



# Spatial Epidemiology of Salmonellosis in Florida, 2009–2018

Xiaolong Li<sup>1†</sup>, Nitya Singh<sup>2†</sup>, Elizabeth Beshearse<sup>2,3</sup>, Jason L. Blanton<sup>4</sup>, Jamie DeMent<sup>3</sup> and Arie H. Havelaar<sup>2\*</sup>

<sup>1</sup> Department of Environmental and Global Health, Emerging Pathogens Institute, University of Florida, Gainesville, FL, United States, <sup>2</sup> Animal Sciences Department, Emerging Pathogens Institute and Food Systems Institute, University of Florida, Gainesville, FL, United States, <sup>3</sup> Food and Waterborne Disease Program, Florida Department of Health, Tallahassee, FL, United States, <sup>4</sup> Bureau of Public Health Laboratories, Florida Department of Health, Jacksonville, FL, United States

## OPEN ACCESS

### Edited by:

Vitali Sintchenko,  
The University of Sydney, Australia

### Reviewed by:

Csaba Varga,  
University of Illinois at  
Urbana–Champaign, United States  
Qinning Wang,  
New South Wales Health  
Pathology, Australia

### \*Correspondence:

Arie H. Havelaar  
ariehavelaar@ufl.edu

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Infectious Diseases–Surveillance,  
Prevention and Treatment,  
a section of the journal  
Frontiers in Public Health

**Received:** 04 September 2020

**Accepted:** 07 December 2020

**Published:** 11 January 2021

### Citation:

Li X, Singh N, Beshearse E,  
Blanton JL, DeMent J and  
Havelaar AH (2021) Spatial  
Epidemiology of Salmonellosis in  
Florida, 2009–2018.  
*Front. Public Health* 8:603005.  
doi: 10.3389/fpubh.2020.603005

Non-typhoidal *Salmonella enterica* infections cause a high disease burden in the United States with an estimated 1.2 million illnesses annually. The state of Florida consistently has a relatively high incidence compared to other states in the United States. Nevertheless, studies regarding the epidemiology of nontyphoidal salmonellosis and its spatial and temporal patterns in Florida were rarely reported. We examined the spatial and temporal patterns of 62,947 salmonellosis cases reported to FL Health Charts between 2009 and 2018. Dominant serotypes circulating in Florida were also explored using whole genome sequencing (WGS) based serotype-prediction for 2,507 *Salmonella* isolates sequenced by the Florida Department of Health during 2017 and 2018. The representativeness of laboratory-sequenced isolates for reported cases was determined by regression modeling. The annual incidence rate of salmonellosis decreased from 36.0 per 100,000 population in 2009 to 27.8 per 100,000 in 2016, and gradually increased in 2017 and 2018. Increased use of culture-independent testing did not fully explain this increase. The highest incidence rate was observed in children, contributing 40.9% of total reported cases during this period. A seasonal pattern was observed with the incidence peaking in September and October, later than the national average pattern. Over these 10 years, the Northeast and Northwest regions of the state had higher reported incidence rates, while reported rates in the Southeast and South were gradually increasing over time. Serotypes were predicted based on WGS data in the Enterobase platform. The top-five most prevalent serotypes in Florida during 2017–2018 were Enteritidis, Newport, Javiana, Sandiego and Braenderup. The highest percentage of isolates was from children under 5 years of age (41.4%), and stool (84.7%) was the major source of samples. A zero-inflated negative binomial regression model showed that the reported case number was a strong predictor for the number of lab-sequenced isolates in individual counties, and the geospatial distribution of sequenced isolates was not biased by other factors such as age group. The spatial and temporal patterns identified in this study along with the prevalence of different serotypes will be helpful for the development of efficient prevention and control strategies for salmonellosis in Florida.

**Keywords:** *Salmonella enterica*, epidemiology, spatial-temporal trends, seasonality, serotypes

## INTRODUCTION

Nontyphoidal salmonellosis is a common foodborne illness caused by a great diversity of *Salmonella enterica* serotypes with typical symptoms of diarrhea, abdominal cramps, vomiting and fever (1). In most cases, salmonellosis is a self-limiting illness and does not require medical treatment or hospitalization, but it is still problematic, especially among specific populations like children under 5 years-old, the elderly, and immunocompromised individuals. Humans get infected with *Salmonella* primarily by eating contaminated food (2). *Salmonella* infection poses a considerable disease burden globally, with an estimate of 93.8 million nontyphoidal *Salmonella* gastroenteritis cases occurring each year. Of these, 80.3 million cases were food-related (3). In the United States, *Salmonella* species are the second leading cause of foodborne illness. According to the estimates by Centers for Disease Control and Prevention (CDC), there are about 1.2 million illnesses, 23,000 hospitalizations, and 450 deaths associated with *Salmonella* infection annually (4). In the United States and other developed countries, a major portion of *Salmonella* illnesses has been attributed to food sources (5, 6). Seeded vegetables, chicken, fruits, pork, beef, eggs and other produce are food categories which contributed to more than 75% of outbreak-related *Salmonella* illnesses in the US (6). Previous source attribution studies conducted in other parts of the world also suggested that contaminated food including eggs, broiler chickens, and pigs were the top sources for human salmonellosis and outbreaks (5). In addition to the food pathway, water (e.g., recreational water/drinking water), person-to-person, animal contact and environmental exposure, such as contact with contaminated mud or soil, also contribute to transmission of *Salmonella* infection. A new study from the US suggests these pathways may contribute more than previously estimated (7).

The relative contribution of each transmission pathway varies among serotypes. Luvsansharav et al. developed a new measure to determine the foodborne relatedness of serotypes by calculating the ratio of proportion of laboratory-confirmed infections caused by each serotype to the proportion of foodborne illnesses associated with each serotype in the recorded outbreaks. Serotypes Saintpaul, Heidelberg and Berta showed the highest foodborne relatedness among the top 20 serotypes in the US; whereas, serotypes Mississippi, Bareilly, and Paratyphi B variant L(+) tartrate(+) had the lowest relatedness (8). Infection with *S. Newport*, the third most common serotype in the US, has been considered to be linked with both the food pathway and environmental sources, such as creek water and sediment (9). While infections from uncommon serotypes are more likely to be attributed to environmental factors and wild animals which are thought to primarily contribute to the variations in serotype diversity (10, 11). Moreover, some serotypes were frequently linked to specific food vehicles attributed to foodborne outbreaks in the US. For example, the salmonellosis outbreak data from CDC's Foodborne Disease Outbreak Surveillance System indicated that 81% of outbreaks with identified source of eggs were associated with *S. Enteritidis* between 1998 and 2015 (12).

Nontyphoidal *Salmonella* has been the leading cause of enteric illness reported in Florida which has consistently been one of the states with the highest incidence rates of salmonellosis in the United States (13). Nevertheless, studies regarding the epidemiology of nontyphoidal Salmonellosis and its spatial and temporal patterns in Florida were rarely reported.

As per the revisions to Rule 64D-3.029 Florida Administrative Code effective in October 20, 2016, laboratories in the state are required to send isolates or specimens of organisms responsible for reportable diseases including salmonellosis to the Florida Bureau of Public Health Laboratories (BPHL) for confirmation or additional characterization (14). Pulsed-field Gel Electrophoresis (PFGE) used to be considered the "gold standard" DNA fingerprinting method used by CDC within PulseNet. However, PulseNet transitioned from PFGE-based serotyping to a whole genome sequencing (WGS) -based method in 2019. BPHL has begun WGS in 2017 and has ceased the identification of PFGE for *Salmonella* isolate typing across the state of Florida and has started increasing the proportion of isolates sequenced with the WGS method, which provides the basis for examining the dominant *Salmonella* serotypes in Florida with WGS data available for 2017 and 2018.

In this study, we aimed to understand the incidence rate, temporal and geographical distribution and the demographics of reported salmonellosis cases in Florida from 2009 to 2018; to elucidate the dominant serotypes circulating in Florida based on the available WGS data in 2017 and 2018 and to determine the representativeness of laboratory-sequenced isolates for reported cases in Florida.

