Implementation of an Online Family Health History Tool using Research Assistants in Rural North Carolina

by

Ashley Nicole Wittmer

Duke Global Health Institute

Duke University

Date:_____

Approved:

_____

John Stanifer, Supervisor

_____

Lawrence Park

_____

Cherry Beasley

Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in the Duke Global Health Institute in the Graduate School
of Duke University

2018

ABSTRACT

Implementation of an Online Family Health History Tool using Research Assistants in Rural North Carolina

by

Ashley Nicole Wittmer

Duke Global Health Institute

Duke University

Date:_____

Approved:

_____

John Stanifer, Supervisor

_____

Lawrence Park

_____

Cherry Beasley

An abstract of a thesis submitted in partial
fulfillment of the requirements for the degree
of Master of Science in the Duke Global Health Institute in the Graduate School of
Duke University

2018

Abstract

Introduction: Chronic diseases have been increasing globally for decades, while the leading chronic diseases worldwide are cancer, cardiovascular disease (CVD), chronic respiratory disease, and diabetes. Behavioral risk factors of chronic diseases that can be modified include physical activity, diet, alcohol consumption and tobacco use. Several guidelines for screening and prevention recommend that family health history (FHH) is collected by primary care providers for disease risk stratification and management. MeTree, developed in 2014, is a computerized, patient-facing program that collects information about family health history and generates decision support for providers and patients. There are several potential barriers to implementation of an online FHH software tool including health literacy, computer skills, and behavioral components. This study collects FHH information through MeTree in a rural population in North Carolina through a unique implementation process using research assistants to manually and verbally assist participants. The aims of this study are to characterize the quality of pedigrees collected and to describe the most common diseases among the families of participants.

Methods: This study enrolled 44 participants from an ongoing study conducted by collaborators from Duke University Health System, Duke Clinical Research Institute,

University of North Carolina Pembroke, and Southeastern Regional Medical Center. To collect FHH information, participants constructed family pedigree in MeTree, one family member at a time with the help of one study research assistant. Once participants created a full family pedigree, an individual risk assessment was generated by MeTree.

Results: More than half of the participants were female (n= 30, 68.2%). The ethnic group that composed the largest part of our study population were Lumbee Indians (n=23, 52.3%) followed by White/Caucasians (n=13, 29.5%) and African Americans (n=7, 15.9%). For quality, the average score across all pedigrees was higher than 65% for all seven components of the criteria. The total number of diseases present among all participants and relatives in the study was 930 (Table 3). Cancer was present in 81.8% of pedigrees and made up 12.2% of all reported diseases. Twenty-five percent of all pedigrees had at least one family member that was diagnosed with lung cancer. Diabetes was also frequently reported and was observed in 75% of all pedigrees. Kidney Disease was reported in at least one or more relatives in 52.3% of pedigrees.

Conclusions: Using a tool such as MeTree could potentially lead to better health outcomes due to risk assessment and individually-targeted prevention strategies. MeTree may be an important tool to use to address the large burden of chronic diseases in regions similar to Robeson County.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisor and committee member, Dr. John Stanifer, for his leadership, efforts, and commitment to this thesis and our summer research project in Robeson County. I would like to thank my committee chair, Dr. Lawrence Park, for his support with this thesis and for his guidance in interpreting the findings of this study. I would also like to recognize my committee member, Dr. Cherry Beasley, for introducing me to several of the concepts that influenced my project. I would like to thank the Director of the Precision Medicine Program in the Center for Applied Genomics and Precision Medicine at Duke University, Dr. Lori Orlando, for her part in the creation of MeTree and for introducing us to this software tool. I would also like to recognize Dr. Rachel Myers, as her help with the organization and data analysis of this project was greatly appreciated. I want to thank Dr. William Puentes and the rest of the staff at the Department of Nursing at The University of North Carolina at Pembroke. I would like to also recognize the staff at the Southeastern Health Clinic in Lumberton and thank them for their help with participant coordination.

1. Introduction

Over the last few decades, chronic diseases have been increasing in number all around the world.[1] By 2030, chronic diseases will be responsible for over 75% of all deaths worldwide.[2] The leading chronic diseases worldwide are cancer, cardiovascular disease (CVD), chronic respiratory disease, and diabetes, which were responsible for 29 million deaths globally in 2002.[1] Nearly half of all deaths caused by chronic diseases are attributable to CVD, with diabetes also presenting a worrying trend due to the fact that it often appears earlier in life.[3]

Some of the non-modifiable risk factors for chronic diseases include age, sex, and genetic susceptibility. Examples of behavioral risk factors that can be modified include physical activity, diet, alcohol consumption and tobacco use. Risk of developing chronic diseases can also be influenced by societal factors such as socioeconomic status and cultural and environmental variables.[3] [4] If an individual is genetically susceptible to a particular disease, that does not guarantee that the disease will develop. However, developments in preventative medicine over the years have encouraged interventions that are based on the probability that disease will occur.[5] Knowledge of genetic susceptibility could prompt an individual to modify lifestyle factors associated with risk of developing that particular chronic disease.

1.1 Family Health History

Family health history (FHH) evaluates risk factors of chronic disease due to genetic susceptibility as well as shared environment and common behaviors. The information obtained from a FHH can range from knowing an immediate family member had a specific disease to analysis of detailed pedigrees with multiple generations of relatives.[4] Many guidelines for screening and prevention recommend that FHH is collected by primary care providers for disease risk stratification and management.[6] [7] For example, one of the first steps in assessing a patient's risk for atherosclerotic CVD is the explanation of a FHH. This cost-effective recommendation can be used to determine the next steps for the patient in preventing or mitigating the impact of the condition.[8]

FHH can be a powerful tool in preventing or mitigating adverse health outcomes; however, in clinical settings, it is often not collected or collected with poor quality due to a number of barriers. Some factors that impede successful collection of FHH in clinical settings include time required to record a valid history, lack of referral guidelines, and difficulty navigating FHH information in a patient's chart.[9] [10] [11] Using an electronic tool that collects FHH information from patients could be used to overcome these barriers by adding information to the patient's electronic medical record (EMR). [10] [12] If these barriers are mitigated, successful collection of FHH could provide insight into treatment and prevention strategies.

