

BMJ Open *Salmonella* infections modelling in Mississippi using neural network and geographical information system (GIS)

Luma Akil, H Anwar Ahmad

To cite: Akil L, Ahmad HA. *Salmonella* infections modelling in Mississippi using neural network and geographical information system (GIS). *BMJ Open* 2016;**6**:e009255. doi:10.1136/bmjopen-2015-009255

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2015-009255>).

Received 15 July 2015
Revised 2 December 2015
Accepted 3 December 2015

ABSTRACT

Objectives: Mississippi (MS) is one of the southern states with high rates of foodborne infections. The objectives of this paper are to determine the extent of *Salmonella* and *Escherichia coli* infections in MS, and determine the *Salmonella* infections correlation with socioeconomic status using geographical information system (GIS) and neural network models.

Methods: In this study, the relevant updated data of foodborne illness for southern states, from 2002 to 2011, were collected and used in the GIS and neural networks models. Data were collected from the Centers for Disease Control and Prevention (CDC), MS state Department of Health and the other states department of health. The correlation between low socioeconomic status and *Salmonella* infections were determined using models created by several software packages, including SAS, ArcGIS @RISK and NeuroShell.

Results: Results of this study showed a significant increase in *Salmonella* outbreaks in MS during the study period, with highest rates in 2011 (47.84±24.41 cases/100 000; p<0.001). MS had the highest rates of *Salmonella* outbreaks compared with other states (36±6.29 cases/100 000; p<0.001). Regional and district variations in the rates were also observed. GIS maps of *Salmonella* outbreaks in MS in 2010 and 2011 showed the districts with higher rates of *Salmonella*. Regression analysis and neural network models showed a moderate correlation between cases of *Salmonella* infections and low socioeconomic factors. Poverty was shown to have a negative correlation with *Salmonella* outbreaks (R²=0.152, p<0.05).

Conclusions: Geographic location besides socioeconomic status may contribute to the high rates of *Salmonella* outbreaks in MS. Understanding the geographical and economic relationship with infectious diseases will help to determine effective methods to reduce outbreaks within low socioeconomic status communities.

INTRODUCTION

Foodborne diseases are a major public health concern. Studies had estimated that each year in the USA alone, 31 pathogens cause 37.2 million illnesses; of these, 9.4 million were foodborne.¹ The prevalence of *Escherichia coli*

Strengths and limitations of this study

- Socioeconomic and demographic indicators may be used to predict which individuals and communities are at an increased risk of acquiring foodborne infections.
- Neural network and geographical information system modelling were shown in this study to be useful tools to predict the correlation of *Salmonella* outbreaks with socioeconomic factors.
- The southern parts of the USA, including Mississippi, are more vulnerable to increase outbreaks of foodborne illnesses due to low socioeconomic status, climatic changes and agricultural practices.
- Research is critically needed in disadvantaged states and areas with low socioeconomic status such as Mississippi.
- The study is limited by the availability of data and the accurate reporting. In addition, data were collected from several sources which may increase the uncertainty of the resulting models.
- Methodologies used in this paper may need further validation.

O157:H7 infection has grown since its first description and, despite the best control measures, *E. coli* O157:H7 remains a serious health concern.²⁻⁴ *E. coli* O157:H7 is responsible for an estimated 73 480 cases of illness, 2168 hospitalisations, and 61 deaths annually in the USA.³ The majority of such *E. coli* O157:H7 outbreaks in the USA are associated with foodborne transmission. Whereas there were approximately 93.8 million human cases of gastroenteritis and 155 000 deaths due to *Salmonella* infection around the world each year.⁵ In the USA alone, *Salmonella* causes an estimated 1.4 million human cases, 15 000 hospitalisations and more than 400 deaths annually.⁶ *Salmonella* rates varied considerably by geographic region. This heterogeneity is likely in part due to differences in reporting. Differences in salmonellosis case rates between geographically and socioeconomically similar US states have been documented,



CrossMark

Department of Biology/
Environmental Science,
Jackson State University,
Jackson, Mississippi, USA

Correspondence to

Dr H Anwar Ahmad;
hafiz.a.ahmad@jsums.edu

with rates differing by as much as 200% between neighbouring states.⁵

The southern parts of the USA are more vulnerable to increase outbreaks of foodborne illnesses due to socio-economic status, climatic changes and agricultural practices. Our previous studies examined the effects of climatic changes on *Salmonella* infections in the southern states, and results showed a significant effect of increased temperature on the rate of outbreaks.⁷

The Mississippi State Department of Health (MSDH) indicated that the most common foodborne illnesses in Mississippi (MS) are salmonellosis, campylobacteriosis and shigellosis. It is not uncommon for the 82 counties of MS to report 200 cases of *Salmonella* a month. In 2011, the MSDH reported a total of 1440 cases of salmonellosis, 239 cases of shigellosis, 71 cases of campylobacteriosis and 15 cases of *E. coli* O157:H7/HUS. In 2009, a consumer report investigation revealed that 67% broiler chicken had tested positive for *Salmonella* and *Campylobacter*.

Limited studies had examined foodborne illnesses in the state of MS. It is critical to determine the extent of these illnesses in the state and their correlation with socioeconomic status.

