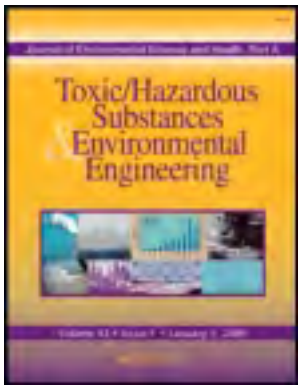
This article was downloaded by: [National Agricultural Library]
On: 09 July 2012, At: 13:17
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances and Environmental Engineering

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lesa20

An analysis of the link between strokes and soils in the South Carolina coastal plains

Thomas F. Ducey [a], Jarrod O. Miller [a], Warren J. Busscher [a], Daniel T. Lackland [b] & Patrick G. Hunt [a]

[a] Coastal Plains Soil, Water, and Plant Research Center, Agricultural Research Service, USDA, Florence, South Carolina, USA

[b] Department of Neurosciences, Medical University of South Carolina, Charleston, South Carolina, USA

Version of record first published: 16 Apr 2012

To cite this article: Thomas F. Ducey, Jarrod O. Miller, Warren J. Busscher, Daniel T. Lackland & Patrick G. Hunt (2012): An analysis of the link between strokes and soils in the South Carolina coastal plains, Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances and Environmental Engineering, 47:8, 1104-1112

To link to this article: http://dx.doi.org/10.1080/10934529.2012.668064

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Journal of Environmental Science and Health, Part A (2012) **47**, 1104–1112
Copyright © Taylor & Francis Group, LLC
ISSN: 1093-4529 (Print); 1532-4117 (Online)
DOI: 10.1080/10934529.2012.668064

Taylor & Francis
Taylor & Francis Group

An analysis of the link between strokes and soils in the South Carolina coastal plains

THOMAS F. DUCEY[1], JARROD O. MILLER[1], WARREN J. BUSSCHER[1], DANIEL T. LACKLAND[2] and PATRICK G. HUNT[1]

[1]*Coastal Plains Soil, Water, and Plant Research Center, Agricultural Research Service, USDA, Florence, South Carolina, USA*
[2]*Department of Neurosciences, Medical University of South Carolina, Charleston, South Carolina, USA*

Downloaded by [National Agricultural Library] at 13:17 09 July 2012

The Stroke Belt is a geographical region of the Southeastern United States where resident individuals suffer a disproportionately higher rate of strokes than the rest of the population. While the "buckle" of this Stroke Belt coincides with the Southeastern Coastal Plain region of North and South Carolina and Georgia, there is a paucity of information pinpointing specific causes for this phenomenon. A number of studies posit that an exposure event–potentially microbial in nature–early in life, could be a risk factor. The most likely vector for such an exposure event would be the soils of the Southeastern Coastal Plain region. These soils may have chemical and physical properties which are conducive to the growth and survival of microorganisms which may predispose individuals to stroke. To this aim, we correlated SC stroke mortality data to soil characteristics found in the NRCS SSURGO database. In statewide comparisons, depth to water table (50 to 100 cm, $R = 0.62$) and soil drainage class (poorly drained, $R = 0.59$; well drained, $R = -0.54$) both showed statistically significant relationships with stroke rate. In a 20 county comparison, depth to water table, drainage class, hydric rating (hydric soils, $R = 0.56$), and pH (very strongly acid, $R = 0.66$) all showed statistically significant relationships with stroke rate. These data should help direct future research and epidemiology efforts to pinpoint the exact exposure events which predispose individuals to an increased stroke rate.

Keywords: Environmental exposure, geostatistical analysis, soil, South Carolina, SSURGO, Stroke Belt, stroke buckle.

Introduction

Higher stroke mortality rates in the southeastern United States, when compared to the rest of the country, have resulted in this region being labeled the "Stroke Belt" (SB). The "buckle" of the SB covers the coastal plain region of Georgia (GA), North Carolina (NC), and South Carolina (SC), and exhibits some of the highest rates of stroke mortality per capita in the United States. Although the underlying factors in this geographic pattern remain largely unresolved,[1] nativity has been implicated as a significant risk marker.[2]

Similarly, additional studies demonstrate that residence in the SB during childhood result in significantly increased risk of stroke,[3,4] regardless of whether or not individuals migrate out of the SB.[4] It has been hypothesized that these findings are indicative of early-life, region-specific exposures to environmental factors, which may include infectious disease processes. This is supported by an additional study that puts forth the hypothesis of acute-exposure events in people within the SB, and strongly considers infectious agents as one of the possible culprits.[5]

Although environmental exposure and infectious disease may serve as potential underlying causes for elevated stroke risk, very little research has been conducted to elucidate which environmental factors or etiological agents may fall into these geographically constrained categories.[1] Recently however, the clarion call has been sounded to delve deeper into elucidating the potential risk markers plaguing the geographical region known as the SB.[5]

Spatial epidemiology recognizes that disease can be limited by geography, where disease clusters are associated with a variety of environmental and socioeconomic factors.[6–8] This field of research is becoming more prevalent as the use of geographic information systems (GIS) increases.[9]

The field has also been advanced through institutional support, newly developed interactive maps, increased amounts of site-specific data, and advanced software that can handle and analyze the data.[10,11] Using GIS and geostatistics, predictions about disease risk have been made

Address correspondence to Thomas F. Ducey, Coastal Plains Soil, Water, and Plant Research Center, Agricultural Research Service, USDA, 2611 West Lucas Street, Florence, SC 29501, USA; E-mail: thomas.ducey@ars.usda.gov
Received September 8, 2011.

