Distribution Agreement

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

Signature:

_____ _____

Jordan L. Kennedy Date

Comparison of Rotavirus Severity Scoring Systems

By

Jordan L. Kennedy

MSPH

Department of Biostatistics

Rollins School of Public Health

Emory University

_____ [Chair's Signature]

Traci Leong

_____ [Member's Signature]

Paul Weiss

Comparison of Rotavirus Severity Scoring Systems

By

Jordan L. Kennedy

B.A. Agnes Scott College, 2012

Thesis Committee Chair: Traci Leong, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Masters of Science in Public Health
in Biostatistics
2015

Comparison of Rotavirus Severity Scoring Systems

By Jordan L. Kennedy

ABSTRACT

Rotavirus severity scoring systems are used for evaluating severity of illness in patients with gastroenteritis. There are two widely used scoring systems (Vesikari and Clark) and a New scoring system, but little research has been done comparing the scoring systems. A dataset including 948 gastroenteritis patients from 5 hospitals was used. The scores were evaluated individually, using Hosmer and Lemeshow goodness of fit tests, as well as against each other. The scores themselves were compared using Cohen's kappa agreement and the scores were then used to create models. The models were compared based on AIC and residual deviance. The kappa agreement showed moderate agreement between each pair of scores, while all other tests showed the New system performing best, with the Vesikari system performing less well and the Clark system performing poorly. Clearly, the New scoring system is worth consideration by investigators, but the Vesikari scoring system is the more accurate of the currently used systems.

Comparison of Rotavirus Severity Scoring Systems

By

Jordan L. Kennedy

B.A. Agnes Scott College, 2012

Thesis Committee Chair: Traci Leong, Ph.D.

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Biostatistics
2015

Table of Contents

Chapter I

Introduction & Background……..……1

Chapter II

Literature Review……..…3

Chapter III

Methodology……7

Chapter IV

Results……..10

Chapter V

Discussion…….13

Reference……..16

Appendix A……...17

Appendix B……..….19

Chapter I

Introduction

Background

Rotavirus is the leading cause of severe diarrhea in children (Glass, 2014). In 2008, more than 450,000 deaths were reported worldwide with about 85 percent of these occurring in low income countries (Glass, 2014). In most cases, rotavirus causes mild dehydration, but in some cases, severe dehydration can occur, requiring intravenous rehydration. There are two licensed vaccines that are currently available in the United States: Rotarix (RV1) and Rotateq (RV5). Several studies have attempted to evaluate the effectiveness of these two vaccines. With any vaccine, the prevention of infection is not guaranteed. Instead, the likelihood of infection is reduced and, if infection occurs, vaccination can reduce illness severity. To evaluate vaccine effectiveness for the rotavirus vaccines, severity of rotavirus symptoms, or similar gastrointestinal symptoms, are measured when they occur. Several scoring systems are currently used, with the most widely used being Vesikari. Other scoring systems include Clark and a new scoring system that combines the Veskari and Clark scoring. An ideal scoring system should be based on easily or routinely recordable variables, well calibrated, have a high level of discrimination, be applicable to all patient populations, usable in different countries, and have the ability to predict functional status or quality of life after discharge (Bouch, 2008).

Problem Statement

Few comparisons between scoring systems are available (Bouch, 2008). While some have tried to compare the Vesikari and Clark scoring systems, but many ultimately only note the differences between them. With the addition of the new scoring system, there needs to be a comparison between these three, not merely a description of the differences. If there is no comparison, more systems could be added without restriction. Additionally, with multiple systems currently in use, different researchers use different scoring systems rendering results that are not comparable.

Purpose Statement

In this paper, the goal is to compare the existing severity scoring systems used for Rotavirus: Vesikari, Clark, and the New system. They each have similar scoring qualifications, so this paper will try to differentiate between the three, as well as compare their outcomes and effectiveness.

Significance Statement

With so many casualties each year, it is vital that the best possible preventative treatments are available and utilized. In order to identify the best methods, studies must correctly evaluate vaccine effectiveness, which hinges on the evaluation of illness severity. There are several scoring systems available and investigators need to be able to compare them and choose the most accurate method. A comparison of the current systems would help investigators choose the best and provide a methodology for evaluating future systems.