MS is one of the leading agriculture states with low socioeconomic status (LSES) and high rates of obesity and associated health disparities. A comprehensive food safety study will delineate the true nature of foodborne illnesses in MS.

The objectives of this paper are to determine the extent of *Salmonella* and *E. coli* infections in MS and compare it with other southern states and with two reference states in the northern USA, and to determine the infections' correlation with socioeconomic status using several modelling approaches including geographical information system (GIS) and neural network (NN).

GIS mapping

A GIS integrates hardware, software and data for capturing, managing, analysing and displaying all forms of geographically referenced information. GIS may be applied to number of disciplines. GIS has been used to visualise, quantify and analyse geographic components of health research. Studies have ranged from analysis of geographic variation in the use of surgical procedures⁸ and examination of the relationships between ethnicity, low birth weight and area socioeconomic status,⁹ to assessing the relation between the respiratory health status and exposure to heavy traffic pollution,^{10 11} and tracing the movement of AIDS epidemic.¹² These studies demonstrate the wide range of uses for GIS.

NN models (using NeuroShell)

Artificial NNs (ANNs) are mathematical constructs that use previously solved examples to build a system of neurons to make a new decision, classify and forecast. ANN models have been applied in diagnosing myocardial infarction, pulmonary emboli and gastrointestinal

(GI) haemorrhage and conditions; in addition to mortality prediction from cardiovascular risk factors.¹³ An ANN modelling technique is based on the observed behaviours of biological neurons, used to mimic the performance of the human system.¹⁴ NeuroShell2 is a program that mimics the human brain's ability to classify patterns or to make predictions or decisions based on past experience. The human brain relies on neural stimuli while the NN uses data sets. It enables the building of sophisticated custom problem-solving applications without programming. These weighted inputs are simply summed inside the neuron, which pass through a suitable threshold (activation). Similarly, the activated outputs from previous layers transfer to the output layer, passing through another activation, produce an output to simulate a desired output (target) at the end. By a learning algorithm, the neural net achieves learning by modifying weights proportional to the difference between the target and the gained output.¹³

METHODS

Data collection

Data of laboratory-confirmed *E. coli* and *Salmonella* cases were collected, from 2002 to 2012, for the selected states including MS, Alabama (AL), Tennessee (TN), Louisiana (LA), Montana (MT) and Michigan (MI). Southern states including AL, TN and LA were selected as neighbouring states to MS, while MT and MI were selected as a reference states in the northern USA based on their geographical and climatic conditions. *E. coli* and *Salmonella* cases were defined by Centers for Disease Control and Prevention (CDC) as 'confirmed' with the isolation of bacteria from a clinical specimen.

MS *E. coli* and *Salmonella* monthly outbreak cases were grouped by year and districts. Data sources for this study include the CDC and the States Department of Health Epidemiology Departments.^{15–18} Data were adjusted to 100 000 of population.¹⁹

In addition to *Salmonella* infections, data for MS, socioeconomic factors for MS counties, categorised by public health districts, for the year 2010–2011, were retrieved from the MSDH County Health Ranking Mississippi Data.²⁰ The selected factors include poverty, uninsured, unemployment and primary care providers' rates,

Data analysis

Data were analysed using SAS V.9.4, ArcGIS V.10, @Risk and NeuroShell2 software packages (Palisade Corporation. @Risk 4.0: a new standard in risk analysis. Ithaca, New York, USA: Palisade Corporation, 2011. <http://www.palisade.com/risk>; SAS Institute Inc. SAS user's guide: statistics Version 9.4 ed. Cary, North Carolina, USA: SAS Institute Inc, 2014; Ward Systems Group. 1993 NeuroShell 2 user's manual. Frederick, Maryland, USA: Ward Systems Group Inc, 1993; ESRI 2011. ArcGIS Desktop: Release 10. Redlands, California, USA: Environmental Systems Research Institute). *E. coli*

and *Salmonella* rates were analysed using analysis of variance (ANOVA). ANOVA (PROC GLM, SAS V.9.4) was carried out to determine the level of significance between the selected variables: years, month, districts and states, which were used as classification variables, while *E. coli* and *Salmonella* outbreaks' rates were used as response variables. TUKEY all pair wise test was followed for further classification.

Regression analysis

Multiple regression analysis was carried out (PROC REG, SAS V.9.4), to test the relationship between *Salmonella* rates and socioeconomic factors: including poverty, uninsured, unemployment and primary care providers' rates. Socioeconomic factors were used as classification variables and *Salmonella* infection rate as a response variable.

GIS mapping

A GIS integrates hardware, software and data for capturing, managing, analysing and displaying all forms of geographically referenced information.

Study area for GIS map

Located in the southern USA, MS (32.9906° N, 89.5261° W) is bordered by TN on the north, Gulf of Mexico on the south, AL on the east, and Arkansas and LA on the west. It covers a total area of 47 689 square miles. GIS allows for the integration and analysis of geographic data, such as coordinates and area perimeters, and tabular data (ie, attributes of geographic data points). Data are imported into mapping software in layers, where each layer represents a different visual component of the map. Shape files are layers which provide visual output of coordinates and area perimeters on the map.

