person Outbreaks

External validity was evaluated by comparing outbreaks with some reported evidence of point-source transmission to outbreaks with no such documentation. Though the majority of outbreaks in the NORS dataset were categorized as person-to-person, outbreaks were considered to be indicative of point-source transmission for this analysis if they had "Environmental contamination other than food/water" as a primary transmission mode, if they included "Food" or "Environmental contamination other than food/water" as secondary transmission modes, or if "Restaurant" was a primary or secondary setting. Other outbreaks from the NORS dataset that did not meet these criteria were considered to be eligible controls, and five control outbreaks were randomly selected for each point-source outbreak, matched by month and year of the initial

outbreak case. Sensitivity was calculated for the model outcomes as a measure of the model's ability to reliably categorize point-source outbreaks.

The control outbreaks were selected based on matching characteristics with the "likely" point source outbreaks and are not expected to be representative of the broader collection of outbreaks without documented evidence of point source transmission. Furthermore, it is impossible to know how representative these outbreaks are of non-point source outbreaks in general, since there is currently no definitive way to verify the modes of transmission in outbreaks reported to NORS. However, these outbreaks are still useful for comparison since they would be expected to be true person-to-person outbreaks more often than the outbreaks identified as likely point source.

Logistic regression was used to calculate odds ratios comparing the mathematical model predictions for the likely point-source outbreaks to those for the control outbreaks, with and without random effects included to account for within-subjects variation based on the outbreak matching by month and year. The ability to include random effects in the model made it possible to detect any temporal trends that might have differentially influenced the prevalence of outcomes in the control outbreaks. Model fit was assessed using the Akaike Information Criterion (AIC).

Assessment of Classification Among Reported Outbreaks

The model was used to classify all outbreaks with sufficient data in the NORS dataset, regardless of etiology. Though the model assumes the pathogen is norovirus, outbreak investigators may not know the specific pathogen at the onset of an outbreak, and these estimates are intended to reflect that practical reality. Estimates for the subset

of outbreaks attributable to norovirus and the subset of outbreaks attributable to other pathogens were also analyzed, though the model would not be expected to reliably detect point source transmission among outbreaks that are not attributable to norovirus. Among the overall dataset and the etiology subsets, the proportions of outbreaks determined to be inconsistent with person-to-person norovirus were based on the simple ratios of outbreaks with likely point source transmission indicated by the model to the total number of outbreaks available for modeling.

Results

Sample of Outbreaks Used in Analyses

Overall, 6,341 (47%) outbreaks had sufficient data for modeling (final case count, size of the affected population, and outbreak duration – Table 4.2). Eleven outbreaks had affected population sizes larger than 5,000 and were excluded, leaving 6,330 outbreaks available for modeling. Of these outbreaks, 5,955 (94%) had a primary mode of transmission reported as person-to-person, 8 (0%) were attributed to environmental contamination, and 367 (6%) were indeterminate (see Figure 4.1). Altogether, 4,163 (66%) had an assigned etiology, including 3,724 (89%) attributed to norovirus, and 3,552 (85%) attributed to person-to-person transmission of norovirus.

Of the 6,330 outbreaks with sufficient data, there were 146 outbreaks with documented evidence indicative of likely point source transmission (based on primary or secondary modes of transmission or a restaurant setting). These 146 outbreaks were reported in NORS with primary modes of transmission of person-to-person (N = 98), environmental contamination (N = 8), and indeterminate/other/unknown (N = 40). There

were 730 matched outbreaks without reported evidence of point source transmission randomly selected as controls, including outbreaks reported in NORS due to person-to-person (N = 679) and indeterminate/other/unknown (N = 51) transmission .

Model Validation and Assessment of Classification Among Reported Outbreaks

When point source events were introduced artificially, the accuracy of the mathematical model predictions increased with the artificial event size, with 7% sensitivity for outbreaks with simulated transmission events twice as large as would be expected for strictly person-to-person transmission, 64% for four times as large, and 78% for six times as large (see Figure 4.2). The specificity of the model was 99.99% for simulated outbreaks without the artificial introduction of a point source event (typical person-to-person R_0).

When applying the model to likely point source outbreaks, the odds of the model indicating some evidence of point source transmission were 3.2 (1.9-5.3) times higher for outbreaks with documented likely point source transmission reported in NORS than for outbreaks without such findings. The sensitivity of the model for detecting these likely point source outbreaks was 87% (Table 4.3). The logistic regression did not perform better with random effects based on outbreak matching included, indicating no temporal trends that needed to be accounted for, so random effects were removed from the final model.

When applied to the full dataset of outbreaks with data sufficient for modeling, the model indicated that at least 70% of person-to-person, environmental contamination,

and indeterminate/other/unknown outbreaks were inconsistent with person-to-person transmission of norovirus. When applied to the subset of 3,552 norovirus outbreaks attributed to person-to-person transmission, the model indicated that 73% of these outbreaks were inconsistent with person-to-person transmission of norovirus, and thus likely had point source transmission occur at some point during the outbreaks. When applied to the 611 outbreaks with confirmed or suspect pathogens other than norovirus or environmental contamination or unknown/indeterminate modes of transmission, only 35% were inconsistent with person-to-person transmission of norovirus according to the model.

Discussion

The mathematical model developed for this research performed adequately on measures of internal and external validity and indicated that nearly three-fourths of outbreaks reported as person-to-person exhibited some evidence of point source transmission. This large number of outbreaks with potential point source transmission identified among AGE outbreaks suggests a need for more reliable methods of transmission mode classification. While little was found in the literature to address this problem, the simple mathematical modeling approach in this paper appears to be adequate for use in public health decision support. The model detected simulated point source events with basic reproduction numbers below four times higher than expected. In contrast, the model sensitivity for the subset of outbreaks with documented likely point source transmission in NORS suggests that outbreak investigators detect point source events with R_0 above six times higher than expected, on average. The sensitivity of the

model was 78% for detecting simulated outbreaks with initial transmission event six times larger than expected versus 87% for detecting outbreaks with evidence of likely point source transmission documented by investigators in NORS. This improvement in detection would increase the number of likely point source outbreaks from 2.3% for the 146 likely point source outbreaks documented by investigators in NORS to 70% total indicated by the model. Such a large increase would be potentially burdensome for responding officials, but the almost one-third of outbreaks that were still identified as consistent with person-to-person transmission of norovirus indicates that the model would still be useful for triage.

There are several limitations associated with the modeling approach used in this study. The Becker method was selected due to practical concerns, but it is unlikely to be the most accurate estimate for the basic reproduction number. Methods that use more information should be more accurate and are advisable when detailed individual-level data are available. Homogeneous mixing is unlikely to accurately represent the true patterns of infectious contact in a population, and efforts to better characterize these patterns, such as distinguishing between residents and staff in settings like nursing homes, may be essential to acceptable model performance in such settings. Similarly, more precise values for the incubation period and duration of symptoms would likely result in more precise model results where such data are available.

Other factors not included in the model, such as new admissions and transfers in healthcare settings and asymptomatic shedding can also have important impacts on outbreak persistence. However, most of the settings covered in this study stop new admissions and transfers during outbreaks and increase environmental disinfection efforts

that would limit the impact of symptomatic and asymptomatic shedding.[3] Likewise, the individual basic reproduction numbers in the model varied according to the standard deviation of the population standard deviation estimates, which underestimates the variance that would be expected to actually occur among individuals. This assumption made the simulated outbreaks more consistent with "low-level" transmission of norovirus and thus more sensitive to events at the super-spreader level selected for the study. However, though this underestimation made the simulations more useful in practice, it also made them somewhat less reflective of reality. Despite these limitations, the model was designed with parsimony in mind, and we felt the disadvantages were outweighed by its portability and simplified communicability and implementation.

There were also limitations in the study designs used for validation. These limitations were mostly driven by the unavailability of data sufficient to readily answer the questions of interest. However, despite data not being systematically collected for animal contact and foodborne outbreaks and the absence of definitive means for identifying an absence of point source transmission, we were able to make assessments of validity by using simulations and secondary investigator findings documented in NORS. However, there were significant limitations to this approach. Simulations of point source transmission due to "super-spreaders" with variable basic reproduction numbers were used, but this is only one type of point source transmission. Structuring the simulations differently to imitate other means of point source transmission may lead to different results. Also, the sensitivity and specificity of the model simulations may have overstated the model performance due to the low variability in the estimates of the intentionally simplified model. Conversely, the odds ratio comparing model performance in outbreaks

with and without likely point source transmission was likely conservative due to the high potential for undetected point source events in the control outbreaks. Also, the model did not detect as many aberrations when applied to outbreaks caused by other pathogens or modes of transmission, likely due to the model assumption that the outbreak was caused by person-to-person transmission of norovirus. This discrepancy is likely due to differences in the prevalence of point source transmission among non-norovirus outbreaks reported to be person-to-person rather than to a failure of the model, which is not designed to discriminate between different types of person-to-person outbreaks. The inclusion of other pathogens in the model validation is another reason to conclude that the validation findings are likely conservative. More research needs to be done on how to improve the model to address this subset of outbreaks or to identify other means of distinguishing between potential pathogens when only limited information is available.

There are significant limitations in the interpretability of the model results. Strictly speaking, the model may only indicate whether or not an outbreak is inconsistent with person-to-person transmission of norovirus. An outbreak with a higher case count than that predicted by the model may be interpreted as having evidence of likely point source transmission, but this finding may also have other explanations. There may have been multiple initial infected individuals instead of one (e.g., if individuals were transferred from one facility to another at the same time), or the date of first onset used for the model may be later than the actual first onset date, which may not be known. If the actual case count falls below the lower limit for the model estimate, the outbreak may be caused by a pathogen other than norovirus or may be consistent with a continuing common source transmission type.[2] Missing cases may also explain this finding, though it

is unknown how often such missing data may occur or how much impact it would have on the model results. It is generally impossible to distinguish outbreaks with complete case counts from those missing affected individuals. The possibility that an outbreak indicated as consistent with person-to-person norovirus may still be caused by other pathogens/modes of transmission is another important limitation that must be factored into any outbreak response decisions informed by the model results. The model is intended to support decision-making, but caution would be required when using any model findings.

The variables necessary for estimating the basic reproduction number were collected in nearly half of outbreaks where it was possible for the data to be entered into NORS, suggesting that there are no substantial barriers preventing the use of the model for decision support in many investigating jurisdictions. Yet there are important considerations concerning the use of the model in public health practice. Becker's method was developed based on the final case count, but there may be little if any need for decision support after an outbreak has concluded. This approach was followed in this paper due to the nature of the NORS data used for validation, but it would have little utility in actual outbreak response. However, Becker suggested a solution to a similar problem regarding when the number susceptible diminishes to zero despite there still being infected individuals in the population. This occurrence would result in a cumulative amount of "wasted infectious period," which Becker recommended could be accounted for by subtracting the number of individuals in the last generation from the final case count in the formulae.[7] An adjustment similar to Becker's recommendation can be applied when the case count is obtained before the outbreak ends, assuming the number

currently symptomatic is known and adequately approximates the last generation of cases. Information on the number currently ill is not required when reporting to NORS, but such monitoring is encouraged in outbreak investigation guidelines.[3]

This adjustment would allow the model to be run at any point during an outbreak investigation, making it useful for decision support as early as the first report if the necessary information for running the model is collected. However, there is a tradeoff, as the smaller sample sizes would impact the reliability of the estimates, especially early in an outbreak. As jurisdictions gain experience using the model in their individual practice settings, thresholds can be developed for outbreak response triage that maximize the effectiveness of the model in those settings. Investigators can use wider percentile limits or cut-points in empirical *p*-values to limit the response or rank outbreaks according to agency capacity or preference.

Conclusion

This study demonstrates the potential utility of mathematical modeling for decision support in routine AGE outbreak response and reporting. A mathematical model developed to distinguish outbreaks caused by person-to-person transmission of norovirus from outbreaks with potential point source transmission performed well on measures of internal and external validity. The large number of outbreaks with previously unidentified potential point source transmission events identified by the model suggests a large potential impact of using the model for decision support in public health practice.

Acknowledgements

The outbreak data used to build and validate the models in this study were provided by the NORS team at CDC.

Table 4.1. List of mathematical model parameter values.

Parameter	Description	Value	Ref.
Δt	Time step	1 day	—
C	Final case count	From NORS	NORS team, 2017
N	Affected population size	From NORS	NORS team, 2017
S	Susceptible population size	$N - 1$	—
t_0	Date of first onset	From NORS	NORS team, 2017
t_N	Date of last onset	From NORS	NORS team, 2017
$t_N - t_0$	Duration of outbreak	$t_N - t_0$	—
$\hat{R}_0$	Estimated basic reproduction number	Estimated using Equation 1	Becker, 1989
R_0	Empirical basic reproduction number	1.64	Simmons et al., 2013
$1/\sigma$	Incubation period	1-2 days	Hall et al., 2011
$1/\gamma$	Duration of symptoms	1-3 days	Hall et al., 2011
B	Probability of contact between individuals	$\gamma \hat{R}_0$	Vynnycky and White, 2010

Table 4.2. Availability of data by mode of transmission for acute gastroenteritis outbreaks reported to the National Outbreak Reporting System, 2009-2014.

Variable	Number of outbreaks with complete data (%)		
	Person-to-Person	**Environmental Contamination**	**Indeterm./ Other/ Unknown**
Total outbreaks	11,447 (100)	55 (100)	1,961 (100)
Final case count	11,447 (100)	55 (100)	1,961 (100)
Date of first onset	11,447 (100)	55 (100)	1,961 (100)
Date of last onset	8,558 (75)	51 (93)	1,819 (93)
Confirmed or suspect pathogen	8,042 (70)	50 (91)	1,146 (58)
Number of exposed guests/staff	6,026 (53)	10 (18)	382 (19)
Adequate data for modeling	5,955 (52)	8 (15)	367 (19)
Outbreaks with > 5,000 exposed	9 (0)	0 (0)	2 (0)

Table 4.3. Mathematical model performance for person-to-person, environmental contamination, and indeterminate transmission mode outbreaks with and without documented evidence of likely point source transmission reported to the National Outbreak Reporting System during 2009-2014.