Chapter II

Literature

Introduction

Thus far, the research regarding rotavirus severity scores has focused on their development and evaluation with a few studies on comparing them. However, little research has specifically addressed the comparison of rotavirus scores and determining a preferable method. We can consider previous studies on score development and evaluation before applying to rotavirus severity scores.

Types of Scoring Systems

According to John Pappachan, "Severity of illness scoring systems aid the case-mix adjusted collection of such data. However, none is perfect and their use to triage individual patients or to compare the quality of care in different ICUs is severely limited" (Pappachan, 2004). Scoring systems were originally created for intensive care units in the 1980s and allowed comparison of outcomes between facilities (Bouch, 2008). These measures of severity often consisted of a score and a calculated probability of mortality. The score is simply a number with a higher number corresponding to a more severe condition. The mortality probability is not always a part of the score, as with rotavirus scoring systems, and typically depicts the risk of in-hospital mortality. Additionally, there are several categories of scoring systems: anatomical scoring, therapeutic weighted scores, organ-specific scoring, physiological assessment, simple scales, and disease specific. The rotavirus scores are disease specific and can only be applied to gastrointestinal symptoms typically associated with rotavirus. Some scores collect data

on first admittance and others collect repetitively throughout the patients hospital stay. Scores can also be considered subjective or objective. Potential uses of these systems include audit and comparison of ICU performance, as well as a mechanism to decide resource allocation (Pappachan, 2004). The best possible scoring system is expected to be based on easily or routinely recordable variables, well calibrated, have a high level of discrimination, be applicable to all patient populations, be able to be used in different countries, and have the ability to predict the functional status or quality of life after discharge (Bouch, 2008).

Comparison

To ensure the scoring systems are effective, they must be validated. In order to evaluate these systems, the score must be used on a different population than was used to create the system. This can be done by randomly splitting the original population into 2 groups (one to make the score, the other to validate the model) or with a completely separate population (Bouch, 2008). If a new set of data is used, it must be collected from consecutive patients (UGH). Once the score is applied to both sets of patients, performance similarity can be measured with a chi-squared test for homogeneity (UGH). Model calibration and discrimination are then assessed (Bouch, 2008). Since no new scores will be developed in this study, all of the data will be used to evaluate three scoring systems.

Model calibration assesses the degree of correspondence between the estimated probability of mortality and that actually observed (Bouch, 2008). In other words, it describes how closely predictions correlate with actual outcomes across the entire range

of risk (Pappachan, 2004). The expected and observed mortality are compared and a p-value is derived. Calibration is considered good if the predicted mortality is close to the observed (Bouch, 2008). It can be tested with a goodness of fit test, most commonly the Hosmer-Lemeshow C statistic (Bouch, 2008). Rotavirus severity scores do not predict mortality, so they cannot be measured against observed mortality. Instead, the scoring systems will be evaluated against a clinical outcome not included in any of the scores.

Model discrimination reviews the ability of the scoring model to discriminate between patients who die and those who survive, based on predicted mortalities (Bouch, 2008). Discrimination is evaluated through the calculation of area under the receiver operating characteristic (ROC) curve or by using a classification matrix. A pair of sensitivity-specificity values produces the ROC curve across the range of mortality prediction scores and the area under the curve (AUC) represents the number of patients who died. The curve is then analyzed to assess discrimination. If the AUC is around 0.5, scoring system is no better than flipping a coin. Typically, AUC is required to be greater than 0.70 (Bouch, 2008). Unfortunately, model discrimination is not assessable in this case because there is no binary outcome to which to compare the scores.

There are other potential problems with assessment of the scoring systems. Developing a model from a sample that is too small can risk the score being unable to distinguish between different patient groups. Scoring systems also have to be modeled and validated against a real cohort of ICU patients, but ICU patients are not always the best representative of all ill patients. Finally, scoring systems often help compare an individual facility's performance over time but comparisons between different units are susceptible to misinterpretation (Bouch, 2008).