MS counties data were pooled and grouped by public health districts. Background map was obtained from ESRI ArcGIS (ESRI 2011. ArcGIS Desktop: Release 10. Redlands, California, USA: Environmental Systems Research Institute) online resources. Maps' layers for *Salmonella*, unemployment and primary care providers' rates were created for the years 2010 and 2011, to visually analyse areas with higher disease rates and socioeconomic status.

NN model

NN models for MS were developed using @RISK (Palisade Corporation. @Risk 4.0: a new standard in risk analysis. Ithaca, New York, USA: Palisade Corporation, 2011. <http://www.palisade.com/risk>) and NeuroShell2 (Ward Systems Group. 1993 NeuroShell 2 user's manual. Frederick, Maryland, USA: Ward Systems Group Inc, 1993) software packages. The network is exposed to the problem being predicted or classified, and NeuroShell2 will be able to 'learn' patterns from training data and be able to make its own classifications, predictions or decisions when presented with new unseen data.

MS districts' *Salmonella* outbreaks and socioeconomic data were used for NN models. Mean and SD were calculated for each variable, including *Salmonella*, poverty, uninsured, and unemployment and primary care providers' rates. Those means and SD were subsequently used to generate data with 500 iterations using @RISK in Risk Normal distribution. The simulated data were then used as training examples for the NN models, while the original data were used for testing with NeuroShell2 software. Advanced NNs were selected and the simulated data files were imported. The network was built by defining poverty, uninsured, unemployment and primary care providers' rates were used as input variables while *Salmonella* outbreaks as output. A General Regression Neural Network (GRNN) model was selected from the software design architecture. This model was trained with the simulated data. The test file containing the original data was imported to the system and applied to previously saved trained NN models. Results were exported to Excel where graphs were created to show the association between actual data and the predicted model.

RESULTS

Results of this study showed highest rates of *Salmonella* outbreaks in MS compared with other states (36±6.29 cases/100 000; $p<0.001$), while the lowest rate was found in MI (9.10±0.65 cases/100 000). No significant change in *Salmonella* rates was observed during the past 10 years in the selected states; however, the highest states' average of *Salmonella* were in 2011 (22.05±14.37; $p>0.05$). Results also showed no significant change in *E. coli* outbreaks from 2002 to 2011 ($p>0.05$). The highest rates of *E. coli* outbreaks were found in MT (3.41±0.67 cases/100 000; $p<0.001$), while the lowest rates were in LA (0.30±0.16 cases/100 000; [figure 1](#)).

In addition, results showed a significant increase in *Salmonella* outbreaks in MS with highest rates in 2011 (47.84±24.41 cases/100 000; $p<0.001$), and the lowest in 2006 (26.69±10.67 cases/100 000). However, no significant change in *E. coli* rates was observed over time ($p>0.05$). The highest outbreaks of *E. coli* were observed in 2010 (2.40±4.06 cases/100 000; [figure 2](#)).

ANOVA showed a significant variation within the different districts of MS. Highest rates of *Salmonella* outbreaks were found in the northeast district of MS with an average of 47.76±21.59 cases/100 000 ($p<0.001$), and the lowest rates were found in the Delta area (17.39±4.93 cases/100 000).

E. coli rates were shown in this study to be higher in the Tombigbee region of the state (1.02±1.58 cases/100 000), and the lowest in the Northwest region of state (0.32±0.35). However, *E. coli* rates were not significantly different between the districts ($p>0.05$).

Regression analysis between poverty, unemployment, uninsured and primary care providers rates showed a moderate correlation with *Salmonella* outbreaks ($R^2=0.34$). A