Model prediction	Likely point source outbreaks*	Control outbreaks	OR (95% CI)
Inconsistent with person-to-person norovirus	127	494	**3.2 (1.9-5.3)**
Consistent with person-to-person norovirus	19	236	1.0 (—)
Total	146	730	
Sensitivity	87%	—	

*Defined as outbreaks with environmental contamination as the primary mode of transmission and person-to-person or indeterminate transmission outbreaks with foodborne or environmental contamination as a secondary transmission mode or restaurant as a primary or secondary setting.

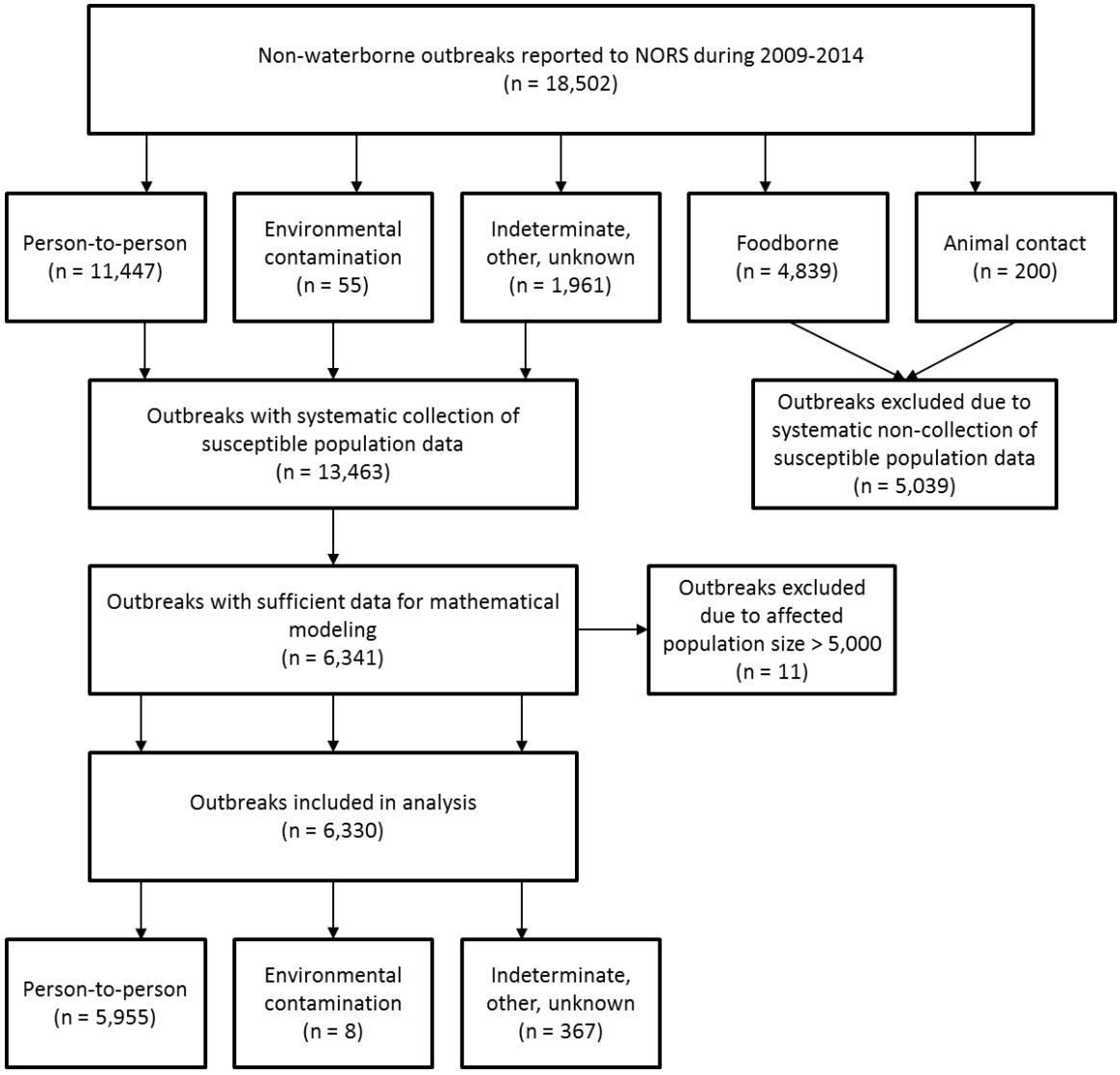


Figure 4.1. Description of nonwaterborne outbreaks from the National Outbreak Reporting System included in the analysis.

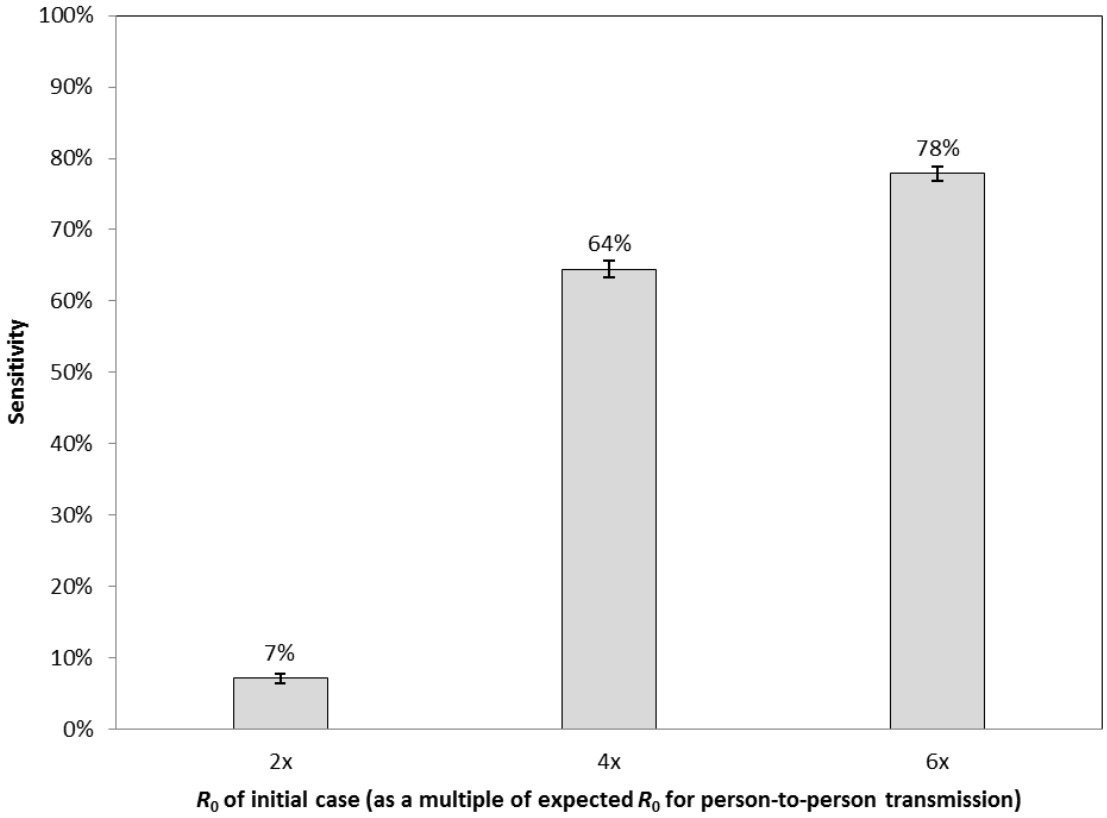


Figure 4.2. Sensitivity of a mathematical model for detecting artificially introduced super-spreader transmission events that were 2, 4, and 6 times the basic reproduction number of 1.64 associated with person-to-person transmission of norovirus for 6,330 acute gastroenteritis outbreaks reported to the National Outbreak Reporting System (NORS) during 2009-2014.

References

1. Hall A, Wikswo M, Manikonda K, Roberts V, Yoder J, Gould L. Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States. *Emerg Infect Diseases* 2013;19(8):1305-9.

2. Council to Improve Foodborne Outbreak Response (CIFOR). Guidelines for Foodborne Disease Outbreak Response. 2nd ed. Atlanta [GA]: Council of State and Territorial Epidemiologists; 2009.

3. Hall A, Vinje J, Lopman B, Park G, Yen C. Updated norovirus outbreak management and disease prevention guidelines. *MMWR Morb Mort Wkly Rep* 2011;60(RR03):1-15.

4. NORS team. National Outbreak Reporting System (NORS) Guidance. Version 2.1. Atlanta [GA]: Centers for Disease Control and Prevention; 2017.

5. Imanishi M, Manikonda K, Murthy B, Gould L. Factors contributing to decline in foodborne disease outbreak reports, United States. *Emerg Infect Diseases* 2014;20(9):1551-3.

6. Siettos C, Russo L. Mathematical modeling of infectious disease dynamics. *Virulence* 2013;4(4):295-306.

7. Becker, N. Analysis of Infectious Disease Data. New York [NY]: Chapman and Hall; 1989, pp. 147-50.

8. Mossong J, Muller C. Estimation of the basic reproduction number of measles during an outbreak in a partially vaccinated population. *Epidemiol Infect* 2000;124(2):273-8.

9. Simmons K, Gambhir M, Leon J, Lopman B. Duration of immunity to norovirus gastroenteritis. *Emerg Infect Diseases* 2013;19(8):1260-7.

10. Lopman B, Simmons K, Gambhir M, Vinje J, Parashar U. Epidemiologic implications of asymptomatic reinfection: a mathematical modeling study of norovirus. *Am J Epidemiol* 2013;179(4):507-12.

11. Vynnycky E, White R. An Introduction to Infectious Disease Modelling. New York [NY]: Oxford University Press; 2010.

12. Turcios R, Widdowson M, Sulka A, Mead P, Glass R. Reevaluation of epidemiological criteria for identifying outbreaks of acute gastroenteritis due to norovirus: United States, 1998-2000. *Clin Infect Diseases* 2006;42(7):964-9.

13. Simmons K, Gambhir M, Leon J, Lopman B. Duration of immunity to norovirus

gastroenteritis. *Emerg Infect Diseases* 2013;19(8):1260-1267.

14. van den Dool C, Bonten M, Hak E, Heijne J, Wallinga J. The effects of influenza vaccination of health care workers in nursing homes: insights from a mathematical model. *PLoS Med* 2008;5(10, article e200):1453-60.

15. R Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing; 2013. Available from: http://www.R-project.org/.

16. Bartsch S, Huang S, Wong K, Avery T, Lee B. The spread and control of norovirus outbreaks among hospitals in a region: a simulation model. *Open Forum Infect Dis* 2014;1(2):1-7.

17. Eden J, Chisholm R, Bull R, White P, Holmes E, Tanaka M. Persistent infections in immunocompromised hosts are rarely sources of new pathogen variants. *Virus Evol* 2017;3(2):1-10.

CHAPTER 5

IMPACT OF DATA AVAILABILITY ON PUBLIC HEALTH DECISION-MAKING DURING ACUTE GASTROENTERITIS OUTBREAKS: IMPLICATIONS FOR DECISION SUPPORT

Introduction

When facing an outbreak of infectious disease, there are several activities routinely pursued in the outbreak response, such as establishing the existence of the outbreak, verifying diagnoses, and developing and testing hypotheses to guide control measures.[1] The availability of information is vital to the completion of each of these steps. Unfortunately, the necessary information to respond appropriately is often unavailable to investigating officials.

Acute gastroenteritis (AGE) outbreaks are a major cause of illness in the U.S., commonly seen and investigated by public health departments.[2] Recommended control measures vary based on the pathogen and mode of transmission.[2,3] For example, environmental health assessments (EHA), or onsite visits to affected facilities by environmental health staff, are commonly utilized during point source outbreaks, where a single source, such as a food vehicle, fomite, or highly infectious individual is the cause of the outbreak. However, less than half of reported gastroenteritis outbreaks are

confirmed to have been caused by specific agents,[4] and the reported mode of transmission is often unconfirmed.[5] Of AGE outbreaks with laboratory-confirmed pathogens, most are caused by person-to-person transmission of norovirus, a highly infectious gastrointestinal pathogen.[2] Current guidelines recommend using Kaplan's Criteria for outbreaks in which such information remains unknown to determine whether the causal pathogen may be norovirus,[6] but the information necessary to apply Kaplan's Criteria has been found to be available in only 37% of outbreaks of unknown etiology.[7]

Several studies have shown that norovirus can be distinguished from other enteric pathogens by clinical attributes such as the incubation period and duration of illness.[7-10] Using Kaplan's Criteria[6] has been found to be relatively sensitive (68%) with very high specificity (99%).[7] Other techniques have also performed well, with successful attribution of outbreaks of unknown etiology to norovirus ranging from 73%[8] to 93%[9] or 94%,[10] though no record of these techniques being used in applied practice was readily found in the literature. Furthermore, each of these methods was applied to subsets of data containing information that a health department may not often have readily available at any point during an outbreak (such as Kaplan's Criteria[7]).