Another use of a FHH is to evaluate recurrence risk, which is defined by the National Cancer Institute (NIH) as the chance that a hereditary disorder or trait that one family member has will occur in other family members.[13] Family aggregation of many diseases is likely to be heavily influenced by genetic factors. [14] One study used recurrence risk ratios in different degrees of relatives to assess the most likely modes of inheritance that influence susceptibility to a certain disease.[15] A recurrence risk ratio is calculated as the risk for a disease in relatives of a random individual who has the disease, divided by the prevalence for that disease in the population. [16] [17]

1.2 MeTree

Patient-facing Health Information Technology (HIT) services are designed to promote the practice of patients playing a more active role in healthcare. Through these services, patients are given more access to professionals and health record information, more support for self-care management, and the ability to conduct health-related transactions.[18] In 2004, MeTree was developed by the collaboration of Duke University, the University of North Carolina Greensboro, and Cone Health System known as the Genomedical Connection. MeTree enables rapid collection of a standardized health history through a simple graphical interface.[12] MeTree is a computerized, patient-facing program that collects information from patients about family health history and then generates decision support for providers and patients.[6] [19] The aim of this tool is to help

primary care providers identify high risk individuals who may require referrals or additional screening.[12]

The capability for MeTree to generate risk assessment of diseases and the construction of pedigrees is largely dependent on the quality of the FHH information provided by the patient. A pedigree provides a visualization of the health-related experiences of the individual and their family.[20] There are a few different attributes that determine the quality of a pedigree. These criteria include three generations of relatives, maternal or paternal lineage, gender of relative, pertinent negatives noted, the age of onset of disease for affected living and deceased relatives, as well as age of death and cause of death for deceased relatives. Pertinent negatives describe a FHH noting common diseases that are not found within the family.[21] [22] [23] [24] [10] In a previous study, these criteria were used to evaluate the quality of pedigrees collected after participants were encouraged to talk to their relatives about their family history. [21]

Another recent study implemented MeTree in four distinct healthcare settings and assessed the implementation outcomes using the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework.[7] [25] This study demonstrated that ethnic and racial minorities were more likely to report enhanced disease awareness, even among those with lower levels of education. As such, MeTree may be an important tool to facilitate reduced potential health disparities that may arise from demographic factors or lower access to risk assessment.[7]

When MeTree was implemented as part of a large study in Greensboro, NC, the quality of pedigrees obtained was very high. [21] However, this study took place in a more urban area than our study in rural Robeson County.

1.3 Robeson County

Robeson County is a rural county in North Carolina and is home to the Lumbee Indians. It is also the poorest county in the state, ranking last in almost every health metric in the state. The state, but not the federal government, recognizes the Lumbee tribe as American Indians. This means that they are ineligible for Indian Health Services (IHS) benefits.[26] Robeson County is ethnically comprised of 38.4% Native residents, 26% white residents, 24.2% black residents, 8.3% Hispanic residents, 2.0% residents who identify as two or more ethnicities.[27]

The estimates for life expectancy at birth in this county are much lower than for the state as a whole. On average, men and women die 6.3 and 4.4 years earlier compared to national sex-specific estimates [28] [29] There are a large number of negative health problems and outcomes among the residents of Robeson County. For example, the Diabetes Prevalence for adults age 20 years and older is 15.7%. In addition, 198 Medicare enrollees received treatment for Congestive Heart Failure in 2010.[27] Of all the counties in North Carolina, Robeson has the highest prevalence of adult smoking at 26.8% and experienced the highest number of deaths due to homicide (20.9 per 100,000 people) from 2009-2015.[27]

1.4 Potential Barriers to Implementation in a Rural Setting

There are several potential barriers to implementation of an online FHH software tool including literacy, health literacy, computer skills, and behavioral components. These barriers can be explained using The Information-Motivation-Behavioral Skills (IMB) model of health behavior (Figure 1). [30] [31] This model proposes that a health behavior can be most successfully changed in an individual who is well informed, motivated, and has the skills and confidence to perform or change the behavior, which leads to better health outcomes.[32] In this case, we want to describe how concepts in this model such as information through literacy and health literacy, and behavioral skills such as self-efficacy impact the health behavior of learning one's family history. Motivation can also be used to describe patients' attitudes on the usefulness of knowing their family history as well as their self-efficacy of being able to use an online FHH tool such as MeTree. Ultimately, we want to describe how knowledge of one's family health history can be elucidated through this tool and potentially lead to better health outcomes. We first need to identify the potential barriers to using this tool through application of the IMB model.

As a part of *Healthy People 2010*, The Centers for Disease Control and Prevention (CDC) sought to improve health literacy. They defined health literacy as the extent to which individuals have the ability to obtain and understand basic health services and understand the information required to make informed health decisions.[33] Low health

literacy and lack of computer access have been identified by previous studies as potential barriers to HIT services.[18] [34] Published research has shown that low health literacy can also be a barrier to education among patients with chronic diseases, leading to increased costs in the health care industry due to inappropriate or inadequate use of medications. Individuals with low health literacy are more likely to delay seeking treatment for a condition until the problem becomes more serious and expensive to treat.[35] [36] Low health literacy is associated with other determinants that also have an impact on health status. In 2003, The National Assessment of Adult Literacy (NAAL) conducted a study on health literacy in the U.S. that surveyed over 19,000 adults and grouped their results into the following categories: proficient, intermediate, basic, or below basic. They found that only 12% of adults surveyed demonstrated proficient health literacy and that uninsured adults along with adults insured by Medicare or Medicaid were more likely to have basic or below basic health literacy.[37] The study also showed that the impact of low health literacy on poorer health outcomes is greater among ethnic and racial minorities. Hispanics, followed by African Americans and American Indians, had the highest population percentage of basic or below health literacy.[37]

Patient-facing online HIT services require that users have some level of computer skills for them to be able to successfully navigate through the software. Although recent research suggests that almost 8 in 10 Americans have internet access and own a

computer, there is an unequal distribution in online capability.[38] [39] Similar to the gaps in service and access that have been present in the U.S. healthcare system for years, there are apparent disparities that impact access and use of the internet. [40] Groups that are commonly underserved such as those with low education and low income are less likely to use or own computers. [41] This uneven distribution in computer access among certain groups is referred to as the digital divide and could cause underserved groups to be excluded from beneficial health outcomes of using an online HIT tool.[42] [39]

A behavioral component that could be a potential barrier to implementing MeTree is lack of self-efficacy, which is defined as an individual's belief in his or her own ability to perform necessary behaviors. This concept was first introduced by Albert Bandura in 1977 as part of a theoretical framework to explain its relationship to behavioral changes.[43]