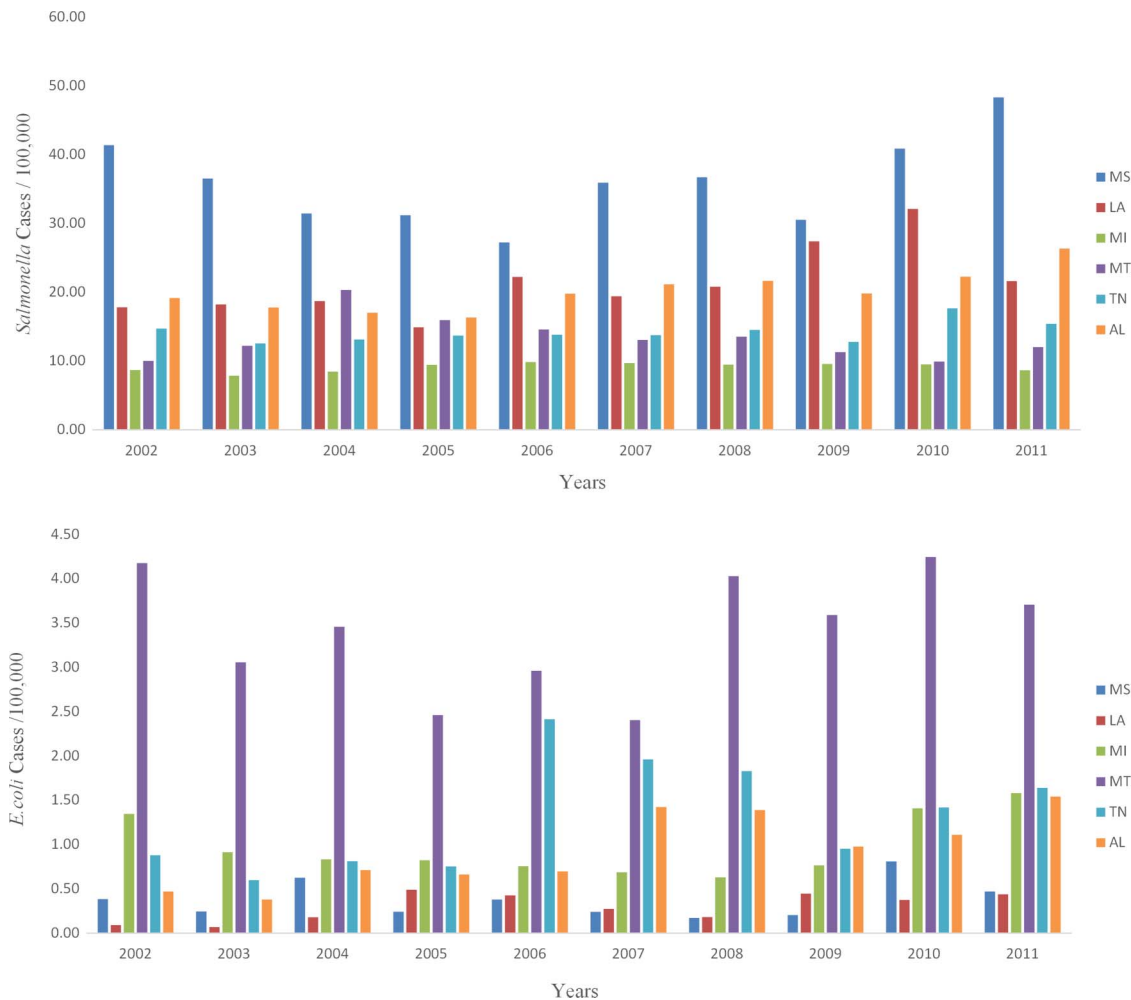


Figure 1 *Salmonella* and *Escherichia coli* rates in Mississippi (MS), Alabama (AL), Tennessee (TN), Louisiana (LA), Montana (MT) and Michigan (MI). The highest rates of *Salmonella* were found in MS, while highest rates of *E. coli* were found in MT from 2002 to 2011.

negative correlation was observed between *Salmonella* and poverty rate ($R^2=0.152$). Areas with high poverty rates were shown to have low rates of *Salmonella* outbreaks. However, a positive correlation was shown between increased *Salmonella* rates and per cent of unemployed, uninsured and primary care providers' rates.

GIS maps of *Salmonella* outbreaks in MS in 2010 and 2011 created by ArcGIS showed higher rates of *Salmonella* outbreaks in the northeast and Tombigbee regions. To determine the relationship between *Salmonella* outbreaks and the socioeconomic factors, GIS maps showed highest rates of unemployment in the northeast, northwest, Tombigbee and Delta, for both years. An average of 42% increase in unemployment rate was observed in these regions in 2011.

Primary care provider rate was shown to be the lowest in the northwest and east-central regions of MS. An average of 17% decrease in primary care provider rates was observed in these regions. On the other hand, highest rate of primary care providers were found in in west-central and southeast regions of the state, with 2% increase from 2010 to 2011 (figure 3).

NN model for *salmonella* and socioeconomic status (including poverty, unemployment, uninsured, unemployment and primary care provider rates including poverty, unemployment, uninsured, unemployment and primary care provider rates) showed a moderate correlation ($R^2=0.41$) between the actual and predicted network (results are shown in table 1 and figure 4).

DISCUSSION

Results of this study showed highest rates of *Salmonella* outbreaks in MS when compared with other selected states, including AL, TN, LA, MT and MI. The average *Salmonella* outbreaks in MS (36 cases/100 000) were more than twice than the average US *Salmonella* outbreaks (16.42 cases/100 000).²¹ A significant variation was observed in this study between the neighbouring southern states and with northern states. Rates were higher in the southern states compared with northern states, with MS having the highest rates of *Salmonella*.

Substantial regional differences in the incidence of *Salmonella* infections have been reported previously.²²