Previous Rotavirus Evaluations

A few studies have addressed Rotavirus scoring systems but, as mentioned, had incomplete conclusions. One study by Givon-Lavi et al. in 2008 compared Vesikari and Clark scoring systems using Cohen's kappa Agreement but only found mild agreement. The Vesikari system classifies patients into two categories: severe and non-severe. The Clark and New scoring systems classify patients into three categories: mild, moderate, and severe. Givon-Lavi et al. first combined the mild and moderate categories of Clark to compare with Vesikari and found kappa of 0.200. When Vesikari was split into three categories to mirror Clark, the kappa was 0.340. With further analysis, the researchers concluded that vaccines and naturally acquired immunity gives moderate protection against rotavirus diarrhea, but high protection against severe episodes (Levi, et al. 2008). They did not, however, conclude if one scoring system was more accurate or preferable. I will split the Clark and New scoring systems and use kappa Agreement to evaluate similarity between the systems, but additional analysis will be utilized to try to determine if one of the scoring systems is better. Another study by Lewis et al. in 2012 compared the two systems and found that Vesikari classified more cases as severe than Clark (Lewis et al., 2012). Lewis et al. also decided that the trials where the data came from were designed to capture mild and moderate cases and therefore the comparison done was not optimal (Lewis et al., 2012).

Chapter III

Methodology

Introduction

A case control study was conducted for children born after March 1, 2009 with acute gastroenteritis from 5 hospitals (3 pediatric hospitals in Atlanta, GA and 2 hospitals in Connecticut). Surveys regarding behavior and demographics were administered, stool specimens collected, and vaccination records obtained. Stool samples were analyzed at the Centers for Disease Control to confirm whether the patient had rotavirus and results were not available during the child's hospital visit. Vaccine records were obtained through state records and medical data was gathered from the patient's chart.

The three scoring systems each assign points based on clinical observations, such as number of days of diarrhea, episodes of diarrhea per day, maximum temperature, whether rehydration was necessary. More points are assigned to more severe circumstances and cutoff points divide the final score into categories (severe/non-severe for Vesikari and mild/moderate/severe for Clark and New). The full scoring system specifications can be found in Appendix A.

Solution

As a preliminary evaluation of how much agreement there is between Vesikari, Clark and the New scoring system, Cohen's kappa agreement will be evaluated. Kappa agreement will measure the degree to which two scores categorize the same patients into the same severe or non-severe category by comparing the proportion of times they agree

to the proportion of times they are expected to agree by random chance. If the two scores categorize patients exactly the same, kappa will be 1. If there is only similar categorization due to random chance, kappa will be 0. As a reference, 0-0.20 is considered no or slight agreement, 0.20-0.40 fair agreement, 0.40-0.60 moderate agreement, 0.60-0.80 substantial agreement, and 0.80 or higher is very good agreement (Mchugh, 2012). Kappa agreement will allow the scores to be compared directly to see how similar they are in their classification of patients. Cohen's kappa is considered more robust than simply calculating agreement because it considers agreement occurring by chance. However, using kappa agreement only compares two scores at a time and looks at how well they concur. It does not provide information on whether one is more accurate than the other.

Each scoring system will be fit with a Poisson regression model using the scores as a predictor and length of hospital stay as the outcome of interest. Length of hospital stay will be used as the outcome because it is an adequate clinical indicator of severe disease and not included in any of the scores. Poisson regression will be useful since the outcome is count data. The data will likely be skewed, but a log transformation is not feasible since not every patient is admitted to the hospital, giving them a length of stay of 0 which cannot be transformed. Comparing the three models directly (such as with AIC or deviance) will allow a clear choice to be made between the three as to which is the best model of severe outcome. The residual deviance is the difference between the deviance of the current model and the maximum deviance of the model that predicts outcomes perfectly. The Akaike Information Criterion (AIC) is often used to compare models and looks at the information lost when the current model is used to generate the data. Smaller

residual deviance and smaller AIC both indicate a better fit. The models will be compared to each other and calibration will be evaluated using Hosmer and Lemeshow goodness of fit tests. The Hosmer and Lemeshow goodness of fit tests will evaluate each scoring system separately in regards to the outcome to determine if it adequately models the data.