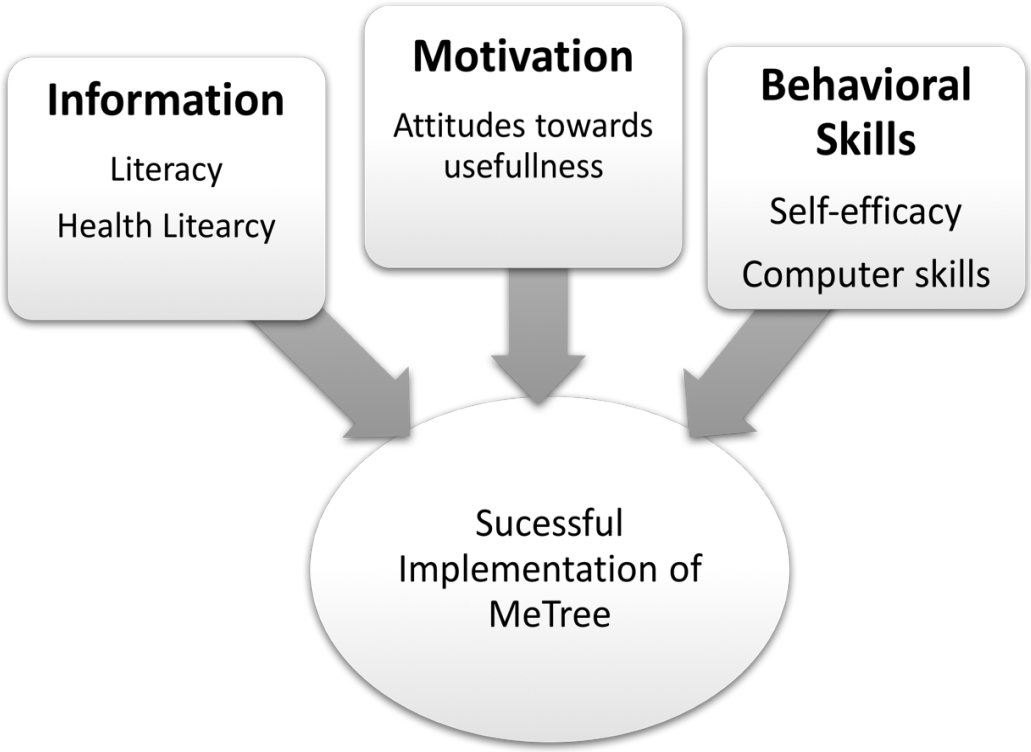

Figure 1 - Barriers to Implementation using the IMB Model

1.5 Study Aims and Rationale

In this study, we wanted to collect FHH information through MeTree in a rural population in North Carolina using a unique implementation process in attempt to eliminate potential barriers, including health literacy, computer skills, attitudes, and self-efficacy. As part of our implementation, we proposed using study research assistants to aid participants in entering FHH information into the patient-facing program. Our rationale behind using the research assistants was, as they manually and verbally assist participants, they may be able to clarify instructions, help define health-

related terms using prompts provided by MeTree, and ensure completion of the pedigrees. The aims of our study are outlined below:

Aim 1: To characterize the quality of pedigrees obtained from MeTree.

Aim 2: To estimate disease prevalence (e.g. diabetes, CKD, hypertension) among ethnic and racial minorities living in Robeson County.

The study aims were designed with the primary focus to assess the quality of pedigrees collected with the help of research assistants. We hypothesized that a majority of pedigrees would satisfy most of the quality criteria, due to our implementation approach. In addition, we hypothesized that we would be able to use information obtained to describe the most common diseases reported in the study population.

2. Methods

2.1 Setting

After the recruitment and consent processes, participants answered surveys in Redcap, filled out their FHH information in MeTree with a research assistant, and then went to the lab to have blood drawn (Figure 2). This study was conducted between May and July of 2018 in rural Robeson County, specifically in the towns of Pembroke and Lumberton. Both of these towns are home to the Lumbee Indians.[26] Our research team conducted interviews from two different locations in order to better accommodate our study population.

The main site in Pembroke was located on the campus of The University of North Carolina at Pembroke (UNCP) in the Weinstein Health Services Building. We called and recruited our participants from this location and also did most of the work with our collaborators there. The participants who lived in or near Pembroke came to this site for interviews. To get their lab specimens drawn for the study, the participants at this site were referred to the Southeastern Express Lab located in Pembroke 1.5 miles from the UNCP campus.

Our other site in Lumberton was located inside the Southeastern Health Mall Clinic. Each day, two out of the four team members would drive to this location from the main site in Pembroke. Many of our participants lived in or near Lumberton, so the majority of the interviews were conducted at this location. They were also able to have their study specimens taken inside the clinic at the Southeastern Express Lab located down the hall from the interview room.

2.2 Kidney NC Study

The participants from this study were recruited from an ongoing study "Improving Outcomes for Chronic Kidney Disease in Southeastern North Carolina: An Evaluation of the Individual-and Community-Level Responses." This study, which started in 2016, is conducted by collaborators from Duke University Health System, Duke Clinical Research Institute, University of North Carolina Pembroke, and Southeastern Regional Medical Center. The aim of this study is to understand the

potential reasons why there is a high prevalence of chronic kidney disease (CKD) in the area. Some information collected in this study includes participants' self-reported medical history, diet, physical activity, and e-health literacy. Data is also collected regarding the mental health of the participants. Such topics include, anxiety and depression, coping mechanisms, self-efficacy, and social support. The overall goal of the study is to alleviate the burden of adverse outcomes caused by CKD in this region. Currently, there are 467 participants in this study. (Stanifer JW, Beasley, CM. KIDNEY NC Study 2018).

2.3 Recruitment

Using non-probability sampling, we recruited participants for this study from the parent study described previously. Study research assistants called participants who consented to being contacted again for future studies and explained the process of this new study. The inclusion criteria included participants who had previously participated in the parent study from 2016 to 2017. Individuals who did not receive medical care at Southeastern Regional Medical Center, who planned to move away from the area in the following 12 months, or who were pregnant were excluded from the study.

2.4 Procedures

Participants who consented to be contacted again in the parent study were called by research assistants using mobile phones provided by the study. Once they agreed to come back in for a family health history interview, research assistants verified that they

met the inclusion/exclusion criteria. In preparation for the interview, research assistants sent out a letter in the mail to each participant approximately one full week before the interview. Similar to the study mentioned previously, we encouraged all participants to speak with their families prior to the MeTree interview.[21] This letter contained a MeTree family health history form that participants were asked to fill out with their families if possible. This form explains the importance of knowing one's family history, a guide on where to start, a list of all diseases that MeTree will ask about, and a page with a list of family members for them to write diseases and current age or age of death next to.

On the day of the interview, participants were greeted by research assistants at either the Pembroke or Lumberton located. Before starting the interview, participants were asked to sign a consent form that was read aloud to them by research assistants. The consent form allowed them to consent to the study, lab testing for the study, or both. Nearly all participants also consented to having labs drawn after the interview. Next, participants answered questions from a brief survey on Redcap that included information on residential information, a short, self-reported medical history, an environmental assessment, and sleep practices and characteristics.