Mathematical modeling is an analytical process whereby "population parameters are described by symbols and linked by algebraic formulae" to "provide a realistic representation of the real world."[11] These models have been widely used in the study of human infectious diseases.[11] Mathematical models have also been widely used for identifying the best control measures to use during outbreaks.[12] The success of techniques that use clinical profiles for norovirus attribution suggests that the parameters for a mathematical model for person-to-person transmission of norovirus (such as incubation

period and duration of illness) can be expected to be relatively consistent between populations. This finding allows the potential for these parameters to be assumed where they cannot be collected as part of an investigation. Such assumptions enable a model to be run using only basic preliminary information such as the current number of cases, size of the at-risk population, and the number of days since the first onset of illness – which may be among the first items of information obtained during an outbreak investigation. Such a mathematical model would be able to predict with reasonable confidence whether the number of outbreak cases over time is consistent with person-to-person transmission of norovirus. The sensitivity and specificity of the model would be expected to approach that of clinical profile-based approaches, but while only requiring data that are available most or all of the time. However, even with high measures of validity, the willingness of public health practitioners to utilize such a tool for decision support is unknown. A clinical decision support system (CDS) with an acceptable user interface and adequate measures of usability would be necessary for mathematical modeling of person-to-person transmission of norovirus to be an effective public health resource in practice.

This study sought to investigate a) how outbreak investigators make decisions in the absence of detailed information, such as epidemic curves and the data required to apply Kaplan's Criteria; b) how the accuracy of those decisions changes when detailed information is available; and c) how decisions based on results from a CDS would compare to those based on more traditional detailed summary data in practice.

Methods

Participants

A convenience sample of participants was selected from two small rural local health departments (LHDs) (population < 100 000, population density < 1 000 per square mile), one medium-sized urban LHD (population = 100 000 to 500 000), one large urban LHD (population > 500 000), and a state department of health. The distribution of participants was intended to approximately reflect the numerical and geographical distribution of public health staff responsible for AGE outbreak response statewide. The study was also designed to support the generalizability of findings to health departments of different sizes and populations (Table 5.1). Participation requests with study information and participant consent forms were sent to supervisory-level staff at each site via email. Two additional AGE epidemiologists were invited to participate, but declined.

Interviews

All interviews were performed onsite at the respective health departments by CG. Only one participant was interviewed at a time, with only the interviewer and the individual being interviewed present.

Interviews were semistructured and consisted of two phases: an initial open-ended phase about general AGE outbreak response in participants' jurisdictions, and a second phase of hypothetical outbreak scenarios based on the literature and participants' immediate conclusions and response decisions in different scenario conditions. Different levels of information were presented for each scenario. Minimal summary data, as might be provided in an initial phone report, results from a prototype CDS developed to

determine the likely mode of transmission for norovirus outbreaks, and detailed summary data including an epidemic curve and data required to apply Kaplan's Criteria[6] were presented. Follow-up and clarifying questions were asked throughout both phases as needed. Detailed instructions were given before each phase, and all notes were recorded in a paper notebook by the interviewer.

Open-ended questions

For the open-ended phase, the participant was asked to consider a generic AGE outbreak report in their jurisdiction and then to answer four standard questions related to typical actions taken, decision-making, and data availability (Figure 5.1). These questions were designed to provide context for interpreting participant answers during the outbreak scenarios. Participant responses relevant to the study objectives were summarized.

Outbreak scenarios

The 10 outbreak scenarios were carefully selected from the literature,[13-22] with the distribution of settings intended to approximately represent that seen in practice[1] (Table 5.2). Outbreaks in restaurants were excluded, since such outbreaks are usually readily attributable to foodborne transmission[3] with little need for decision support. Other outbreak settings, such as petting zoos and water parks, were similarly excluded. Outbreaks were considered for inclusion if they were attributed at least in part to norovirus (the focus of the CDS) and if the daily case count (clearly identifiable in a table or epidemic curve in the outbreak report), number of individuals in the affected population, and outbreak duration were available (for running the CDS). Participants

were not given any information indicating the pathogen beyond the summary data contained in the scenarios. Data were summarized at the midpoint of the outbreaks to simulate the timing of outbreak reporting in practice.

Participants were instructed that each simulated outbreak report was for symptoms consistent with AGE. The data availability levels presented for each scenario included:

- **Minimal summary data.** A paragraph of text describing the setting, current case count (with separate numbers for residents/patients and staff where appropriate), size of the affected population, attack rate (*case count*/*size of the affected population*), number of days since first onset, and number currently ill. (All scenarios at this level also included the statement, "No food handlers are known to have been ill, and there are no other known common exposures");
- **CDS results.** A chart of predicted cumulative case counts (Figure 5.2), including the upper, median, and lower bounds for each simulated day of the outbreak under the assumption of person-to-person transmission of norovirus; a block of text listing the 95% confidence interval for the current day of the outbreak; the actual case count reported on that day; and a statement of "No evidence of point source transmission" or "Some evidence of point source transmission" depending on whether the actual case count fell within the 95% confidence interval for the predicted case count;
- **Detailed summary data.** A table including Kaplan's Criteria for determining if an outbreak is likely caused by norovirus[6] and outbreak summary data for

each criterion and a chart depicting the epidemic curve for the outbreak up to the current date (with separate numbers for residents/patients and staff where appropriate).

The minimal summary data were intended to represent the information that might be provided during an initial report of an outbreak from the public.[23-24] The statement about the absence of known food handlers or other common exposures was not taken from the outbreak reports in the literature, but was intended to represent the common situation faced by outbreak investigators where such information is unknown.

The order in which the outbreak scenarios and data availability levels were presented to participants was assigned using a random number generator such that four participants received the detailed summary data first and five received the CDS results first. Participants were given the scenarios as a stack of 30 sheets of paper, with one data availability level per sheet and three sheets per scenario. Each sheet included the text or figures for each scenario and data availability level as described above, together with three standard questions (Figure 5.1). The participant proceeded through the interview questions by flipping through the pages one sheet at a time. The minimal summary data was always presented first, followed by the mathematical model and detailed summary data in the assigned order from the randomization. Participants were told they could allow information on the previous one or two sheets for the same scenario to influence their decision-making when answering questions for subsequent data availability levels.

Clinical Decision Support System (CDS)

A detailed description of the mathematical model used to generate transmission mode and EHA predictions in the CDS is included in Chapter 4.

The model has been found to have sensitivity as high as 87% and specificity as high as 99% for outbreaks in the settings covered in this paper (Chapter 4). However, though it was validated on a population of outbreaks including those caused by norovirus and other pathogens, the model explicitly assumes norovirus as the causal pathogen and may not be appropriate for outbreaks caused by other pathogens. Despite this restriction to norovirus outbreaks and other characteristics specific to the CDS design, many of the study findings should still be loosely applicable to other public health decision support systems with similar performance and display of results.

The importance of the model in the context of this study concerns how the data required to run the model, including only the case count, size of the affected population, number currently ill, and number of days since the first onset were all included in the minimal summary data provided to the study participants. Thus any comparability to performance based on the detailed summary data would indicate an improvement in data efficiency, with more or equal information obtained from fewer data.

Data Analysis

Differences by data availability level and accuracy of the assumed pathogen, mode of transmission, and recommendation of EHA were evaluated using generalized linear mixed effects models (GLMM) logistic regression to account for within-subjects variation at the outbreak scenario and participant levels. Effect modification from the

study design was assessed by including interaction terms for the order of data level presentation (mathematical model data first vs. detailed summary data first) and the order of scenario presentation (presented in the first five scenarios vs. presented in the last five scenarios) to see if the inclusion of these variables significantly impacted the results. *P*-values less than 0.05 were considered statistically significant.

Human Participant Compliance Statement

The Institutional Review Board at the University of Utah reviewed the procedures outlined above and determined that the study did not meet the criteria for human subjects research.

Results

Participant Responses to Open-Ended and Scenario Questions

Participant responses regarding outbreak investigation activities, decision-making, and data covered all the "epidemiologic steps of an outbreak investigation."[25] However, there were several responses that gave insight into results specifically pertinent to this study, including pathogen identification, determining the mode of transmission, and deciding when to conduct an EHA.

Participant statements indicated an assumption of norovirus until more detailed information is available ("Treat as if it's norovirus," "Default to noro"), and there appeared to be some mental shortcuts used to determine the mode of transmission ("If we got a confirmed *Salmonella*, then it would be point source," "Long number of days, so probably person-to-person," "If there's only one event, it's likely a point source").

Decisions on when to conduct EHAs varied widely, even within participants from the same health department, with some recommending them very frequently ("We always send out EH," "Policy says we have to, because it has a kitchen," "I would let them know – might be time to inspect anyway") and some being much more hesitant ("Not until we know more," "It's always based on food handlers," "We usually wouldn't send them out for this"). In some cases, participants were deferential, especially for certain facilities ("I would offer, since it's a hospital," "Might not do an inspection if it's the last day of school," "It's the military, so I would only recommend it").

With regard to the CDS, some participants were skeptical at first ("This wouldn't change what I would do"), though participants seemed to become more comfortable with the CDS as the interview progressed ("I didn't notice the scale at first," "Okay, so it doesn't help with determining the pathogen," "Nice that it fits on this one"). There were no direct statements made about the user interface or any kind of heuristic assessment.

Outbreak Scenarios

The odds of attributing an outbreak to a pathogen other than norovirus ($p = 0.03$), of determining the mode of transmission to be point source ($p = 0.04$), and of conducting an EHA ($p = 0.001$) were all statistically significantly higher when participants had access to detailed summary data than when participants only had access to minimal summary data. The odds were also higher for determining the mode of transmission to be point source ($p = 0.002$) and conducting an EHA ($p = 0.003$) when participants had the CDS results, compared to the minimal summary data (Table 5.3). However, the odds of *correctly* identifying the reported mode of transmission were *lower* ($p = 0.03$) when the

detailed summary data were available than when only the minimal summary data were available, and neither level of data was associated with the determination of the mode of transmission or EHA reported in the literature. Both the CDS results and the detailed summary data had higher odds of agreeing with the CDS-predicted mode of transmission ($p < 0.001$ and $p = 0.004$, respectively) and recommendation to conduct an EHA ($p = 0.003$ and $p = 0.03$, respectively) (Table 5.4). When interaction terms were included in the model for scenario order and data level order, no interaction terms were statistically significant, and these terms were not included in the final model.

Discussion

This study identifies several important associations between data availability and AGE outbreak response decision-making. Outbreak investigators were more likely to attribute outbreaks to person-to-person transmission of norovirus and to not recommend an EHA when data availability was at the level of an initial report, indicating some bias on the part of investigators in such circumstances away from assumptions that may require more intensive interventions.[2] Though investigators were more likely to take action with more data availability, they were not more likely to identify the same mode of transmission or EHA determination as reported in the literature, and they were significantly less likely to identify the pathogen correctly in the case of the detailed summary data. The pathogen misclassification may be attributable in part to an assumed connection between point source transmission and pathogens other than norovirus indicated in the participants' open-ended responses. The lack of association with reported findings may be due to the absence of information used to make such decisions in the

detailed summary data or CDS results, such as ill food handlers or common exposures. This finding is consistent with the participants' open-ended responses, but the associations with the CDS predictions suggest that it may also be explained in part by possible misclassification of the outbreaks upon which the scenarios were based. Three outbreaks classified as person-to-person in the literature were classified as having some evidence of point source transmission by the CDS.

Participants with the CDS results or detailed summary data had higher odds of correctly determining the CDS-predicted mode of transmission and EHA determination. This finding might seem intuitive, since the CDS predictions were included in the mathematical model results given to participants, but the order in which the participants received the CDS data did not significantly modify the associations. This lack of interaction suggests that the CDS was able to steer decision-making among participants in ways similar to the detailed summary data (though the effect sizes were larger for the CDS) while only using data that are much more likely to be available in practice. It also suggests that participants were similarly likely to use the CDS results as they were the detailed summary data for decision support, and to use them to make decisions consistent with the CDS predictions, even when those results had not yet been presented. This apparent acceptance of the CDS results is an important finding for the CDS, but it also has implications for the future of decision support in outbreak response.

Limitations

While participants were not randomly selected, they represented 5 of 14 health departments in Utah and accounted for approximately one third of all AGE epidemiology

staff in the state. The study findings should be generalizable to similar public health officials in similar settings. Since the study was limited to health departments from Utah, results may be more useful in states with a decentralized public health infrastructure similar to Utah's, though other unknown idiosyncrasies in the Utah public health system may exist that would make some findings applicable only to Utah. The reliability of the expert interviews may have been affected by the varying levels of expertise of the study participants and the diverse populations they serve, but since they were all active AGE outbreak investigators, their perceptions of decision-making during outbreaks would necessarily be reflective of decisions made by public health personnel in the field.

The outbreak scenarios were also not randomly selected, but were intended to be representative of norovirus outbreaks in settings where the mode of transmission may not be easily observed, such as long-term care facilities, schools, and other institutions. However, it is possible that bias in the selection of outbreaks from the literature influenced the study findings. Also, only norovirus outbreaks were included in the scenarios due to the nature of the CDS used. Results may vary for outbreaks caused by other pathogens. However, there were no associations between the CDS results and the pathogen attribution or accuracy of the pathogen attribution.

Small sample size is an important limiting factor for this study. While several associations were statistically significant, other potential associations may not have been observed due to inadequate statistical power. Increasing the number of participants or the number of outbreak scenarios would have improved the statistical power, but were determined to be infeasible for this study.

Conclusion

This study demonstrates the impact of data availability on outbreak response, the potential for CDS to contribute meaningfully when data are sparse, and the acceptability of CDS to public health practitioners. More CDS development and research are needed in this area.

Implications for Policy and Practice

- Acute gastroenteritis outbreak investigators have diverse approaches to pathogen identification, transmission mode determination, and conducting environmental health assessments.
- When only minimal data are available, acute gastroenteritis outbreak investigators assume reported outbreaks are caused by norovirus, are spread by person-to-person transmission, and do not require an environmental health assessment.
- Mathematical modeling of acute gastroenteritis outbreaks may perform similarly to detailed summary data (epidemic curves, symptoms, etc.) for determining the mode of transmission and whether to conduct environmental health assessments using only data that are available earlier and more often.
- Data from decision support systems may be treated similarly to data from traditional sources by acute gastroenteritis outbreak investigators.

