Appendix B.

Local Health Survey - Methodology

Selecting the survey questions

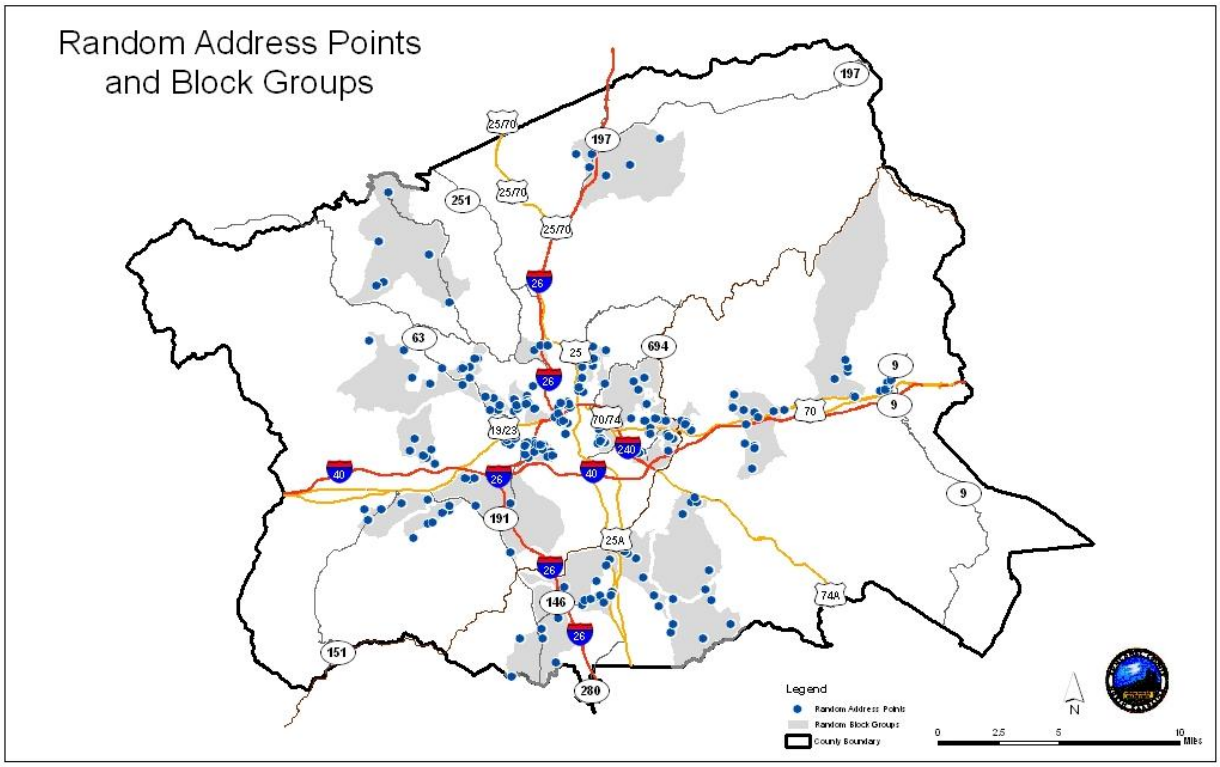
Approximately 60 community members were involved in development of the local survey instrument. Survey Monkey (online survey instrument) was used to gather input about questions to be included / excluded in the survey (compared to 2000 and 2005 surveys). Using the online survey tool expanded community involvement and input. A **Data Team** met to finalize the selection of the questions, using input from the online survey, and their own expertise and/or prior experience with health assessments. Emphasis was placed on creating a local survey that included questions that were comparable to prior surveys (didn't change the wording) and was similar to BRFSS questions.

Health Survey Preparation for GPS handheld computers

The health survey was developed and formatted by Buncombe County Department of Health and then converted by Buncombe County Information Technology to a Trimble Data Dictionary using Trimble GPS Analyst and Trimble Terra Sync software. Once the Data Dictionary was created, the health survey could be collected digitally and a GPS location could be collected for each survey. Each survey and GPS location was stored as a file on the handheld computer. To make data entry as easy as possible, dropdown menus and other ease of use functions were created. Freehand data entry of text was limited as much as possible because the stylus/keyboard system on the GPS units is not user friendly. Also, the interpret-tation and analysis of freehand text is much more complicated.