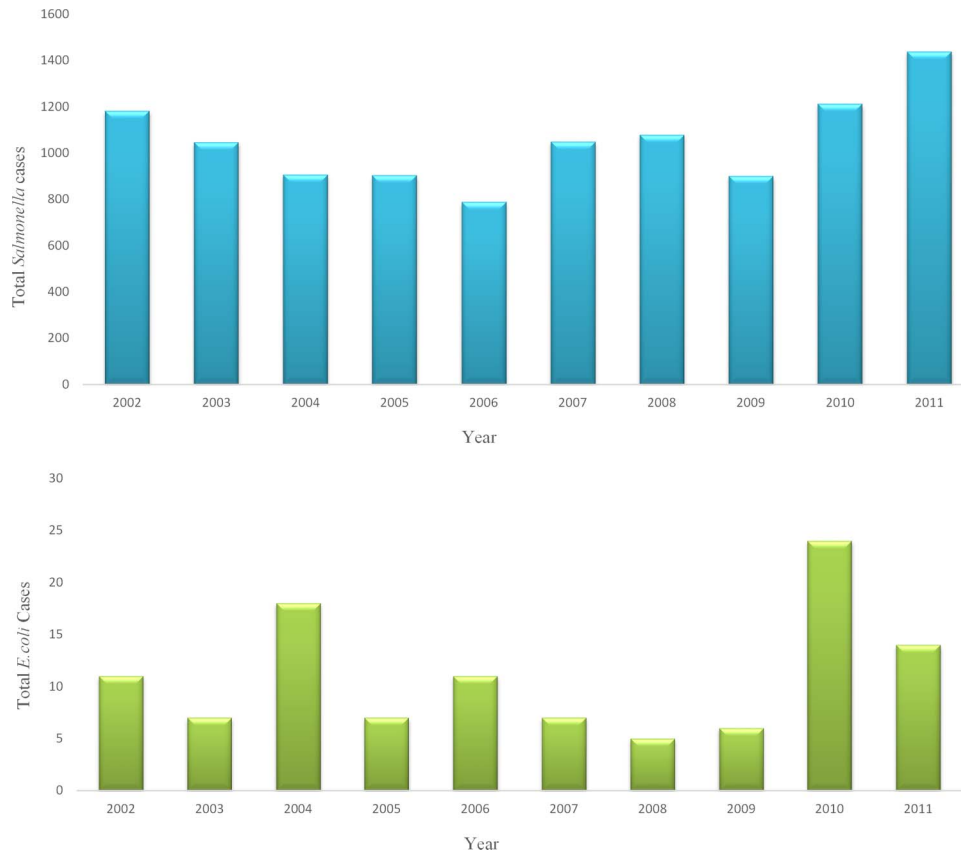


Figure 2 Total number of *Salmonella* and *Escherichia coli* cases in Mississippi 2002–2011. Highest rates of *Salmonella* were observed during 2011. Highest rates of *E. coli* were observed in 2010.

The study reported particularly large relative increases in incidence occurred at sites in the southern USA, and a northeastern state had the highest mean annual incidence in FoodNet.

On the other hand, MS had low rates of *E. coli* outbreaks during the study period. An average of 0.38 cases/100 000 were observed, which was lower than the US average of 1.12 cases/100 000.²¹ A geographical variation was also observed for *E. coli* rates with MT having the highest rates. Cattle and fresh produce are the major source of *E. coli* outbreaks.

In MS, 1440 cases of salmonellosis were reported to MSDH in 2011. This marked a continued increase in the rate and number of reported cases in MS, since 2009. The CDC reported that for every case of *Salmonella* and *E. coli* reported, there are 29 and 26 cases that are not diagnosed or reported, respectively.²¹

Targeted studies of regional factors, such as egg or chicken suppliers, state egg quality assurance programmes, and consumer and food handler educational initiatives, might help clarify reasons for the regional incidence variability.²²

A significant variation was observed in *Salmonella* and *E. coli* outbreaks among the MS districts. GIS mapping, regression analysis and NN models were used to determine the relationships of outbreaks with socioeconomic factors. GIS allows for the integration and analysis of geographic