Next, research assistants prepared for the MeTree portion of the interview by helping participants set up their own account and login. After creating a portal in MeTree, they were asked to first provide answers to questions related to lifestyle, followed by an individual health history. Once these steps were competed, participants

began constructing a family pedigree in MeTree, one family member at a time. MeTree requires that they at least provide basic information on each parent, both maternal grandparents, and both paternal grandparents. From here, they added aunts, uncles, siblings, cousins, children, and grandchildren based on what they knew about their health. Lastly, after participants created their full family pedigree, MeTree provided an individual risk assessment for them, based on the information entered. Research assistants walked through the information on the assessment with participants in order to ensure that they comprehended all of the recommendations. The research assistants encouraged participants to visit their doctor if they had any concerns based on the risk assessment. After the interview was completed, participants were given direction to the lab closest to them to have their blood drawn and given the study lab requisition form.

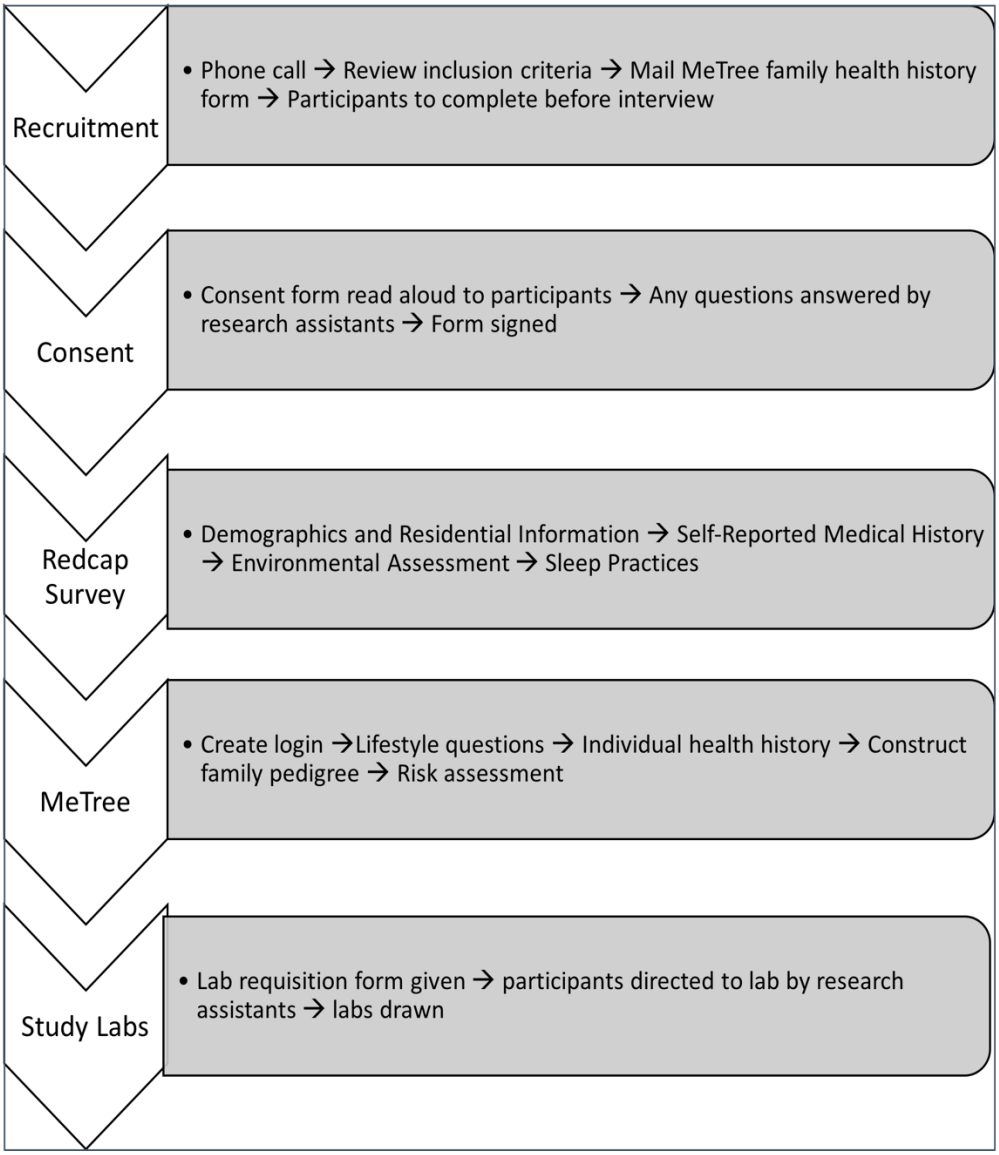

Figure 2 - Study Flow Diagram.

2.5 Measures

In order to measure pedigree quality, we used established criteria.[10] The criteria are: 1) Three generations of relatives; 2) Distinguished lineage; 3) Gender of relatives; 4)

Pertinent Negatives Noted; 5) Updated Information; 6) Age of Death; 7) Cause of Death; 8) Age of onset of disease. The first five components (i.e. three generations of relatives, distinguished maternal and paternal lineage, gender of relative, age of death, cause of death, and pertinent negatives noted are required by MeTree. The tool automatically generates a pedigree with the individual, their parents, and both maternal and paternal sets of grandparents. It requires the participant to go into each of these relatives' profiles and select their current age or, if deceased, their age of death. If deceased, it requires the participant to enter that relative's cause of death and this also gives an option to state that this information is unknown. Another component of the quality criteria, information that is updated, depends on whether they reported knowing a relative's history. There is an option to select whether or not they know each family member's history and if they checked the box, they were considered to have met this metric for the given relative. Some participants were able to provide several details about the person's lifestyle factors and background along with a full health history, while some only knew the health history and not much else about the family member. Similar to a previous study, we considered a pedigree to be high quality if at least one family member met all components of the specified criteria.[21]

In order to describe the most common diseases reported among family health histories collected, we used the information provided by the participant for each family member. Each disease for each relative that was entered by the participant was stored

into MeTree, so we were able to count the frequency of each type of disease as they were recorded.

2.6 Analysis

First, we used counts to determine how many relatives per pedigree met or did not meet the given requirement. From there, we used proportions to compare the number of relatives for which the metric was met in each pedigree. We used those numbers to calculate percentages that corresponded to the quality of each individual pedigree for each metric. After we had an average value for each pedigree listed as a percentage, we calculated the mean of these averages for each of the metrics. Cross-tabulations between variables were used to examine possible associations and one-way frequencies described how disease values were distributed across pedigrees. We also conducted a sub-analysis to determine if diseases distribution would be different when comparing index participants (n=44) to their relatives included in the pedigrees (n=686). To determine if there were significant differences in proportions of disease prevalence between these two groups, we used Fisher's Exact test at the 5% significance level. We used RStudio v1.2 to analyze pedigree quality and information obtained from the data collection.

