WOMEN'S INTERAGENCY HIV STUDY
SECTION 40: GEOCODING PROTOCOL

A. BACKGROUND & STUDY PURPOSE

This study will investigate how socioeconomic context and health care coverage contribute to health outcomes among women in the Women's Interagency HIV Study (WIHS). There is mounting evidence that socioeconomic factors (e.g., poverty and income inequality) and health care coverage impact health outcomes. However, we still do not know exactly how these factors interact with each other or how they impact an individual's health. Based on what we know, we hypothesize that neighborhood poverty and income inequality worsens people's health by depriving them of important resources, such as health care, education, and affordable food and housing.

We plan to use data already collected during WIHS visits about health insurance coverage, estimated annual income, as well as contextual factors derived from census data, to look at changes in health outcomes as people change their insurance. Because women are most affected by both poverty and HIV, it is especially relevant to understand how socioeconomic context and health care coverage impacts health, as this information will inform interventions and policies to reduce HIV health disparities.

B. SCIENTIFIC AIMS

The overall aim of this study is to evaluate the relationship between virologic suppression and serious AIDS-defining events (both among HIV+ women), blood pressure control (among all women diagnosed with hypertension), and mortality (among all women) and:

- Proportion of population below the federal poverty line in each participant's census block group

- Income inequality in each participant's county

- Participant's health insurance status

C. STUDY DESIGN

All of the information needed to complete this aim is routinely obtained during the course of WIHS study visits. Study staff will record the address of each participant at even study visits starting in visit 40, geocode the address, match it to a FIPS code which contains the participant's census block group and county of residence. Standard social and medical history, physical examination, and laboratory examination during baseline and follow-up study visits routinely capture: the participant's insurance status; medical diagnoses, including hypertension; interval medical complications since the last visit; medication use, blood pressure (BP), and determination of plasma HIV RNA levels.

Each WIHS site will use ArcGIS – a geographic information system- to assign to each participant the location of where they live by assigning geospatial coordinates (e.g., longitude and latitude) to their street address using an automatic geocoding method. The coordinates of the geocoded locations will then be assigned to a census block group (CBG). A number of contextual variables that are associated with the CBG will be obtained from the most current,

available US Bureau of the Census data. Linkage of addresses to census block group will be done at the site and shared/analyzed data will only contain the contextual variables and **WILL NOT contain addresses, CBG groups or other identifiers.**

D. DATA SECURITY

1. SECURITY FOR WIHS ADDRESSES

These data will reside at your site ONLY. Each site has protocols approved by their IRB that outline procedures for proper storage of these sensitive data. At UNC, these files are stored on a dedicated server which contains extensive protections and securities. Addresses are kept separately from the de-identified data under additional password security and encryption.

2. SECURITY FOR WIHS CENSUS BLOCK GROUPS

A note about block groups: these census divisions contain between 600 and 3,000 people and encompass geographical areas that vary in size. In highly populated places (like New York City), the geographical area may be 1-3 city blocks. In more rural areas, the geographical area may be as large as 100-200 square miles.

Your file containing block group information on each WIHS participant will be uploaded using CoreFTP, the standard for uploading sensitive files at the University of North Carolina. By using the FTP for both authentication and data transfer, WIHS data will be fully encrypted. Files uploaded to this secure server will be placed in a folder that has access limited to a few study staff. These data will be moved to a secure server and password protected with access to limited study staff.

Once the data have been merged with the census data for each block group, the block group variable will be deleted from the dataset to leave a de-identified dataset.

E. GEOCODING PROTOCOL

Please visit the video tutorial for this protocol on YouTube: http://youtu.be/QQCRtBoUEj8 or on the WIHS Admin Website.

1. SOFTWARE REQUIREMENTS

- **ArcGIS for geocoding with ESRI Data and Maps 10.2 disc** (if your site does not have ArcGIS the UNC Site can provide access to a license)

- **Microsoft Excel for capturing participant addresses**
  - o **NOTE:** ArcGIS can import data from Excel but not Access directly, so if you use Access to keep track of addresses, export the table containing addresses to Excel. ArcGIS requires that variable/column names do not contain any spaces. If any of your column headings in Excel contain spaces, remove the spaces before uploading to ArcGIS. We've included an Excel addresses template if your site does not use Access and prefers to keep addresses in Excel.