Downloaded by [National Agricultural Library] at 13:17 09 July 2012

using environmental vectors, vertebrate reservoirs, or actual human cases; of these, associating actual human cases with surrounding environmental variables produces better predictive maps.[7]

These tools have been used to create risk maps to predict hantavirus pulmonary syndrome, develop projections for future distribution of falciparum malaria, as well as identify geographical features that would serve as risk factors for mosquito-borne, La Crosse encephalitis.[7] Likewise, a GIS study of Lyme disease correlated disease risk with slope, soil type, and forest edges,[12] yet another study determined that a 1% increase in soil clay particle content resulted in an 8.9% increased risk of prion infection.[13]

Infectious disease processes have been previously linked to stroke. A study by Wimmer et al. demonstrated that the presence of IgA antibodies and circulating IgG immune complexes, directed towards *Chlamydia pneumoniae*, were more prevalent in people who had suffered a stroke.[14] The presence of *C. pneumoniae* has also been detected in soils and a wide-range of non-human hosts, indicating the possibility for human infection by environmental exposure.[15,16] To our knowledge however, very little work has been done to look at the environmental conditions conducive to *C. pneumoniae* propagation, or the bacterial load of *C. pneumoniae* in various geographic regions. If *C. pneumoniae*, or another infectious organism which is capable of being acquired through the environment, can serve as risk factors for stroke, then approaches need to be taken to help researchers target geographic locations where such hypotheses can be effectively tested.

To this aim we have taken a preliminary look at the soils of SC and compared them to stroke data collected from 1996–2006. Our objective was to link strokes to some aspect in the environment in SC by comparing stroke mortality rates within the region. We do not believe that any particular soil type *per se* will be linked to stroke; rather strokes would be linked to some aspect of soil(s). Our hypothesis therefore is that certain soil characteristics may serve as risk markers for stroke and can be determined by geostatistical analysis. Once the general relationship has been identified, more specific field sites can be established to quantify these relationships and possibly determine more specific causes.

Materials and methods

Study area

The state of South Carolina was used in this research although the methodology is applicable to other geographical regions. Initial analysis was statewide and included all forty six counties. Additional analysis focused on the 10 highest and 10 lowest counties in regard to stroke rate diagnoses (Table 1).

Table 1. High and low stroke rate counties used in the 20-county analysis.

Stroke rate	*Counties*
High	Bamberg, Clarendon, Darlington, Dorchester, Florence, Lee, Marion, Orangeberg, Sumter, Williamsburg
Low	Aiken, Anderson, Beaufort, Calhoun, Chesterfield, Edgefield, Greenville, McCormick, Oconee, York

Data sources

South Carolina (SC) inpatient and emergency room discharge surveillance data according to county for the years 1996–2006, with a primary diagnosis of stroke (ICD-9 codes 430-438), were obtained from the Centers for Disease Control and Prevention (CDC; http://wonder.cdc.gov).

Discharges were unduplicated by patient and year, with each patient only counting once. Geographic soil properties and taxonomic data are available for the United States through the Soil Survey Geographic (SSURGO) database maintained by the United States Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS). The SSURGO data for SC was obtained from the NRCS Geospatial Data Gateway. Soil survey spatial data, in tabular and Microsoft Access 2000 (Microsoft Corporation, Redmond, WA) templates, was downloaded from Soil Data Mart for each county in SC.

Data preparation and analysis

Using an SC template, all county data were combined into a statewide soil survey database. Survey data was incorporated into ESRI ArcView 9.3 geographic information systems (GIS) program. Uploading soil survey data was done by adding the county level SSURGO spatial maps into ArcView and combining them into one statewide map using the merge function in ArcToolbox. Then using Soil Data Viewer (SDV; http://soils.usda.gov/sdv/download.html), cation exchange capacity (CEC), hydric rating, hydrologic soil group, sand, silt, clay, organic matter, saturated hydraulic conductivity (ksat), and depth to water maps were made based on the surface layer of each soil map unit. Maps of septic suitability, soil surface pH, drainage class, and flooding frequency had to be created by SDV using two separate eastern and western maps of SC due to memory issues in ArcView.

These maps were then merged into one state map. Using the dissolve tool in ArcToolbox, each map was "dissolved" by map unit symbol (MUSYM), area symbol, and the soil class property. In the SSURGO data bases there are often unknown values for some soil properties listed as "null", which is the same as areas mapped as water bodies (W). After the initial dissolve, all of the areas listed as "W" under

Table 2. Ranges of sand, silt, clay, organic matter, Ksat, depth to water, and pH setup in ArcView for SSURGO data.

	Sand		*Silt*		*Clay*	
Map Type	*%*	*Class*	*%*	*Class*	*%*	*Class*
	0 to 20	Silty	0 to 20	Sandy	0 to 18	Low
	20 to 45	Loamy	20 to 40	Loamy	18 to 35	Loamy
	45 to 85	Sandy	40 to 80	Silty	35 to 60	Fine
	85–100	Sand	80–100	Silt	60 to 100	Very Fine
	Organic matter		*Ksat (water conductivity)*		*Depth to water*	
	%	*Class*	*cm s^{-1}*	*Class*	*cm*	*Class*
	0 to 0.5	Very Low	0 to 0.01	Very Low	0 to 25	0 to 25
	0.5 to 1.5	Low	0.01 to 0.1	Low	25 to 50	25 to 50
	1.5 to 5	Mod Low	0.1 to 1	Mod Low	50 to 100	50 to 100
	5 to 10	Mod High	1 to 10	Mod High	100 to 150	100 to 150
	10 to 20	High	10-100	High	150 to 200	150 to 200
	>20	Very High	>100	Very High	>200	>200

pH	
pH	*Class*
<3.5	Ultra Acid
3.5 to 4.5	Extremely Acid
4.5 to 5.0	Very Strongly Acid
5.0 to 5.5	Strongly Acid
5.5 to 6.0	Moderately Acid
6.0 to 6.5	Slightly Acid
6.5 to 7.3	Neutral
7.3 to 7.8	Slightly Alkaline
7.8 to 8.5	Moderately Alkaline
8.5 to 9.0	Strongly Alkaline
>9.0	Very Strongly Alkaline

Downloaded by [National Agricultural Library] at 13:17 09 July 2012

MUSYM were deleted so they would not be included in the "unknown" soil category.