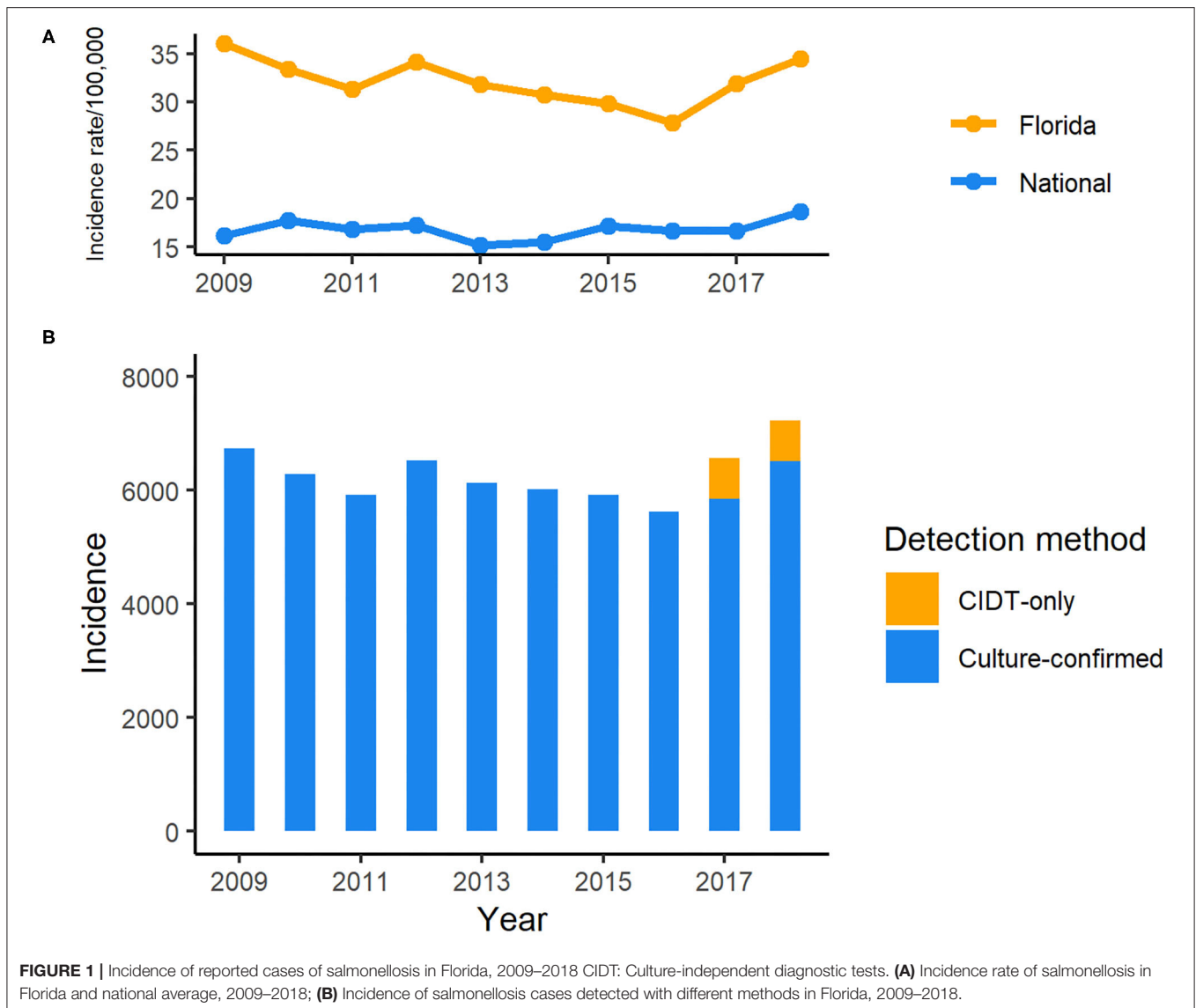
## MATERIALS AND METHODS

### Data Sources and Study Setting

The state of Florida is located in the southeast of the United States. It consists of 67 counties with an estimated population of 21,477,737 in 2019 (<https://www.census.gov/quickfacts/FL>, accessed November 2, 2020).

The definition of salmonellosis case was according to the Florida Department of Health's Surveillance Case Definitions for Select Reportable Diseases in Florida (15). Case data of salmonellosis (including outbreak-related but not travel-related cases) in Florida were downloaded from <http://www.flhealthcharts.com/> (accessed October 31, 2019) and included reported incidence and incidence rates and demographic information such as age and gender, from January 2009 to December 2018 in 67 counties of Florida. National data on salmonellosis incidence rate were obtained from the National Notifiable Diseases Surveillance System (<https://www.cdc.gov/nndss/>, accessed May 15, 2019) for comparison.

Annual estimates of county resident population by age and the estimated population by ethnicity in Florida were obtained from the U.S. Census Bureau (<https://data.census.gov/cedsci/>, accessed October 31, 2019). As per ethical requirements for the purpose of visualization, counties with a population <20,000 were merged with other counties based on contiguous and comparable geographical locations (coastal or inland) and median household



income, a total number of 58 (merged) counties were used in this study (16).

The land area data used for calculation of population density were derived from the Census 2010 summary file (<https://www.census.gov/data/tables/2010/dec/2010-summary-file-1.html>, accessed October 31, 2019).

Median household income for each county in Florida between 2017 and 2018 was obtained from the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program (<https://www.census.gov/programs-surveys/saipe.html>, accessed October 31, 2019).

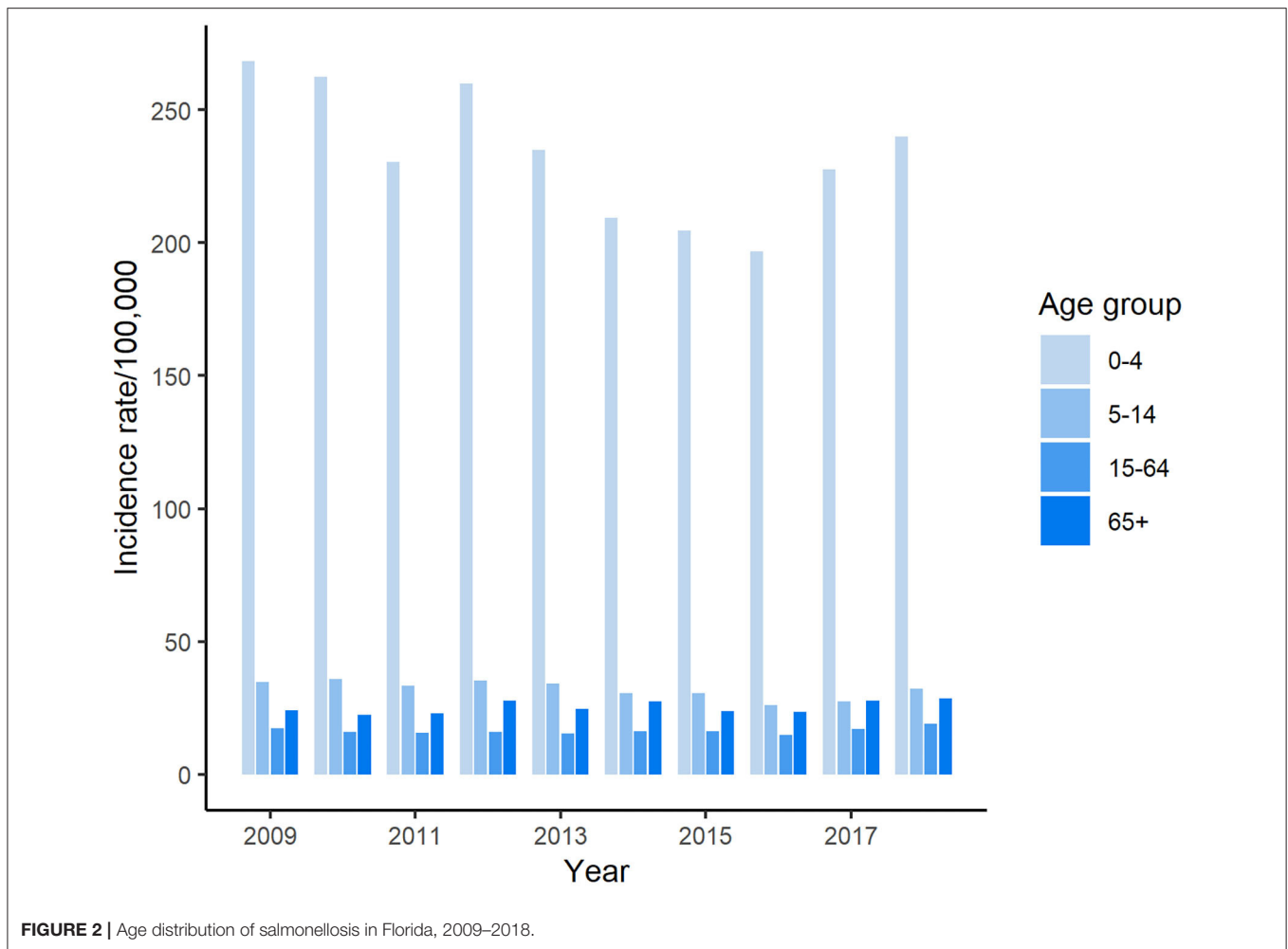
Metadata of sequenced isolates from the BPHL Bionumerics database and the FDOH Merlin database were merged by a surveillance epidemiologist at FDOH and deidentified data were shared with UF under a data sharing agreement, resulting in a set of 2,507 complete records, 1,353 isolates from 2017 and 1,154 isolates from 2018. Sequencing of included isolates was

performed at BPHL in Jacksonville who follows a standardized sequencing protocol and shares all sequenced data with CDC for surveillance purposes. These data are submitted to the NCBI BioSample database (<https://www.ncbi.nlm.nih.gov/biosample/>), under Bioproject number PRJNA230403 with limited metadata information. More detailed standardized protocol can be found in **Supplementary Material**.

All datasets were cleaned, organized and analyzed in the R statistical software version 3.6.0 (17).

### Epidemiological Analysis of Surveillance Data, 2009–2018

Annual incidence, incidence rate and monthly cumulative incidence from 2009 to 2018 in Florida were plotted to illustrate trends and seasonal patterns. The incidence rate was age-adjusted using direct standardization at both county and region level with the age distribution in the state of Florida as a reference (18).