Table 5.1. Characteristics of study participants (*N* = 9).

Experience (years)	**Median (Range)**
Public health experience (years)	10 (5-29)
AGE epidemiology experience (years)	8 (2-20)
Professional characteristics	**Count (%)**
Supervises other staff	5 (56)
Degree/professional license*	9 (100)
Master of public health degree	7 (78)
Environmental health scientist	2 (22)
Registered nurse	1 (11)
Health Department size	
Small rural LHD (pop. < 100,000, pop. density < 1000/sq. mi.)	2 (22)
Medium urban LHD (pop. 100,000-500,000)	2 (22)
Large urban LHD (pop. > 500,000)	3 (33)
State health department	2 (22)

*Participants were allowed multiple responses.

Table 5.2. Characteristics of acute gastroenteritis outbreak scenarios (*N* = 10).

Characteristic	**Mean (Range)**
Case count	51.5 (18-139)
Number exposed	314 (75-1800)
Attack rate	0.21 (0.02-0.45)
Outbreak duration (days)	13 (6-17)
Setting	**Count (%)**
Long-term care facility	5 (50)
School (elementary school, university)	2 (20)
Other healthcare setting (hospital, psychiatric center)	2 (20)
Military training facility	1 (10)
Reported mode	
Person-to-person	5 (50)
Point source	5 (50)
Environmental contamination	2 (20)
Foodborne	2 (20)
Waterborne	1 (10)
Reported response	
EHA	6 (60)
No EHA	4 (40)
Model-predicted mode/response	
Point source/EHA	8 (80)
Person-to-person/No EHA	2 (20)

Table 5.3. Odds ratios (OR) for participant-predicted pathogen, mode of transmission, and response by level of information availability (*N* = 90 for each).

Data level	**Pathogen**		**OR (95% CI)**	***P*-value**
	Non-norovirus	Norovirus		
Minimal	13	77	1.0 (—)	—
Math model	16	74	1.4 (0.6-3.6)	0.47
Detailed	23	67	2.7 (1.1-6.9)	0.03*
	Mode			
	Point source	Person-to-person		
Minimal	39	51	1.0 (—)	—
Math model	59	31	2.8 (1.5-5.4)	0.002*
Detailed	52	38	1.9 (1.0-3.6)	0.04*
	Response			
	EHA	No EHA		
Minimal	51	39	1.0 (—)	—
Math model	65	25	3.7 (1.6-9.6)	0.003*
Detailed	67	23	4.6 (1.9-12.2)	0.001*

*Statistically significant at the 0.05 level

Table 5.4. Odds ratios (OR) for accuracy of participant-selected pathogen, mode of transmission, and response compared to reported epidemiologic findings and mathematical model predictions by level of information availability (*N* = 90 for each).

Data level	**Agreement with reported pathogen**		**OR (95% CI)**	***P*-value**
	Same	Different		
Minimal	77	13	1.0 (—)	—
Math model	74	16	0.7 (0.3-1.8)	0.47
Detailed	67	23	0.4 (0.1-0.9)	0.03*
	Agreement with reported mode			
	Same	Different		
Minimal	52	38	1.0 (—)	—
Math model	60	30	1.5 (0.8-2.8)	0.21
Detailed	57	33	1.3 (0.7-2.4)	0.44
	Agreement with reported response			
	Same	Different		
Minimal	51	39	1.0 (—)	—
Math model	53	37	1.1 (0.6-2.0)	0.76
Detailed	54	36	1.2 (0.6-2.1)	0.65

Table 5.4 continued.

Data level	**Agreement with model-predicted mode**		**OR (95% CI)**	***P*-value**
	Same	Different		
Minimal	41	49	1.0 (—)	—
Math model	69	21	4.1 (2.2-8.0)	< 0.001*
Detailed	60	30	2.5 (1.3-4.6)	0.004*
	Agreement with model-predicted response			
	Same	Different		
Minimal	47	43	1.0 (—)	—
Math model	65	25	2.7 (1.4-5.4)	0.003*
Detailed	60	30	2 (1.1-3.9)	0.03*

*Statistically significant at the 0.05 level

Open-ended questions for each participant

- "What specific actions would you take? Please describe a typical outbreak investigation in your jurisdiction from beginning to end."
- "What decisions would you have to make? Please describe your thought processes in detail, including what practices and policies would impact your decision-making."
- "What data would you need to make those decisions? What data would you usually expect to have available?"
- "What other factors might influence your decision-making process?"

Questions repeated for each outbreak scenario and data availability level for each participant

- "Based on these data, do you believe that the outbreak is more likely to be caused by norovirus or by some other pathogen? (Norovirus/Other pathogen)"
- "Based on these data, do you believe that the outbreak is more likely to be primarily caused by point source transmission or person-to-person transmission? (Point source/Person-to-Person)"
- "Based on whether or not you believe there is evidence that the outbreak is caused by point source transmission or a pathogen other than norovirus, would you advise conducting an onsite environmental health assessment at this time? (Yes/No)"

Figure 5.1. Interview questions.

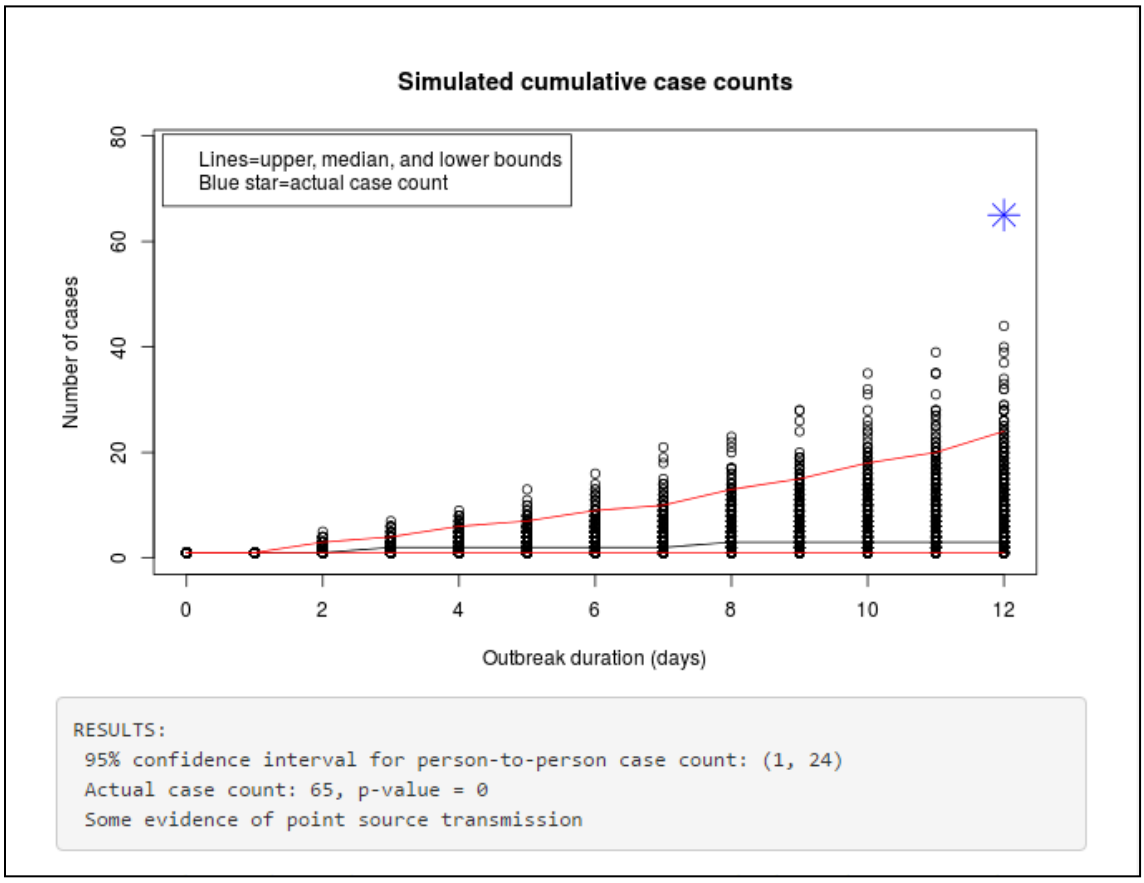


Figure 5.2. Example of decision support system results (additional background information is available in the Appendix).

References

1. Gregg M. Field Epidemiology. 3rd ed. New York [NY]: Oxford University Press; 2008.

2. Hall A, Vinje J, Lopman B, Park G, Yen C. Updated norovirus outbreak management and disease prevention guidelines. *MMWR Morb Mort Wkly Rep* 2011;60(RR03):1-15.

3. Council to Improve Foodborne Outbreak Response (CIFOR). Guidelines for Foodborne Disease Outbreak Response. Atlanta [GA]: Council of State and Territorial Epidemiologists; 2009.

4. Hall AJ, Wikswo ME, Manikonda K, Roberts VA, Yoder JS, Gould L. Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States. *Emerg Infect Diseases* 2013;19(8):1305-9.

5. NORS team. National Outbreak Reporting System (NORS) Guidance. Centers for Disease Control and Prevention. Version 2.1. Atlanta [GA]: Centers for Disease Control and Prevention; 2017.

6. 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.

7. Turcios RM, Widdowson M, Sulka AC, Mead PS, Glass RI. Reevaluation of epidemiological criteria for identifying outbreaks of acute gastroenteritis due to norovirus: United States, 1998-2000. *Clin Infect Diseases* 2006;42(7):964-9.

8. Domínguez A, Broner S, Torner N, Martínez A, Jansà JM, Alvarez J, et al. Utility of clinical-epidemiological profiles in outbreaks of foodborne disease, Catalonia, 2002 through 2006. *J Food Protect* 2010;73(1):125-31.

9. Hedberg CW, Palazzi-Churas KL, Radke VJ, Selman CA, Tauxe RV. The use of clinical profiles in the investigation of foodborne outbreaks in restaurants: United States, 1982-1997. *Epidemiol Infect* 2008;136(1):65-72.

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

11. Vynnycky E, White RG. An Introduction to Infectious Disease Modelling. New York [NY]: Oxford University Press; 2010.

12. Siettos C, Russo L. Mathematical modeling of infectious disease dynamics.

Virulence 2013;4(4):295-306.

13. Mellou K, Sideroglou T, Potamiti-Komi M, Kokkinos P, Ziros P, Georgakopoulou T, et al. Epidemiological investigation of two parallel gastroenteritis outbreaks in school settings. *BMC Public Health* 2013;13:241.

14. Ludwig A, Sato K, Schirmer P, Maniar A, Lucero-Obusan C, Fleming C, et al. Concurrent outbreaks with co-infection of norovirus and *Clostridium difficile* in a long-term-care facility. *Epidemiol Infect* 2013;141:1598-603.

15. Lai C, Wanga Y, Wub C, Hung C, Jiang D, Wu F. A norovirus outbreak in a nursing home: Norovirus shedding time associated with age. *J Clin Virol* 2013;56:96–101.

16. Honish L, Talbot J, Dragon D, Utgoff D. Outbreak of norovirus gastroenteritis at a university student residence – Edmonton, Alberta, 2006. *Can Commun Dis Rep Weekly* 2008;34(4).

17. de Menezes F, Correa V, Franco F, Ribeiro MI, dos Santos Cardoso MF, Morillo SG, et al. An outbreak of norovirus infection in a long-term care facility in Brazil. *Einstein (Sao Paolo)* 2010;8(4 Pt 1):410-3.

18. Ohwaki K, Nagashima H, Aoki M, Aoki H, Yano E. A foodborne norovirus outbreak at a hospital and an attached long-term care facility. *Jpn J Infect Dis* 2009; 62:450-4.

19. Rondy M, Koopmans M, Rotsaert C, Van Loon T, Beljaars B, Van Dijk G, et al. Norovirus disease associated with excess mortality and use of statins: a retrospective cohort study of an outbreak following a pilgrimage to Lourdes. *Epidemiol Infect* 2011;139:453-63.

20. Odelin M, Ruel N, Berthelot P, Diana M-C, Blanchon M-A, Omar S, et al. Investigation d'une épidémie de gastro-entérite à norovirus dans un hôpital gériatrique. *Ann Biol Clin* 2006;64(2):141-7.

21. Cui C, Pan L, Wang Y, Xue C, Zhu W, Zhu L, et al. An outbreak of acute GII.17 norovirus gastroenteritis in a long-term care facility in China: the role of nursing assistants. *J Infect Public Health* 2017;10(6):725-9.

22. Ramos M, Soto G, Loayza L, Valle R, Luna G, Gonzalez S, et al. Brote de infección por norovirus del grupo II en un centro de entrenamiento military, Perú 2013. *Rev Peru Med Exp Salud Publica* 2015;32(1):87-92.

23. GI Initial Outbreak Report Form [Internet]. Placer County Public Health Department [cited 2017 Nov]. Available from http://www.placer.ca.gov/departments/hhs/~/~/media/AC8FCDE842D54AEA90E

3962B8A11885F.ashx.

24. Norovirus Outbreak Summary Form: Long Term Care Facilities. Williamson County and Cities Health District [cited 2017 Nov]. Available from http://www.wcchd.org/docs/Norovirus_Outbreak_Summary_Form.pdf.

25. Career Development Division. Principles of Epidemiology in Public Health Practice. 3rd ed. Atlanta [GA]: Centers for Disease Control and Prevention; 2006, pp. 6-8.

CHAPTER 6

CONCLUSION

Overview

Several important findings were uncovered through our research. First, there is an excessive amount of variation in state guidelines for norovirus outbreaks in healthcare settings, despite the release of national guidelines in 2011. This variation appears to be associated with outbreak outcomes, with states with guidelines that conflict with the national guidance having more outbreaks and higher case counts than other states. The amount of variation in stool specimen collection alone is an impactful finding that warrants a national conversation.