3. Results

3.1 Participant Characteristics

We enrolled 44 participants with a median age of 68.5 (Table 1). More than two-thirds of the participants were female (n= 30, 68.2%). The ethnic group that composed the largest part of our study population were Lumbee Indians (n=23, 52.3%) followed by White/Caucasians (n=13, 29.5%) and African Americans (n=7, 15.9%). About one-fourth of participants reported having a high school education (n=12, 27.3%) and an equal number of participants did not report their education level at all (n=12, 27.3%). Most participants reported having some college education and one-fifth of participants attended graduate school (n=9, 20.4%).

We collected 44 pedigrees representing 730 individuals in total. The mean pedigree size was 16.6 (SD 2.4), ranging from 7-39 family members. Most pedigrees took about two hours to collect and for some of the larger ones, up to three hours were spent inputting information with participants.

Table 1 - Study Index Participant Demographics (N=44)

	Study Participants (N=44)
Gender (% Female)	30 (68.2%)
Age (years) Median	68.5
Range	32 – 82
Ethnicity/Race	
White	13 (29.5%)
African American	7 (15.9%)
Lumbee Indian	23 (52.3%)
Other	1 (2.3%)
Covered by Insurance	42 (95.5%)
Median Family Size (SD)	16.6 (2.4)
Range of Family Members	7 – 39
Education	
High School	12 (27.3%)
Community College	5 (11.4%)
4 Year University	6 (13.6%)
Graduate School	9 (20.4%)
Not Reported	12 (27.3%)
Occupation	
Professional	15 (34.0%)
Office Worker	3 (6.8%)
Other	17 (38.6%)
Not Reported	9 (20.4%)

3.2 Pedigree Quality

Most (n=36, 81.8%) pedigrees met the criteria for high quality, having one or more relatives that satisfied all metrics.[21] [10] Each quality metric score was high among the individuals and relatives and age of onset was the quality metric that had the lowest average value of 65.7% (Table 2). On average, cause of death was known for 74.7% all deceased family members across all pedigrees while age of death was known for 73.0% of all nonliving relatives.

Table 2 – Percentage of Individuals (n=730) Across Pedigrees Meeting Each Quality Criterion

Quality Metrics	*Percentage of Pedigrees*
Three generations of relatives	**100%**
Maternal or Paternal Lineage	**100%**
Gender of Relatives	**100%**
Age of Onset of Disease	**65.7%**
Age of Death	**73.0%**
Cause of Death	**74.7%**
Known History	**72.7%**

3.3 Prevalence of Common Diseases

The total number of diseases present among all participants (n=730) and relatives in the study was 930 (Table 3). Cardiovascular disease (CVD) was the most common disease group reported and was present in at least one or more family members in 88.6% of the pedigrees and accounted for 17.1% of all diseases reported. Within the CVD group, Heart Attack/Coronary Artery Disease was the most reported with 104 occurrences across all family members and present in 75% of all pedigrees. Hypertension, listed as a separate category, accounted for 12.6% of all diseases reported but was present in just as many pedigrees as CVD at 88.6%.

Some form of cancer was present in 81.8% of pedigrees and totaled up to 12.2% of all reported diseases. The three types of cancer that occurred most frequently among all relatives were prostate cancer, breast cancer, and lung cancer. Twenty-five percent of all pedigrees had at least one family member that was diagnosed with lung cancer, with 13 total counts in the study population. Prostate cancer was reported in 27 relatives across all pedigrees while breast cancer was reported in 18 relatives. However, breast cancer was present in 30% of all pedigrees while prostate cancer was only reported in 27.3% of the pedigrees.

Diabetes was also frequently reported, observed in 75% of all pedigrees. Type 2 diabetes occurred more often than Type 1 diabetes and composed 70.5% of the Diabetes category. Type 1 diabetes only accounted for 1.7% for all diseases reported while Type 2 diabetes itself was responsible 9.8% of all disease occurrences and was present in 70.5% of all pedigrees.

Another commonly reported disease group was Kidney Disease, which was reported in at least one or more relatives in 52.3% of pedigrees. The most common type of kidney disease in the study was the subtype known as Other Kidney disease which was noted in 29 relatives accounting for 3.1% of total diseases reported. Within kidney disease, the two other most reported types were Diabetic Kidney disease and Kidney Nephrosis, while Cystic Kidney disease and Nephritis were less common. Diabetic

Kidney disease and Nephrosis were both noted 5 times across all relatives in the study and were reported at least once in 9.1% of all pedigrees.

Table 3 - Commonly Reported Conditions

Disease Group	Sub-type	Frequency N=930 (%)	Present in Pedigrees N=44 (%)
Cardiovascular Disease	**Heart Attack/Coronary Artery Disease**	104 (11.2%)	33 (75.0%)
	Other Heart Disease	53 (5.7%)	20 (45.5%)
	Atrial Fibrillation	2 (0.22%)	1 (2.3%)
	Total CVD	**159 (17.1%)**	**40 (88.6%)**
Cancer	**Bone**	2 (0.22%)	2 (4.5%)
	Brain	4 (0.43%)	4 (9.1%)
	Breast	18 (1.9%)	13 (30.0%)
	Colon	7 (0.75%)	6 (13.6%)
	Kidney	3 (0.32%)	3 (6.8%)
	Leukemia	5 (0.54%)	5 (11.4%)
	Liver	2 (0.22%)	1 (2.3%)
	Lung	13 (1.4%)	11 (25.0%)
	Other	13 (1.4%)	11 (25.0%)
	Ovarian	3 (0.3%)	3 (6.8%)
	Pancreatic	1 (0.11%)	1 (2.3%)
	Prostate	27 (2.9%)	12 (27.3%)
	Skin	7 (0.8%)	6 (13.6%)
	Esophageal	1 (0.11%)	1 (2.3%)
	Stomach	5 (0.54%)	5 (11.4%)
	Lipoma	2 (0.22%)	1 (2.3%)
	Total Cancer	**113 (12.2%)**	**36 (81.8%)**
Hypertension	**Total Hypertension**	**117 (12.6%)**	**40 (88.6%)**
Diabetes	**Type 1 Diabetes**	16 (1.7%)	8 (18.2%)
	Type 2 Diabetes	91 (9.8%)	31 (70.5%)
	Total Diabetes	**107 (11.5%)**	**33 (75.0%)**
Kidney Disease	**Cystic**	1 (0.11%)	1 (2.3%)
	Diabetic	5 (0.5%)	4 (9.1%)
	Nephrosis	5 (0.5%)	4 (9.1%)
	Nephritis	1 (0.11%)	2 (4.5%)

As part of a sub-analysis, we evaluated the conditions reported by index participants (n=44) compared to conditions reported for their relatives (n=686). We found that there were 198 diseases among the index participants and 732 diseases among their relatives. This corresponds to a mean of 4.5 diseases per person among index participants. Along with having a higher average number of diseases per person, the index participants also had some differences in most reported conditions. Hypertension was the most reported condition and present in 72.2% of these index participants. High cholesterol was more frequently reported in participants compared to all relatives, reported by 56% of this group. Similar to what we observed among all participants, Diabetes accounted for 11.6% of all diseases (n=198) reported by index participants. We found that relatives (n=686) had much higher counts of Cardiovascular Disease and Cancer compared to index participants (Figure 4). Among the relatives, we also found smaller percentages of Psychological Disorders, Rheumatoid Arthritis, and Obesity. However, obesity was self-reported in 22.7% of index participants, while it was only reported for less than 1% of all relatives (n=686).