Many scoring systems predict mortality, allowing for a direct evaluation of their effectiveness by looking at which patients ultimately died. In this case, there is no binary outcome of interest. Length of hospital stay should be a helpful indicator in evaluation but the goal of these scores is not to predict hospital stay, only to categorize severity. It is possible that the scores will not accurately predict length of stay. This could potentially lead to less than desirable observations in the goodness of fit tests.

Chapter IV

Results

Introduction

A total of 948 patients were recorded (702 from Atlanta, GA and 246 from Connecticut). A summary table of the data can be found in Table 1 in Appendix B where the data was separated by location and case and control and the distribution of several characteristics compared. The Clark scoring system and age showed significant differences between the patients who were Rotavirus positive and those that were Rotavirus negative. The Clark scoring system also saw a difference in distribution of the Rotavirus positive patients between the Georgia hospitals and the Connecticut hospitals. Length of stay was severely skewed with 812 patients (approximately 85%) not being admitted. A histogram of the data can be found in Figure 1.

To compare two patient groups (severe and non-severe) among all three scores, Vesikari (the most widely used system), Clark, and New we will reconfigure scores if necessary. The cutoff points were chosen as the median values observed: 12 for Clark and 4.5 for New. Any value below these in their respective scoring systems will be considered non-severe and these values or higher will be considered severe. Note that these cut-off points were used in Table 1.

Agreement

Using the newly divided scoring systems, kappa agreement was calculated between each of the three scores and each found a significant p-value ($p<0.0001$). The

kappa statistic tests the null hypothesis that the agreement is due to chance only. Each of these found "moderate agreement". Between Vesikari and Clark, the kappa estimate was 0.454 (95% CI 0.3964, 0.5127). For Vesikari and the New system, the kappa estimate was 0.5799 (95% CI 0.5280, 0.6318). Finally, for Clark and the New system, the kappa estimate was 0.5442 (95% CI 0.4908, 0.5976). While there is no strong agreement between any of these three systems, there is some agreement, with Vesikari and the New system having the most patients classified the same.

Modeling

A Poisson regression was fit to each of the scoring systems, using the factors of each as the predictors and length of stay as the outcome. All three scores include number of diarrhea episodes, number of vomiting episodes, number of days with diarrhea, number of days with vomiting, and the maximum temperature. The Clark and New systems include number of days with fever. Vesikari and New include whether the patient required rehydration and whether they were admitted to the hospital. Additionally, Clark considers the patient's behavior (was the child less playful, lethargic, or experiencing seizures) and the New system considers whether the patient died (no patients in this data set died). A table summarizing the distribution of patients in each category can be found in Table 2 in Appendix B (Scoring Components table – rephrase this).

The Vesikari model had an AIC of 1855.8 and a residual deviance of 1493.2 on 946 degrees of freedom. The Clark model had an AIC of 1928.2 and a residual deviance of 1565.6 on 946 degrees of freedom. The New model had an AIC of 1731.7 and a residual deviance of 1369.1 on 946 degrees of freedom. Looking at either the AIC or the residual

deviance, the New model had much smaller values, and therefore a much better fit to the data. The Vesikari model did not fit as well as the New model and the Clark model's fit was the worst. It should be noted that there are fewer differences in the models for Vesikari and New because these scores are calculated similarly. They look at the same variables, with the only difference being that the New system considers the number of days the patient had a fever and whether or not death occurred. However, the Vesikari and New scoring systems do weight the variables differently.

Goodness of Fit

Each of the scores were tested using a Hosmer and Lemeshow goodness of fit test. Note that these tests were done with the Clark and New scoring systems separated into the two categories instead of their typical three. All three tests had 8 degrees of freedom and p-values less than 0.0001, indicating that all three poorly predict length of hospital stay. Vesikari had the lowest chi-squared value at 1120.9. The Clark scoring system had the second lowest chi-squared value at 2027.4 and the New scoring system had a chi-squared of 2626.3. The differences in chi-squared values should be noted, but all three scores were still found to not adequately predict the outcome. This is likely at least partly due to the outcome chosen. As previously mentioned, these scoring systems were not designed to predict any outcome, only to classify patients into severity categories.