The mathematical model was found to predict point source transmission with high levels of internal and external validity. The difference in the modes of transmission predicted by the model and those reported by outbreak investigators indicates the high potential for improved classification if public health officials used such a tool for decision support and seek additional information to accurately classify outbreaks. The potential for the model to perform well at the beginning of an outbreak gives investigators the opportunity to make evidence-based decisions earlier and more accurately. The sizeable proportion of outbreaks that were still found to be consistent

with person-to-person transmission further suggests that the model would be useful for triage in distinguishing between outbreaks likely caused by point source transmission from those that are not. These findings are important for the field of applied epidemiology.

Finally, the epidemiologist interviews confirmed the diversity in responses when only limited data are available, with participants more likely to assume norovirus, person-to-person transmission, and not conduct an EHA when only minimal summary data were available. These findings suggest that many of an outbreak investigator's first decisions may be based more on predisposition and experience than on the data at hand. This opportunity to better leverage initial data supports the importance of CDS to help improve those early decisions, guide information gathering, and ultimately prevent more illness in the community. The lack of an interaction based on the type of detailed information presented (traditional or CDS) further supported the potential for CDS to play an increased role in outbreak response. This potential was confirmed by the ability of the CDS to perform similarly to traditional data based on much less information than would likely be available earlier in the outbreak. The potential was also confirmed by the similarity between participants' use of traditional data and the CDS results, suggesting the acceptability of CDS data to current practitioners.

Altogether, this dissertation research supported the growing importance of CDS in public health practice. Complicated policies, missing information, and investigator predispositions necessitate the development and application of ever-more-sophisticated CDS in AGE outbreak response and public health in general.

Uniqueness/Significance to Biomedical Informatics

The research we performed applied mathematical modeling techniques to an important use case in applied public health, and no similar studies were readily found in the literature. Outbreak management is a focus area of public health informatics,[1] and the proposed research aimed to contribute meaningfully to the management of acute gastroenteritis outbreaks, which constitute a sizeable portion of the public health burden caused by outbreaks. Similar methods can be used in the management of other communicable disease outbreaks, such as respiratory outbreaks or contagious skin conditions. Though this mathematical modeling approach could be useful in a variety of settings at different levels of government, this research is intended to be implemented where it would be most impactful – at the local health department level where actions first occur in response to outbreaks. The validation of the model against national data and the evaluation of the model as it is implemented in local health departments could further enable the results to have a broader impact as they are intended to be generalizable and portable to health departments across the country.

Future Directions

The results of the research we undertook are promising, but preliminary. Much remains to be understood about each of the research areas, and there are great opportunities for further work.

The next steps regarding variation in state guidelines covered efforts to better understand how local jurisdictions develop guidelines and to re-emphasize the importance of adhering to national guidelines. However, there is also a need to better

understand the relationship between the guidelines and the disparities in outbreak outcomes that were observed. Efforts need to be made to investigate if these associations remain after potential confounders are accounted for, and the causality of any remaining associations needs to be assessed. The potential for adjustments to the national guidelines should be explored, or even the potential impact of developing guidelines on how to develop guidelines or enhance national guidelines based on local outcomes.

With regard to the mathematical model, there is a similarly wide range of opportunities for continued research. The validation of the model was limited by the data available, but richer data could lead to more robust conclusions about the effectiveness of the model. It could also enable improved evaluation of the settings for which the model is most accurate. There could also be modifications to the model itself. The model was simple by design, but the incorporation of more complicated techniques to further optimize limited data has great potential to further improve the results. More sophisticated simulations, perhaps at the agent level, could likewise make the results more generalizable to the affected population and thus more accurate and useful.

The outbreak scenarios were conducted due to the inability to collect adequate real-world data to see how epidemiologists actually functioned in practice, which is an important limitation of the study we conducted. The real-world impact of data availability on outbreak response is still unknown, and our results only loosely approximated what that impact would be. Collecting and analyzing real-world data would help to clarify how these decisions are made in practice, and how data, or the lack thereof, may bias the results. The impact of CDS on decision-making should likewise be explored, and is suitable for further research.

References

1. Public Health Informatics [Internet]. American Medical Informatics Association [cited 2017 Jul]. Available from https://www.amia.org/applications-informatics/public-health-informatics.

APPENDIX: INTERVIEW TEXT

Interview Introductory Text

Hello _____,

The purpose of this study is to better understand the decision-making processes followed by public health officials during acute gastroenteritis (AGE) outbreaks. Please reflect upon your own experience as an AGE outbreak investigator as you respond to the following questions.

Consider the following scenario:

You receive a call from a concerned resident reporting an outbreak of acute gastroenteritis.

(Q1) What specific actions would you take? Please describe a typical outbreak investigation in your jurisdiction from beginning to end.

(Q2) What decisions would you have to make? Please describe your thought processes in detail, including what practices and policies would impact your decision-making.

(Q3) What data would you need to make those decisions? What data would you usually expect to have available?

(Q4) What other factors might influence your decision-making process?

Now you'll be asked to consider some scenarios with data provided from real acute gastroenteritis outbreaks. You will be given three levels of data for each outbreak:

minimal data, equivalent to basic information that might be reported on an initial telephone report of an outbreak; full data, including summary information based on detailed case-level interviews from the outbreak; and results from a mathematical model that simulates the outbreaks and predicts whether the observed case counts are consistent with person-to-person transmission of norovirus. The model has been found to have 87% sensitivity and an estimated specificity of nearly 100%. The statistical tests used are two-tailed, meaning that statistical significance is indicated when *p*-values are less than 0.025.

Please consider the provided information and describe how you would respond to each outbreak if it occurred in your jurisdiction. You may consider previous levels of data that pertain to the same outbreak in your decisions. Describe your initial impressions and what interventions you would implement, including the likely mode of transmission and whether or not you would conduct an onsite environmental health assessment. For the purposes of this study, an outbreak is considered to have a point source if any transmission other than person-to-person occurred, and an environmental health assessment is defined as any in-person visit by environmental health staff related to the outbreak investigation. Please don’t try to guess what the actual pathogens or modes of transmission are. Just respond as you would normally if these outbreaks occurred as part of your job.

Scenarios

You receive a phone call from an elementary school nurse. He tells you there are 51 ill students out of 113 students in the affected grades (attack rate = 45%). There were 2 initial cases, and there have been 7 days since they first became ill. There are 3 children currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.1. Minimal Data (Scenario ID=8346).

Table A.1. Kaplan's Criteria (Scenario ID=8346).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Median illness duration = 48 hours
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	77% of cases with vomiting
No bacterial agent found	No bacterial agent found
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

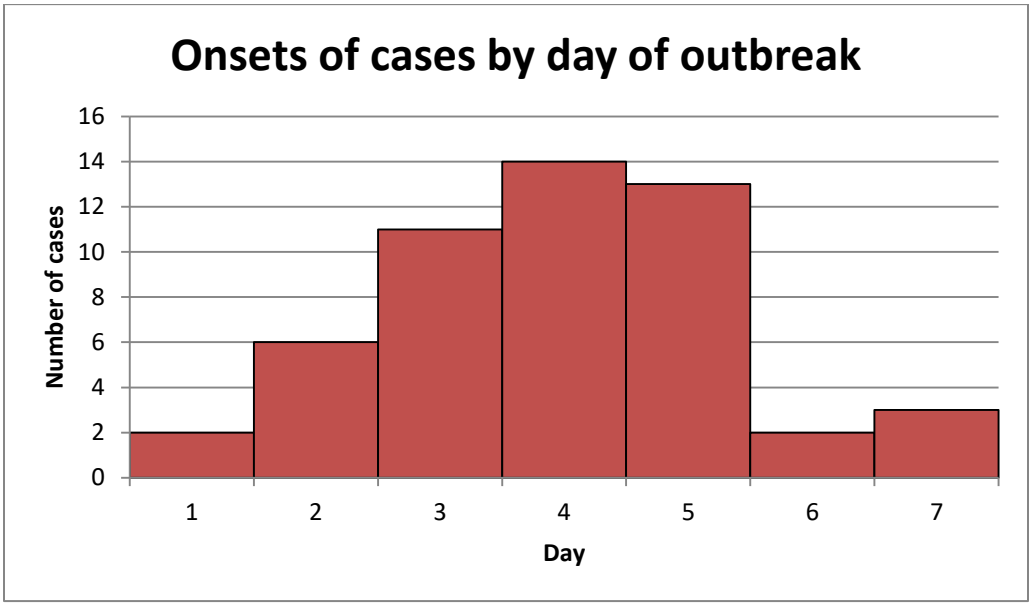
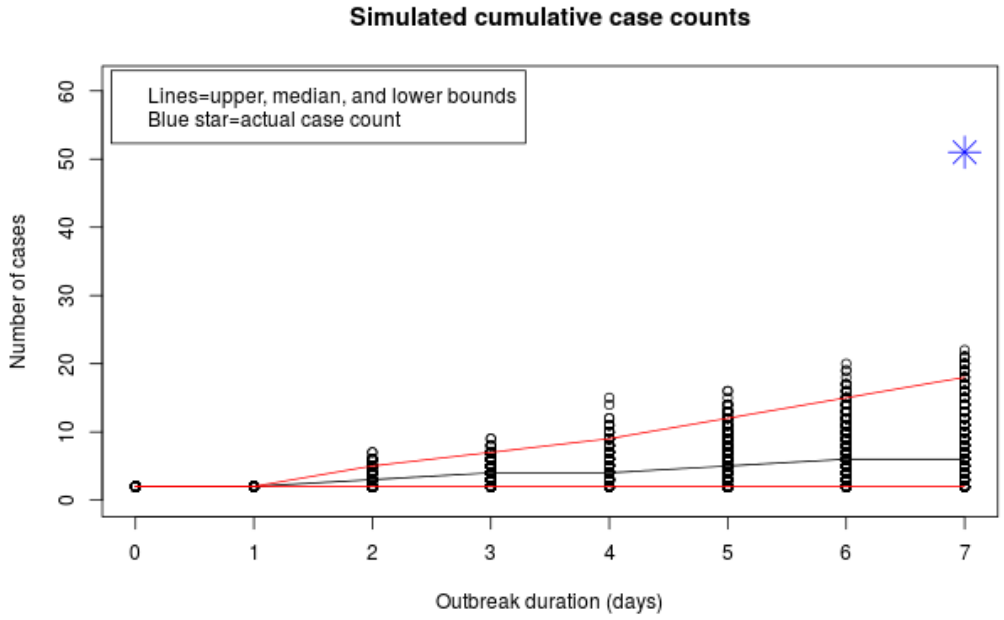


Figure A.2. Epidemic Curve (Scenario ID=8346).




```
RESULTS:
 95% confidence interval for person-to-person case count: (2, 18)
 Actual case count: 51, p-value = 0
 Some evidence of point source transmission
```

Figure A.3. Clinical Decision Support Results (Scenario ID=8346).

You receive a phone call from a nursing home administrator. He tells you there are 25 ill residents out of 195 total in the facility (attack rate = 13%) and 29 ill staff out of 97 total in the facility (attack rate = 30%). There were 2 initial cases, and it has been 19 days since they first became ill. There is 1 resident currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.4. Minimal Data (Scenario ID=8276).

Table A.2. Kaplan's Criteria (Scenario ID=8276).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Median illness duration = unknown
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	23% of cases with vomiting
No bacterial agent found	*C. difficile* isolated
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

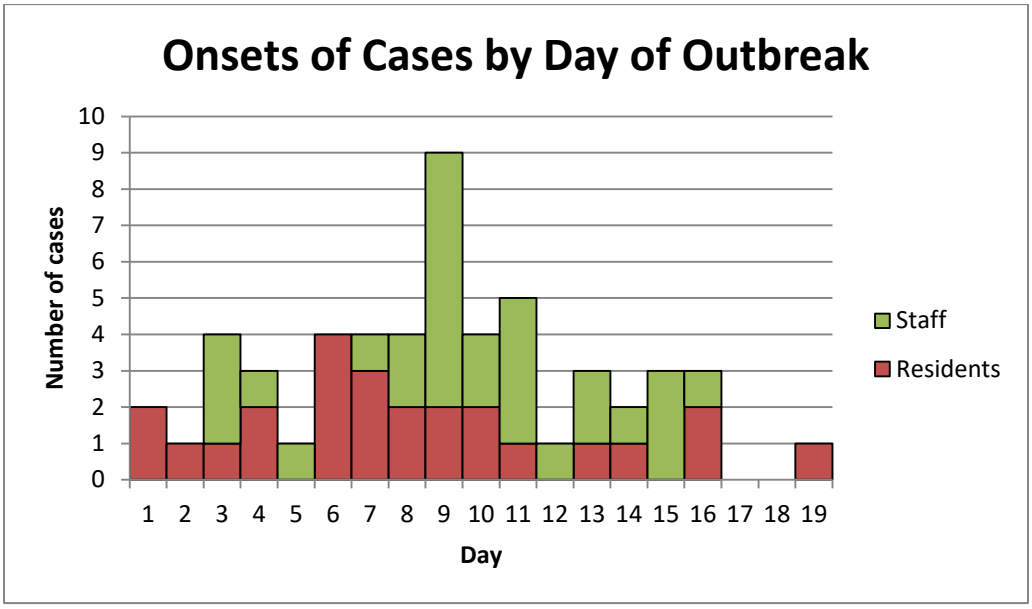
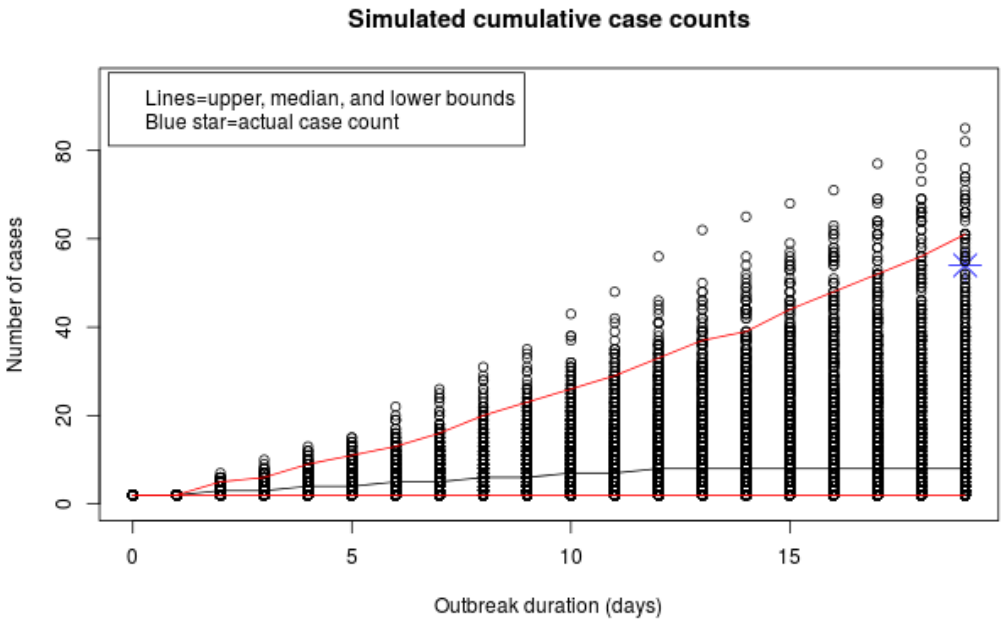


Figure A.5. Epidemic Curve (Scenario ID=8276).




```
RESULTS:
 95% confidence interval for person-to-person case count: (2, 61)
 Actual case count: 54, p-value = 0.028
 No evidence of point source transmission
```

Figure A.6. Clinical Decision Support Results (Scenario ID=8276).

