

# **ISGC/SiGN – REGARDS Risk Factor Study**

**George Howard, DrPH**

**(a stranger in a strange land)**

# Outline

- 1. The general idea**
- 2. Overview of REGARDS (you already know about ISGC/SiGN)**
- 3. Analytic approaches**
- 4. Discussion of potential for expansion and logistics**

# General Idea

- **Studies**

- **SiGN/ISGC has**

- **Compiled a remarkable cohort of stroke cases**
    - **Already phenotyped major stroke risk factors (hypertension, diabetes, AFIB, CAD, smoking)**
    - **Likely has (indirect) access to medical records to expand phenotyping**

- **REGARDS has**

- **Evaluated (phenotype)  $\approx$  28,000 stroke-free, black and white, participants from 1,855 counties across the US**

# General Idea

- **Together**
  - We can do a case-control study of massive size to assess risk factors for stroke
  - This large size sample size will allow
    - Remarkable precision to detect small effects for risk factors
    - Assessment of risk factors in small subgroups (young stroke, stroke in blacks, etc.), or assessment of effect modification between strata
    - Assessment of risk factors for stroke subtypes
- **Examples of papers are really pretty obvious and numerous**

# Outline

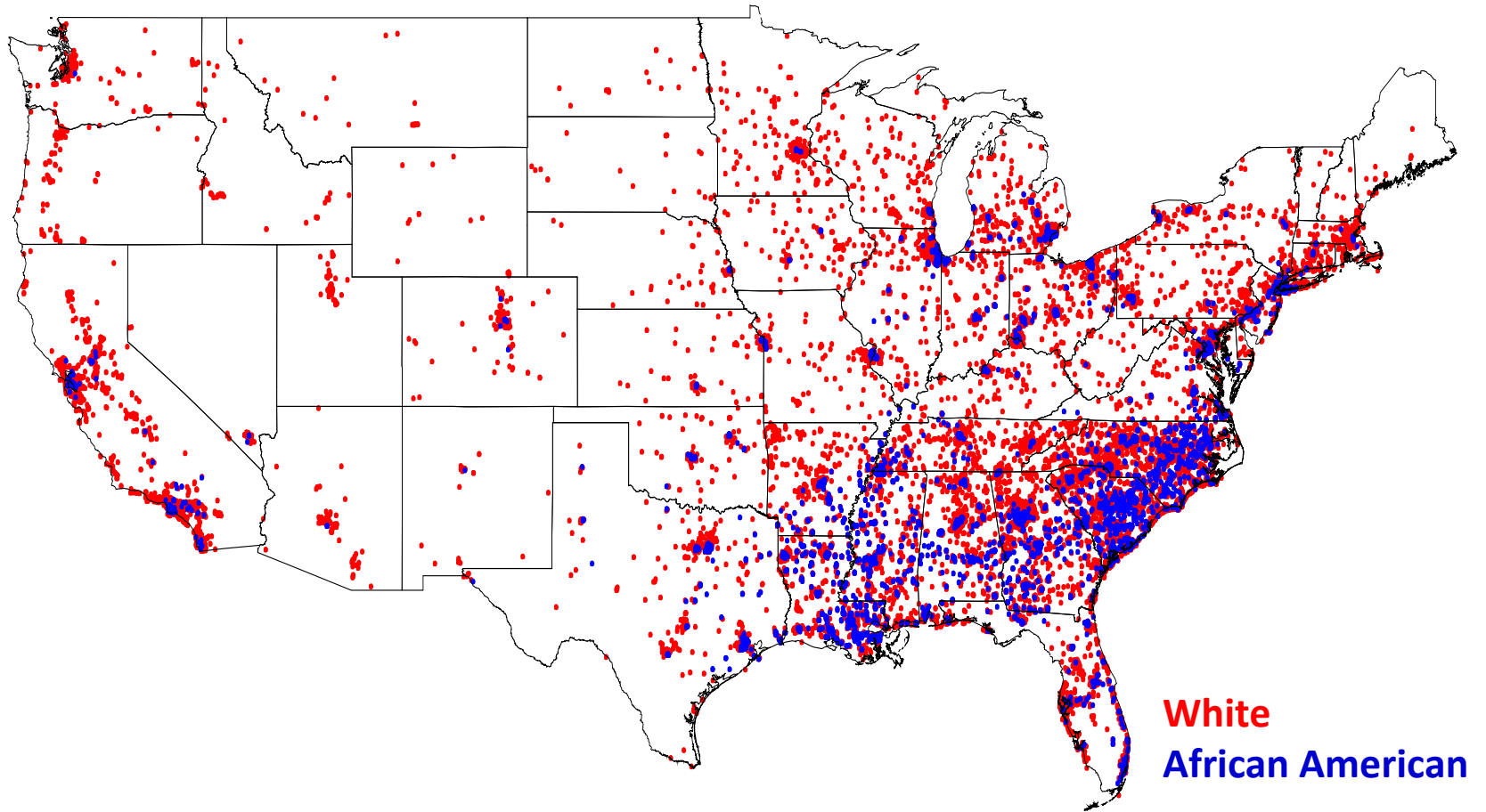
1. The general idea
2. Overview of REGARDS (you already know about ISGC/SiGN)
3. Analytic approaches
4. Discussion of potential for expansion and logistics