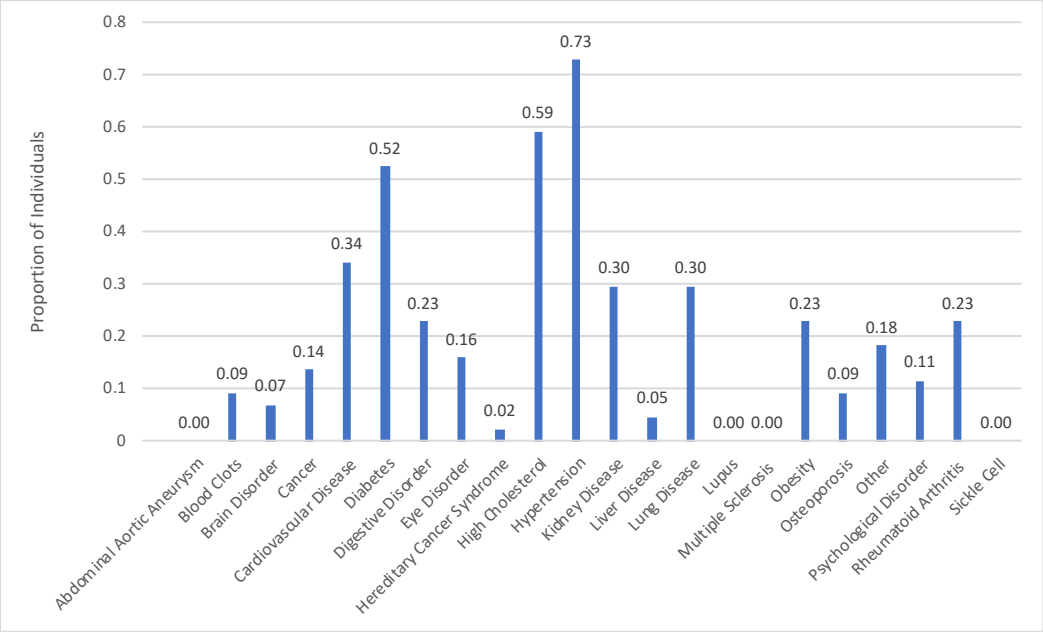

Figure 3 - Proportion of Individuals with Common Diseases Among Index Participants (N=44)

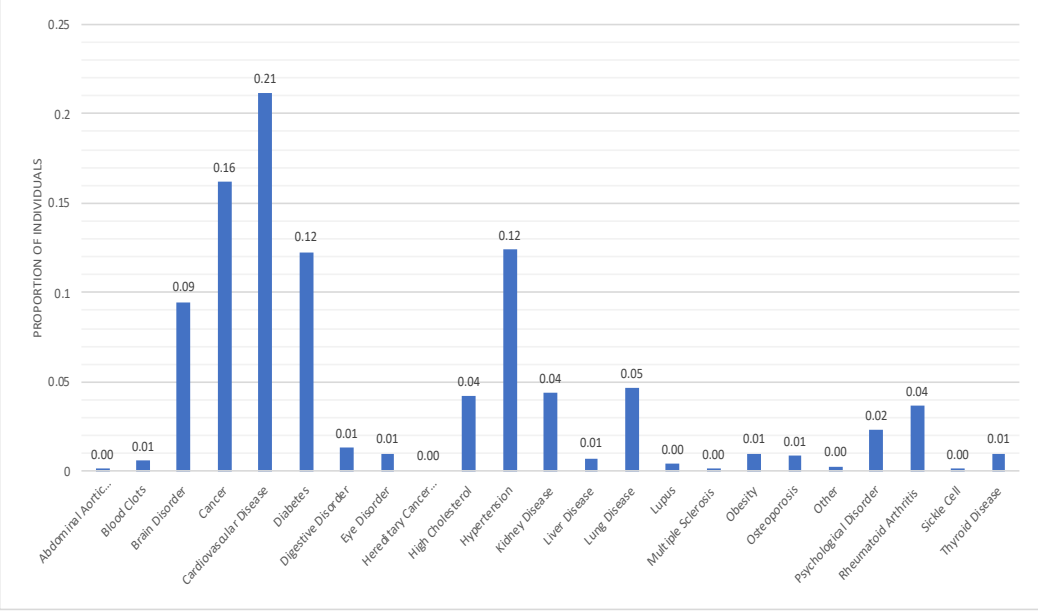

Figure 4 - Proportion of Individuals with Common Diseases Among Relatives (N=686)

The proportion of relatives with cancer and CVD was higher than the proportion of index participants but this difference was not statistically significant (Table 4). However, we found that there was a significant difference in the reporting of Hypertension, High Cholesterol, Obesity, Kidney Disease, and Psychological Disorders between index participants and relatives.

Table 4 -Differences in Distribution of Diseases Between Index Participants and Relatives

Disease	Relatives (n=686)	Index Participants (n=44)	Fisher's Exact P-value (p<0.05)
Cancer	111	6	0.83
CVD	144	14	0.13
High Cholesterol	29	26	**<0.001**
Hypertension	85	32	**0.01**
Kidney Disease	30	13	**<0.001**
Obesity	7	10	**<0.001**
Psychological Disorder	16	4	**0.03**

4. Discussion

4.1 Findings of Pedigree Quality

Overall, using our modified approach to address implementation barriers, the pedigree quality obtained by MeTree was high. [7] In the study that implemented MeTree in Greensboro, they used the same quality criteria and reported 99.8% of collected histories to be high quality.[21] They stated that a factor that contributed to high quality was talking to relatives before using MeTree. During the enrollment process of our study, we talked to participants over the phone about the importance of talking to

relatives before the MeTree interview. We mailed forms that nearly every single participant brought into the interview already filled out. The discrepancy in quality between our study and the previous study could be due to barriers that exist in a rural setting such as health literacy or computer access.[39] [35]

4.2 Findings of Reported Diseases in Pedigrees

The FHH collected by MeTree found a large number of CVD, cancer, diabetes, and several other conditions among individuals and families in this study. However, behavioral risk factors that can be modified include alcohol consumption and tobacco use, diet, and physical activity. Societal factors such as socioeconomic status and cultural and environmental variables also contribute to risk of these conditions. [3] [4]

Robeson County is among the highest burdened counties in North Carolina by lung cancer, which may be in large part due to the high prevalence of cigarette smoking.[27] [44] [45] In a county where almost a quarter of all adults smoke, we identified lung cancer as one of the most prevalent diseases reported in family histories. Providing a risk assessment for lung cancer, which may be otherwise overlooked by primary care providers, may be an important tool to address the large burden of lung cancer in this region.