- **Internet Access**

2. **COLLECTING PARTICIPANT ADDRESSES**

   1. Ask for participant contact address; collect:

      a. Address (primary residence, or where she spends the majority of her living time, please do not use PO boxes because they do not give information about where the participant lives)

      b. Any alternate names by which their street is known

      c. Nearest major intersection to their house (NOTE: this information will only be used when ArcGIS is having difficulty finding a match)

   2. Show the participant the address and intersection and ask if they are correct

   NOTE: If changing practices around address collection is too onerous for your site, it is acceptable to collect information about alternate street names and nearest major intersection only if the geocoding is not successful.

3. **GEOCODING PARTICIPANT ADDRESSES**

   1. Install ArcGIS. We will work with sites to get this software. You need a single use license for ArcInfo (also known as Desktop Advanced or Intermediate).

   2. Create a folder to contain the files used for geocoding and the geocoding output. We called ours "WIHS geocoding."

   3. Save the WIHS addresses, in Excel format, from your site into the "WIHS geocoding" folder.

   4. In your web browser, navigate to:
      http://www.arcgis.com/home/item.html?id=1c924a53319a491ab43d5cb1d55d8561

      This site has the current USA census block group map (called a shapefile) for ArcGIS.

   5. Click the "Open" button and select "Download." Next, select "Save file" and "OK." Once the file has downloaded, move the file from your download folder to the "WIHS geocoding" folder.

   6. Navigate to the "WIHS geocoding" folder, and double-click on the "USA_Block_Group_Boundaries" shapefile; this should automatically launch ArcMap. In ArcMap, click the add data button , navigate to the "WIHS geocoding" folder. You may have to click the connect to folder button if your do not see your folder among the selections. Select the address file, and select the appropriate sheet from the Excel file, and click "Add."

   7. Insert the Data and Maps 10.2 (disc 2) disc into your computer. In the left-hand Table of Contents pane, right-click on your address file. Select "Geocode Addresses" from the drop-down menu. Click "Add" in the upper right corner; and navigate to the location of the disc. You may have to click the connect to folder button if your do not see the disc among the selections. Select "streetmap_NA," then "data," then

"Street_Addresses_US." Click "Add." Back in the "Choose an address locator" pane, make sure the "Street_Addresses_US" option is selected, and click "OK."

8. A window will open asking which variable names from your Excel file correspond to which input fields. The address, city, state, and zip need to have corresponding variables from the Excel file (Zip4 and country are not necessary). Check for correctness and make necessary changes.

9. In the bottom half of this pane is the dialog box that controls the output. ArcMap automatically places files into a default geodatabase inside the ArcGIS folder on your computer. From this point on, ArcMap should automatically save output in the location you choose here.

    a. If you do not want to change the default destination, just click "OK."

    b. If you would like to change the destination of your file (to point to your server for example), you first need to click on the browse file button on the right of the output shapefile or feature class

    Output shapefile or feature class:
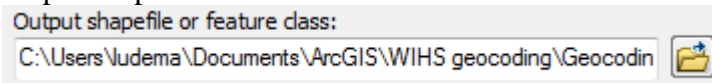
    C:\Users\ludema\Documents\ArcGIS\WIHS geocoding\Geocodin 🗁 .

    Navigate to your "WIHS geocoding folder." Click "Save" and then click "OK."