# REasons for Geographic And Racial Differences in Stroke (REGARDS)

## *Study Population*

- Longitudinal population-based cohort of 30,239 volunteers age 45 and older
- Simple random sampling with geographic representation
  - 21% from the buckle of the stroke belt (goal was 20%)
  - 35% from the stroke belt (goal was 30%)
  - 44% from the rest of the contiguous US (goal 50%)
- Racial representation: 42% African American / 58% white (goal was 50/50%)
- Sex representation: 45% male / 55% female (goal was 50/50%)

# REGARDS Participants



**N = 30,239**

# Types of Data Available in REGARDS

- **Examples of “Expected” data**
  - Questionnaire data on major CVD risk factors, HRQoL, depression, stress, etc.
  - Lab data: lipids, glucose, insulin, CRP, creatinine, albumin, cystatin C, etc.
  - ECG for MI, atrial fibrillation, LVH
  - Self administered data for diet, family history
  - Follow-up data for incident stroke/MI events
- **Examples of “Novel” data**
  - Remarkable geographic heterogeneity
  - Longitudinal cognitive assessments
  - Residential history
  - Medication inventory
  - Links to satellite data on pollution weather
  - Direct measures of physical activity
  - Childhood assessment of SES
- **Examples of data not present**
  - Some interview data
  - Anything requiring complex equipment (carotid ultrasound, coronary calcium, etc)

# Outline

1. The general idea
2. Overview of REGARDS (you already know about ISGC/SiGN)
3. **Analytic approaches**
4. Discussion of potential for expansion and logistics

# ISGC Stroke Cases (US Only)

Study ID	ISGC Cases
GASROS- Boston area	686
GCNKSS- Cincinnati-Kentucky	642
GEOS- Balt-Washington	891
ISGS -5 medical ctrs	675
MCISS- New Jersey	876
MIAMISR- Miami	331
NHS	470
NOMAS- Manhattan	578
REGARDS	555
SWISS	407
SPS3	30
WHI	840
WUSTL- Washington Univ	643
<i>Subtotal (US)</i>	<i>7,594*</i>

# Combined Studies

(VERY tight geographical matching)

Study ID	ISGC Cases	REGARDS
GASROS- Boston area	686	149
GCNKSS- Cincinnati-Kentucky	642	194
GEOS- Balt-Washington	891	252 + 527
ISGS -5 medical ctrs	675	?
MCISS- New Jersey	876	720
MIAMISR- Miami	331	258
NHS	470	?
NOMAS- Manhattan	578	720
REGARDS	555	?
SWISS	407	?
SPS3	30	?
WHI	840	?
WUSTL- Washington Univ	643	315
<i>Subtotal (US)</i>	<i>7,594*</i>	<i>3,135 (already)</i>