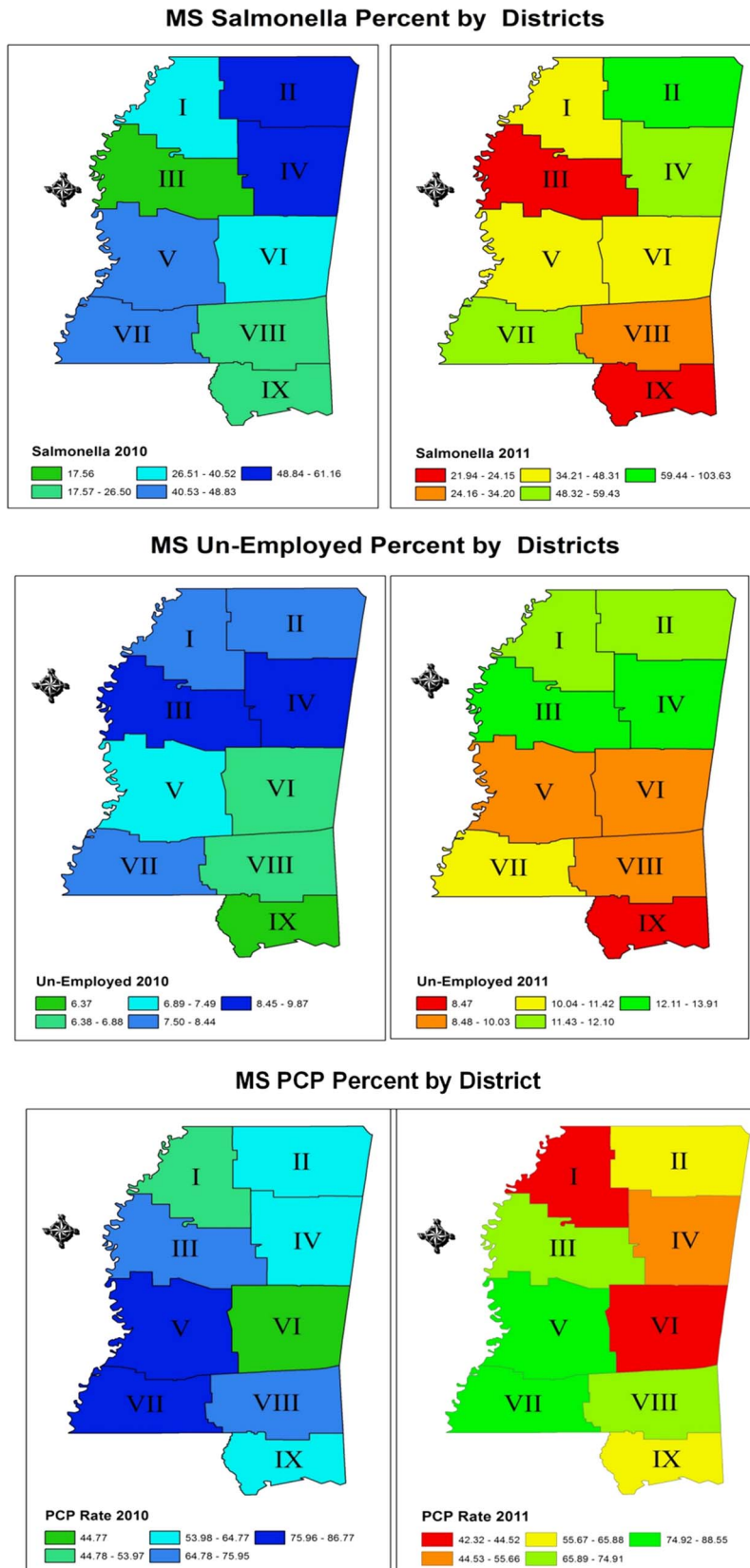
data, such as coordinates and area perimeters, and tabular data (ie, attributes of geographic data points).

The northern region of the state including northeast, northwest, Tombigbee and Delta district had the highest rates of unemployment as well. An average of 42% increase in unemployment rate was observed in the region in 2011. Primary care provider rate was shown to be the lowest in the northwest and east-central regions of MS. An average of 17% decrease in primary care provider rates was observed in these regions. On the other hand, highest rates of primary care providers were found in west-central and southeast regions of the state, with 2% increase from 2010 to 2011.

Socioeconomic and demographic indicators can be used to predict which individuals and communities are at an increased risk of acquiring infections. Generally, LSES is an important predictor of several poor health outcomes including chronic diseases, mental illnesses and mortality, which is the case in MS as shown in this study.

In this study, regression analysis results showed a positive correlation between low socioeconomic factors and increased rates of *Salmonella* infections, with the exception of poverty rates, which were negatively correlated with *Salmonella* outbreaks. Poverty and the availability of physician care showed the highest correlation with *Salmonella* outbreaks.

Figure 3 GIS maps of *Salmonella* outbreak, unemployment rate and PCP in MS districts 2010–2011. Substantial regional differences in the incidence of *Salmonella* infections were found. Higher rates of *Salmonella* outbreaks were found in the northeast and Tombigbee regions. The northern region of the state including northeast, northwest, Tombigbee and Delta district had the highest rates of unemployment. Primary care provider rate was shown to be the lowest in the northwest and east-central regions. The highest rates of primary care providers were found in west-central and southeast regions. GIS, geographical information system; MS, Mississippi; PCP, primary care providers rate.



Under-reporting is an important issue in disease surveillance systems, especially for enteric infections. Generally only those patients with severe symptoms go to see the doctor and are notified to health authorities. As of 2011,

more than 22.6% of MS populations are living under poverty line. According to the United States Department of Agriculture (USDA) Economic Research Service, the average percapita income for MS residents in 2011 was

Table 1 Results of neural network model

Output	GRNN
R^2	0.4169
r^2	0.4169
Mean squared error	175.872
Mean absolute error	11.545
Min. absolute error	4.381
Max. absolute error	22.137
Correlation coefficient r	0.6457

Results of GRNN model for *Salmonella* and socioeconomic factors.
GRNN, General Regression Neural Network.

\$32 000, although rural percapita income lagged at \$29 550. Moreover, there are 96 hospitals in MS, 163 rural health clinics and 21 federally qualified health centres that provide services at 170 sites in the state. An average of 19.0% of MS residents lack health insurance.^{19 20}

A geographical variation of poverty rates was also observed in different districts of the state. In the Delta region of MS, for example, the poverty rate was 44.2%. The lowest *Salmonella* rates were observed in this region as well. With high rates of poverty, many individuals cannot afford to seek medical care, which may result in under-reporting of the disease.