Chapter V

Conclusions, Implications, and Recommendations

Summary

Rotavirus severity scoring systems are used for evaluating severity of illness in patients with gastroenteritis. There are two widely used scoring systems (Vesikari and Clark) and a New scoring system, but little research has been done comparing the scoring systems. Here, the scores were evaluated individually, using Hosmer and Lemeshow goodness of fit tests, as well as against each other. The scores themselves were compared using Cohen's kappa agreement and the scores were then used to create models. The models were compared based on AIC and residual deviance. The kappa agreement showed moderate agreement between each pair of scores, while all other tests showed the New system performing best, with the Vesikari system performing less well and the Clark system performing poorly.

Conclusion

This project attempted to evaluate each of these three scoring systems individually and against each other. A comparison of fit has not been done since these scores do not predict an outcome, but length of hospital stay is a logical choice. Using this outcome, the scores were able to be evaluated and compared directly with each other, in terms of effectiveness. While there are some problems that could be resolved with further study, this project is a good step towards looking more closely at rotavirus severity scoring systems. Here, the New scoring system performed better than the others, with the Vesikari scoring system performing second best and Clark the worst in each of

the tests. Currently, Vesikari is the most commonly used system and a shift towards using Vesikari only would be preferable to using Clark. However, this study has shown that the New scoring system is competitive with the Vesikari system and worth consideration by investigators.

Implications for Practice

A direct comparison of Rotavirus severity scoring systems is an important aspect of evaluating illness severity. These scores are often used in studies or to evaluate an institution's performance over time. If the scores are not accurate, the conclusions drawn from them are not accurate. Moreover, a single universal scoring system would benefit researchers and medical professionals in evaluating and comparing all forms of treatment available. This study shows the Vesikari system would be a better choice as a universal system than the Clark system. However, the New scoring system is worth consideration due to its superior performance here.

Recommendations

In the future, more investigation could be made into how the scores are split when they are being compared. Additionally, when the models were fitted with the variables of each scoring system, the program omitted patients with missing values. A method for avoiding this would be preferable in order to get as much information from the data as possible. A dataset without missing information is not likely, so it needs to be managed. Study would also be more informative if there were more locations and patients. The most challenging problem when evaluating these scoring systems is finding a relevant

outcome. It is important that the scoring systems not only have their differences enumerated, but are evaluated and compared to each other.

References

Bouch, D., & Thompson, J. (2008). Severity scoring systems in the critically ill. *Continuing Education in Anaesthesia, Critical Care & Pain,* 181-185.

Chaikitamnuaychok, R., & Patumanond, J. (2013). Upper Gastrointestinal Hemorrhage: Validation of the Severity Score. *Gastroenterology Research, 6*(2), 56-62.

Givon-Lavi, N., Greenberg, D., & Dagan, R. (2008). Comparison between two severity scoring scales commonly used in the evaluation of rotavirus gastroenteritis in children. *Vaccine,* 5798-5801.

Glass, R., Parashar, U., Patel, M., Gentsch, J., & Jiang, B. (2014). Rotavirus vaccines: Successes and challenges. *Journal of Infection,* S9-S18.

Leteurtre, S., Martinot, A., Duhamel, A., Gauvin, F., Grandbastien, B., Nam, T., Proulx, F., Lacroix, J., & Leclerc, F. (1999). Development of a Pediatric Multiple Organ Dysfunction Score: Use of Two Strategies. *Medical Decision Making*, 399-410.

Lewis, K., Dallas, M., Victor, J., Ciarlet, M., Mast, T., Ji, M., Armah, G., Zaman, K., Ferraro, A., & Neuzil, K. (2012). Comparison of two clinical severity scoring systems in two multi-center, developing country rotavirus vaccine trials in Africa and Asia. *Vaccine,* A159-A166.

McHugh, M. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica.* 22(3), 276-282.

Pappachan, J. (2004). Severity of illness scoring systems. *Anaesthesia & Intensive Care Medicine,* 5(1), 29-31.