Site selection

The main goal in creating the survey sample was to randomly identify household locations to interview people that represent the community as a whole. For this process, a Two-Stage Cluster Sampling method was used because of its popularity and successful results around the world. This scheme was developed by the World Health Organization (WHO) with the aim of calculating the prevalence of immunized children. In the first stage, 38 census block groups were randomly selected from all census block groups within Buncombe County. In statistical terms, census block groups were selected through a method known as "probability proportionate to size," which means that a census block group with more households is more likely to be selected than one with fewer households. This first stage was completed by using a Survey Sites Selection Toolkit in combination with the Buncombe County Civic Address Street Locator. The map below shows the 38 randomly selected block groups. In the second stage six interview sites were randomly selected for every selected block group. This task was completed by running a python script. Python is an open source programming language that can run scripts that perform geoprocessing function on GIS data. It is an automated way of executing GIS processes. While the WHO commonly uses a "30x7", we are using a "38x6" sampling method because adding additional census block groups provides more information than selecting additional points within a block group.



For more information on the methodology used by the WHO, visit:
http://whqlibdoc.who.int/bulletin/1982/Vol60-No2/bulletin_1982_60(2)_253-260.pdf

Conducting the health survey with help from community and college student volunteers

Over 65 students and community volunteers signed up to help conduct 228 survey interviews, along with 12 health department staff, including bilingual staff (Spanish and Russian). Several community health classes at UNC-Asheville and Mars Hill College dedicated a segment of their class learning requirements to the health survey. Most students were excited about using GPS features and handheld computers.

Training for volunteers was mandatory and focused on student / volunteer interviewing skills, selection of households when no one was home, use of handheld units and GPS features, safety protocols and use of sheriff radios for communication, familiarity of survey questions, and commitment to the project.

Students were often paired with adult volunteers or staff until they gained experience. Many student volunteers became quite "expert" and volunteered time far exceeding class requirements. Gift cards were offered to students who worked more than required to meet class / teacher expectations. There were a few adults who were without work who enjoyed working with the project in exchange for a gift card for food or gas.

Safety of Volunteers – Buncombe County Sheriff Department

Surveying began in October 2009, on the heels of media coverage related to the census worker who died while gathering census surveys. This motivated health department staff to significantly increase safety precautions for volunteers who were expected to go to unknown homes to conduct the health surveys. We partnered with *Buncombe County Sheriff Department* because the sheriff department has an active *Sheriff Reserve program*. The Sheriff Reserve program enrolls deputies that are not on payroll because they are retired or have other jobs but receive the same annual training as paid sheriff deputies and can serve just as sheriff deputies. The Sheriff Reserves are required to participate in activities such as street barricades, parade and bike routes, large even security, etc. They also have access to patrol cars and handheld radios for communication.

The Sheriff Department dedicated staff time for Lieutenant Calhoun, Coordinator of the Sheriff Reserve Program, to coordinate training and scheduling Sheriff Reserve volunteers, communication equipment and patrol cars, mapping / driving routes for volunteers, and on several occasions served as headquarters for survey teams. This was a significant commitment of resources and an affirmation of our commitment to the safety of our volunteers.

The sheriff reserve deputies were assigned to teams of volunteers and were in constant radio contact using a designated radio channel, also monitored by 911. The deputies conducted general security checks of the selected households based on 911 call histories which enabled them to direct volunteers away from potentially dangerous households. They also assisted with dogs and other security issues, such as access to gated communities which would have otherwise been excluded from the surveys. Their presence was a comfort to the volunteers who frequently "got lost", and their presence often increased the validity of the survey among household occupants being asked to participate in the survey. There was only one safety issue – a minor car accident. The deputy assigned was promptly on the scene and took on securing the safety of those involved until Asheville Police Department arrived.

The Sheriff Reserve deputies worked over 150 hours on the days when we deployed survey teams, which would not have been feasible if paid sheriff department staff were involved. The coordinator, Lt. Calhoun, was present and helped patrolling on the 15 scheduled times teams were surveying, while also providing additional hours of assistance to review maps & driving directions for volunteers, plans to deploy volunteers so they would be in close proximity to a deputy, organize cars and radios, and provide overall safe implementation of the survey project. We wouldn't want to this type of survey without these men and women!

GPS Unit training (how to use the handheld computers)

The survey was performed using handheld Trimble Juno GPS units. Volunteers were trained to capture both the survey answers and a location where the survey was taken. Training on the GPS units happened in a couple of different settings. The majority of volunteers and students who conducted the surveys attended a two-hour training session to learn about the survey and the protocol to ensure reliability, role play conducting the survey with residents, and test out the technology using the

handheld devices. We also performed a refresher course and orientation for any volunteers that were not in the volunteer training.

Data Post Processing

Once surveys were collected, the data was synthesized and converted to a format that is usable for mapping. This process involved several steps.