10. A window will appear that reports the match rate of your addresses. Click "Rematch" and a rematching pane will appear after a few second delay (see below for example). The address that you entered will be on the left side of the pane and the candidate match information will be on the right side (under "Candidate details").
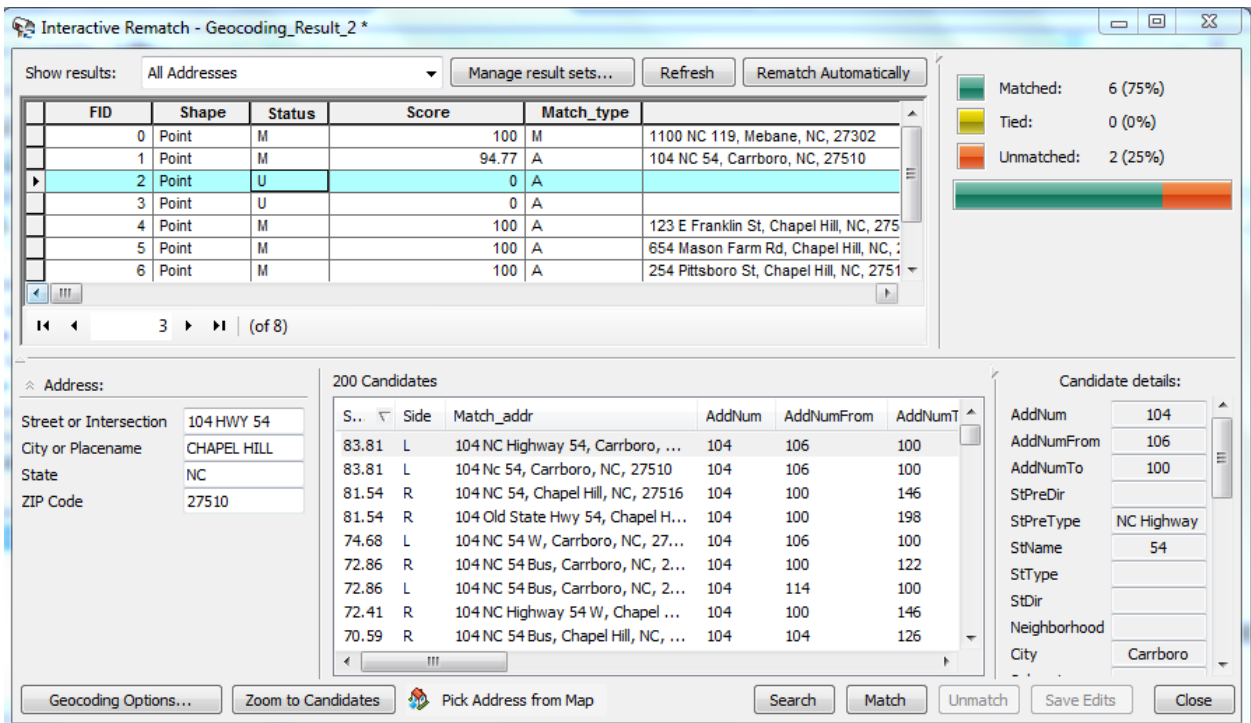
11. To ensure data quality, please follow these steps:

    a. If you have any matches with a match "Score" less than 100, manually go through the address(es) to see if see if there are misspellings or improper formatting. You can correct the data in the bottom left of the Interactive Rematch pane where it lists "Address:" or you can correct the data in your Excel spreadsheet (if ArcMap is still open while you make these changes, you will have to save a copy of the Excel file) and start the geocoding process again with the new file from step 5.

    b. If you have any ties, indicated by a "T" in the "Status" column, select that address and choose the address in the middle pane that most closely matches the contact information, and click "Match." Once you do this, the "Status" should change to "M" and the "Match_type" should change to "M" (for a manual match).

    c. If after these steps you still have any addresses with a "Status" of "U" (unmatched), use the nearest intersection data. To do this, open the address Excel file and copy the "Nearest_Intersection" into the "Address_Street_number_and_name" column ONLY for the problematic address. Start the geocoding process over from step 5 with the new file. When the below pane comes up, choose "Geocoding Options" and enter the connector between the two street names into the "Connectors" field under "Intersections" (for example, if your intersections are formatted "First Ave and Second St," then you'd enter "and"; for "First Ave & Second St," you'd enter "&"; if you had addresses formatted multiple ways, you'd enter all connectors with a space between each one, e.g., "and &").
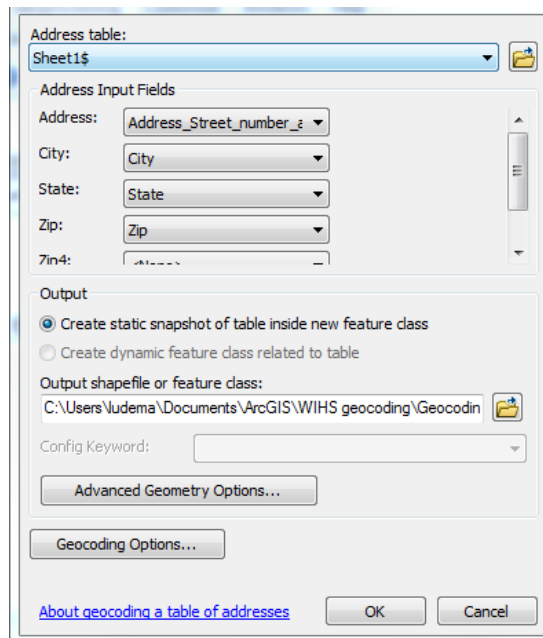
    d. If any addresses remain unmatched, leave them in the Excel document (do not delete any entries), and close the "Interactive Rematch" pane.