Appendix A

Vesikari Severity Scoring System

Point Value	1	2	3
Duration of diarrhea	1-4	5	≥6
Maximum number of diarrhea stools/24 hours	1-3	4-5	≥6
Duration of vomiting (days)	1	2	≥3
Maximum number of vomiting episodes/24 hours	1	2-4	≥5
Temperature (°C)	37.1-38.4	38.5-38.9	≥39.0
Dehydration*	-	Mild	Moderate to Severe
Treatment	Rehydration	Hospitalization	

Total Score (0-20)	<11 (non-severe)	≥11 (severe)

*Mild-Moderate Dehydration = Oral/IV rehydration

Severe Dehydration = IV+Hospitalization

Clark Severity Scoring System

Point Value	1	2	3
Diarrhea	2-4	5-7	≥8
- Number of stools/day - Duration in days	1-4	5-7	≥8
Vomiting	1-3	4-6	≥7
- Number of emeses/day - Duration in days	2	3-5	≥6
Rectal Temperature	38.1-38.2	38.3-38.7	≥38.8
- Temperature (°C) - Duration in days	1-2	3-4	≥5
Behavioral symptoms/signs	Irritable/less playful	Lethargic/listless	Seizure
- Description - Duration in days	1-2	3-4	≥5

Total Score (0-24)	<9 (mild)	9-16 (moderate to severe)	>16 (severe)

New Severity Scoring System

Point Value	0.5	1	2	4
Diarrhea - Number of stools/day	2-4	5-7	≥8	
- Duration in days	1-4	5-7	≥8	
Vomiting - Number of emeses/day	1-3	4-6	≥7	
- Duration in days	1-2	3-5	≥6	
Oral Temperature (°C)	37.5-38.2	38.3-38.7	≥38.8	
- Duration in days	1-2	3-4	≥5	
Treatment/Outcome	Oral/IV rehydration	Hospitalization	IV+Hospitalization	Death

Total Score (0-32.5)	<5 (mild)	6-13 (moderate)	>14 (severe)

Appendix B

Table 1. Population Characteristics.

Variables	All (n=948)	Rotavirus Positive (n=295)		Rotavirus Negative (n=653)		p-value*
		Georgia	Connecticut	Georgia	Connecticut	
Severity (Vesikari)						
Non-Severe	610 (64.3%)	127 (57.5%)	41 (55.4%)	333 (69.2%)	109 (63.4%)	0.862
Severe	338 (35.7%)	94 (42.5%)	33 (44.6%)	148 (30.8%)	63 (36.6%)	0.189
						0.228
Severity (Clark)						
Non-severe	535 (56.4%)	121 (54.8%)	20 (27.0%)	298 (62.0%)	96 (55.8%)	<0.001
Severe	413 (43.6%)	100 (45.2%)	54 (73.0%)	183 (38.0%)	76 (44.2%)	0.186
						<0.001
Severity (New)						
Non-Severe	475 (50.1%)	83 (37.6%)	32 (43.2%)	275 (57.2%)	85 (49.4%)	0.465
Severe	473 (49.9%)	138 (62.4%)	42 (56.8%)	206 (42.8%)	87 (50.6%)	0.096
						0.393
Gender						
Male	548 (57.8%)	134 (60.6%)	30 (40.5%)	277(57.6%)	107 (62.2%)	0.004
Female	400 (42.2%)	87 (39.4%)	44 (59.5%)	204 (42.4%)	65 (37.8%)	0.334
						0.481
Age						
< 3 months	29 (3.1%)	6 (2.7%)	0 (0.0%)	19 (4.0%)	4 (2.3%)	0.450
3- <6 months	119 (12.6%)	16 (7.2%)	4 (5.4%)	79 (16.4%)	20 (11.6%)	0.012
6-<9 months	161 (17.0%)	26 (11.8%)	8 (10.8%)	100 (20.8%)	27 (15.7%)	<0.001
9-<12 months	131 (13.8%)	21 (9.5%)	3 (4.1%)	77 (16.0%)	30 (17.4%)	
12-<24 months	349 (36.8%)	95 (43.0%)	36 (48.6%)	142 (29.5%)	76 (44.2%)	
≥ 24 months	159 (16.8%)	57 (25.8%)	23 (31.1%)	64 (13.3%)	15 (8.7%)	
Race						
White	292 (30.8%)	57 (25.8%)	32 (43.2%)	131 (27.2%)	72 (41.9%)	<0.001
Black	454 (47.9%)	133 (60.2%)	14 (18.9%)	274 (57.0%)	33 (19.2%)	<0.001
Other	170 (17.9%)	26 (11.8%)	24 (32.4%)	60 (12.5%)	60 (34.9%)	0.860
None/Unknown	32 (3.38 %)	5 (2.3%)	4 (5.4%)	16 (3.3%)	7 (4.1%)	
Ethnicity[o]						
Hispanic or Latino	259 (27.3%)	29 (13.1%)	37 (50.0%)	105 (21.8%)	88 (51.2%)	<0.001
Not Hispanic or Latino	687 (72.5%)	192 (86.9%)	37 (50.0%)	375 (78.0%)	83 (48.3%)	<0.001
Unknown	2 (0.2%)	0 (0.0%)	0 (0.0%)	1 (0.2%)	1 (0.6%)	0.025
Insurance[o]						
Public	555 (58.5%)	118 (53.4%)	57 (77.0%)	269 (55.9%)	111 (64.5%)	0.002
Private	310 (32.7%)	77 (34.8%)	16 (21.6%)	162 (33.7%)	55 (32.0%)	0.003
None	63 (6.7%)	19 (8.6%)	1 (1.4%)	40 (8.3%)	3 (1.7%)	0.890
Unknown	20 (2.1%)	7 (3.2%)	0 (0.0%)	10 (2.1%)	3 (1.7%)	