A 3-month moving average was used to smooth the monthly salmonellosis data to assess the overall trend. A series of thematic maps, based on age-adjusted incidence rate data at the county level, were used to analyze spatial and temporal distribution of salmonellosis in counties of Florida.

Median polish was used to explore the variation of salmonellosis age-adjusted incidence rate among six regions in Florida as defined by the FDOH Weekly Morbidity Statistics Report (19) and years and was represented by the following formula:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

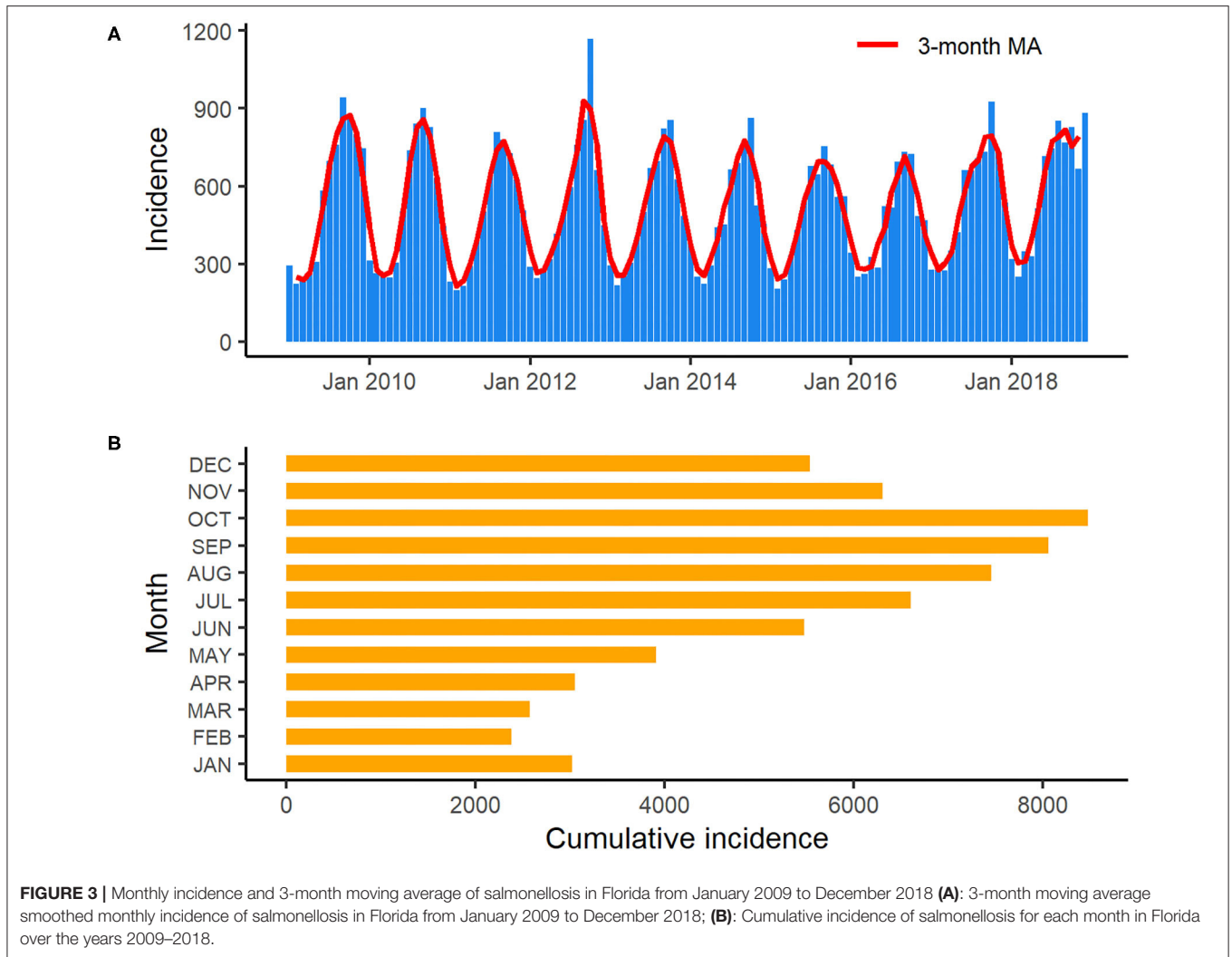
where  $y_{ij}$  is the response variable (incidence rate of salmonellosis) for row  $i$  and column  $j$ ,  $\mu$  is the common value (median),  $\alpha_i$  is the row (region) effect,  $\beta_j$  is the column (year) effect and  $\varepsilon_{ij}$  is the residual (20).

The row/column effects reflect the influence of each level of region/year on the incidence rate of salmonellosis, while the residuals indicate how incidence rates of salmonellosis from different regions and in different years are distinct from each other after removing the median value and corresponding row and column effects.

### Analysis of Laboratory-Sequenced Isolates, 2017–2018

Incidence rates of laboratory-sequenced isolates were calculated at the county level for the years 2017 and 2018 separately. For 2017, 56 isolates were dropped (6 with missing zip code information, 16 with PO Box zip codes and 34 from out of Florida residents), and 1,297 isolates were available for analysis. For 2018, 38 isolates were dropped (12 with missing zip code information, 16 with PO Box zip codes and 10 from out of Florida residents), and 1,116 isolates were available for analysis. Five-digit zip code was recorded for each sequenced isolate. We assigned zip code areas crossing county boundaries to one county based on the geometry center of the zip code shapefile polygon in ArcGIS (ERSI, Redlands, CA).

Serotype prediction was performed using Salmonella *in silico* Typing Resource (SISTR ver.1) algorithm built-in the EnteroBase online server (<https://enterobase.warwick.ac.uk/> accessed October 31, 2019) (21). Of 2,507 lab-sequenced isolate records, 1,709 records were available in the EnteroBase server and were used for serotype prediction in this study. SISTR version 1 was not able to resolve the serotypes for 68 isolates. We submitted these sequences to the developer group of SISTR (21),



who resolved the serotypes for these samples using an advanced version of SISTR (version 2) which is not yet released on any public platform.

To investigate whether sequenced isolates were a non-biased sample of reported cases, a zero-inflated negative binomial regression model was used including other putative explanatory variables [population density (person per square miles), median household income, and the percentage of children under 5 years and adults over 65 years, respectively, in the overall population].

## RESULTS

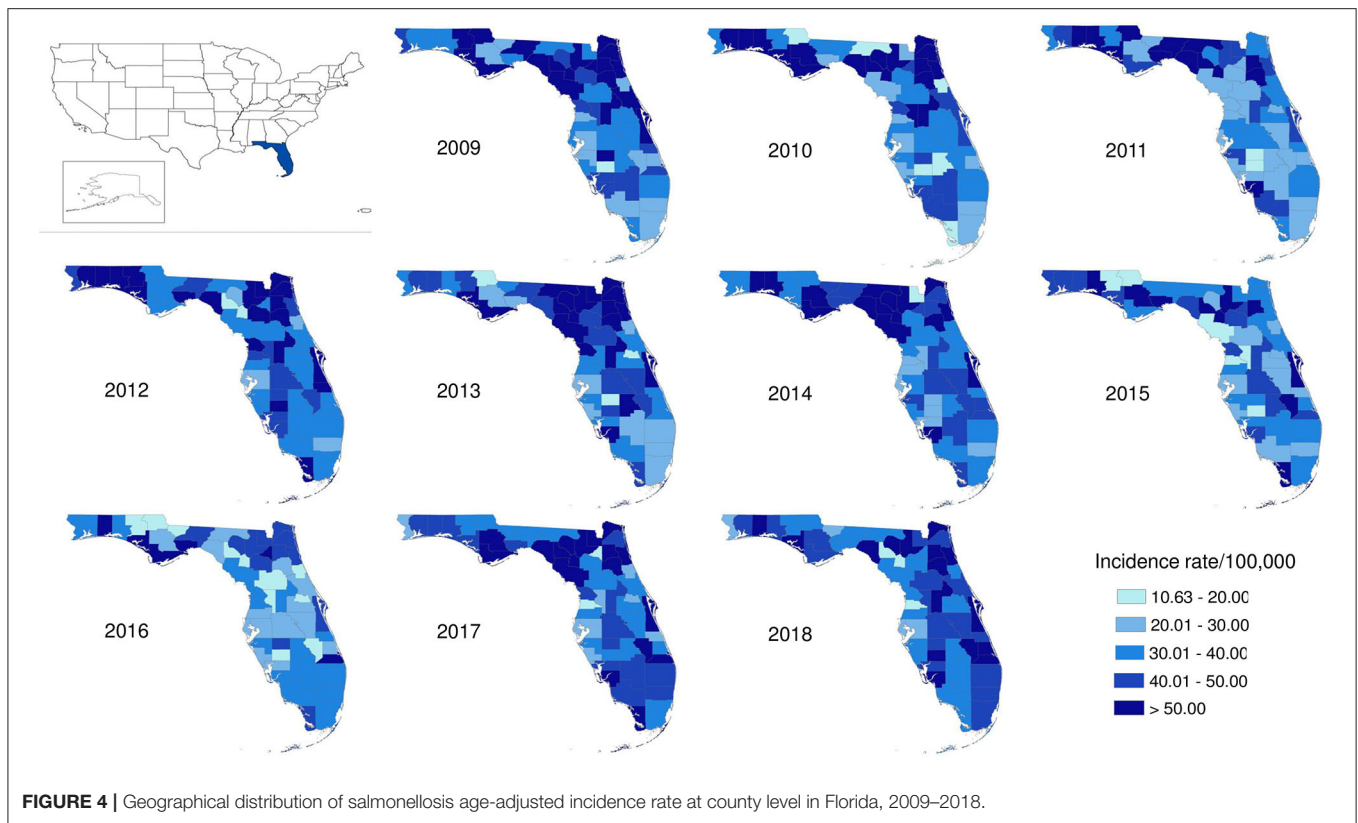
### Epidemiology and Demographics of Salmonellosis Reported Cases, 2009–2018