Creating a spatial dataset: Data was transferred from the Juno units in their native .ssf file format. It was then converted to a shapefile using the Trimble GPS Analyst extension for ArcGIS. At this point in the process, each individual survey point was a separate dataset. These

separate datasets were then merged into a single contiguous spatial dataset using tools in ArcGIS. This process was complicated by the fact that there were small changes in the survey after the first couple of surveys were performed so we had to take that into account as we pulled them together.

Dataset cleanup: Once the dataset was created there was quite a bit of work put into cleaning the dataset to make it easy to use in our analysis. Due to the use of a rotating cast of volunteers, answers were not always entered into the GPS unit uniformly. For instance, while one group would enter A to identify the first answer of a multiple choice question another group might type out the entire answer. Some of this clean up could be automated but much of it was a manual process. Another step in the process of cleaning the survey-wide dataset was the removal of extraneous data points. These were often created in the mornings during training before surveys were conducted. They became very obvious because of the time they were recorded and the fact that few or none of the survey questions were answered.

Data Conversion: Data conversion took place in a couple of phases as well. First, we exported data from a GIS format for ease of use from a statistical perspective. The format we agreed would be best from everyone was a Microsoft Excel (.xls) file. We also made some changes to the format of the answers to try to ease analysis. The survey had several questions where a respondent could pick multiple answers from a list of possibilities, as well as, an 'other' option where the respondent could add an answer not already listed in the survey. For analysis we found it much more convenient to have these questions broken out to several Yes/No questions.

Health survey sample demographics compared to Buncombe County population - 2009 estimates

GENDER	Total	% of known	*Bunc*
Female	132	58%	52%
Male	96	42%	48%
Grand Total	228		

AGE GRP	Total	% of known	
18-24	22	10%	9%
25-44	69	30%	27%
45-64	72	32%	28%
65-79	42	18%	11%
80+	23	10%	5%
Declined	3		
Grand Total	228		

ETHNICITY	Total	% of all	
Hispanic	10	4%	4%
Russ/Ukr/Mold	4	2%	
Neither	214	94%	
Grand Total	228		

RACE	Total	% of all	
Black	13	6%	7%
White	200	88%	89%
Multi-White/American Indian	1	0%	
Multi-White/Asian	1	0%	
Other-Hispanic	4	2%	
Other-Mediterranean	3	1%	
Other-unspecified	1	0%	
Declined	1	0%	
Not sure	1	0%	
Unknown	3	1%	
Grand Total	228		

EDUCATION	Total	% of known	*Bunc*
Gr 1-11	31	14%	14%
HS Grad / GED	61	27%	26%
Some coll	51	22%	29%
Coll grad	84	37%	32%
Unknown	1		
Grand Total	228		

Household INCOME	Total		
<$15,000	51	22%	15%
$15 <$25,000	38	17%	14%
$25 <$50,000	55	24%	29%
$50 <$75,000	27	12%	18%
$75,000 +	39	18%	24%
Declined	10	4%	
Don't know	7	3%	
Unknown	1		.
Grand Total	228		

"Buncombe" Population Data

Source: US Census, Buncombe, 2009 population estimates

American Community Survey (2009)

The survey sample was created by randomizing selection of households using double cluster randomizing method (explained on pg 109 – Site selection). The randomized selection aligns with the actual Buncombe County population in most categories. We find that lower income (<$25,000) is somewhat over represented in the survey, as well people over the age of 65, attributable to the survey design (going to households). Although were conducted on Fridays and Saturdays, busy, working families were less likely to be at home between 9 – 6 on either day, a limitation of this type of survey. All in all, the randomized method helped create a sample of buncombe's population that is quite representative, especially regarding race and ethnicity.