*The p-values in the table above compare Georgia and Connecticut data for Rotavirus Positive in the first row, Georgia and Connecticut for Rotavirus Negative in the second row, and Rotavirus Positive and Rotavirus Negative in the third row.

Table 2. Scoring Components

Scoring System	Characteristic	Georgia	Connecticut	Total
All	Duration of Diarrhea (Days)	3.51 (2.08)	3.49 (2.18)	3.50 (2.10)
	Max # Diarrhea Stools/24 h	7.59 (5.55)	7.24 (5.28)	7.50 (5.48)
	Duration of vomiting (Days)	2.89 (1.98)	2.86 (1.92)	2.88 (1.96)
	Max # Vomiting Episodes/24 h	5.50 (4.28)	5.53 (4.45)	5.51 (4.32)
	Temperature (°F)	102.0 (1.69)	102.3 (1.67)	102.1 (1.69)
Vesikari	Dehydration			
	Mild (Oral Rehydration)	426 (60.7%)	155 (63.0%)	581 (61.3%)
	Moderate to Severe (IV Fluids)	103 (14.7%)	90 (36.6%)	193 (20.4%)
	Treatment			
	Rehydration	426 (60.7%)	155 (63.0%)	581 (61.3%)
	Hospitalization	73 (10.4%)	62 (25.2%)	135 (14.2%)
Clark	Behavioral Symptoms/Signs			
	Irritable/Less Playful	445 (63.4%)	189 (76.8%)	634 (66.9%)
	Lethargic/Listless	303 (43.2%)	161 (65.4%)	464 (48.9%)
	Seizure	1 (0.001%)	2 (0.008%)	3 (0.003%)
	Duration of Behavior (Days)			
	Irritable/Less Playful	3.28 (1.84)	2.51 (1.95)	3.03 (1.91)
	Lethargic/Listless	3.07 (1.73)	2.01 (1.88)	2.63 (1.86)
	Seizure	0.50 (0.71)	0.01 (0.10)	0.02 (0.12)
Clark and New	Duration of Temp (Days)	3.34 (9.33)	3.65 (9.91)	3.43 (9.49)
New	Treatment/Outcome			
	Oral/IV Rehydration	426 (60.7%)	155 (63.0%)	581 (61.3%)
	Hospitalization	73 (10.4%)	62 (25.2%)	135 (14.2%)
	IV+Hospitalization	53 (7.5%)	35 (14.2%)	88 (9.3%)
	Death	0	0	0

Figure 1. Length of Hospital Stay Distribution