You receive a phone call from a nursing home nurse supervisor. She tells you there are 12 ill residents out of 42 total in the facility (attack rate = 29%) and 6 ill staff out of 33 total in the facility (attack rate = 18%). It has been 6 days since the first resident became ill, and there are 7 residents and staff currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.7. Minimal Data (Scenario ID=4158).

Table A.3. Kaplan's Criteria (Scenario ID=4158).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Mean illness duration = 2.1 days
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	79% of cases with vomiting
No bacterial agent found	No bacterial agent found
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

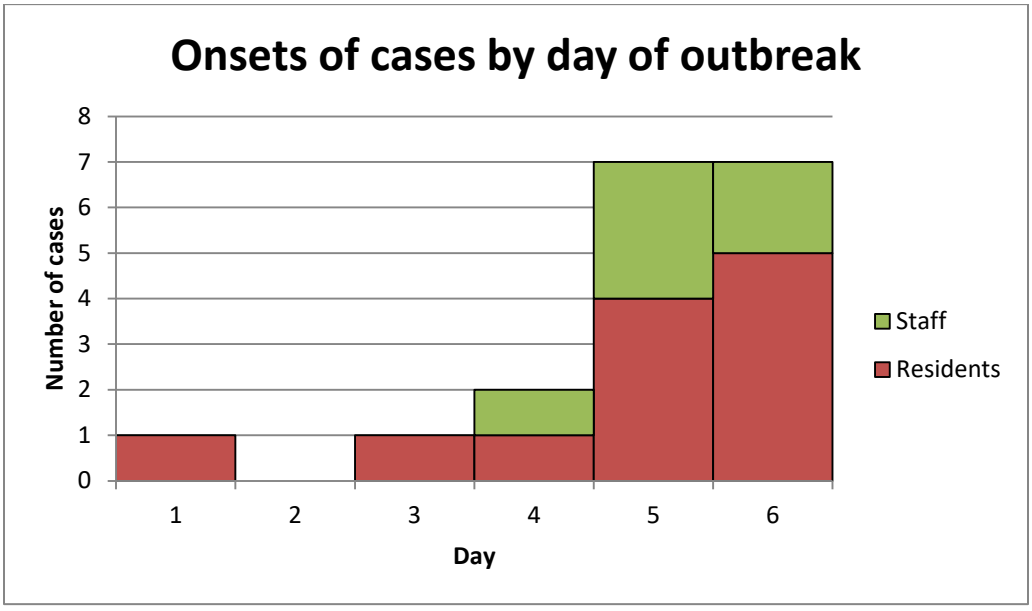
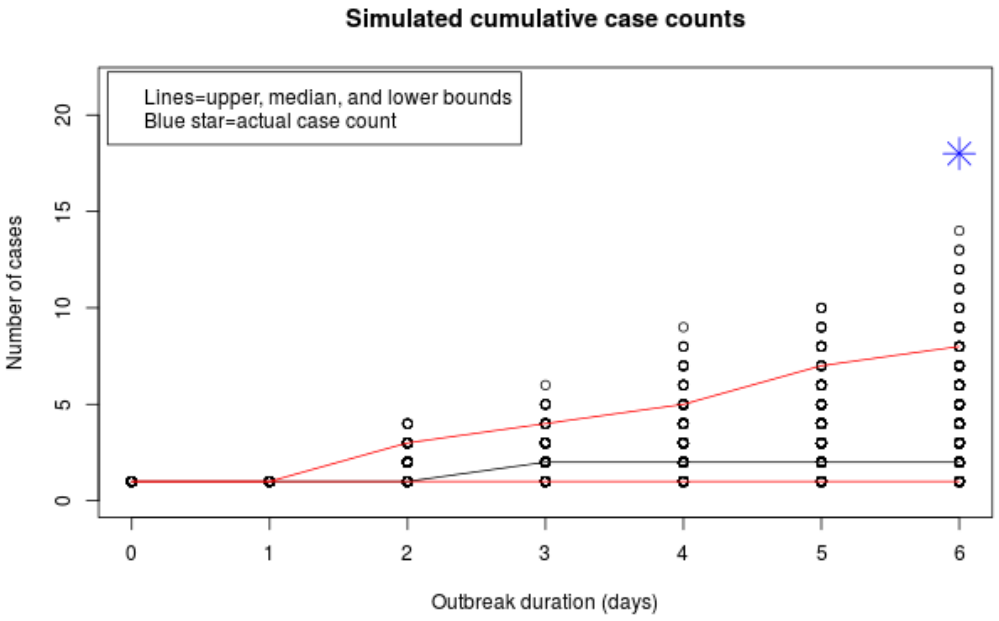


Figure A.8. Epidemic Curve (Scenario ID=4158).




```
RESULTS:
 95% confidence interval for person-to-person case count: (1, 8)
 Actual case count: 18, p-value = 0
 Some evidence of point source transmission
```

Figure A.9. Clinical Decision Support Results (Scenario ID=4158).

You receive a phone call from a university administrator. She tells you there are 43 ill students out of about 1800 total at the school (attack rate = 2%). It has been 17 days since the first resident became ill, and there are 3 students currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.10. Minimal Data (Scenario ID=7020).

Table A.4. Kaplan's Criteria (Scenario ID=7020).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Median illness duration = 48 hours
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	37% of cases with vomiting
No bacterial agent found	No bacterial agent found
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

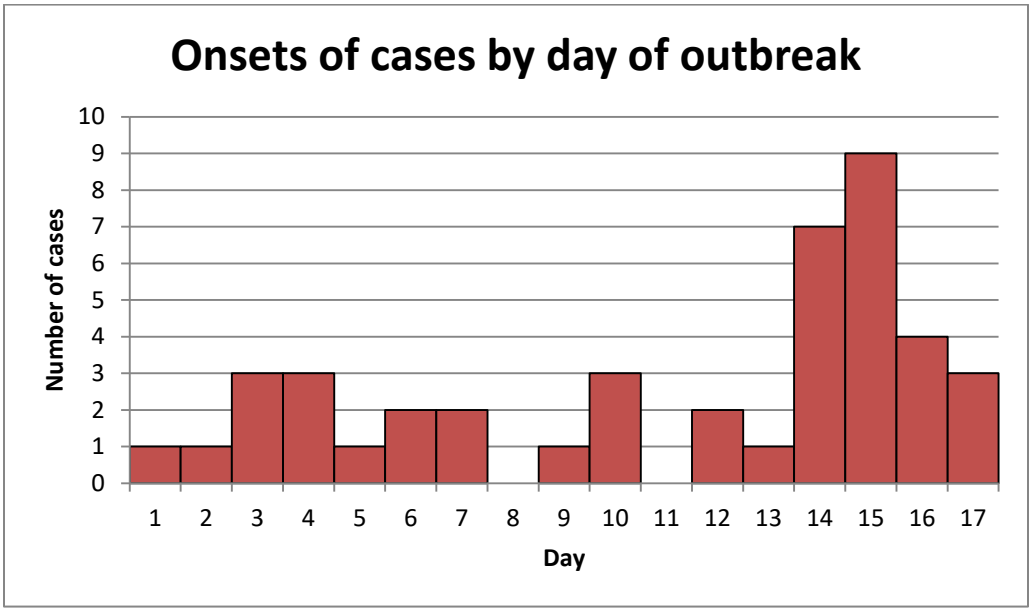
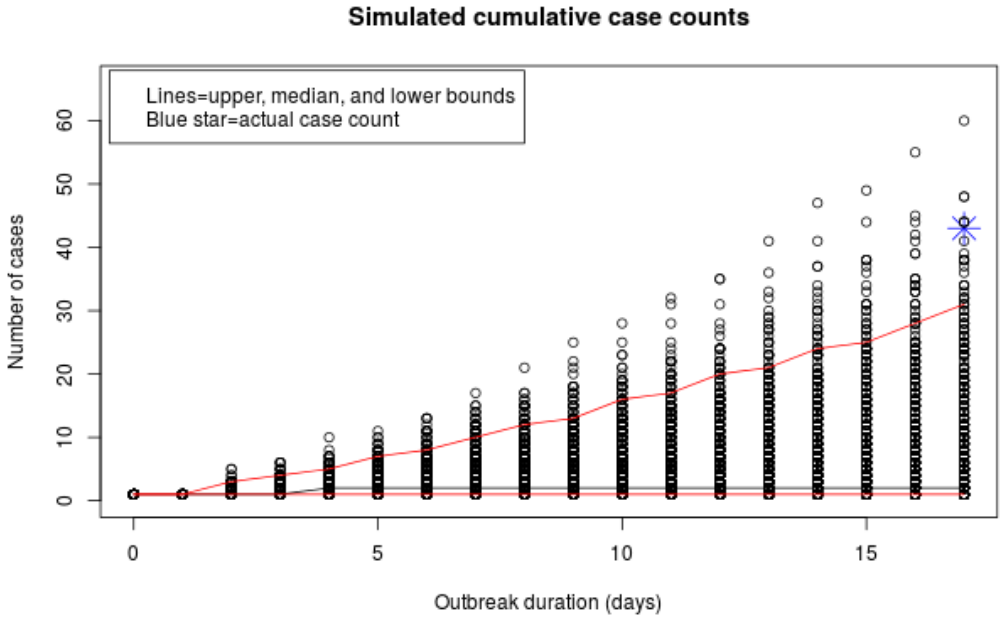


Figure A.11. Epidemic Curve (Scenario ID=7020).




```
RESULTS:
 95% confidence interval for person-to-person case count: (1, 31)
 Actual case count: 43, p-value = 0.003
 Some evidence of point source transmission
```

Figure A.12. Clinical Decision Support Results (Scenario ID=7020).

You receive a phone call from a nursing home manager. He tells you there are 41 ill residents out of 150 total in the facility (attack rate = 27%) and 24 ill staff out of 40 total in the facility (attack rate = 60%). It has been 12 days since the first resident became ill, and there are 2 residents and staff currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.13. Minimal Data (Scenario ID=1722).