Studies suggested that high socioeconomic status (HSES) groups may be over-represented in incidence statistics. It is possible that since LSES groups tend not to have health insurance or financial means to seek medical care in the event of illness, the ratio of HSES to LSES cases may be skewed in the opposite direction. Access to healthcare may be an important influence on rates of reported bacterial infection. In an economy without universal healthcare coverage, propensity to seek care for GI infection has been associated with having health insurance.^{1 23} However, the new Affordable Care Act (ACA) is expected to expand

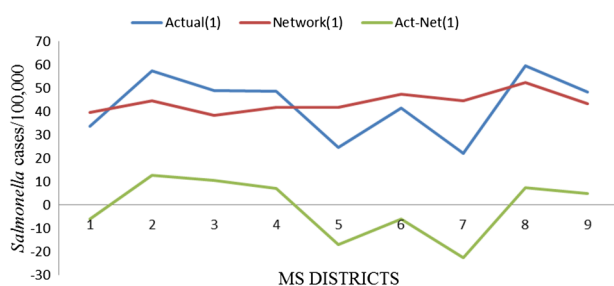


Figure 4 General Regression Neural Network model for *Salmonella* outbreaks and socioeconomic factors. The network was built by defining poverty, uninsured, unemployment and primary care providers' rates were used as input variables while *Salmonella* outbreaks as output. The blue line represents the actual *Salmonella* cases in the nine districts of Mississippi (MS). The orange line represented the *Salmonella* cases predicted by the network, while the grey line represents the difference between the models.

insurance coverage for millions of people in the USA. As a result, rates of reported cases of diseases and infections are expected to increase. In future projects, we will try to understand the impact of ACA on diseases reporting, especially among minority and LSES groups.

In MS, the west-central region of the state showed higher rates of *Salmonella* infections and lower poverty rates (36%), when compared with the Delta region. However, more medical facilities are available in west-central region, resulting in higher identification and reporting of diseases. In addition, in 2011, 20% of the populations in west-central region are college graduate, with 10% unemployment rate, while only 14% of populations at the Delta region are college graduate and 13% are unemployed.

A study had identified lower rates of shigellosis and salmonellosis in communities with high rates of unemployment. The authors speculate that the reduction in access to healthcare due to lack of employment may lead to underdetection of disease in unemployed individuals.²⁴

Other studies had similarly utilised GIS to examine the relationships between area-based socioeconomic measures and incidence of salmonellosis.^{23 25} Their results showed higher incidences of salmonellosis in the groups with high education compared with the less educated groups. They suggested that education may play a significant role in health-seeking behaviour and the predisposition for *Salmonella* infections at the population level.²⁵

NN modelling was shown in this study to be a useful tool to predict the correlation of *Salmonella* outbreaks with socioeconomic factors. A moderate correlation between actual and network-predicted output was observed of 41% which is shown to be an acceptable level considering the biological system.

ANNs are non-linear mapping structures based on the function of the human brain. They have been shown to be universal and highly flexible function approximates for any data. ANNs were developed initially to model biological functions.^{13 14 26 27} NN molding has been used previously for prediction of T-cell epitopes,²⁸ prediction of cancer using gene expression profiling,²⁹ temperature prediction,³⁰ diabetes prediction,¹⁴ poultry growth modelling,³¹ egg price forecasting,³² in addition to predicting the relation between obesity and high blood pressure.²⁷

Our results are different from reported individual-level epidemiological studies that have found a higher level of foodborne infections among low-education and low-income groups.

In the USA, MS ranked 50th among all the states for healthcare, according to the Commonwealth Fund, a non-profit foundation working to advance performance of the healthcare system. For the past 3 years, more than 30% of MS residents, and 22.8% of the state's children, have been classified as obese. On top of obesity, MS had the highest rates in the nation for high blood pressure, diabetes and adult inactivity.²⁷

Social and economic conditions underpin poverty and can directly and indirectly affect health status and health outcomes. Major epidemics emerge and chronic conditions cluster and persist wherever poverty is widespread.³³

CONCLUSION

Human foodborne illnesses are significant public health concerns. In the current study, foodborne illnesses in southern USA with particular emphasis to MS were examined. This study showed a significant correlation between socioeconomic status and the increased rates of *Salmonella* especially in MS which had higher rates than other neighbouring states and some of the northern states. A significant increase in *Salmonella* outbreaks for the past 3 years in MS were observed, with no change in *E. coli* outbreaks. A correlation between increase in outbreaks of *Salmonella* and the LSES was also observed. NN models were shown to be a useful tool to model and predict outbreaks. The model was created using four input variables and one output. NN models accounting for non-linearity predicted better association than regression models. GIS mapping was also shown to be a very useful instrument to map and visualise the areas and districts of highest *Salmonella* outbreaks in addition to socioeconomic status. Our results showed that Northeast and Tombigbee regions of MS had the highest rates of *Salmonella* outbreaks. The northern region also had the highest rate of unemployment, and primary care provider rate was shown to be the lowest in the northwest and east-central.

Understanding the geographical and economic relation with infectious diseases will help to determine effective methods to reduce outbreaks within these communities.

Contributors LA and HAA designed the paper, wrote and reviewed the manuscript. LA collected and analysed the data.