For maps of sand, clay, silt, organic matter, Ksat, depth to water, and pH (Table 2) values were split into property classes for better aggregation to regress against in their ArcView property table (e.g., sand content was grouped into silty (0–20%), loamy (20–45%), sandy (45–85%), and sand (85–100%), based on the soil textural triangle). This had to be done so that there were classes to regress stroke rates against. Some maps, such as drainage class or hydric grouping, already were split into classes by SSURGO, while others (depth to water) had to be reassigned their SSURGO values in ArcView. All maps were dissolved again by their property classes so that each map table only contained one value for each county, and then Hawth's Tools (http://www.spatialecology.com/htools/tooldesc.php) was used to calculate the total area for each class property by county.

For example, there were several map units within Florence County, SC that were labeled as "0 to 25 cm" on the depth to water map, and the dissolve tool combined all of those polygons into one unit in the table. This allowed Hawth's Tools to calculate a total area of the "0 to 25 cm" depth to water class for each county. This area was then exported into a Microsoft Excel table to regress against age adjusted stroke rate by county. Correlations, and linear and stepwise regressions were performed using Proc Corr and Proc Reg respectively in SAS version 9.2 (SAS Institute Inc., Cary, NC).

Results

Statewide analysis

Visual inspection of both the age-adjusted stroke rate and major land resource area (MLRA) maps highlight the SC portion of the "stroke buckle" (Fig. 1a), which extends eastward from the Southeastern Coastal Plain region and is composed of the Southern Coastal Plain, Atlantic Coast Flatwoods, and Tidewater MLRA's (Fig. 1b). The highest rates of stroke are contained primarily within the Southern Coastal Plain (NRCS MLRA 133A), which comprises 17.7% of SC's total land mass.

The Southern Coastal Plain extends down from Virginia, through the Carolinas, Georgia, and the Florida panhandle. It then heads west and northwards into Alabama, Mississippi, Louisiana, and Kentucky. All of these regions are recognized as part of the "Stroke Belt" and, like SC, have unusually high incidences of stroke. Since SC is located in

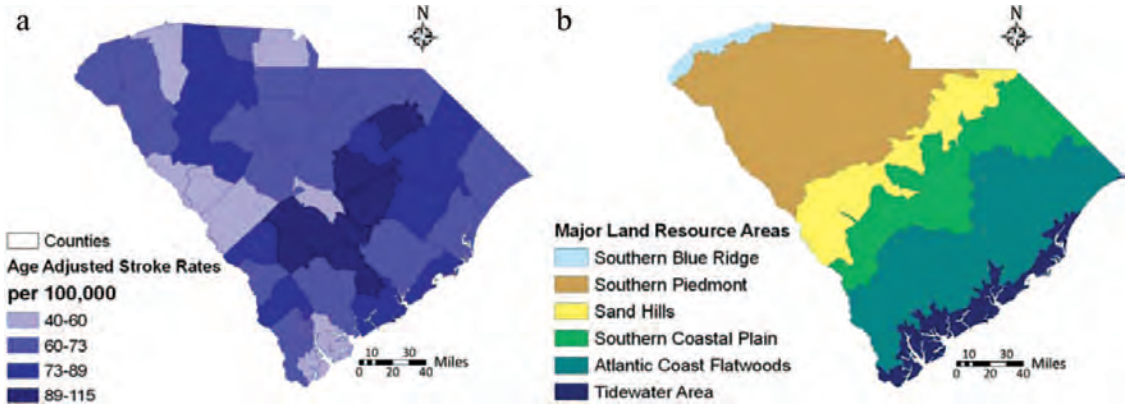


Fig. 1. Maps of (a) age adjusted stroke rate, and (b) major land resource areas in South Carolina (color figure available online).

Downloaded by [National Agricultural Library] at 13:17 09 July 2012

the "Stroke Belt" a majority of the counties in the state do have high rates of stroke (>60 cases per 100,000), which is evident by looking at the age adjusted stroke rate map (Fig. 1a).

Statewide examination identified relationships when stroke rate was correlated against soil survey data. Specifically, properties relating to water table depth and drainage class corresponded to stroke rates. For depth to water, 50–100 cm ($R = 0.62$; $P < 0.01$), and 100–150 cm ($R = 0.62$; $P < 0.01$) were statistically significant. For soil drainage class, well drained soils ($R = -0.54$; $P = 0.05$) and poorly drained soils ($R = 0.59$; $P = 0.01$) were statistically significant. Soils with water tables less than 50 cm deep are usually near streams, rivers, or wetlands, and it will be rare to find residences here, as they may also frequently flood. The positive relationship of water table depths of 50–150 cm is most likely related to the shallow water tables that are more prevalent in the Coastal Plain compared to the Piedmont.

Twenty-county analysis

The low number of statewide relationships is most likely due to the fact that the data were averaged over counties reducing the size of the data pool and the inherent nature of the SSURGO database, which is not meant for regional scientific analysis. This resulted in a limited size data base that was possibly diluted by transitional counties as soil properties change across the different MLRA. Geographic differences are not always constrained to individual counties, and people readily move among these differences as well. To counterbalance these possibilities, the data was refined to focus on the ten highest and ten lowest counties (Table 1), in terms of stroke rates, to see what soil properties may be different across these regions. The 10 low stroke counties all fell into the Blue Ridge/Piedmont regions, while the high stroke counties were all located within the Coastal Plain. When only using these twenty counties, several relationships were observed.