During the period of 2009–2018, 62,947 salmonellosis cases were reported in Florida, accounting for 11.8% of the total number in U.S. cases for these 10 years. The annual incidence rates ranged from 27.8 to 36.0 per 100,000 population, much higher than the national average incidence (15.2–18.64/100,000; **Figure 1**). There was a decreasing trend in incidence rate, except in 2012 when a peak appeared, with the incidence rate dropping from

36.0 per 100,000 population in 2009 to 27.8 per 100,000 in 2016. However, the incidence rate of salmonellosis started to increase again in 2017. Of 6,557 reported cases in 2017, 720 cases were confirmed only by culture-independent diagnostic tests (CIDTs) that was included as one of the criteria in the revised case definition of salmonellosis in Florida since 2017. The incidence continued to rise in 2018, and 718 cases were confirmed by CIDT only.

The highest number of salmonellosis cases was detected in the under 5-years-old age group, which accounted for 40.9% (25,737/62,947) of the total number of reported cases in Florida during these 10 years. Among age groups, the incidence rate in the under 5-years-old age group reached above 250 per 100,000 population in some years, which was more than 7 times higher than that in any other age groups (**Figure 2**). The average incidence rate of salmonellosis in 15–64-years-olds was the lowest of all age groups. The trend of incidence rate among children under 5 years of age displayed a similarity with the overall pattern of all reported cases during the decade, decreasing from 2009 through 2016 and then ascending again from 2017 onwards.





**FIGURE 4** | Geographical distribution of salmonellosis age-adjusted incidence rate at county level in Florida, 2009–2018.

## Seasonal Pattern

An obvious seasonality was observed from the distribution of salmonellosis cases from January 2009 to December 2018 (Figure 3). Salmonellosis occurred throughout the whole year in Florida, but the incidence gradually rose through spring and summer and peaked in September and October. After October, the incidence decreased notably, and the lowest incidence was typically observed in February. A total of 26.3% (16,547/62,947) of salmonellosis cases occurred in September and October on aggregate in the decade. Data from the Laboratory-based Enteric Disease Surveillance for 2006–2015 show that the peak at national level typically occurred in August, i.e., 1–2 months before the peak in Florida. The lowest national incidence was typically observed in February, similar to Florida (22).

## Geographical Distribution of Reported Salmonellosis Cases

The salmonellosis cases occurred in almost every county in Florida throughout 2009–2018 (Figure 4). The age-adjusted incidence rate in many counties could be as high as 50–60 per 100,000 population whereas in some years, it even went up to 80–90 per 100,000 population.

The age-adjusted incidence rate of reported cases by region as used in the Florida Department of Health Weekly Morbidity Statistics Report is presented in Table 1. On average, the Northeast and the Northwest regions had the highest incidence rate with 47.0 per 100,000; while the Southwest and the Southeast regions had the lowest. To examine the variation of incidence

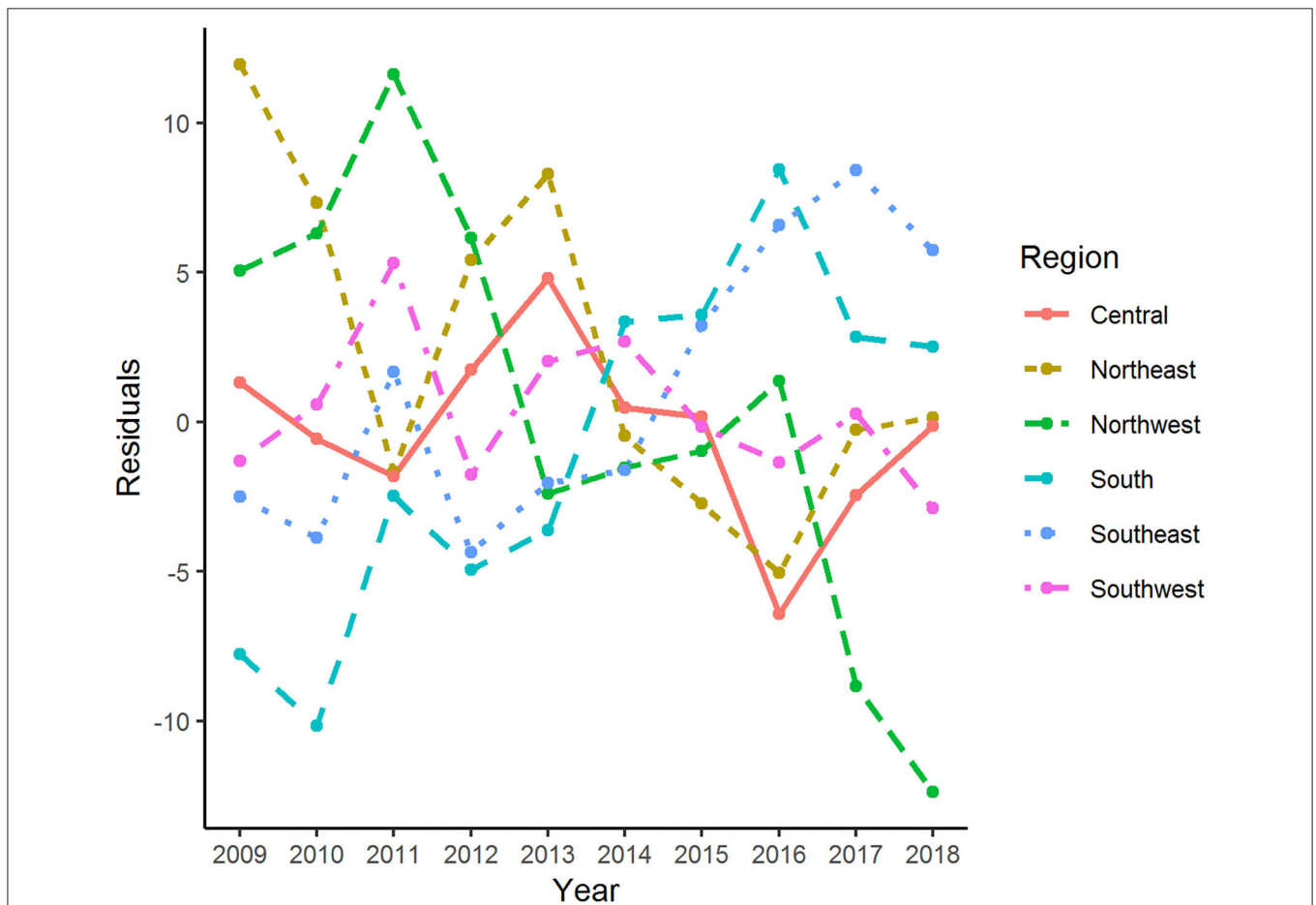
rates between regions across these 10 years, Tukey's median polish analysis was implemented with the data listed in Table 1. The overall row (region) effect showed that the Northwest region had the highest incidence rate of salmonellosis with 10.9 per 100,000 higher than overall median, respectively; while the Southeast region had the lowest incidence rate at 4.3 per 100,000 below the overall median. Similarly, the effects from years (column effect) showed a higher incidence rate of salmonellosis in 2018 (with an effect value of 5.3) and lower incidence rates in 2015 (with an effect value of  $-3.7$ ). A final residual table indicating the residuals of incidence rate for each region in a specific year after the row medians and column medians being removed iteratively was obtained and plotted as shown in Figure 5. It suggested that, although the incidence rates were higher in the Northeast and Northwest region than any other regions at the beginning of this decade, there was a decreasing trend on incidence rate observed in these two regions. On the contrary, the incidence rates in the Southeast and South regions were continually increasing during this time period.

## Descriptive Analysis of Lab-Sequenced Isolates, 2017–2018

The male: female ratio of sequenced isolates was 0.90:1.00 (Table 2). Based on the estimate by US Census Bureau in 2018, 26.1% of the Florida population were Hispanic and Latino, while 29.7% (617/2,076) of lab-sequenced isolates with known ethnicity were reported from this group between 2017 and 2018 (Table 2).