12. Right-click on "Geocoding Result" in the Table of Contents pane (make sure this is the most recent geocoding result if you had to rematch in Step 11) and select "Join

and Relates," and then "Join." Select the appropriate options depicted below, and "OK."
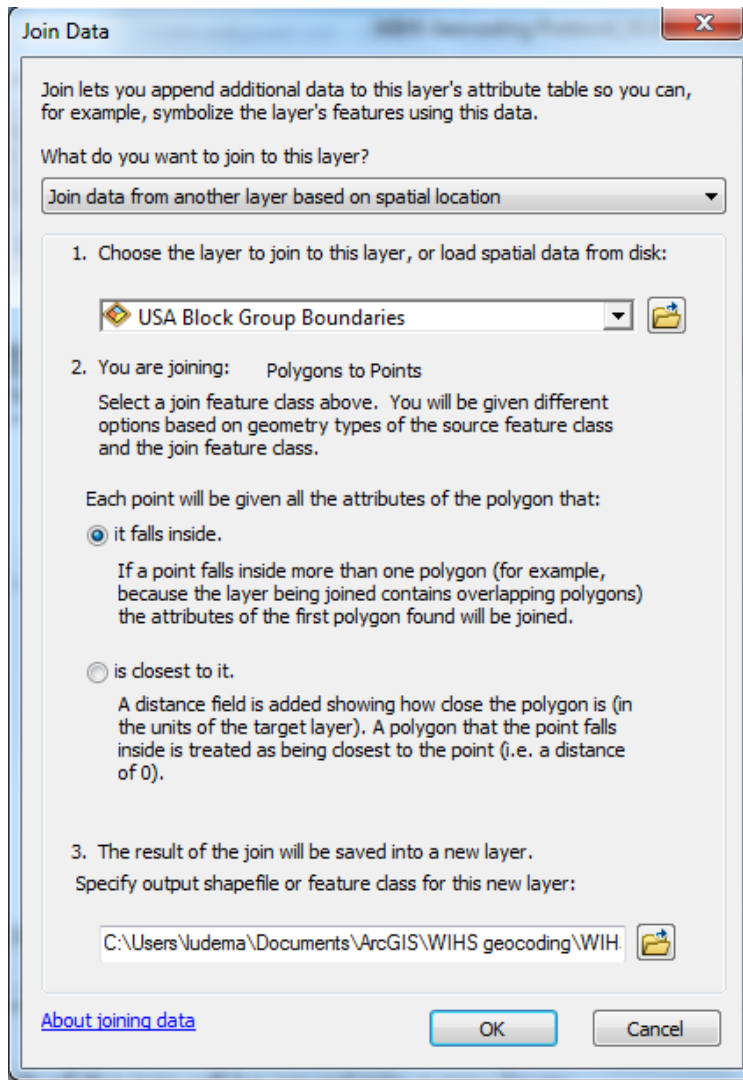
13. Right-click on the joined file (Join_Output) in the Table of Contents pane and select "Open Attribute Table."