Robeson County also has a very high prevalence of Type 2 diabetes, diagnosed in nearly one-fifth of all adults . [27] The findings of this small study mirrored the trend of Type 2 diabetes in the larger population, present in almost three quarters of all family

histories. Because lifestyle factors such as diet and physical activity play such a large role in developing Type 2 diabetes, this tool could benefit many who are genetically susceptible. Using the results of this tool's risk assessment, primary care providers and patients may be able to work together to prevent some of the burden caused by Type 2 diabetes.

To be included in the parent study, participants did not have to be diagnosed with CKD, so not all of those who also took part in this study were diagnosed with CKD. However, CKD was one of the most commonly reported conditions across all pedigrees. Along with risk factors discussed previously for other chronic conditions, some risk factors for CKD include age, hereditary factors, and ethnicity with African Americans being disproportionately affected in the U.S. [46]

After exploring the distribution in diseases reported for index participants (n=44) compared to their relatives (n=686), we found some similarities and differences. The index participants reported a larger percentage of hypertension and high cholesterol, but a lower percentage of cancer and CVD. These observed differences between disease prevalence in the index participants and overall family members could be due to ascertainment bias. This type of bias is likely to have occurred due to the way we collected data. There could be a misrepresentation of the true frequency of some of these diseases among relatives. Because individuals are likely to know their own medical

history better than their relatives' histories, some conditions could be under-reported or over-reported for family members of the index participants.

4.3 Implications for Practice and Further Research

Unlike other studies that collect FHH information using MeTree, this study adds to the literature by implementing this tool using research assistants to guide participants through the process. This study was able to identify common conditions among participants and families and explore how these findings reflected trends in the larger population.

Currently, more information is needed to describe familial aggregation of the most prevalent diseases reported. We found that prostate cancer was present in more relatives than breast cancer; yet breast cancer was reported in more pedigrees overall. This could be due to the way prostate cancer is aggregated within the families in the study that indicated one or more occurrence. Family aggregation of many diseases is likely to be heavily influenced by genetic factors. [14] Recurrence risk can also be used to assess the most likely modes of inheritance for certain diseases. Knowing how a disease clusters in families could potentially be used to identify individuals who are at a higher risk of developing the condition. As a result, appropriate prevention guidelines could be implemented into that patient's treatment regimen. Identifying familial aggregation patterns could lead to increased awareness of disease susceptibility among the patient

and all other family members who may be at risk, leading to possible modification of certain risk factors.

Another powerful component of FHH studies is that they are able to provide insight into heritability patterns and gene-environment interaction. The phenotypic expression of some genes is largely impacted by the environment. In families, a shared environment is typically inherited among family members which means that they are vulnerable to the same exposures.[47] These types of studies can help patients and providers explore possible pathways between toxins and gene expression as well as potential routes of exposure.

These types of family studies also have the potential to raise awareness of disease risk among families and the community. If one family member finds out that they are genetically susceptible to a disease, they will most likely share this information with their relatives. Increasing awareness among entire families could benefit communities as a whole and encourage collaboration in altering modifiable risk factors.

Further research is needed on how to effectively implement MeTree in a clinical setting in rural areas. Strategies centered on better preparing patients for MeTree should be planned as part of the implementation. Patients with lower levels of health literacy, lack of access to a computer, poor computer skills, and lower levels of self-efficacy may need to prepare in a different manner and time frame than patients who are not influenced by these barriers. Targeting the patients' perceived usefulness and attitudes

towards MeTree could also contribute to better implementation, which could ultimately lead improved health outcomes in the population.

4.4 Study Strengths and Limitations

The primary strength of this study is that it demonstrated feasibility of using MeTree with research assistants in a rural and underserved setting. In addition, all participants in this study were recruited from the parent study and were already familiar with the goals and purpose of the larger project. Most of them remembered which location Pembroke or Lumberton, that they visited previously and indicated that they were eager to come back. For most, there was already a sense of trust between the participants and research assistants.

Although we were able to collect fairly high-quality pedigrees, we noted a few limitations to this study. All diseases were self-reported; as such inaccurate or missing values in some pedigrees may have led to reporting bias. While this is also a major strength, having research assistants present during FHH collection could potentially lead to response bias, particularly among sensitive topics such as tobacco or alcohol use.

5. Conclusion

Using tools such as MeTree that help facilitate FHH collection could potentially lead to better health outcomes due to risk assessment and individually-targeted prevention strategies. Further research is need on how to successfully implement this

software tool in clinical settings within rural areas. Research is needed to identify how to overcome some potential barriers to use in a rural setting such as health literacy, computer access, and patient self-efficacy.

References

1. Yach, D., Hawkes, C., Gould, C. L. & Hofman, K. J. The Global Burden of Chronic Diseases: Overcoming Impediments to Prevention and Control. *JAMA* **291**, 2616–2622 (2004).

2. Geneau, R. *et al.* Raising the priority of preventing chronic diseases: a political process. *The Lancet* **376**, 1689–1698 (2010).

3. *Diet, nutrition, and the prevention of chronic diseases: report of a WHO-FAO Expert Consultation ; [Joint WHO-FAO Expert Consultation on Diet, Nutrition, and the Prevention of Chronic Diseases, 2002, Geneva, Switzerland]*. (World Health Organization, 2003).

4. Can family history be used as a tool for public health and preventive medicine? | Genetics in Medicine. Available at: https://www.nature.com/articles/gim200244. (Accessed: 13th October 2018)

5. Davison, C., Macintyre, S. & Smith, G. D. The potential social impact of predictive genetic testing for susceptibility to common chronic diseases: a review and proposed research agenda. *Sociol. Health Illn.* **16**, 340–371 (1994).

6. Orlando, L. A. *et al.* Development and validation of a primary care-based family health history and decision support program (MeTree). *N. C. Med. J.* **74**, 287–296 (2013).

7. Wu, R. R. *et al.* Implementation, adoption, and utility of family health history risk assessment in diverse care settings: evaluating implementation processes and impact with an implementation framework. *Genet. Med. Off. J. Am. Coll. Med. Genet.* (2018). doi:10.1038/s41436-018-0049-x

8. Members, W. C. *et al.* 2010 ACCF/AHA Guideline for Assessment of Cardiovascular Risk in Asymptomatic Adults: Executive Summary. *Circulation* (2010).