Depth to water table continued to show a strong relationship with stroke rates, with 50–100 cm ($R = 0.54$; $P = 0.01$) and 100–150 cm ($R = 0.53$; $P = 0.03$) positively correlated to stroke rate, and >200 cm depths negatively correlated ($R = -0.58$; $P < 0.01$). Likewise, soil drainage class also continued to demonstrate a correlation with stroke rates. Soils with impeded water removal rates all showed a positive correlation with stroke rates, with moderately well drained ($R = 0.54$, $P = 0.05$), somewhat poorly drained ($R = 0.45$, $P = 0.04$), and poorly drained soils ($R = 0.60$; $P < 0.01$) being statistically significant. Inversely, well drained soils demonstrated a negatively correlation ($R = -0.50$, $P = 0.02$) to stroke rates. Additionally, hydric soils ($R = 0.56$; $P = 0.01$) and non-hydric soils ($R = -0.56$, $P = 0.01$), demonstrated positive and negative correlations respectively, while very strongly acidic soils (pH 4.5 to 5.0) positively correlated with stroke rate ($R = 0.66$, $P < 0.001$).

Discussion

The results of this study demonstrate that relationships exist in South Carolina between soil characteristics and stroke rates. Depth to water table, drainage class, hydric rating, and pH all correlated with stroke rates; the first three (depth to water table, drainage class, and hydric rating) all show positive correlations between the environment and stroke rate when water was found for prolonged periods of time in the soil profile. Inversely, when SC soils were well drained, rated as non-hydric, and had water tables >200 cm, there was a negative correlation with stroke rate.

Direct measurements of soil moisture are seasonally varied; these three characteristics however represent the average state of soil moisture over extended periods of time. Incidentally, soil moisture and pH–the fourth characteristic identified in this study–have both been previously demonstrated to play significant roles in bacterial community composition.[17,18]

A study by Griffiths et al., where they performed an extensive spatial microbial fingerprinting assessment of Great Britain (GB), revealed that pH and moisture were two of the four strongest relationships between environmental

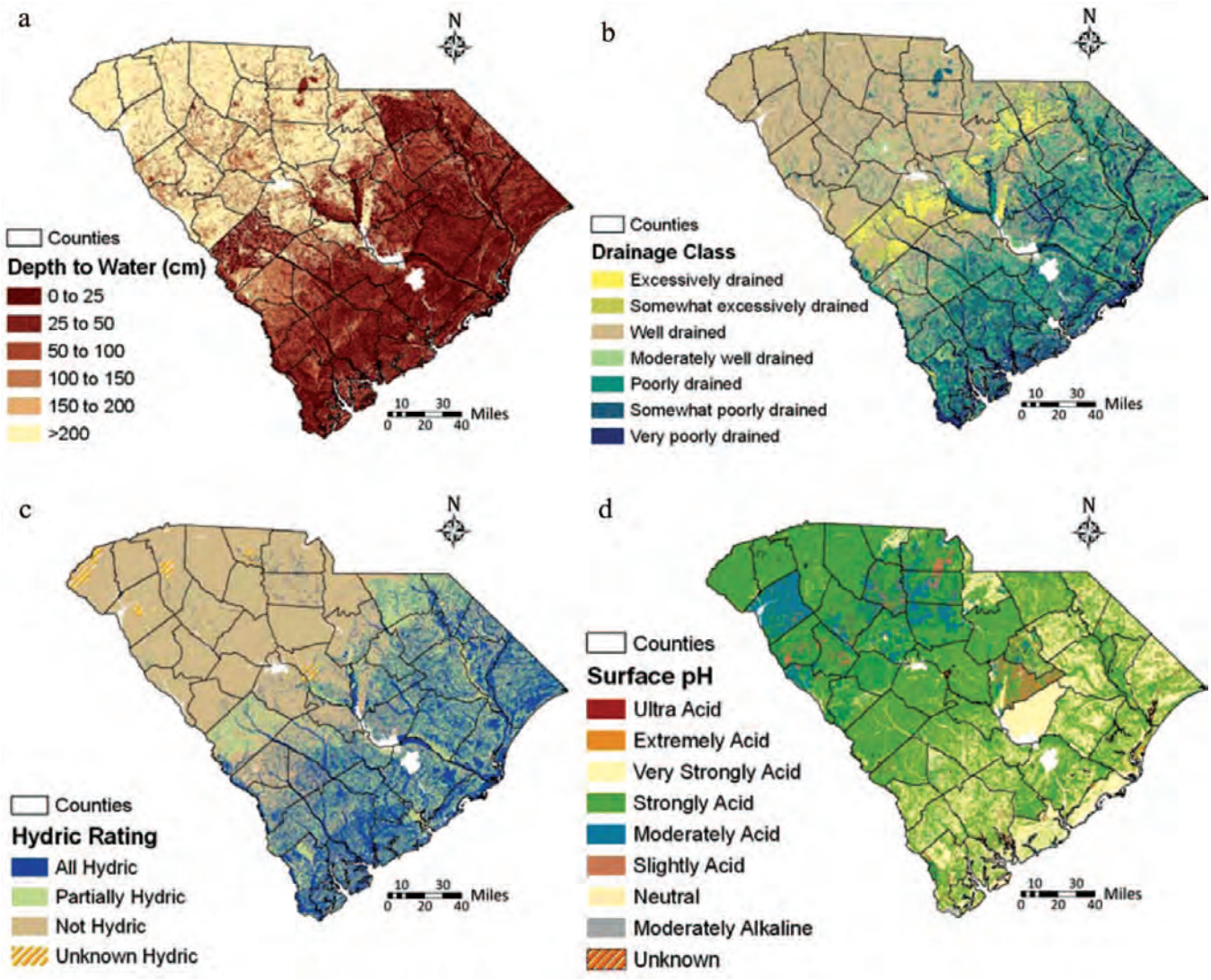


Fig. 2. Statewide SSURGO derived maps of (a) depth to water, (b) drainage class, (c) hydric rating, and (d) pH for South Carolina (color figure available online).

