

# Data Collection Methods & Limitations

WNC **HEALTHY** IMPACT



## Primary Data

### WNC Healthy Impact Community Health Survey

#### **Survey Methodology**

The 2018 WNC Healthy Impact Community Health Survey was conducted from March to June. The purpose of the survey was to collect primary data to supplement the secondary core dataset, allow individual counties in the region to collect data on specific issues of concern, and hear from community members about their concerns and priorities. The survey was conducted throughout the entire WNC Healthy Impact region, which includes the following 16 counties: Buncombe, Cherokee, Clay, Graham, Haywood, Henderson, Jackson, Macon, Madison, McDowell, Mitchell, Polk, Rutherford, Swain, Transylvania and Yancey.

Professional Research Consultants, Inc. (PRC) designed and implemented the survey methodology with review and input from WNC Healthy Impact, which included a combination of telephone (both landline and cell phone) interviews, as well as an online survey. The survey methodology was designed to achieve a representative sample of the regional population that would allow for stratification by certain demographic characteristics, while also maximizing data collection timeliness and efficiency. Survey sampling and implementation methodology is described in greater detail below.

#### **Survey Instrument**

The survey instrument was developed by WNC Healthy Impact's data workgroup, consulting team, and local partners, with assistance from PRC. Many of the questions were derived from the CDC Behavioral Risk Factor Surveillance System (BRFSS) and other validated public health surveys. Other questions were developed specifically by WNC Healthy Impact, with input from regional and local partners, to address particular issues of interest to communities in western North Carolina. Each county was given the opportunity to include three additional questions of particular interest to their county, which were asked only of their county's residents.

#### **Sampling Approach & Design**

PRC designed the survey methodology to minimize sample bias and maximize representativeness by using best practice random-selection sampling techniques. They also used specific data analysis techniques, including poststratification, to further decrease sample bias and account for underrepresented groups or nonresponses in the population. Poststratification involves selecting demographic variables of interest within the population (here, gender, age,

race, ethnicity, and poverty status) and then applying “weights” to the data to produce a sample which more closely matches the actual county populations that make up the region.

In order to determine WNC regional data estimates, survey responses were weighted in proportion to the demographics of the adult population distribution within each county. After all data were collected, counties were weighted in proportion to one another based on actual population distribution to yield aggregate results representative of western North Carolina as a whole. Since the sample design and quality control procedures used in the data collection ensure that the sample is representative, the findings may be generalized to the region with a high degree of confidence (maximum error rate of  $\pm 1.7\%$  at the 95% confidence interval).

### ***Survey Administration***

PRC piloted the survey through 30 interviews across the region and consulted with WNC Health Network staff to resolve substantive issues before full implementation. PRC used trained, live interviewers with the help of CATI (computer-aided telephone interviewing) system to administer the survey region-wide. Survey interviews were conducted primarily during evening and weekend hours, with some daytime weekday attempts. Interviewers made up to five call attempts per telephone number. Interviews were conducted in either English or Spanish, as preferred by respondents. The final sample included 29 percent cell phone-based survey respondents and 71 percent landline-based survey respondents. Including cell phone numbers in the sampling algorithm allowed better representation of demographic segments that might otherwise be under sampled in a landline-only model.

PRC also worked with Research Now SSI, a third-party provider to identify and invite potential respondents from a pre-selected panel for an online survey for a small proportion (20%) of the sample population. The online survey was identical to the telephone survey instrument and allowed better sampling of younger and more urban demographic segments.

### **Survey Limitations and Information Gaps**

#### ***Limitations***

The survey methodology included a combination of telephone (both landline and cell phone) interviews, as well as an online survey. Limitations exist for these methods. For example, potential respondents must have access to a landline or a cell phone to respond to the telephone survey. In addition, the telephone survey sample included a large proportion of landlines (versus cell phones), which may further skew responses to individuals or households with landlines.

The online survey also has inherent limitations in recruitment and administration. Respondents were recruited from a pre-identified panel of potential respondents. This panel may not be representative of the overall population. In addition, online survey respondents must interpret survey questions themselves, rather than have them explained by a trained, live interviewer. This may change how they interpret and answer questions.

Lastly, the technique used to apply poststratification weights helps preserve the integrity of each individual's responses while improving overall representativeness. However, this technique can also exaggerate an individual's responses when demographic variables are under-sampled. Demographic variables not meeting WNC Healthy Impact reporting criteria have been suppressed in regional data reports.

