Linking the National Health Interview Survey (NHIS) and the Florida Cancer Data System (FCDS)

Laura A. McClure\(^1\), Monique Hernandez\(^2\), Jill A. Mackinnon\(^2\), Brad Wohler\(^2\), Donna M. Miller\(^3\), Youjie Huang\(^4\), Tara Hylton\(^4\), Recinda Sherman\(^2\), Cristina A. Fernandez\(^5\), Lora E. Fleming\(^2,5\), David J. Lee\(^2,5\)

\(^1\)Sylvester Comprehensive Cancer Center, University of Miami, Miami, FL; \(^2\)Florida Cancer Data System (FCDS), Sylvester Comprehensive Cancer Center, University of Miami, Miami, FL; \(^3\)Special Projects Branch, National Center for Health Statistics (NCHS), Hyattsville, MD; \(^4\)Chronic Disease Epidemiologic Group, Florida Dept of Health, Tallahassee, FL; \(^5\)Department of Epidemiology and Public Health, University of Miami, Miami, FL; \(^6\)European Centre for Environment and Human Health (ECEHH), Peninsula College of Medicine and Dentistry, Truro, Cornwall, UK

### Purpose
- Assess the feasibility and logistics of linking population-based survey data from the National Center for Health Statistics (NCHS) with the CDC National Program of Cancer Registries (NPCR) and Surveillance, Epidemiology and End Results (SEER) Cancer Registries.
- Develop a model and documentation for conducting such linkages.
- Enrich cancer registry data with sociodemographics, health conditions, health status, cancer screening, cancer history, and cancer risk factors from the NHIS and its linked data sources.

### Methods & Preliminary Results

#### Matched based on:
- Social Security Number
- Name
- Date of Birth
- Sex
- City of Residence

- **1986, 1988-2009 National Health Interview Survey**
  - N=1,914,210 Eligible Records
  - 26,177 False Matches
  - 38,946 Potential Matches
  - 9,096 True Matches
  - 3,772 Questionable Matches
  - 2,462 False Matches
  - 10,406 Total Matches

- **1981-2010 Florida Cancer Data System**
  - N=2,520,333 Records
  - 3,772 Questionable Matches
  - 2,462 False Matches
  - 1,310 True Matches

### Proposed Analyses
- After adjusting for the complex survey design of the NHIS and recalculating survey weights accounting for linkage-ineligibility, use descriptive analyses and Structural Equation Modeling to evaluate:
  - Cancer screening compliance & cancer diagnosis.
  - Socioeconomic factors & late stage cancer diagnosis.
  - Activity limitations after cancer diagnosis.

### Collaboration
- Ultimate goal: Recruit other central cancer registries into a Linkage Consortium to conduct a linkage with the NHIS and all NPCR and SEER registries.
- Establish a Consortium Advisory Board and Consortium website to provide members with standardized information and protocols, aid in linkage-related activities, disseminate research findings, and recruit other interested registries.
- Collaboration will result in an unparalleled data resource for evaluating cancer risk factors, screening behaviors, health conditions, and healthcare access and utilization in a large sample of cancer patients.

### Background
Previously, our study team completed a trial linkage between data from the 1987 National Health Interview Survey (NHIS) and the Florida Cancer Data System (FCDS) database. This linkage resulted in 712 linked Florida cancer cases that participated in the 1987 NHIS. This linked dataset was deposited at the secure NCHS Research Data Center (RDC). Using this dataset, we are currently conducting analyses to look at the impact of cancer screening and insurance status on stage at diagnosis, among other analyses.

### Conclusions
The linkage resulted in a dataset of matched cases that represent NHIS participants diagnosed with cancer in Florida prior to or subsequent to their NHIS interview. The results from both the one-year linkage and the full linkage indicate this is a feasible and worthwhile research endeavor to enrich cancer registry data. The de-identified, linked dataset will be deposited at the secure NCHS Research Data Center (RDC) and can be analyzed by approved researchers. We hope to engage with other interested registries to soon begin development of the Linkage Consortium.