**TABLE 1** | Age-adjusted incidence rates of salmonellosis in Florida by region, 2009–2018.

Region	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Average
Northwest	53.0	52.7	54.6	56.2	43.0	45.1	41.9	44.4	39.3	39.5	47.0
Northeast	58.0	52.0	39.5	53.7	51.9	44.4	38.3	36.2	46.0	50.2	47.0
Central	40.4	37.0	32.3	43.0	41.4	38.3	34.2	27.8	36.8	42.9	37.4
Southwest	32.2	32.6	33.9	34.0	33.1	35.0	28.3	27.3	34.0	34.6	32.5
Southeast	30.2	27.4	29.5	30.6	28.2	29.9	30.9	34.5	41.3	42.5	32.5
South	27.2	23.3	27.6	32.2	28.9	37.1	33.5	38.6	38.0	41.5	32.8



**FIGURE 5** | Residuals of age-adjusted incidence rate of salmonellosis for various regions and years after median polish.

Similar to reported cases, most lab-sequenced isolates were found in children under 5-years of age (Table 3).

**Representativeness of Lab-Sequenced Isolates**

To statistically test the representativeness of count of lab-sequenced isolates (spread of frequency counts included as Supplementary Figure 1) for the reported cases in counties of Florida, we fitted a zero-inflated negative binomial (ZINB) regression model with the number of reported cases from the surveillance system, population density (persons per square mile), median household income, the percentage of children

under 5 years in the overall population and the percentage of adults over 65 years in the overall population as putative explanatory variables. The overall model performance statistics with statistically significant predictors are enumerated in Table 4, and the plot showing the model fit with observed data is included in Supplementary Figure 3. The Vuong test suggested that ZINB model was a significant improvement over the standard negative binomial (NB) model (detailed model diagnostics are enumerated in Supplementary Tables 1–3, and model fit with observed data is depicted in Supplementary Figure 2). In the logit model predicting excessive zeros, reported case number was

**TABLE 2** | Demographics of 2,507 laboratory-sequenced salmonellosis cases with complete metadata in Florida, 2017–2018.

Category	Number of cases
<b>Gender</b>	
Male	1,186
Female	1,313
Unknown	8
<b>Ethnicity</b>	
Hispanic	617
Non-Hispanic	1,459
Unknown	431

the only (marginally) significant variable. ZINB model correctly predicted 19.79% excess zero for original proportion of 20.15%. The negative association suggested that the greater the reported case number, the less likely it was that no isolate was sequenced in a county. Reported case number, population density, and median household income were significant predictors in the negative binomial model predicting the count of the sequenced isolates. The prediction performance of these three predictors was further compared after standardization, demonstrating that the number of reported cases was the strongest predictor of the number of lab-sequenced isolates (Figure 6).

### Serotypes of Salmonella Isolates

Based on the 1,709 isolates that both were available in Enterobase and had complete metadata information available, a considerable variety of *Salmonella* serotypes was detected in Florida. Of the top-20 most frequently detected serotypes, 19 belonged to subspecies I *S. enterica* ssp. *enterica*, while one belonged to subspecies IV *S. enterica* ssp. *houtenae*. Serotypes Enteritidis, Javiana, Newport, Sandiego, and Braenderup were the top five most commonly detected serotypes (Table 5). Serotype Enteritidis had the highest incidence rate (0.90 per 100,000), followed by Newport and Javiana with incidence rates of 0.79 per 100,000 and 0.60 per 100,000, respectively. Of 1,709 isolates, ~46% belonged to less frequently detected serotypes (including 798 isolates that could not be assigned a serotype by SISTR).

### Sample Sources and Serotypes

Sources of all 2,507 samples delivered to state laboratories were summarized by serotype, and the source distribution for the top 20 serotypes was displayed in Table 5. For all top 20 serotypes, the percentage of samples from stool ranged from 2.9 to 92.6% with an overall percentage of 84.7%. Overall percentage for blood was 5.2%; while, for serotypes Oranienburg (13.2%), Sandiego (12.2%), IV 50:z4,z23:- (Flint; 11.7%), and Poona (10.3%), at least 10% of isolates were from blood. Overall, 3.0% of isolates were from urine but 97.1% of isolates from serotype I 4,[5],12:i:-, a monophasic variant of Typhimurium, followed by 17.1% of Anatum isolates were from urine.

**TABLE 3** | Age distribution of reported and laboratory-sequenced salmonellosis cases in Florida, 2017–2018.

Age	Reported cases	Lab-sequenced isolates
0–4 years	5,287 (38.4%)	1,037 (41.4%)*
5–14 years	1,381 (10.0%)	240 (9.6%)
15–64 years	4,778 (34.7%)	824 (32.9%)
65+ years	2,329 (16.9%)	405 (16.2%)
Unknown	0	1 (0.0%)
Total	13,775 (100%)	2,507 (100%)

\*Percent of cases with known age.

### Geospatial Distribution of Lab-Sequenced Isolates

The age-adjusted incidence rates of lab-sequenced *Salmonella* isolates in counties of Florida in 2017 and 2018 were illustrated in Figure 7. The origin of the *Salmonella* isolates appears to be distributed over the whole state. Generally, the incidence rates of lab-sequenced isolates were notably lower than that of reported cases, with most counties having an incidence rate <15 per 100,000 population. The incidence rates ranged from 0 to 41.1 per 100,000 population in 2017 and from 0 to 7.5 per 100,000 population in 2018. The incidence rates across counties were relatively higher in 2017 than 2018, and some counties with higher incidence rates were observed in the Northwest and Northeast regions in 2017.

## DISCUSSION

Approximately 11,000–14,000 cases of enteric illness are reported each year in Florida, which makes Florida consistently one of the states with highest incidence rate of enteric diseases across the country (23). Among those, salmonellosis is the most common enteric illness in Florida with 6,000–7,000 cases reported annually. Generally, young children, especially children under 5 years old, are disproportionately affected by enteric illnesses (24, 25). Salmonellosis is not an exception. The CDC's Foodborne Diseases Active Surveillance Network (FoodNet) reported that children aged <5 years old had the highest incidence rate of salmonellosis (56.3 cases per 100,000 population) among age groups in 2015 (26). In Florida, the incidence rate in this age group ranged from 196.7 to 268.2 cases per 100,000 population during the years of 2009–2018. Such higher incidence rate in children may be partly due to the increased exposure to enteric pathogens through fecal-oral contaminations in different exposure settings (25, 27). Several risk factors like contact with pet reptiles and amphibians or live poultry (28, 29), riding in a shopping cart next to meat or poultry (30), and attending a day care center with a *Salmonella*-infected child have been linked to salmonellosis in children (30). It is not clear which proportion of *Salmonella* infections among children in Florida that can be attributed to these risk factors.

Despite the high reported incidence of salmonellosis in Florida, the *Salmonella* Annual Reports issued by CDC suggest a very low incidence rate in the state. For example, while the incidence rate of reported cases was 27.8 per 100,000 population



**TABLE 4** | Zero-inflated negative binomial regression model for lab-sequenced isolates per county.

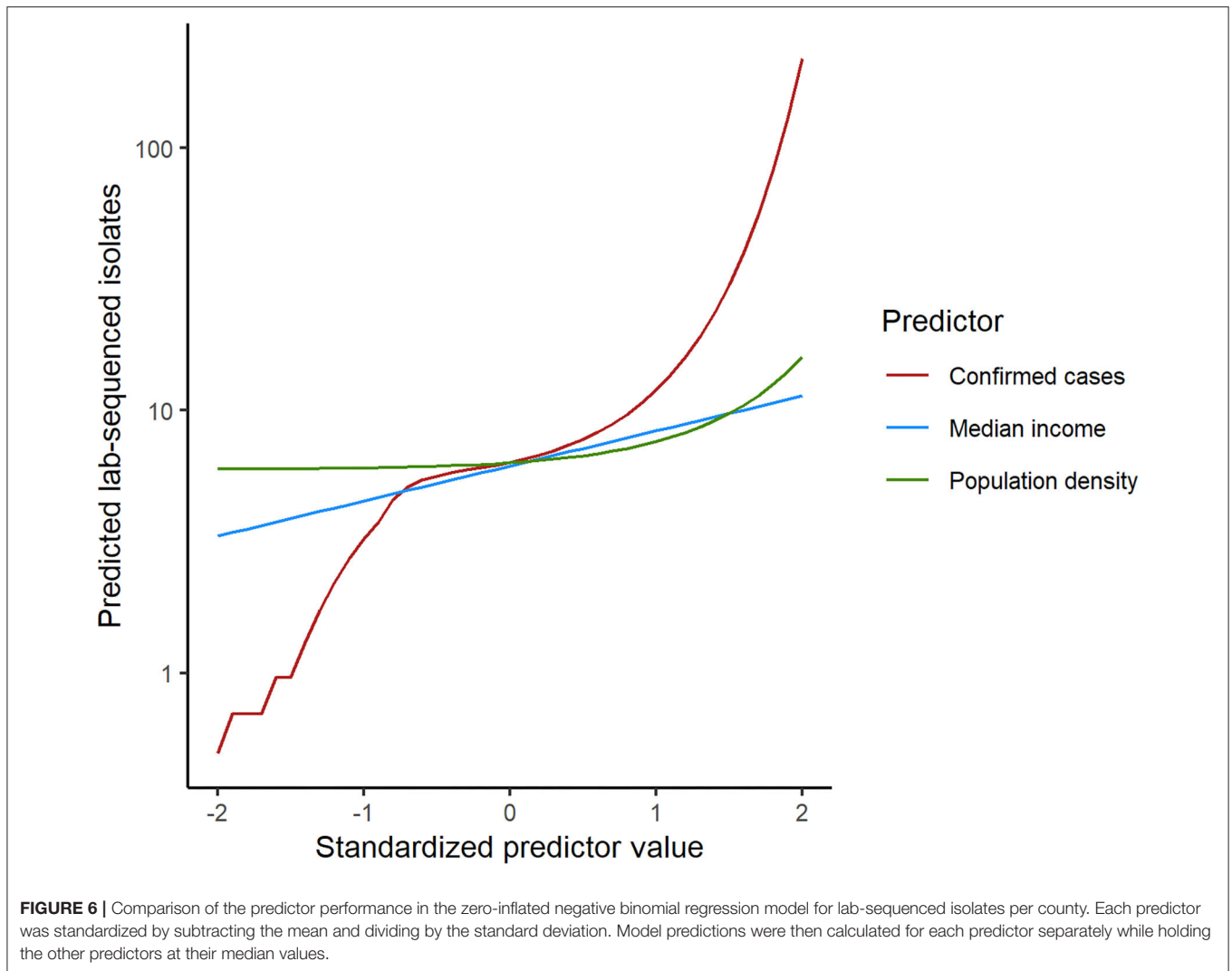
Parameter	Logit			Negative binomial		
	Estimate	95% CI*	P-value	Estimate	95% CI	P-value
Intercept	2.66	(0.14, 5.17)	0.04	-15.80	(-24.49, -7.12)	<0.001
Reported case No.	-0.38	(-0.77, 0.01)	0.06	0.005	(0.004, 0.006)	<0.001
Population density				0.0004	(0.0002, 0.0006)	<0.001
Ln(median household income)				1.61	(0.81, 2.41)	<0.001
Dispersion	0.93					
Log-likelihood	-378.84; df = 7	<0.001				
AIC	771.68; df = 7					
BIC	791.97; df = 7					
RMSE <sup>‡</sup> (LOO <sup>§</sup> cross-validation average)	15.29					
Vuong test of ZINB versus NB <sup>#</sup>			<0.002			