Table A.5. Kaplan's Criteria (Scenario ID=1722).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Median illness duration = 1 day
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	50.4% of cases with vomiting
No bacterial agent found	No bacterial agent found
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

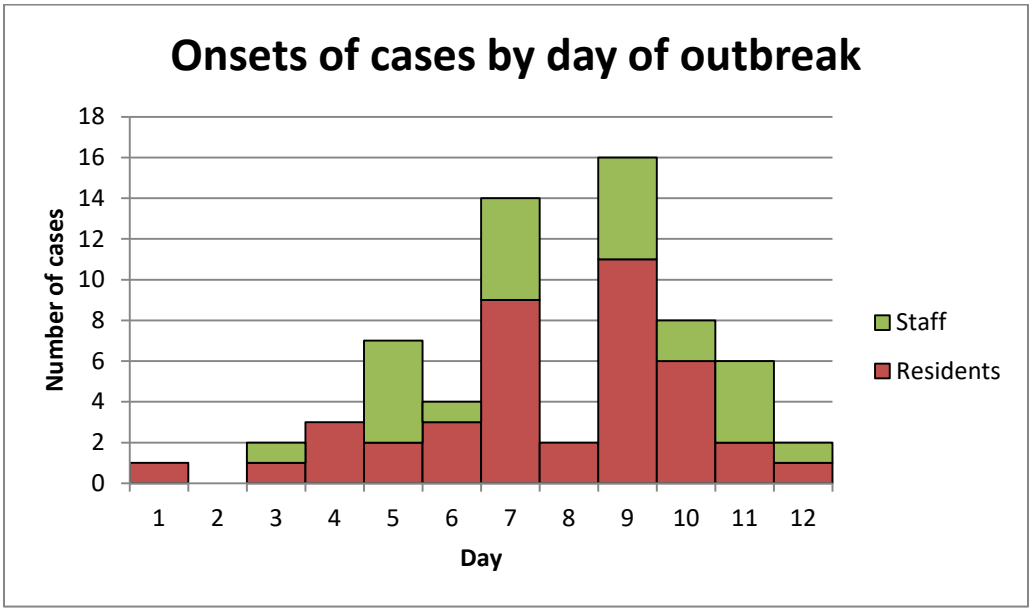
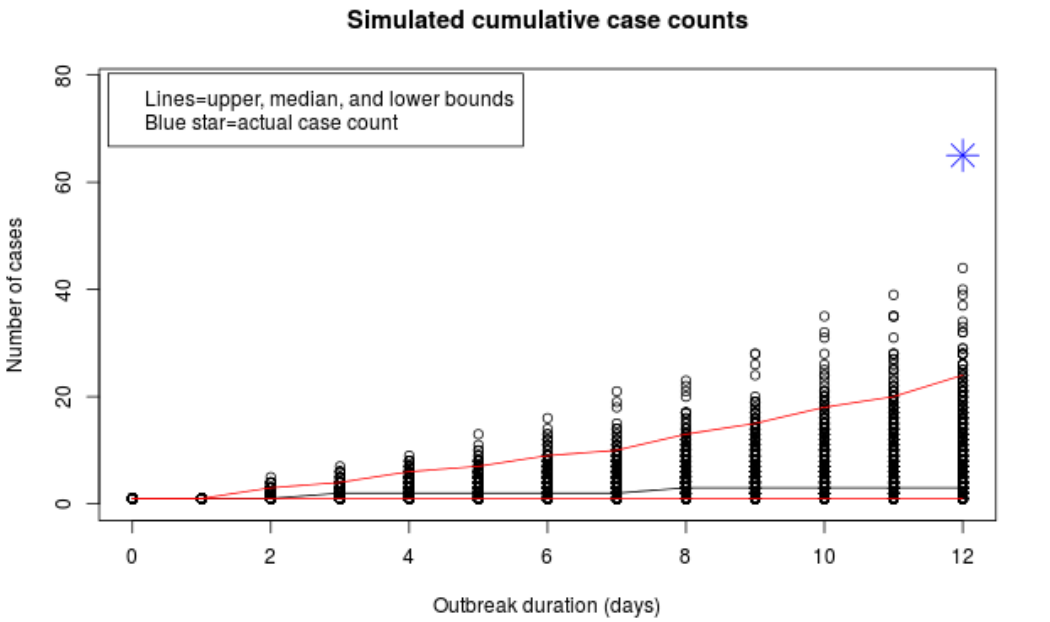


Figure A.14. Epidemic Curve (Scenario ID=1722).




```
RESULTS:
 95% confidence interval for person-to-person case count: (1, 24)
 Actual case count: 65, p-value = 0
 Some evidence of point source transmission
```

Figure A.15. Clinical Decision Support Results (Scenario ID=1722).

You receive a phone call from the infection preventionist (IP) at a local hospital. He tells you there are 81 ill patients out of 413 total in the facility (attack rate = 20%) and 58 ill staff out of 285 total in the facility (attack rate = 20%). There were 2 initial cases, it has been 7 days since they first became ill, and there are 5 residents and staff currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.16. Minimal Data (Scenario ID=6041).

Table A.6. Kaplan's Criteria (Scenario ID=6041).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Median illness duration = unknown
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	62% of patients with vomiting, 57% of staff with vomiting
No bacterial agent found	No bacterial agent found
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

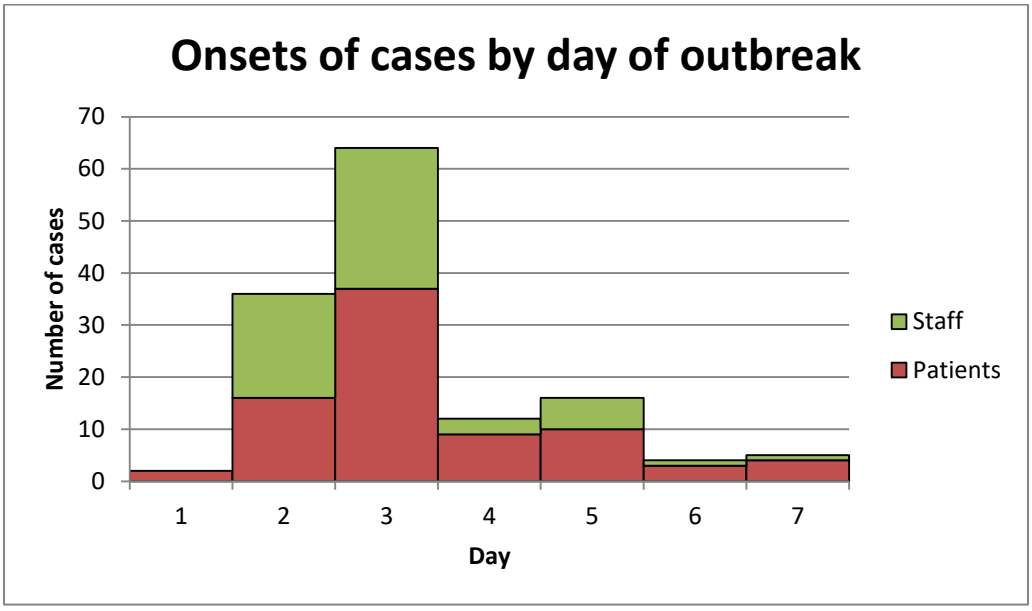
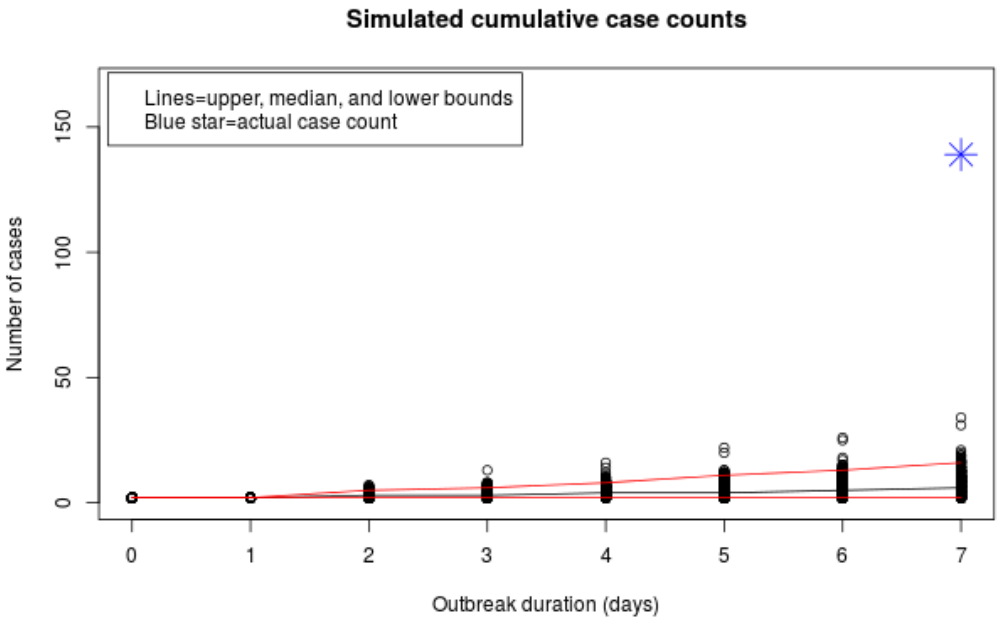


Figure A.17. Epidemic Curve (Scenario ID=6041).




```
RESULTS:
 95% confidence interval for person-to-person case count: (2, 16)
 Actual case count: 139, p-value = 0
 Some evidence of point source transmission
```

Figure A.18. Clinical Decision Support Results (Scenario ID=6041).

You receive a phone call from a psychiatric center administrator. She tells you there are 19 ill residents out of 60 total in the facility (attack rate = 32%) and 4 ill staff out of 49 total in the facility (attack rate = 8%). It has been 17 days since the first resident became ill, and there are no residents or staff currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.19. Minimal Data (Scenario ID=5438).

Table A.7. Kaplan's Criteria (Scenario ID=5438).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Median illness duration = 3 days for patients, unknown for staff
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	48% of patients with vomiting, unknown for staff
No bacterial agent found	No bacterial agent found
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

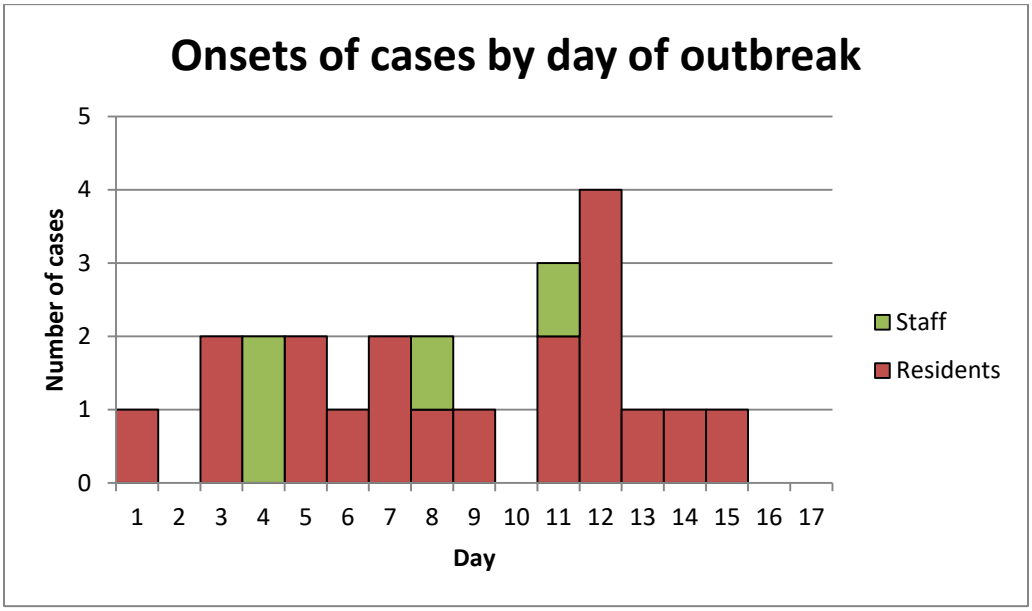
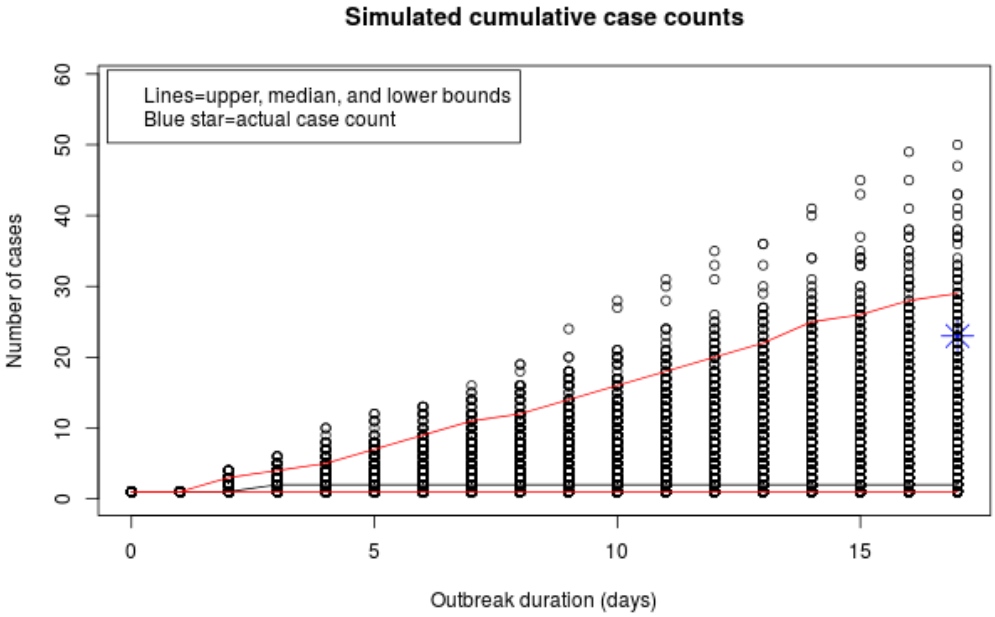


Figure A.20. Epidemic Curve (Scenario ID=5438).




```
RESULTS:
 95% confidence interval for person-to-person case count: (1, 29)
 Actual case count: 23, p-value = 0.058
 No evidence of point source transmission
```

Figure A.21. Clinical Decision Support Results (Scenario ID=5438).

You receive a phone call from a nursing home administrator. She tells you there are 43 ill residents out of 182 total in the facility (attack rate = 24%) and 9 ill staff out of 154 total in the facility (attack rate = 6%). It has been 17 days since the first resident became ill, and there are 8 residents and staff currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.22. Minimal Data (Scenario ID=4474).