### ***Information Gaps***

This assessment was designed to provide a comprehensive and broad picture of the health of the community overall. It does not measure all possible aspects of health in the community, nor does it represent all possible populations of interest. For example, due to low population numbers, members of certain racial/ethnic groups may not be identifiable or represented in numbers sufficient for independent analyses. In these cases, information gaps may limit ability to assess the full array of the community's health needs.

## **[Online Key Informant Survey](#)**

### **Survey Methodology**

#### ***Survey Purpose and Administration***

WNC Healthy Impact, with support from PRC, implemented an Online Key Informant Survey to solicit input from local leaders and stakeholders who have a broad interest in the health of the community. WNC Healthy Impact shared with PRC a list of recommended participants, including those from each county. This list included names and contact information for physicians, public health representatives, other health professionals, social service providers, and a variety of other community leaders. Potential participants were chosen because of their ability to identify primary concerns of the populations with whom they work, as well as of the community overall.

Key informants were contacted through an email that introduced the purpose of the survey and provided a link to take the survey online. Reminder emails were sent as needed to increase participation.

#### ***Survey Instrument***

The survey provided respondents the opportunity to identify critical health issues in their community, the feasibility of collaborative efforts around health issues, and what is helping/hurting their community's ability to make progress on health issues.

#### ***Participation***

In all, 329 community stakeholders took part in the Online Key Informant Survey for the WNC region, as outlined below:

<b>Local Online Key Informant Survey Participation</b>		
<b>Key Informant Type</b>	<b>Number Invited</b>	<b>Number Participating</b>
Community Leader	288	141
Other Health Provider	149	76
Physician	44	16
Public Health Representative	55	41
Social Services Provider	80	55

Through this process, input was gathered from individuals whose organizations work with low-income, minority populations, or other medically underserved populations.

**Online Survey Limitations**

The Online Key Informant Survey was designed to gather input from participants regarding their opinions and perceptions of the health of the residents in the area. Thus, these findings are based on perceptions, not facts.

To collect this data, purposive sampling (a type of non-probability sampling which targets a specific group of people) was used. Unlike the random sampling technique employed in the telephone survey, the purpose is not to make generalizations or statistical inferences from the sample to the entire population, but to gather in-depth insights into health issues from a group of individuals with a specific perspective.

**Secondary Data**

**Secondary Data Methodology**

To learn about the specific factors affecting the health and quality of life of residents of WNC, the WNC Healthy Impact data workgroup and data consulting team identified and tapped numerous secondary data sources accessible in the public domain. For data on the demographic, economic and social characteristics of the region sources included: the US Census Bureau; Log Into North Carolina (LINC); NC Department of Health and Human Services; NC Office of State Budget and Management; NC Department of Commerce; Employment Security Commission of NC; UNC-CH Jordan Institute for Families; NC Department of Public Instruction; NC Department of Justice; NC Division of Medical Assistance; NC Department of Transportation; and the Cecil B. Sheps Center for Health Services Research. The WNC Healthy Impact data consultant team made every effort to obtain the most current data available at the time the WNC Healthy Impact Data Workbook was prepared. It is not possible to continually update the data past a certain date; in most cases that end-point is August 2018. Secondary data will only be updated every August in between Community Health Assessment (CHA) years.

The principal source of secondary health data for the WNC Healthy Impact Data Workbook is the NC State Center for Health Statistics (NC SCHS), including its County Health Data Books, Behavioral Risk Factor Surveillance System, Vital Statistics unit, and Cancer Registry. Other health data sources included: NC Division of Public Health (DPH) Epidemiology Section; NC Division of Mental Health, Developmental Disabilities and Substance Abuse Services; the Centers

for Disease Control and Prevention; National Center for Health Statistics; NC DPH Nutrition Services Branch; and NC DETECT.

Environmental data were gathered from sources including: US Environmental Protection Agency; US Department of Agriculture; and NC Department of Environment and Natural Resources.

The WNC Healthy Impact data workbook contains only secondary data that are : (1) retrieved directly from sources in the public domain; and (2) are available for all 16 counties in the WNC Healthy Impact region. All secondary data included in the workbook are the most current available, but in some cases may be several years old. Names of organizations, facilities, and geographic places presented in the tables and graphs are quoted exactly as they appear in the source data. In some cases, these names may not be those in current or local usage; nevertheless, they are used so readers may track a particular piece of information directly back to the source.