\*95% confidence interval.

# ZINB, zero-inflated negative binomial; NB, negative binomial.

‡ Root Mean Squared Error.

§ Leave one out.



**TABLE 5** | Top 20 *Salmonella* serotypes and corresponding sample sources in Florida based on the lab-sequenced salmonellosis cases between 2017 and 2018.

Serotype	Number reported	Percent	Incidence/100,000	Blood (%)*	Stool (%)	Urine (%)	Unknown (%)
Enteritidis	185	10.83	0.9	14 (7.6)	153 (82.7)	4 (2.2)	14 (7.6)
Newport	163	9.54	0.793	7 (4.3)	137 (84.1)	4 (2.5)	15 (9.2)
Javiana	123	7.2	0.598	3 (2.4)	106 (86.2)	7 (5.7)	7 (5.7)
Sandiego	123	7.2	0.598	15 (12.2)	99 (80.5)	6 (4.9)	3 (2.4)
Braenderup	91	5.32	0.443	0	79 (86.8)	6 (6.6)	6 (6.6)
Typhimurium	83	4.86	0.404	6 (7.2)	68 (80.7)	2 (2.4)	8 (9.6)
IV 50:z4,z23:-	77	4.51	0.375	9 (11.7)	55 (71.4)	4 (5.2)	9 (11.7)
Muenchen	60	3.51	0.292	0 (0)	54 (90)	1 (1.7)	5 (8.3)
Rubislaw	59	3.45	0.287	3 (5.1)	49 (83.1)	2 (3.4)	5 (8.5)
Bareilly	54	3.16	0.263	0	50 (92.6)	1 (1.9)	3 (5.6)
Saintpaul	46	2.69	0.224	1 (2.2)	42 (91.3)	2 (4.4)	1 (2.2)
Montevideo	44	2.57	0.214	1 (2.3)	37 (84.1)	1 (2.3)	5 (11.4)
Poona	39	2.28	0.19	4 (10.3)	31 (79.5)	1 (2.6)	3 (7.7)
Oranienburg	38	2.22	0.185	5 (13.2)	28 (73.7)	1 (2.6)	4 (10.5)
Anatum	35	2.05	0.17	0	29 (82.9)	6 (17.1)	0
I 4,[5],12:i:-	35	2.05	0.17	0	1 (2.9)	34 (97.1)	0
Paratyphi B var. Java	25	1.46	0.122	0	22 (88)	1 (4)	2 (8)
Baildon	24	1.4	0.117	2 (8.3)	20 (83.3)	1 (4.2)	1 (4.2)
Thompson	23	1.35	0.112	2 (8.7)	21 (91.3)	0	0
Infantis	22	1.29	0.107	0	18 (81.8)	1 (4.6)	3 (13.6)
Other Serotypes	360	21.06	1.751	25 (6.9)	286 (78.3)	18 (5.0)	34 (9.7)
Serotype not available	798	46.69	3.882	32 (4.0)	705 (88.7)	4 (0.5)	54 (6.8)
All isolates	2,507	146.69	12.197	129 (5.2)	2,123 (84.7)	74 (3.0)	181 (7.2)

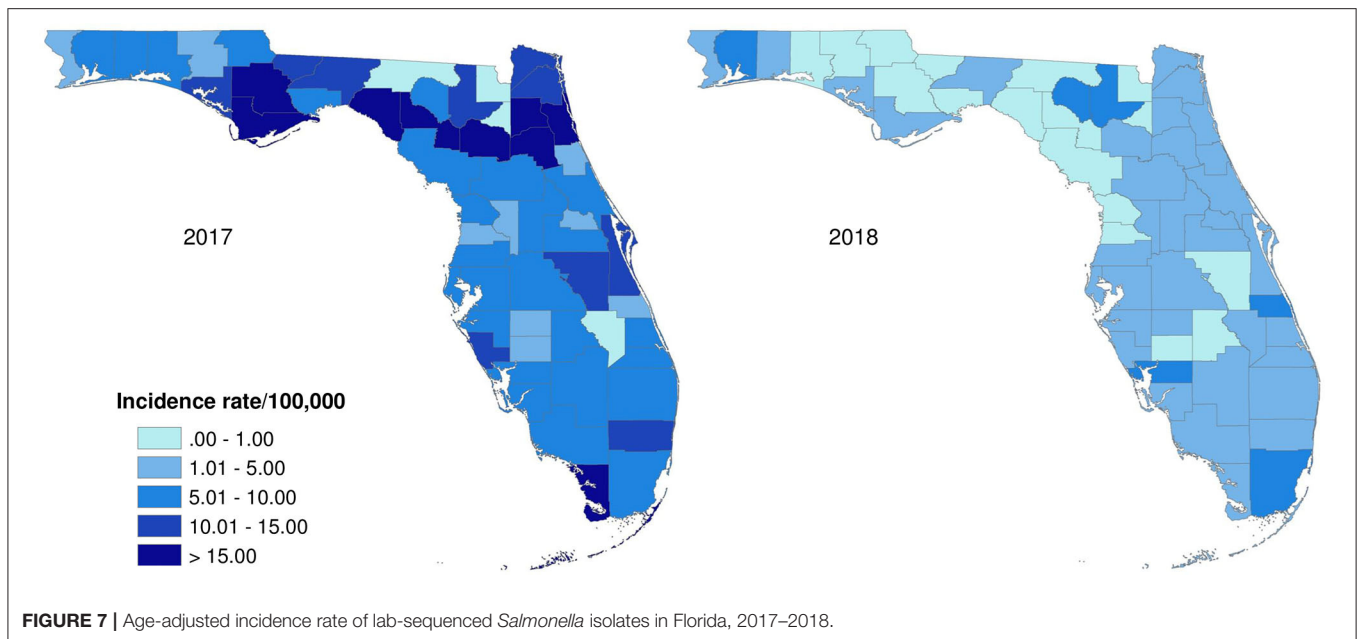
\*Source Percentages are row-wise for each serotype per sample group.

in 2016 (the latest year for which the CDC report is currently available), the *Salmonella* Annual Report indicated an incidence rate of 5.07 per 100,000 population, the lowest in the country (22). Infectious disease surveillance results are reported from Florida to the National Notifiable Diseases Surveillance System (NNDSS), which accurately reflects the incidence reported by Florida. In contrast, the *Salmonella* Annual Reports are based on the Laboratory-based Enteric Disease Surveillance (LEDS) system, which collects data from state and regional public health laboratories. The seemingly disparate results are related to the limited capacity the state had to access and fully type isolates from all confirmed cases. Before 2016, when new requirements to submit isolates or samples were included in the Florida Administrative Code, BPHL only received between 1,900 and 2,500 specimens per year. The number of isolates submitted to BPHL has increased from about 4,500 in 2017 to about 7,200 in 2019. The current capacity of BPHL for sequencing *Salmonella* isolates is approximately 8,000. In 2016, Florida reported 1,007 isolates to LEDS, and this has increased to 5,810 in 2018. Thus, new regulations and the WGS workflow have greatly increased coverage of surveillance.