Table A.8. Kaplan's Criteria (Scenario ID=4474).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Mean illness duration = 48 hours among healthcare workers
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	64% of healthcare workers with vomiting
No bacterial agent found	No bacterial agent found
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

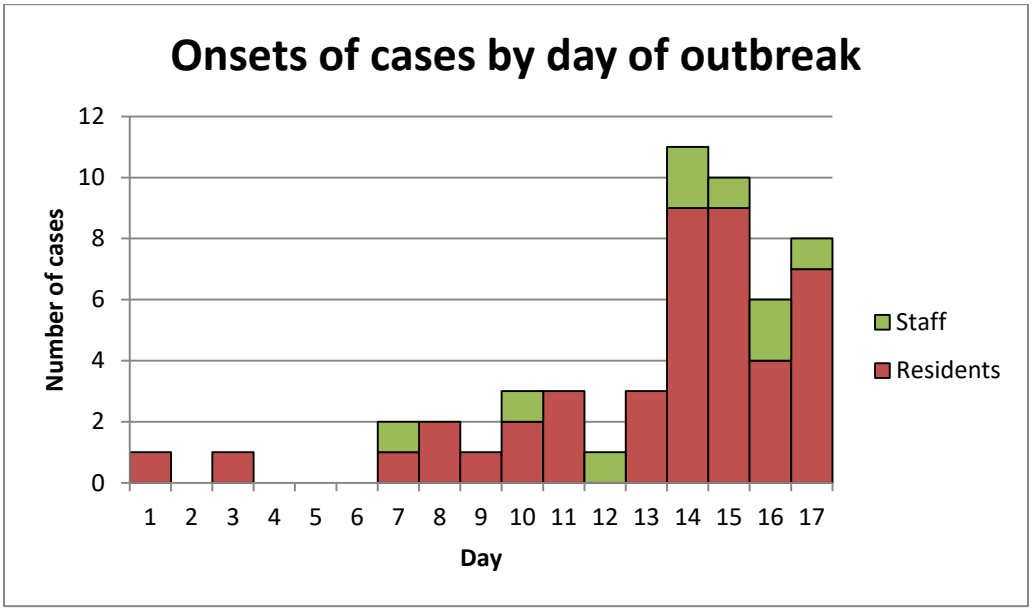
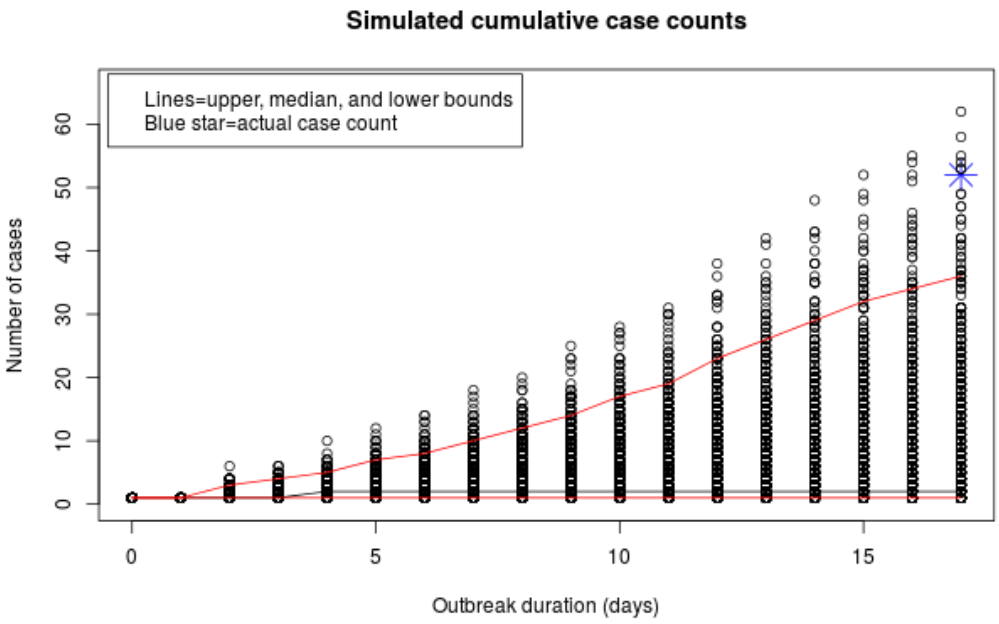


Figure A.23. Epidemic Curve (Scenario ID=4474).




```
RESULTS:
 95% confidence interval for person-to-person case count: (1, 36)
 Actual case count: 52, p-value = 0.002
 Some evidence of point source transmission
```

Figure A.24. Clinical Decision Support Results (Scenario ID=4474).

You receive a phone call from a nursing home nurse supervisor. She tells you there are 15 ill residents out of 295 total in the facility (attack rate = 5%) and 7 ill staff out of 65 total in the facility (attack rate = 11%). There were 2 initial cases, it has been 6 days since they first became ill, and there are no residents or staff currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.25. Minimal Data (Scenario ID=5897).

Table A.9. Kaplan's Criteria (Scenario ID=5897).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Median illness duration = unknown
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	41% of cases with vomiting
No bacterial agent found	No bacterial agent found
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

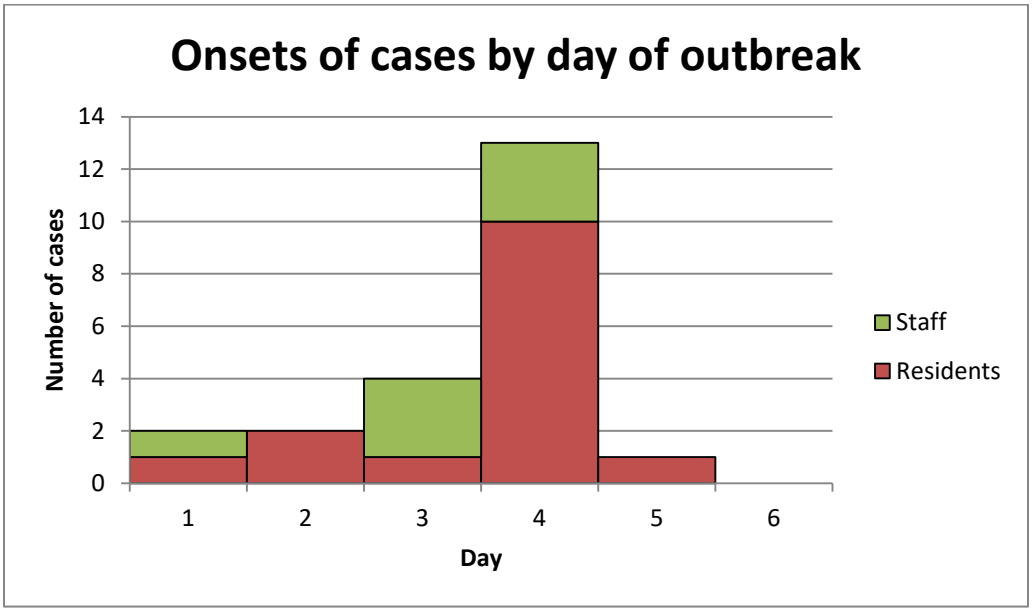
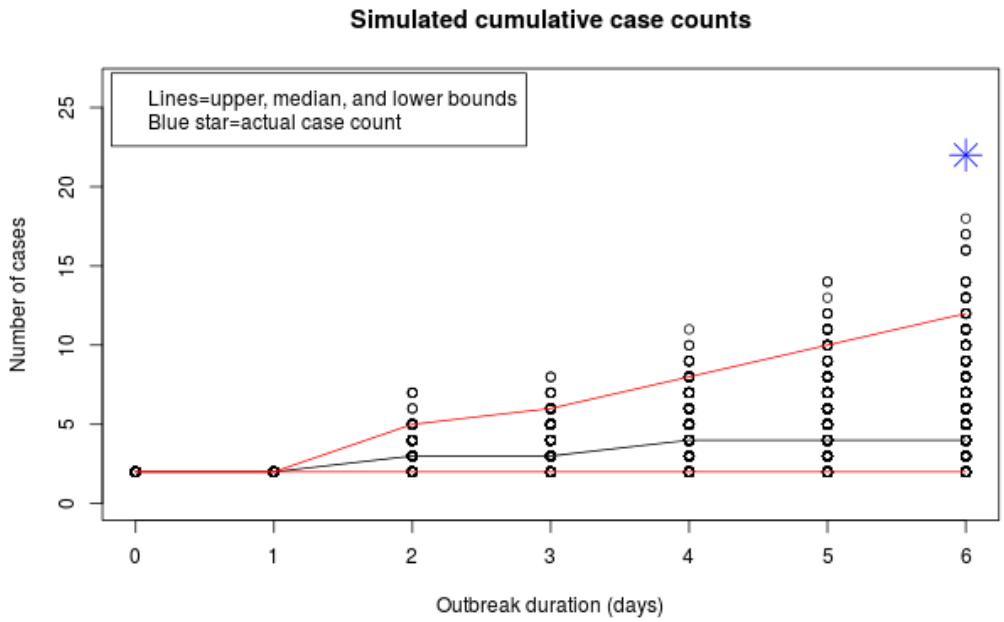


Figure A.26. Epidemic Curve (Scenario ID=5897).




```
RESULTS:
 95% confidence interval for person-to-person case count: (2, 12)
 Actual case count: 22, p-value = 0
 Some evidence of point source transmission
```

Figure A.27. Clinical Decision Support Results (Scenario ID=5897).

You receive a phone call from an officer at a military training facility. He tells you there are 121 ill personnel out of 363 total in the facility (attack rate = 33%). It has been 14 days since the first case became ill, and there are 7 personnel currently ill. No food handlers are known to have been ill, and there are no other known common exposures.

Figure A.28. Minimal Data (Scenario ID=1919).

Table A.10. Kaplan's Criteria (Scenario ID=1919).

Kaplan Criteria for determining if an outbreak is likely caused by norovirus[1]	**Outbreak findings**
Mean (or median) illness duration of 12 to 60 hours	Median illness duration = 48 hours
Mean (or median) incubation period of 24 to 48 hours	Incubation period = unknown
More than 50% of cases with vomiting	34% of cases with vomiting
No bacterial agent found	No bacterial agent found
[1]CDC. Responding to Norovirus Outbreaks. https://www.cdc.gov/norovirus/php/responding.html	

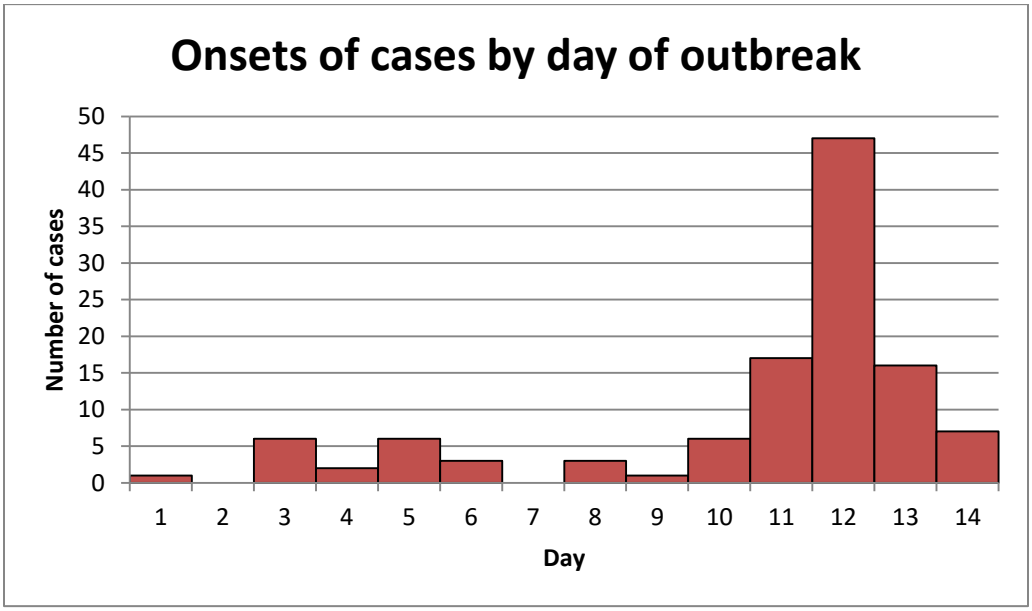
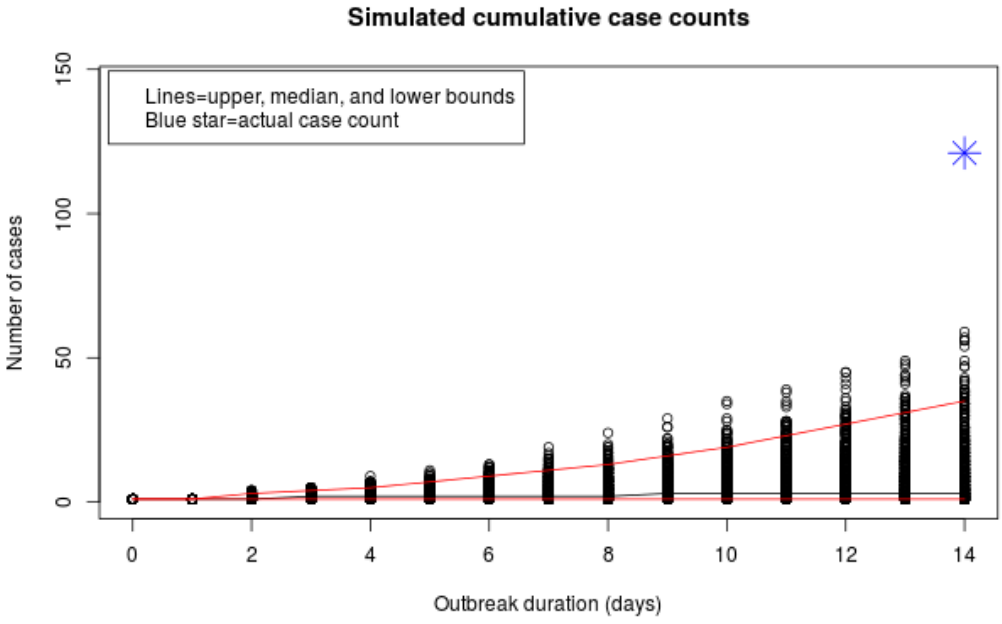


Figure A.29. Epidemic Curve (Scenario ID=1919).




```
RESULTS:
 95% confidence interval for person-to-person case count: (1, 35)
 Actual case count: 121, p-value = 0
 Some evidence of point source transmission
```

Figure A.30. Clinical Decision Support Results (Scenario ID=1919).