14. Select the Table Options icon in the upper left ⬚ ▾, and go down to the Export option.

15. An Export Data window will come up. Choose "Export All Records," and then select the browse icon 📂 to choose the export file type. Navigate to where you would like the file to be stored. From the drop-down Save As menu, choose "dBASE Table." Name the file "Geocoding with IDs." Your file name should have a .dbf file extension. Click "Save" and "OK." If it asks if you want to add the new table to the map, select "No."
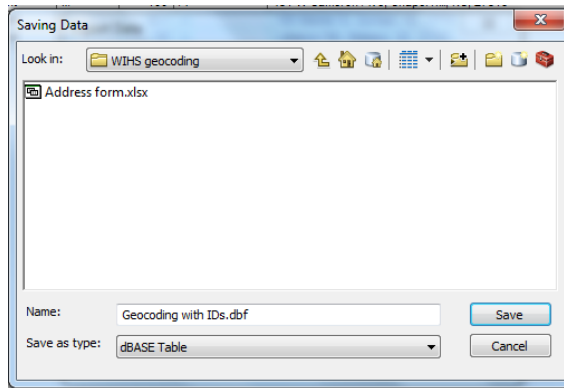
16. Exit the Attribute Table and ArcMap. There is no need to save the project.

17. Open Excel and choose the "Open File" option [Open icon] under the File menu. Navigate to the folder with the "Geocoding with IDs" file. Select "All Files" in the drop down file type menu next to the File name (see below). Select your "Geocoding with IDs.dbf" file. Select "Open."
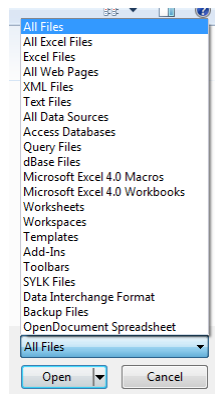
18. To maintain privacy in the file you will be sharing, delete all variables except Status, Score, Match_type, WIHS ID, Visit number, and FIPS. This can be accomplished by copying and pasting into a new spreadsheet the 7 variables mentioned above. Choose "Save As" from the File menu and save this file as "Block Group Only-(Site Name)" (where "Site Name" is your city, e.g. "Chapel Hill."). Save the file as an Excel file (.xlsx). This is the file you will be sending to UNC. Before you upload the file as described in step 15, please make sure this file does not contain any address, latitude, longitude, nearest intersection, or other specific geocoded information.

4. TRANSFERRING DATA TO UNC VIA SECURE FILE TRANSFER

1. Download and install CoreFTP (http://www.coreftp.com/).

   a. For first time use, open CoreFTP, and configure your settings as depicted below. For uploading, select the anonymous box next to the Username field. Click "Connect," and select "Accept Once" when prompted.