Funding The project described was supported by grant number G12RR013459 from the National Center of Research Resources, G12MD007581 from NIH/NIHMHD and PGA-P210944 from the US Department of State.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

- Scallan E, Hoekstra RM, Angulo FJ, *et al.* Foodborne illness acquired in the United States—major pathogens. *Emerg Infect Dis* 2011;17:7–15.
- Pennington H. *Escherichia coli* O157. *Lancet* 2010;376:1428–35.
- Mohawk KL, O'Brien AD. Mouse models of *Escherichia coli* O157:H7 infection and Shiga toxin injection. *J Biomed Biotechnol* 2011;2011:258185.
- Scheutz E, Møller N, Frimodt-Møller J, *et al.* Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing *Escherichia coli* O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. *Euro Surveill* 2011;16:pii:19889.
- Hoelzer K, Moreno Switt AI, Wiedmann M, *et al.* Animal contact as a source of human non-typhoidal salmonellosis. *Vet Res* 2011;42:34.
- Callaway TR, Edrington TS, Anderson RC, *et al.* Gastrointestinal microbial ecology and the safety of our food supply as related to *Salmonella*. *J Anim Sci* 2008;86(Suppl 14):E163–72.
- Akil L, Reddy RS, Ahmad HA. Effects of climate change on *Salmonella* infections. *Foodborne Pathog Dis* 2014;11:974–80.
- McPherson M, Arango P, Fox H, *et al.* A new definition of children with special health care needs. *Pediatrics* 1998;102Pt 1):137–40.
- Pearl M, Braveman P, Abrams B. The relationship of neighborhood socioeconomic characteristics to birthweight among 5 ethnic groups in California. *Am J Public Health* 2001;91:1808–14.
- Sarno G, Tonini F, Maio S, *et al.* Use of GIS technology to analyse the health effect of exposure to heavy traffic pollution in an Italian general population sample. *Eur Respir J* 2013;42(Suppl 57):P1550.
- Svensden ER, Gonzales M, Mukerjee S, *et al.* GIS-modeled indicators of traffic-related air pollutants and adverse pulmonary health among children in El Paso, Texas. *Am J Epidemiol* 2012;176(Suppl 7):S131–41.
- Needle RH, Richard H, Robert T, *et al.* Rapid assessment of the HIV/AIDS crisis in racial and ethnic and ethnic minority communities: an approach for timely community interventions. *Am J Public Health* 2003;93:970–9.
- Lin B. Resilience in agriculture through crop diversification: adaptive management for environmental change. *Bioscience* 2011;61:183–93.
- Park J, Edington DW. A sequential neural network model for diabetes prediction. *Artif Intell Med* 2001;23:277–93.
- TN Department of health. Communicable Disease Interactive Data. http://health.state.tn.us/ceds/WebAim/WEBAim_criteria.aspx
- State of Louisiana Department of Health and Hospitals, Infectious diseases reports. <http://new.dhh.louisiana.gov/index.cfm/newsroom/archives/64>
- Montana Department of Public Health and Human Services Surveillance. <http://www.dphhs.mt.gov/publichealth/cdepi/surveillance/index.shtml>
- Division of Communicable Disease, Bureau of Epidemiology, Michigan Department of Community Health, reportable infectious diseases in Michigan 2002–2011 Summaries. http://www.michigan.gov/documents/mdch/2005_CD_Epi_Profile_177917_7.pdf, http://www.michigan.gov/documents/mdch/2009_CDEpiProfile_337815_7.pdf, http://www.michigan.gov/documents/mdch/2011_CDEpiProfile_400563_7.pdf
- United States Census Bureau. <http://www.Census.gov>.
- MSDH, County Health Ranking Mississippi. http://msdh.ms.gov/msdhsite/_static/31,0,211.html
- Centers for Disease Control and prevention, Food Safety. <http://www.cdc.gov/foodsafety/> (accessed 15 Mar 2014).
- Chai SJ, White PL, Lathrop SL, *et al.* *Salmonella enterica* Serotype Enteritidis: increasing incidence of domestically acquired infections. *Clin Infect Dis* 2012;54(Suppl 5):488–97.
- Darcey VL. *GIS mapping of retail food access to assess risks of (chronic and acute) illness in populations of different socioeconomic status.* Drexel University, 2011.
- Chang M, Groseclose SL, Zaidi AA, *et al.* 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.
- Younus M, Hartwick E, Siddiqi AA, *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.
- Sovan L, Guégan JF. Artificial neural network as a tool in ecological modelling, an introduction. *Ecol Model* 1999;120:65–73.
- Akil L, Ahmad HA. Relationships between obesity and cardiovascular diseases in four southern states and Colorado. *J Health Care Poor Underserved* 2011;22:61–72.
- Nielsen M, Lundegaard C, Worning P. Reliable prediction of T-cell epitopes using neural network with novel sequence representations. *Protein Sci* 2003;12:1007–17.

29. Khan J, Wei JS, Ringnér M. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat Med* 2001;7:673–9.
30. Smith BA, Hougrenboom G, McClendon RW. Artificial neural networks for automated year-round temperature prediction. *Comput Electron Agric* 2009;68:52–61.
31. Ahmad HA. Poultry growth modeling using neural networks and simulated data. *J Appl Poult Res* 2009;18:440–6.
32. Ahmad HA, Mariano M. Comparison of forecasting methodologies using egg price as a test Case. *Poult Sci* 2006;85:798–807.
33. World Health Organization Food Safety. <http://www.who.int/foodsafety/en/> (accessed 25 Jan 2014).