The distribution of salmonellosis cases in Florida displayed a strong seasonality, with cases spiking in late summer/early fall (September–October). Florida has a distinct climatic pattern with high average temperature and precipitation, especially in summer (the rainy season extends from May through October). Whether this specific climatic pattern contributes to the seasonality of

salmonellosis, especially infections caused by several *Salmonella* serotypes related to non-food and environmental exposure needs to be further studied. It would be interesting to compare with seasonality of salmonellosis in other states with similar climates. More tropical storms taking place in the rainy season is another typical feature of Florida's climate, and nearly one third of these storms occurred in September. The influences of extreme weather events like hurricanes on food- and water-borne diseases have been previously reported (31, 32). The association between increasing salmonellosis cases and the effects of extreme weather events is also an interesting topic to address in the future study.

Though the incidence rate of salmonellosis has been continually high in Florida compared to the national level, it decreased by 23% in 2016 compared to 2009 but rebounded after 2016 and reached a peak in 2018. With the increasing use of CIDTs by clinical laboratories to diagnose enteric pathogens in recent years nationwide, CIDT was included in the criteria for suspected cases as a part of the expansion of salmonellosis case definition by Florida Department of Health in January 2017 (15). The CIDT positive-only cases contributed to 76 and 45% of the increase in total case number in 2017 and 2018, respectively, compared with 2016 in Florida. Interestingly, the incidence rate of salmonellosis did not seem to be affected obviously by the wide adoption of CIDTs and remained relatively stable during these years (Figure 1). Nevertheless, Marder et al. reported the increasing use of CIDTs had complicated the interpretation of the enteric infection surveillance data (33). In some cases,



the changes in laboratory testing practices had significantly increased the incidence rate of infections caused by enteric pathogens. For example, based upon the data from ten U.S. sites of FoodNet between 2013 and 2016, there was a significant increase in the incidence rates of both confirmed Shiga toxin-producing *Escherichia coli* (STEC) infection alone and confirmed or CIDT positive-only infections in 2016 compared with the average level of 2013–2015. For *Salmonella*, that increase was slight but not significant, although the proportion of clinical laboratories within the network of FoodNet that used CIDTs to detect *Salmonella* increased markedly to 14% in 2016 from <1% in 2013 (33). It is worth noting that the proportion of *Salmonella* infections identified by CIDTs without culture confirmation was only 8% within the ten sites of FoodNet, the lowest among other enteric pathogens monitored by FoodNet in 2016.

Underreporting is one of the major limitations for most surveillance systems of notifiable diseases; enteric disease surveillance system is not an exception. These surveillance data are heavily based upon notification of ill cases from health care providers and laboratories that can only capture affected cases who seek medical care after getting sick, whose specimens are tested by laboratory and results are being reported to the public health authorities. These laboratory-confirmed illnesses are only a fraction of the true disease burden across the nation. CDC estimates that for every case of salmonellosis reported to the national surveillance system, there are 29.3 cases in the population (4). Such estimates are not available for individual states. Regardless of the limitations, surveillance data are still helpful for determining the trend or patterns of disease occurrence and burdens (34).

Population data from floridahealthcharts.com do not account for annual fluctuations in population numbers affected by tourism and seasonal residence (snowbirds, seasonal agricultural workers). These areas are mainly found in Central and Southern Florida. Therefore, the incidence rates of salmonellosis may have

been overestimated for these regions. Nevertheless, we did find that incidence rates were higher in the North and Northeast than in other regions. We did find an increasing incidence rate over the years in the Southern part of the state which might partly be explained by seasonal migration that is unaccounted for in population numbers.

Epidemiological studies have suggested that socioeconomic status (SES) indicators including educational attainment, household income, and poverty level are determinants that have been linked to the risk of salmonellosis at both individual and community level (35–40). Lower educational level of parents in a household, always along with a higher level of poverty, may be related to improper practice of food handling and poor personal hygiene, which contributes to the high occurrence of salmonellosis among their children living in this environment (36, 37). Therefore, adults with lower educational attainment should be the target population of food safety interventions in the future to decrease the incidence rate of salmonellosis among children.

Household income was another potential risk factor responsible for high incidence rate of salmonellosis. Previous studies have reported that families with higher income were more likely to seek healthcare with higher healthcare accessibility, eat outside, keep pets at home, and have more frequent international travels, which had been identified as risk factors that increase the exposure to enteric pathogens (36, 39). Besides, Varga et al. suggested that lower household income was also associated with increased *S. Enteritidis* infection in Canada, which can be partially explained by the universal access to provincially-funded healthcare in Canada, consumption of poor microbial quality food in low-income households, as well as the frequent violation of food safety procedures by retailers in the neighborhoods of low-income households (39). To better understand the role of household income in the increased risk of exposure to enteric pathogens, further serotype-specific

analysis studying these associations needs to be conducted in the future.

Since there are a large variety of serotypes of *Salmonella*, and related reservoirs and transmission pathways could differ between serotypes (7, 41), it is of significance to continuously monitor the prevalence of *Salmonella* serotypes globally and locally. We report this distribution in Florida based on WGS data during 2017–2018. More detailed serotype sequence characterizations need further investigation. The sheer number of cases reported in Florida make it impossible to sequence all the salmonellosis cases reported to FDOH. A relevant question was therefore how representative the sequenced isolates were for all reported cases. Our results suggested that the number of laboratory-sequenced isolates was significantly associated with the reported case number but also with median income and population density per county. However, the predictive value of these two additional variables was much lower than the number of reported cases. As discussed above, enteric illnesses including salmonellosis disproportionately affect young children. When fitting the regression model, we also included the proportion of children under 5 years old in the whole population as a predictor in the initial model. This indicator did not fit significantly in the model, suggesting the distribution of sequenced isolates was not biased by the influence of age group. We therefore conclude that the lab-sequenced isolates are a relatively unbiased sample of all reported cases.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/>, PRJNA230403. Additional data have been deposited in Harvard Dataverse: doi: 10.7910/DVN/KBZLMV.

## REFERENCES

- Chen HM, Wang Y, Su LH, Chiu CH. Nontyphoid salmonella infection: microbiology, clinical features, and antimicrobial therapy. *Pediatr Neonatol.* (2013) 54:147–52. doi: 10.1016/j.pedneo.2013.01.010
- Vally H, Glass K, Ford L, Hall G, Kirk MD, Shadbolt C, et al. Proportion of illness acquired by foodborne transmission for nine enteric pathogens in Australia: an expert elicitation. *Foodborne Pathog Dis.* (2014) 11:727–33. doi: 10.1089/fpd.2014.1746
- Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal *Salmonella* gastroenteritis. *Clin Infect Dis.* (2010) 50:882–9. doi: 10.1086/650733
- Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, et al. Foodborne illness acquired in the United States—major pathogens. *Emerg Infect Dis.* (2011) 17:7–15. doi: 10.3201/eid1701.P11101
- Pires SM, Vieira AR, Hald T, Cole D. Source attribution of human salmonellosis: an overview of methods and estimates. *Foodborne Pathog Dis.* (2014) 11:667–76. doi: 10.1089/fpd.2014.1744
- Interagency Food Safety Analytics Collaboration. *Foodborne Illness Source Attribution Estimates for 2017 for Salmonella, Escherichia coli O157, Listeria monocytogenes, Campylobacter Using Multi-year Outbreak Surveillance Data, United States.* GA D.C.: U.S. Department of Health and Human Services, CDC, FDA, USDA-FSIS (2019).
- Beshearse E, Bruce BB, Nane GF, Cooke RM, Aspinall W, Hald T, et al. Attribution of illnesses transmitted commonly by food and water in the

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Florida Institutional Review Board. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

XL, NS, and AH conceived this manuscript. JB led the laboratory sequencing work. JD led the surveillance work. XL and NS analyzed the data. XL, NS, EB, and AH wrote the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This analysis was funded by the Florida Department of Health through the Florida Integrated Food Safety Center of Excellence.

## ACKNOWLEDGMENTS

The authors thank Dr. R. Michael Hoekstra for reviewing the manuscript and providing helpful suggestions.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2020.603005/full#supplementary-material>

- United States to comprehensive transmission pathways using structured expert judgment. *Emerg Infect Dis.* (accepted).
- Luvsansharav UO, Vieira A, Bennett S, Huang J, Healy JM, Hoekstra RM, et al. Serotypes: a novel measure of association with foodborne transmission. *Foodborne Pathog Dis.* (2020) 17:151–5. doi: 10.1089/fpd.2019.2641
- Crim SM, Chai SJ, Karp BE, Judd MC, Reynolds J, Swanson KC, et al. *Salmonella enterica* serotype Newport infections in the United States, 2004–2013: increased incidence investigated through four surveillance systems. *Foodborne Pathog Dis.* (2018) 15:612–20. doi: 10.1089/fpd.2018.2450
- Judd MC, Hoekstra RM, Mahon BE, Fields PI, Wong KK. Epidemiologic patterns of human *Salmonella* serotype diversity in the USA, 1996–2016. *Epidemiol Infect.* (2019) 147:e187. doi: 10.1017/S09502688190000724
- Simpson KMJ, Hill-Cawthorne GA, Ward MP, Mor SM. Diversity of *Salmonella* serotypes from humans, food, domestic animals and wildlife in New South Wales, Australia. *BMC Infect Dis.* (2018) 18:623. doi: 10.1186/s12879-018-3563-1
- Snyder TR, Boktor SW, M'ikanatha NM. Salmonellosis outbreaks by food vehicle, serotype, season, and geographical location, United States, 1998 to 2015. *J Food Prot.* (2019) 82:1191–9. doi: 10.4315/0362-028X.JFP-18-494
- Adams DA, Thomas KR, Jajosky RA, Foster L, Baroi G, Sharp P, et al. Summary of notifiable infectious diseases and conditions - United States, 2015. *MMWR Morb Mortal Wkly Rep.* (2017) 64:1–143. doi: 10.15585/mmwr.mm6453a1