   b. The next time you open CoreFTP, you can connect by selecting the quick connect button [lightning icon].
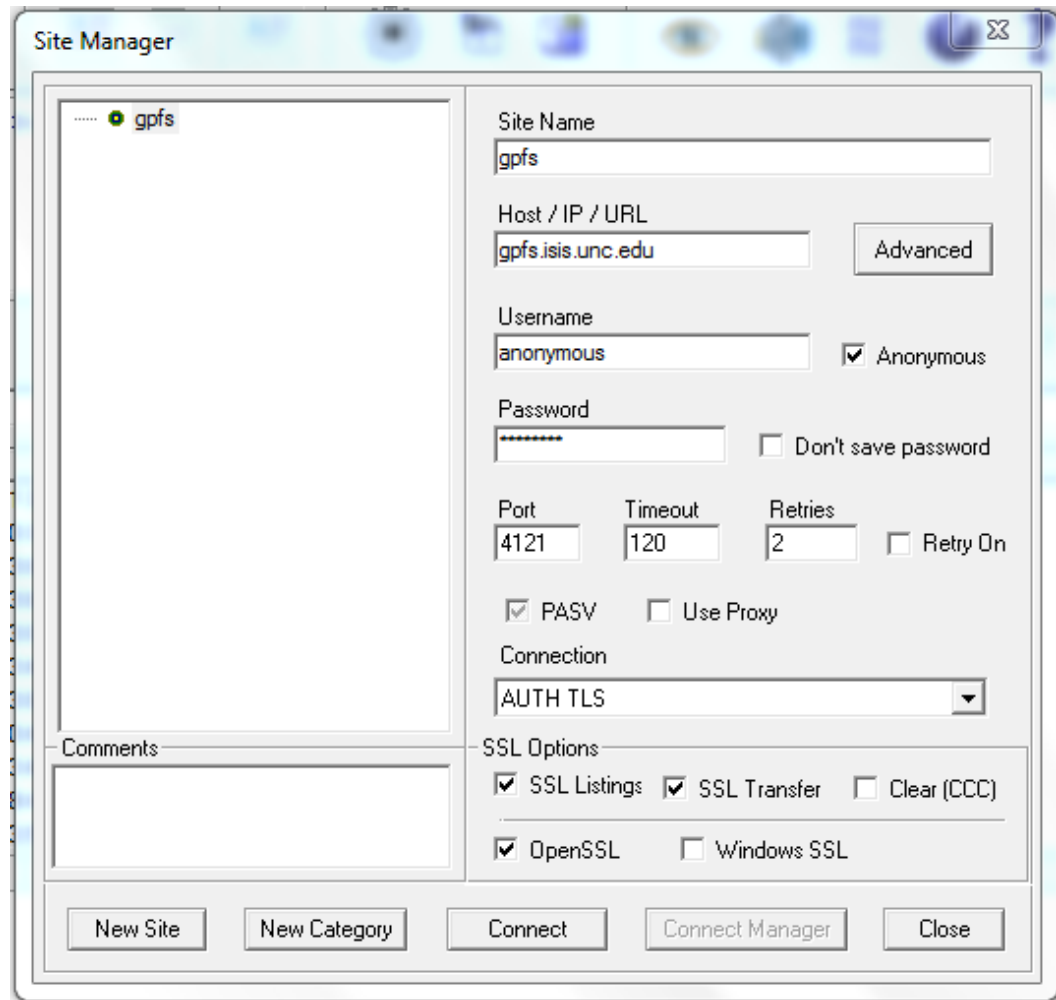
2. Double-click on the "upload" folder in the right hand panel.

3. Browse to your Block Group Only file in the left hand pane; drag and drop your Block Group file from the left hand pane to the right hand pane (i.e., "upload" folder).

4. Wait until the "Status" column at the bottom says "Success" before closing CoreFTP

5. Send an e-mail to research@unc.edu, and cc: aedmonds@email.unc.edu to notify them that you have uploaded data. Title the e-mail "file on SoNAS" and use the following as the text:

> I uploaded "Block Group Only-(Site Name).xlsx" to the SoNAS upload folder. Please deliver this to Andrew Edmonds' "WIHS" folder.
>
> Name: (Your name)
>
> Institution: (Your institution)

NOTE: Most of the files that are created in this geocoding process contain participant's addresses and specific geographic information. **Please make sure that these files are stored in a secure location as per your IRB.**

CONTACT FOR QUESTIONS ABOUT THIS PROTOCOL

Christina Ludema, PhD

130 Mason Farm Rd. CB 7030

Chapel Hill, NC 27599

(919) 597-1422

ludema@email.unc.edu

Consent to Participate in a Research Study
Adult Participants

Consent Form Version Date: Version 2.0, September 04, 2013

IRB Study:

Title of Study: The Affordable Care Act, Socioeconomic Context, Spatial Mapping, and the WIHS
Principal Investigator: Adaora Adimora, MD, MPH
Principal Investigator Department: Medicine – Infectious Diseases
Principal Investigator Phone number: (919) 843-0723
Principal Investigator Email Address: adimora@med.unc.edu

Study Contact: Catalina Ramirez, MPH
Study Contact telephone number: 919-966-0082
Study Contact email: catalina_ramirez@med.unc.edu

Funding Source and/or Sponsor: National Institutes Of Health (NIH)
_____

What are some general things you should know about research studies?
You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty. Research studies are designed to obtain new knowledge. This new information may help people in the future. You may not receive any direct benefit from being in the research study. There also may be risks to being in research studies.

Deciding not to be in the study or leaving the study before it is done will not affect your relationship with the researcher or your health care provider. If you are a patient with an illness, you do not have to be in the research study in order to receive health care.

Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study. You will be given a copy of this consent form. You should ask the researchers named above, or staff members who may assist them, any questions you have about this study at any time.