Downloaded by [National Agricultural Library] at 13:17 09 July 2012

factors and bacterial community composition.[19] Their results demonstrated that bacterial diversity and community composition follow along a general north-south gradient, with pH serving as the predominant driving environmental factor (Fig. 2a).

Like GB's north-south bacterial community composition gradient, GB soil pH also decreases as one heads in a northerly trajectory;[20] this is of import because the SB of the United States is not unique; a similar belt exists in GB.[21] The gradient of cerebrovascular disease rates in GB increase in a northerly direction (Fig. 2b), changes which run concurrent with both the soil bacterial fingerprinting profiles reported by Griffiths et al., and soil pH. These findings make GB a potentially intriguing case study. Not only have they extensively characterized their soils on a national level, similar to the United States, but they now have a thorough landscape-scale map of bacterial distributions across the nation. Combined, these repositories of data would make any examination of the relationship between environmental factors and stroke rate more robust. Such a scenario does not currently exist, to our knowledge, for any U.S. state, let alone the entire country.

This of course raises the question as to what further approaches should be taken to better define the risk factors for stroke in SC and the SB in general. Both a review by Howard,[1] and a study by Shrira et al.,[5] consider infectious agents as one of the most appealing, but currently unproven, risk markers associated with increased stroke rates in the SB.

Therefore, considering this hypothesis–that these risk factors are somehow microbiological in nature–there are a number of tools which can be used to analyze the issue. Such an analysis would need to rely on extensive sampling, which would likely rule out, at least in exploratory studies,

Downloaded by [National Agricultural Library] at 13:17 09 July 2012

the culturing of microbial populations. Instead, such analysis would most likely rely on molecular biological tools which have been honed to be performed rapidly, inexpensively, and in a high-throughput manner. Such methods can all be readily adapted to analyze not only bacteria (which we will consider further), but also fungi, parasites, protozoa, and viruses.

The first approach, as taken by Griffiths et al., would be to capture a "fingerprint" of the bacterial community using a technique such as terminal restriction fragment length polymorphism (T-RFLP), or a related approach such as ribosomal intergenic spacer analysis (RISA). These methods are known for the ability to examine large numbers of samples–in the case of the Griffiths et al. study, over 1000–accurately, rapidly, and inexpensively.[22] To replicate the study of Griffiths et al. in the state of SC based on land area, would require somewhere between 300 and 400 soil samplings and, consequently, T-RFLP profiles. The downside to such an approach however is that such profiles generally are capable of identifying only the major inhabitants of the tested microbial community,[23] which means if the agent in question is found in low-abundance within the bacterial community, it may go undetected.

A second approach would be to utilize deep 16S DNA sequencing, microarrays carrying phylotype signature-sequences, or metagenomic analysis. Such approaches have been used in both environmental and medical studies to analyze, in-depth, the complex interactions of microbial communities. Such studies may have already indirectly provided insight into one of the observed phenomenon about the SB: that being the link between nativity and stroke rates. Lackland et al. reported that children born in the SB carried an increased risk to strokes in their lifetime, as compared to individuals living outside this region.[2] This risk was also determined as following the individual even after they left the SB later in life.[2]

This once again, points to an exposure event in the first three years of an individual's life that increases their chances of having a stroke later in life. A study by Palmer et al., in which they used microarrays designed to carry phylotype signature-sequences–in this case species-specific segments of the bacterial 16S gene–determined that it takes approximately a year for an infant's gastrointestinal tract to resemble that of an adults.[24] This maturation is linked, at least in part, to ingestion of microbes derived from the environment. Soil no doubt serves at least as another part of that environmental reservoir.

Reports have carefully linked the transit of soil organisms via aerosolization and dust particles.[25,26] These aerosols could then be deposited onto items that will eventually find their way into an individual's mouth, thereby being ingested. Ingestion of microorganisms that could pose problematic to an individual's health also warrants examination of the oral flora.[27] Trauma to, and inflammation of, the oral mucosa has been extensively linked to cardiovascular disease,[28] and infants and young children are not immune to such trauma (i.e., teething). Like T-RFLP analysis, deep 16S sequencing, microarray analysis, and metagenomic analysis can be done rapidly after sample collection and will provide an in-depth analysis of the microbial community composition and structure, well beyond what T-RFLP can provide, especially in terms of identifying low-frequency inhabitants of the community. The drawback to these methods however is that even with improvements to sequencing technology, per sample, they continue to cost considerably more than an individual T-RFLP. Therefore, one must weigh the need to identify low-frequency organisms versus assay cost.

A third approach is the use of quantitative Real-Time PCR (qPCR) to detect the presence and abundance of particular infectious agents in the environment of question. The use of qPCR to determine the abundance of groups of bacterial organisms has been used in environmental studies ranging from air, to soil and water.[29–32] Several bacteria have been discussed as underlying risk factors for stroke;[14, 33–35] this could lead to studies to specifically identify their presence–and if confirmed, determine their abundance–in soils which provide the environmental conditions identified in this paper that correlate to stroke rate.

This approach could also be used in conjunction with the other above mentioned methods, whereupon identifying suitable microbial targets, determining their presence and abundance in the environment, or in humans. As with T-RFLP, the benefit to this approach is low cost, as well as rapid and high-throughput turnaround. The specific nature of this approach however dictates that hypotheses would need to be formed directly targeting microorganisms believed to be related to increased stroke rates, which may require, as mentioned earlier, some of the aforementioned molecular tools to identify probable targets.

In order to more accurately identify which organisms should be targeted, to determine their role in causing strokes, changes to the current tools, both medical- and soils-related, used to direct these searches should be considered. From a medical perspective, while performing these studies, it became apparent that the stroke data which is readily accessible to the research community proves problematic for use over small geographical spaces. For example, the use of zip codes to assign patients to geographical areas is fraught with issues.