Data Analysis for creating charts, graphs, written analysis

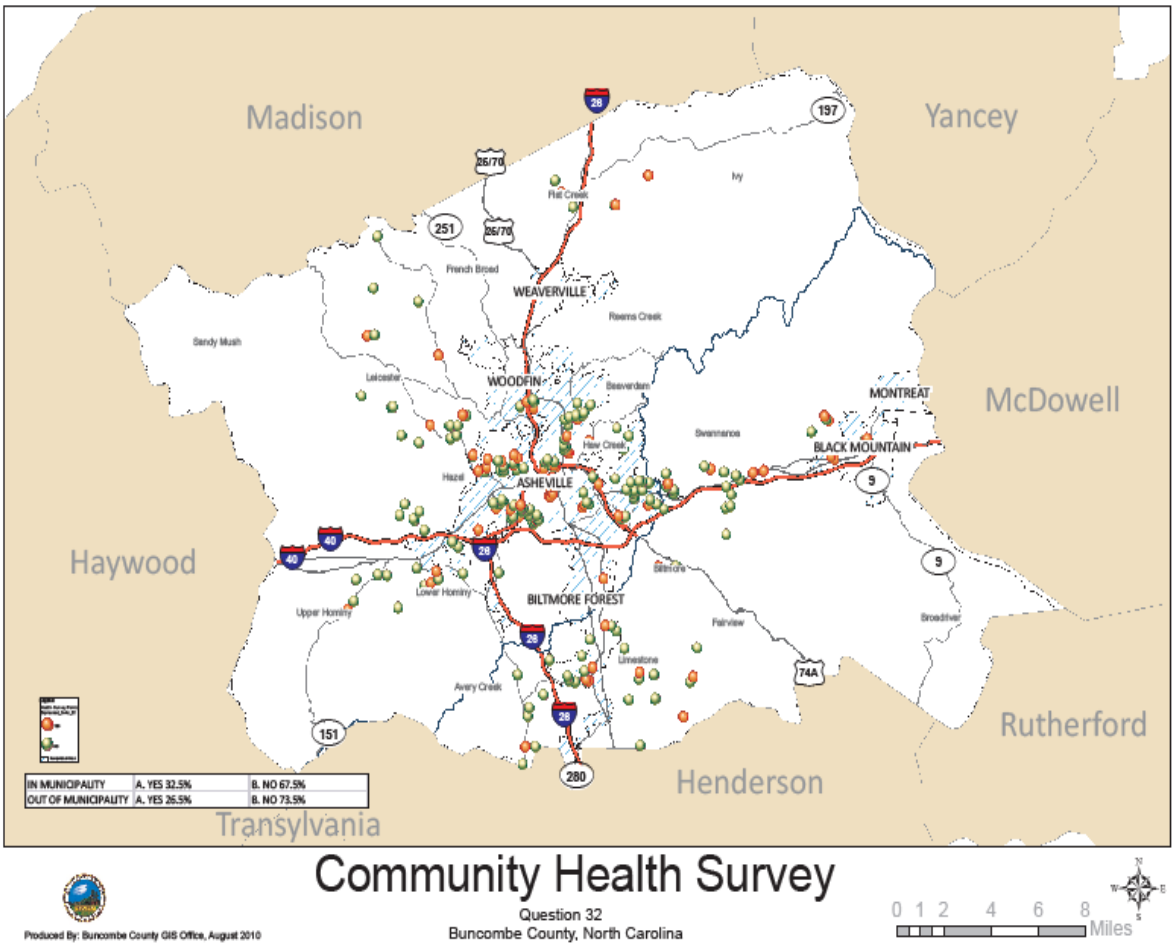
With the help of a paid consultant, Dr. Jill Fromewick of Summit Research Associates, the local health survey data was prepared for graphic displays and written analysis. Frequency tables were created for all questions, broken down in detail according to selected demographics. Then the demographic categories were further formatted to make it easier to display break down of data. For example, age was collapsed into categories of 18 – 44 and 44+, although it is available in much greater detail in frequency tables. Similar formatting was created for income, education, race and ethnicity. We used comparable formatting to NC BRFSS and/or US Census data.

The consultant created bar charts or pie charts of selected local survey variables. Selection of variables were based on community input (via an online survey) telling us the data that is most important to find in the report, as well as knowledge of the variables that have typically appeared in Buncombe's previous health assessment reports.

Data for spatial analysis (mapping results)

With some exceptions, much of what we found in our spatial analysis (mapping) was the limitations of mapping results of a randomized survey. The small sample size and the fact that the entire county was not represented due to the Clustered Random Sample made it difficult to draw too many generalizations. One way in which I think spatial analysis will be helpful is when our survey points are mapped in comparison to demographic data. This process could help to predict health service needs in areas that were not surveyed. For instance, if we see a correlation between and age group or race and health issue (lack of exercise, depression, etc) we could use the demographics from census data to identify areas that have not been surveyed to which we might target specific populations or services. It could allow us to make educated guesses on how best to deploy our resources.

Example of spatial analysis of survey respondents *Reporting feelings of depression for two or more weeks in past year* (yes = orange and no = green).



A limitation of mapping randomized survey responses is the danger of misinterpreting the data, leading some to believe that the points on the map indicate the geographic locations of a specific health issues. We can correctly interpret from this spatial analysis that slightly more people living in a municipality report feelings of depression than those living out of municipalities. If we layered income data we would also find that the red dots (those reporting depression) are more likely to be located in areas with lower income. If the actual data subsets support that, then we could map low income areas and generalize where depression might more likely be found.