9. Suther, S. & Goodson, P. Barriers to the provision of genetic services by primary care physicians: A systematic review of the literature. *Genet. Med.* **5**, 70–76 (2003).

10. Powell, K. P. *et al.* Collection of Family Health History for Assessment of Chronic Disease Risk in Primary Care. **74**, 9

11. Christianson, C. A. *et al.* The Use of a Family History Risk Assessment Tool within a Community Health Care System: Views of Primary Care Providers. *J. Genet. Couns.* **21**, 652–661 (2012).

12. Orlando, L. A. *et al.* Protocol for implementation of family health history collection and decision support into primary care using a computerized family health history system. *BMC Health Serv. Res.* **11**, 264 (2011).

13. Definition of recurrence risk - NCI Dictionary of Cancer Terms. *National Cancer Institute* (2012). Available at: https://www.cancer.gov/publications/dictionaries/genetics-dictionary. (Accessed: 15th October 2018)

14. Khoury, M. J., Beaty, T. H. & Kung-Yee, L. CAN FAMILIAL AGGREGATION OF DISEASE BE EXPLAINED BY FAMILIAL AGGREGATION OF ENVIRONMENTAL RISK FACTORS? *Am. J. Epidemiol.* **127**, 674–683 (1988).

15. Brown, M. A., Laval, S. H., Brophy, S. & Calin, A. Recurrence risk modelling of the genetic susceptibility to ankylosing spondylitis. *Ann. Rheum. Dis.* **59**, 883–886 (2000).

16. Rybicki, B. A. & Elston, R. C. The Relationship between the Sibling Recurrence-Risk Ratio and Genotype Relative Risk. *Am. J. Hum. Genet.* **66**, 593–604 (2000).

17. Risch, N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am. J. Hum. Genet.* **46**, 222–228 (1990).

18. Ahern, D. K., Woods, S. S., Lightowler, M. C., Finley, S. W. & Houston, T. K. Promise of and potential for patient-facing technologies to enable meaningful use. *Am. J. Prev. Med.* **40**, S162-172 (2011).

19. Orlando, L. A. *et al.* Clinical utility of a Web-enabled risk-assessment and clinical decision support program. *Genet. Med.* **18**, 1020–1028 (2016).

20. Bennett, R. L. The family medical history as a tool in preconception consultation. *J. Community Genet.* **3**, 175–183 (2012).

21. Wu, R. R. *et al.* Quality of family history collection with use of a patient facing family history assessment tool. *BMC Fam. Pract.* **15**, 31 (2014).

22. Beadles, C. A. *et al.* Providing patient education: impact on quantity and quality of family health history collection. *Fam. Cancer* **13**, 325–332 (2014).

23. Guttmacher, A. E., Porteous, M. E. & McInerney, J. D. Educating health-care professionals about genetics and genomics. *Nat. Rev. Genet.* **8**, 151- (2007).

24. Bennett, R. L. The family medical history. *Prim. Care Clin. Off. Pract.* **31**, 479–495 (2004).

25. Glasgow, R. E., Vogt, T. M. & Boles, S. M. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am. J. Public Health* **89**, 1322–1327 (1999).

26. Bryant, A. *et al.* Health Differences Among Lumbee Indians Using Public and Private Sources of Care. *J. Rural Health* **20**, 231–236 (2004).

27. Robeson County, NC. *Data USA* Available at: https://datausa.io/profile/geo/robeson-county-nc/#demographics. (Accessed: 22nd October 2018)

28. U.S. Census Bureau QuickFacts: Robeson County, North Carolina. Available at: https://www.census.gov/quickfacts/robesoncountynorthcarolina. (Accessed: 13th October 2018)

29. Stanifer, J. W. & Hall, Y. N. Are County Codes More Indicative of Kidney Health Than Genetic Codes? *Am. J. Kidney Dis.* **72**, 4–6 (2018).

30. DiClemente, R. J., Crosby, R. A. & Kegler, M. C. *Emerging Theories in Health Promotion Practice and Research: Strategies for Improving Public Health*. (John Wiley & Sons, 2002).

31. Osborn, C. Y. & Egede, L. E. Validation of an Information–Motivation–Behavioral Skills model of diabetes self-care (IMB-DSC). *Patient Educ. Couns.* **79**, 49–54 (2010).

32. *Social Psychological Foundations of Health and Illness*. (Wiley-Blackwell, 2003).

33. National Action Plan to Improve Health Literacy - health.gov. Available at: https://health.gov/communication/initiatives/health-literacy-action-plan.asp. (Accessed: 13th October 2018)

34. Jimison, H. *et al.* Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. *Evid. ReportTechnology Assess.* 1–1422 (2008).

35. Nutbeam, D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot. Int.* **15**, 259–267 (2000).

36. Low Health Literacy Skills Increase Annual Health Care Expenditures by $73 Billion. Available at: https://hpi.georgetown.edu/agingsociety/pubhtml/healthlit.html. (Accessed: 13th October 2018)

37. Vernon, J. A., Trujillo, A., Rosenbaum, S. & DeBuono, B. Low Health Literacy: Implications for National Health Policy. 18

38. NW, 1615 L. St, Washington, S. 800 & Inquiries, D. 20036 U.-419-4300 | M.-419-4349 | F.-419-4372 | M. Networked Families | Pew Research Center. (2008).

39. Utilization of Internet Technology by Low-Income Adults: The Role of Health Literacy, Health Numeracy, and Computer Assistance - Jakob D. Jensen, Andy J. King, LaShara A. Davis, Lisa M. Guntzviller, 2010. Available at: http://journals.sagepub.com/doi/full/10.1177/0898264310366161?utm_source=summon&utm_medium=discovery-provider. (Accessed: 23rd October 2018)

40. MC, G. A historical overview of health disparities and the potential of eHealth solutions. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/16403714. (Accessed: 23rd October 2018)

41. Schmeida, M. & McNeal, R. S. The telehealth divide: disparities in searching public health information online. *J. Health Care Poor Underserved* **18**, 637–647 (2007).

42. Viswanath, K. & Kreuter, M. W. Health Disparities, Communication Inequalities, and e-Health: A Commentary. *Am. J. Prev. Med.* **32**, S131–S133 (2007).

43. Bandura, A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychol. Rev.* **84**, 191–215 (1977).

44. Hecht, S. S. Tobacco Smoke Carcinogens and Lung Cancer. *JNCI J. Natl. Cancer Inst.* **91**, 1194–1210 (1999).

45. BLOT, W. Cancers of the lung and pleura. *Cancer Epidemiol. Prev.* (1996).

46. Taal, M. W. 22 - Risk Factors and Chronic Kidney Disease. 31

47. Furrow, R. E., Christiansen, F. B. & Feldman, M. W. Environment-Sensitive Epigenetics and the Heritability of Complex Diseases. *Genetics* **189**, 1377–1387 (2011).