First, zip codes cross county lines, and they can change periodically as the postal service improves its routes or reduces work force. In addition, given that they have been in use only for the last five decades, placing older patients into zip coded areas can be inaccurate because while zip codes are assigned to a particular city, the patient's residence and treatment facility might be in different areas. Given the rapid variability of soil series over short distances, this can lead to inaccurate associations of residence or area of exposure with soil series.

Additionally, while the NIH has established the Human Microbiome Project, tasked with performing metagenomic

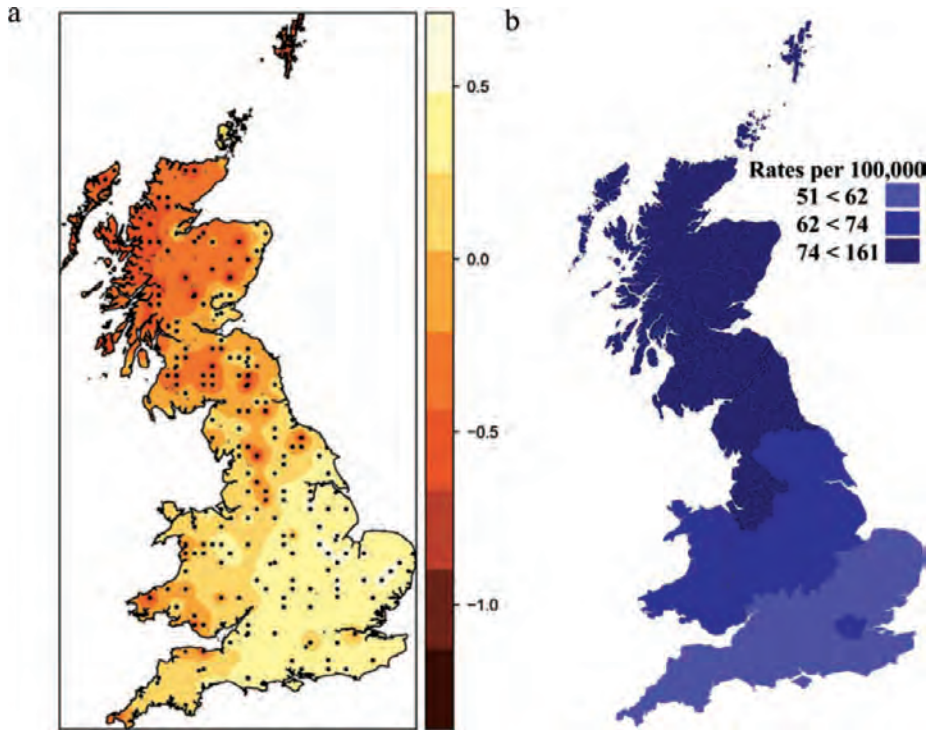


Fig. 3. (a) Spatial map of soil bacterial communities across Great Britain as published in Griffiths et al.[19] reprinted with permission of John Wiley & Sons, Ltd. (b) Age adjusted stroke rate mortality in Great Britain for men (age group 45–74; year 2000), reproduced from data found in Muller-Nordhorn et al.[22] (color figure available online).

Downloaded by [National Agricultural Library] at 13:17 09 July 2012

analysis of healthy humans, and then comparing their microbial communities to individuals suffering from a variety of diseases, stroke is not one of the diseases targeted for analysis.[36,37] We should also point out that while we and others have focused on the possibility that individual exposures to certain microorganisms can lead to increased stroke risk, there is also the possibility that certain microorganisms may not find the environmental conditions of Coastal Plain soils suitable for survival. The absence of this organism in the soil bacterial community could then result in a failure to colonize individuals.

Without this organism to serve as an environmentally derived probiotic,[38] individuals lacking this organism could have increased susceptibility to strokes. Such a scenario is highlighted in a series of reports that demonstrate that patients suffering from Crohn's disease had lower abundances and diversity of certain classes of commensal bacteria.[39–41] The presence of these microorganisms was found to exhibit anti-inflammatory effects that resulted in a decrease in disease in animal models.[42] Care should therefore be taken to examine the presence, as well as the absence of, specific microorganisms.

Likewise, on the soils-related portions of the analyses, a couple of issues also presented themselves while using the SSURGO database. This can most easily be demonstrated by looking at Figure 3 and other examples in SSURGO where soil characteristics stop right at county and state lines. This demonstrates differences in interpretation by soil surveyors and in some cases the differences are very great. For example, the soils for the entirety of Clarendon County are mapped as very strongly acid (Fig. 3d), while soils in surrounding counties, even soils near the borders, are a heterogeneous mix.

A future issue is also that while the NRCS currently tracks all of these soil characteristics across the United States the one important characteristic that is not monitored are the soil microbial communities. Such data, when it is accumulated, should most likely be stored in a central, easily accessible repository alongside the appropriately relevant data; it is our recommendation that the SSURGO database is the most logical choice.

Conclusion

The objective of this investigation was to identify soil characteristics which may serve as risk markers for stroke. Using data provided by the NIH and NRCS, we were able to identify several soil characteristics that correlate with stroke rate in the state of South Carolina. Soil properties related to wet soil conditions, as indicated by depth to water table, soil drainage class, and hydric soil rating all positively correlated with an increased stroke rate. Additionally, soil acidity as measured by pH positively correlated with an

increased stroke rate. These soil properties have been previously reported to have an effect on microbial population structure and composition. If the hypothesized exposure events that predispose individuals residing in the Stroke Belt to strokes are microbial in nature, the correlations identified in this report will be useful for directing future research and epidemiology efforts.

Acknowledgments

We would like to thank Dean Evans for assistance with the preparation of the data used in this manuscript, and Patricia Donovan for GIS advice and manuscript review. The mention of firm names or trade products does not imply that they are endorsed or recommended by the U.S. Department of Agriculture over other firms or similar products not mentioned.