14. Florida Department of Health. *Laboratory Reporting Guidelines for Reportable Diseases and Conditions in Florida*. (2016). Available online at: [http://www.floridahealth.gov/diseases-and-conditions/disease-reporting-and-management/\\_documents/guidelines-laboratories.pdf](http://www.floridahealth.gov/diseases-and-conditions/disease-reporting-and-management/_documents/guidelines-laboratories.pdf) (accessed December 15, 2019).
15. Florida Department of Health. *Surveillance Case Definitions for Select Reportable Diseases in Florida*. (2017). Available online at: [http://www.floridahealth.gov/diseases-and-conditions/disease-reporting-and-management/disease-reporting-and-surveillance/\\_documents/2017-surveillance-case-definitions.pdf](http://www.floridahealth.gov/diseases-and-conditions/disease-reporting-and-management/disease-reporting-and-surveillance/_documents/2017-surveillance-case-definitions.pdf) (accessed December 15, 2019).
16. Li X, Sapp AC, Singh N, Matthias L, Bailey C, DeMENT J, et al. Detecting foodborne disease outbreaks in Florida through consumer complaints. *J Food Prot.* (2020) 83:1877–88. doi: 10.4315/JFP-20-138
17. R Core Team. *R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing*. Vienna: R Core Team (2019). Available online at: <http://www.R-project.org/> (accessed October 31, 2019).
18. Kahn HA, Sempos CT. *Statistical Methods in Epidemiology*. New York, NY: Oxford University Press (1989).
19. Florida Department of Health. *Weekly Morbidity Statistics Report*. Available online at: <http://www.floridahealth.gov/diseases-and-conditions/disease-reporting-and-management/disease-reporting-and-surveillance/data-and-publications/weekly-morbidity-reports/index.html> (accessed December 15, 2019).
20. Hoaglin DC, Mosteller F, Tukey JW. *Exploring Data Tables, Trends, and Shapes*. New Jersey, NJ: John Wiley & Sons (2011).
21. Yoshida CE, Kruczkiewicz P, Laing CR, Lingohr EJ, Gannon VP, Nash JH, et al. The Salmonella in silico typing resource (SISTR): an open web-accessible tool for rapidly typing and subtyping draft Salmonella genome assemblies. *PLoS ONE.* (2016) 11:e0147101. doi: 10.1371/journal.pone.0147101
22. Centers for Disease Control and Prevention (CDC). *National Enteric Disease Surveillance: Salmonella Annual Report, 2016*. Atlanta, GA: Centers for Disease Control and Prevention (2018).
23. Florida Department of Health. *Florida Annual Morbidity Statistics Report*. (2017). Available online at: [http://www.floridahealth.gov/diseases-and-conditions/disease-reporting-and-management/disease-reporting-and-surveillance/data-and-publications/\\_documents/2017-annual-morbidity-statistics-report.pdf](http://www.floridahealth.gov/diseases-and-conditions/disease-reporting-and-management/disease-reporting-and-surveillance/data-and-publications/_documents/2017-annual-morbidity-statistics-report.pdf) (accessed December 15, 2019).
24. Cheng LH, Crim SM, Cole CR, Shane AL, Henao OL, Mahon BE. Epidemiology of infant salmonellosis in the United States, 1996–2008: a foodborne diseases active surveillance network study. *J Pediatric Infect Dis Soc.* (2013) 2:232–9. doi: 10.1093/jpids/pit020
25. Scallan E, Mahon BE, Hoekstra RM, Griffin PM. Estimates of illnesses, hospitalizations and deaths caused by major bacterial enteric pathogens in young children in the United States. *Pediatr Infect Dis J.* (2013) 32:217–21. doi: 10.1097/INF.0b013e31827ca763
26. Centers for Disease Control and Prevention. *Foodborne Diseases Active Surveillance Network (FoodNet): FoodNet 2015 Surveillance Report (Final Data)*. (2017). Available online at: <https://www.cdc.gov/foodnet/pdfs/FoodNet-Annual-Report-2015-508c.pdf> (accessed December 15, 2019).
27. Bula-Rudas FJ, Rathore MH, Maraqa NF. Salmonella infections in childhood. *Adv Pediatr.* (2015) 62:29–58. doi: 10.1016/j.yapd.2015.04.005
28. Wells EV, Boulton M, Hall W, Bidol SA. Reptile-associated salmonellosis in preschool-aged children in Michigan, January 2001–June 2003. *Clin Infect Dis.* (2004) 39:687–91. doi: 10.1086/423002
29. Basler C, Forshey TM, Machesky K, Erdman CM, Gomez TM, Brinson DL, et al. Notes from the field: multistate outbreak of human Salmonella infections linked to live poultry from a mail-order hatchery in Ohio—February–October 2014. *MMWR Morb Mortal Wkly Rep.* (2015) 64:258.
30. Jones TF, Ingram LA, Fullerton KE, Marcus R, Anderson BJ, McCarthy PV, et al. A case-control study of the epidemiology of sporadic Salmonella infection in infants. *Pediatrics.* (2006) 118:2380–7. doi: 10.1542/peds.2006-1218
31. Hulland E, Subaiya S, Pierre K, Barthelemy N, Pierre JS, Dismar A, et al. Increase in reported cholera cases in Haiti following Hurricane Matthew: an interrupted time series model. *Am J Trop Med Hyg.* (2019) 100:368–73. doi: 10.4269/ajtmh.17-0964
32. Roca MA, Brown RS, Solo-Gabriele HM. Fecal indicator bacteria levels at beaches in the Florida keys after Hurricane Irma. *Mar Pollut Bull.* (2019) 138:266–73. doi: 10.1016/j.marpolbul.2018.09.036
33. Marder EP, Cieslak PR, Cronquist AB, Dunn J, Lathrop S, Rabatsky-Ehr T, et al. Incidence and trends of infections with pathogens transmitted commonly through food and the effect of increasing use of culture-independent diagnostic tests on surveillance - foodborne diseases active surveillance network, 10 U.S. Sites, 2013–2016. *MMWR Morb Mortal Wkly Rep.* (2017) 66:397–403. doi: 10.15585/mmwr.mm6615a1
34. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, et al. Food-related illness and death in the United States. *Emerg Infect Dis.* (1999) 5:607–25. doi: 10.3201/eid0505.990502
35. Chang M, Groseclose SL, Zaidi AA, Braden CR. An ecological analysis of sociodemographic factors associated with the incidence of salmonellosis, shigellosis, and *E. coli* O157:H7 infections in US counties. *Epidemiol Infect.* (2009) 137:810–20. doi: 10.1017/S0950268808001477
36. Younus M, Hartwick E, Siddiqi AA, Wilkins M, Davies HD, Rahbar M, et al. The role of neighborhood level socioeconomic characteristics in Salmonella infections in Michigan (1997–2007): assessment using geographic information system. *Int J Health Geogr.* (2007) 6:56. doi: 10.1186/1476-072X-6-56
37. Simonsen J, Frisch M, Ethelberg S. Socioeconomic risk factors for bacterial gastrointestinal infections. *Epidemiology.* (2008) 19:282–90. doi: 10.1097/EDE.0b013e3181633c19
38. Quinlan JJ. Foodborne illness incidence rates and food safety risks for populations of low socioeconomic status and minority race/ethnicity: a review of the literature. *Int J Environ Res Public Health.* (2013) 10:3634–52. doi: 10.3390/ijerph10083634
39. Varga C, Pearl DL, McEwen SA, Sargeant JM, Pollari F, Guerin MT. Evaluating area-level spatial clustering of Salmonella Enteritidis infections and their socioeconomic determinants in the greater Toronto area, Ontario, Canada (2007–2009): a retrospective population-based ecological study. *BMC Public Health.* (2013) 13:1078. doi: 10.1186/1471-2458-13-1078
40. Hadler JL, Clogher P, Libby T, Wilson E, Oosmanally N, Ryan P, et al. Relationship between census tract-level poverty and domestically acquired Salmonella incidence: analysis of foodborne diseases active surveillance network data, 2010–2016. *J Infect Dis.* (2020) 222:1405–12. doi: 10.1093/infdis/jiz605
41. Boore AL, Hoekstra RM, Iwamoto M, Fields PI, Bishop RD, Swerdlow DL. *Salmonella enterica* infections in the United States and assessment of coefficients of variation: a novel approach to identify epidemiologic characteristics of individual serotypes, 1996–2011. *PLoS ONE.* (2015) 10:e0145416. doi: 10.1371/journal.pone.0145416

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Li, Singh, Beshearse, Blanton, DeMent and Havelaar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.