What is the purpose of this study?
There is mounting evidence that socioeconomic factors (e.g., poverty and income inequality) and health care coverage impact health outcomes. However, we still do not know exactly how these factors interact with each other or how they impact an individual's health. Based on what we know, we hypothesize that neighborhood poverty and income inequality worsens people's health

by depriving them of important resources, such as health care, education, and affordable food and housing.

Dr. Adimora and colleagues at the WIHS site at University of North Carolina-Chapel Hill would like to use data already collected during WIHS visits about health insurance coverage and estimated annual income, and census data to look at changes in health outcomes as the ACA goes into effect across the United States. Because women are most affected by both poverty and HIV, it is especially relevant to understand how socioeconomic context and health care coverage impacts health, as this information will inform interventions and policies to reduce HIV health disparities.

You are being asked to participate in this study because you are a current WIHS participant.

How many people will take part in this sub-study?
If you decide to be in this sub-study, you will be one of over 4,000 women who will participate in this study across all nine WIHS sites.

How long will your part in this sub-study last?
If you choose to participate, you may be asked to provide additional information about the location of your place of residence. Your participation will take a few minutes.

What will happen if you take part in the study?
We are asking you to give WIHS authorization to use your address so that we can examine how you and your neighborhood influence your health and wellness. A member of the WIHS staff will match your address to a geographic area called a census block group (CBG)- which is large enough so that your exact address cannot be determined. CBG's are used by the US Census to report on demographic characteristics of entire communities. The community demographic data – and not your address- will be used by researchers to help them better understand how neighborhood and community factors relate to health and wellness. **Your actual address will not be reported, published or shared with anyone else.**

What are the possible benefits from being in this study?
You will not receive any direct benefit from participation in this study. Research is designed to benefit society by gaining new knowledge and may provide researchers with additional information that will inform interventions and policies to reduce HIV health disparities.

What are the possible risks or discomforts involved from being in this study?
There are no additional risks beyond those associated with participating in the WIHS cohort.

What if we learn about new findings or information during the study?
You will be given any new information gained during the course of the study that might affect your willingness to continue your participation.

How will information about you be protected?
Records of your participation in the study will be kept as confidential as possible. Your identity will not be used in any reports or publications generated by the study. A member of the WIHS staff will

match your address to the matching census block group (CBG). Your CBG and your WIHSID will be entered into a password-secured spreadsheet on a secured computer, which is encrypted and kept in locked offices at our site. We will not use any records that contain your name, address or any other potentially identifiable information.

Data from our site will be transmitted to the UNC-WIHS site via a secure website and contains no identifying information; it can be accessed only with private passwords.

What if you want to stop before your part in the study is complete?
You can withdraw from this study at any time, without penalty.

Will you receive anything for being in this study?
You will not receive any payment or financial benefit from participating in this study.

Will it cost you anything to be in this study?
There are no costs to you as a result of taking part in this study.

Who is sponsoring this study?
This research is funded by the National Institutes of Health (NIH). This means that the research team is being paid by the sponsor for doing the study. The researchers do not, however, have a direct financial interest with the sponsor or in the final results of the study.

What if you have questions about this study?
You have the right to ask, and have answered, any questions you may have about this research. If you have questions about the study (including payments), complaints, concerns, or if a research-related injury occurs, you should contact the researchers listed on the first page of this form.

What if you have questions about your rights as a research participant?
All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject, or if you would like to obtain information or offer input, you may contact the UNC-CH Institutional Review Board at 919-966-3113 or by email to IRB_subjects@unc.edu.

Title of Study: The Affordable Care Act, Socioeconomic Context, Spatial Mapping, and the WIHS

Principal Investigator: Adaora A. Adimora, MD, MPH

Participant's Agreement:

I have read this consent or have had it read to me. I have had a chance to ask any questions, and these questions have been answered to my satisfaction. I freely and voluntarily agree to participate in this research study.

_____ _____
Signature of Research Participant   Date

_____
Printed Name of Research Participant

_____ _____
Signature of Research Team Member Obtaining Consent   Date

_____
Printed Name of Research Team Member Obtaining Consent