Authors T.F.D. and J.O.M. contributed equally to the writing of this manuscript.

Downloaded by [National Agricultural Library] at 13:17 09 July 2012

References

[1] Howard, G. Why do we have a stroke belt in the southeastern United States? A review of unlikely and uninvestigated potential causes. Am. J. Med. Sci. **1999**, *317*, 160–167.

[2] Lackland, D.T.; Egan, B.M.; Jones, P.J. Impact of nativity and race on "Stroke Belt" mortality. Hypertension **1999**, *34*, 57–62.

[3] Fang, J.; Madhavan, S.; Alderman, M.H. The association between birthplace and mortality from cardiovascular causes among black and white residents of New York City. N. Engl. J. Med. **1996**, *335*, 1545–1551.

[4] Glymour, M.M.; Avendano, M.; Berkman, L.F. Is the 'stroke belt' worn from childhood? Risk of first stroke and state of residence in childhood and adulthood. Stroke **2007**, *38*, 2415–2421.

[5] Shrira, I.; Christenfeld, N.; Howard, G. Exposure to the US Stroke Buckle as a risk factor for cerebrovascular mortality. Neuroepidemiology **2008**, *30*, 229–233.

[6] Elliott, P.; Wartenberg, D. Spatial epidemiology: Current approaches and future challenges. Environ. Health Persp. **2004**, *112*, 998–1006.

[7] Ostfeld, R.S.; Glass, G.E.; Keesing, F. Spatial epidemiology: an emerging (or re-emerging) discipline. Trends Ecol. Evol. **2005**, *20*, 328–336.

[8] Rezaeian, M.; Dunn, G.; St Leger, S.; Appleby, L. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. J. Epidemiol. Commun. H. **2007**, *61*, 98–102.

[9] Ryan, L. Spatial epidemiology some pitfalls and opportunities. Epidemiology **2009**, *20*, 242–244.

[10] Goovaerts, P. Geostatistical analysis of disease data: Visualization and propagation of spatial uncertainty in cancer mortality risk using Poisson kriging and p-field simulation. Int J. Health Geogr. **2006**, *5*, art. no. 7, pp. 26.

[11] Guan, W.W.; Burns, B.; Finkelstein, J.L.; Blossom, J.C. Enabling geographic research across disciplines: Building an institutional infrastructure for geographic analysis at Harvard University. J. Map Geogr. Libr. **2011**, *7*, 36–60.

[12] Glass, G.E.; Schwartz, B.S.; Morgan, J.M.; Johnson, D.T.; Noy, P.M.; Israel, E. Environmental risk factors for Lyme disease identified with geographic information systems. Am. J. Public Health **1995**, *85*, 944–948.

[13] Walter, W.D.; Walsh, D.P.; Farnsworth, M.L.; Winkelman, D.L.; Miller, M.W. Soil clay content underlies prion infection odds. Nature Commun. **2011**, *2*, art. no. 200, pp. 6.

[14] Wimmer, M.L.J.; Sandmann-Strupp, R.; Saikku, P.; Haberl, R.L. Association of chlamydial infection with cerebrovascular disease. Stroke **1996**, *27*, 2207–2210.

[15] Matsuo, J.; Kobayashi, M.; Nakamura, S.; Mizutani, Y.; Yao, T.; Hirai, I.; Yamamoto, Y.; Yamaguchi, H. Stability of Chlamydophila pneumoniae in a harsh environment without a requirement for acanthamoebae. Microbiol. Immunol. **2010**, *54*, 63–73.

[16] Bodetti, T.J.; Jacobson, E.; Wan, C.; Hafner, L.; Pospischil, A.; Rose, K.; Timms, P. Molecular evidence to support the expansion of the hostrange of *Chlamydophila pneumoniae* to include reptiles as well as humans, horses, koalas and amphibians. Syst. Appl. Microbiol. **2002**, *25*, 146–152.

[17] Drenovsky, R.E.; Vo, D.; Graham, K.J.; Scow, K.M. Soil water content and organic carbon availability are major determinants of soil microbial community composition. Microb. Ecol. **2004**, *48*, 424–430.

[18] Lauber, C.L.; Hamady, M.; Knight, R.; Fierer, N. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl. Environ. Microbiol. **2009**, *75*, 5111–5120.

[19] Griffiths, R.I.; Thomson, B.C.; James, P.; Bell, T.; Bailey, M.; Whiteley, A.S. The bacterial biogeography of British soils. Environ. Microbiol. **2011**, *13*, 1642–1654.

[20] Skinner, R.J.; Todd, A.D. Twenty-five years of monitoring pH and nutrient status of soils in England and Wales. Soil Use Manage. **1998**, *14*, 162–169.

[21] Morris, R.W.; Whincup, P.H.; Emberson, J.R.; Lampe, F.C.; Walker, M.; Shaper, A.G. North-South gradients in Britain for stroke and CHD: Are they explained by the same factors? Stroke **2003**, *34*, 2604–2609.

[22] Hartmann, M.; Frey, B.; Kolliker, R.; Widmer, F. Semi-automated genetic analyses of soil microbial communities: Comparison of T-RFLP and RISA based on descriptive and discriminative statistical approaches. J. Microbiol. Meth. **2005**, *61*, 349–360.

[23] Orcutt, B.; Bailey, B.; Staudigel, H.; Tebo, B.M.; Edwards, K.J. An interlaboratory comparison of 16S rRNA gene-based terminal restriction fragment length polymorphism and sequencing methods for assessing microbial diversity of seafloor basalts. Environ. Microbiol. **2009**, *11*, 1728–1735.

[24] Palmer, C.; Bik, E.M.; DiGiulio, D.B.; Relman, D.A.; Brown, P.O. Development of the human infant intestinal microbiota. PLoS Biol. **2007**, *5*, 1556–1573.

[25] Zweifel, U.L.; Fahlgren, C.; Bratbak, G.; Sandaa, R.A.; Thyrhaug, R. Diversity of airborne bacteria in samples collected using different devices for aerosol collection. Aerobiologia **2011**, *27*, 107–120.

[26] Pakarinen, J.; Hyvarinen, A.; Salkinoja-Salonen, M.; Laitinen, S.; Nevalainen, A.; Makela, M.J.; Haahtela, T.; von Hertzen, L. Predominance of Gram-positive bacteria in house dust in the low-allergy risk Russian Karelia. Environ. Microbiol. **2008**, *10*, 3317–3325.

[27] Lazarevic, V.; Whiteson, K.; Huse, S.; Hernandez, D.; Farinelli, L.; Osteras, M.; Schrenzel, J.; Francois, P. Metagenomic study of the oral microbiota by Illumina high-throughput sequencing. J Microbiol Methods **2009**, *79*, 266–271.

[28] Beck, J.; Garcia, R.; Heiss, G.; Vokonas, P.S.; Offenbacher, S. Periodontal disease and cardiovascular disease. J. Periodontol. **1996**, *67*, 1123–1137.

[29] Peccia, J.; Hernandez, M. Incorporating polymerase chain reaction-based identification, population characterization, and quantification of microorganisms into aerosol science: A review. Atmos. Environ. **2006**, *40*, 3941–3961.

[30] Noble, R.T.; Blackwood, A.D.; Griffith, J.F.; McGee, C.D.; Weisberg, S.B. Comparison of rapid quantitative PCR-based and conventional culture-based methods for enumeration of

Enterococcus spp. and *Escherichia coli* in recreational waters. Appl Environ Microbiol **2010**, *76*, 7437–7443.

[31] Ducey, T.F.; Shriner, A.D.; Hunt, P.G. Nitrification and denitrification gene abundances in swine wastewater anaerobic lagoons. J. Environ. Qual. **2011**, *40*, 610–619.

[32] Fierer, N.; Bradford, M.A.; Jackson, R.B. Toward an ecological classification of soil bacteria. Ecology **2007**, *88*, 1354–1364.

[33] Chiang, C.H.; Huang, C.C.; Chan, W.L.; Chen, Y.C.; Chen, T.J.; Lin, S.J.; Chen, J.W.; Leu, H.B. Association between mycoplasma pneumonia and increased risk of ischemic stroke: A nationwide study. Stroke **2011**, *42*, 2940–2943.

[34] Elkind, M.S.V. Infectious burden: A new risk factor and treatment target for atherosclerosis. Infect. Disord. Drug Targets **2010**, *10*, 84–90.

[35] Melnick, S.L.; Shahar, E.; Folsom, A.R.; Grayston, J.T.; Sorlie, P.D.; Wang, S.P.; Szklo, M. Past infection by *Chlamydia pneumoniae* strain TWAR and asymptomatic carotid atherosclerosis. Am. J. Med. **1993**, *95*, 499–504.

[36] Peterson, J.; Garges, S.; Giovanni, M.; McInnes, P.; Wang, L.; Schloss, J.A.; Bonazzi, V.; McEwen, J.E.; Wetterstrand, K.A.; Deal, C.; Baker, C.C.; Di Francesco, V.; Howcroft, T.K.; Karp, R.W.; Lunsford, R.D.; Wellington, C.R.; Belachew, T.; Wright, M.; Giblin, C.; David, H.; Mills, M.; Salomon, R.; Mullins, C.; Akolkar, B.; Begg, L.; Davis, C.; Grandison, L.; Humble, M.; Khalsa, J.; Little, A.R.; Peavy, H.; Pontzer, C.; Portnoy, M.; Sayre, M.H.; Starke-Reed, P.; Zakhari, S.; Read, J.; Watson, B.; Guyer, M. The NIH Human Microbiome Project. Genome Res. **2009**, *19*, 2317–2323.

[37] Turnbaugh, P.J.; Ley, R.E.; Hamady, M.; Fraser-Liggett, C.M.; Knight, R.; Gordon, J.I. The human microbiome project. Nature **2007**, *449*, 804–810.

[38] Ordovas, J.M.; Mooser, V. Metagenomics: the role of the microbiome in cardiovascular diseases. Curr. Opin. Lipidol. **2006**, *17*, 157–161.

[39] Manichanh, C.; Rigottier-Gois, L.; Bonnaud, E.; Gloux, K.; Pelletier, E.; Frangeul, L.; Nalin, R.; Jarrin, C.; Chardon, P.; Marteau, P.; Roca, J.; Dore, J. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut **2006**, *55*, 205–211.

[40] Sokol, H.; Seksik, P.; Furet, J.P.; Firmesse, O.; Nion-Larmurier, I.; Beaugerie, L.; Cosnes, J.; Corthier, G.; Marteau, P.; Dore, J. Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. Inflamm. Bowel Dis. **2009**, *15*, 1183–1189.

[41] Sokol, H.; Seksik, P.; Rigottier-Gois, L.; Lay, C.; Lepage, P.; Podglajen, I.; Marteau, P.; Dore, J. Specificities of the fecal microbiota in inflammatory bowel disease. Inflamm. Bowel Dis. **2006**, *12*, 106–111.

[42] Sokol, H.; Pigneur, B.; Watterlot, L.; Lakhdari, O.; Bermudez-Humaran, L.G.; Gratadoux, J.J.; Blugeon, S.; Bridonneau, C.; Furet, J.P.; Corthier, G.; Grangette, C.; Vasquez, N.; Pochart, P.; Trugnan, G.; Thomas, G.; Blottiere, H.M.; Dore, J.; Marteau, P.; Seksik, P.; Langella, P. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. P. Natl. Acad. Sci. USA **2008**, *105*, 16731–16736.

Downloaded by [National Agricultural Library] at 13:17 09 July 2012