# Analytic Approaches

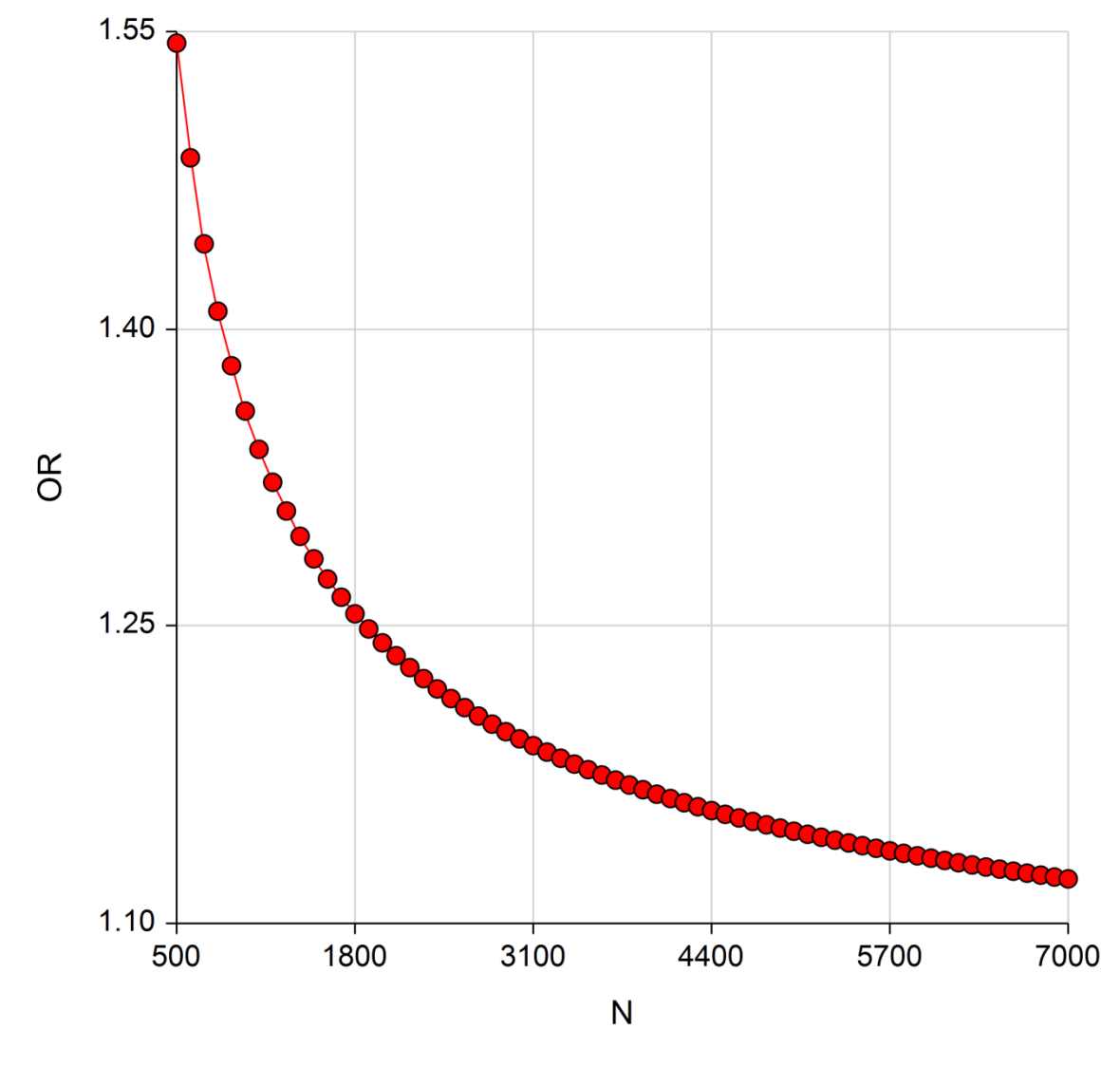
- Ordinary, straightforward, case-control methods
- Unlike most case/control studies, we don't have to do a single comparison cohort
  - If tight geographical matching is desired
    - Matching is reasonable, but not great
    - Do unmatched, covariate adjusted analysis
  - If tight geographic matching is not needed
    - Matched analysis possible

# Limitations

- **Clearly, the big limitation is that risk factors are not evaluated the same in the two studies**
- **Matching may be harder than it appears on the surface**
- **(At least I would suggest that) analysis needs to be limited to whites and blacks, aged 45+ in the US**

# Detectable Odds Ratio if We Do Match

*(90% power; Phi of 0.1)*



# Outline

1. The general idea
2. Overview of REGARDS (you already know about ISGC/SiGN)
3. Analytic approaches
4. Discussion of potential for expansion and logistics

# Opportunities for Expansion

## *(Expanding the Exposures)*

- Current list of risk factors is somewhat limited
- Is it possible to go back to charts?
  - Broaden available “traditional” risk factors
- Can we geocode home locations of stroke patients? IRB issues?
  - Neighborhood exposures
- Can we contact patients or family members for additional information?
  - Non-traditional risk factors
- Other ideas?

# Opportunities for Expansion

- This morning, Brad Worrall spoke of many additional stroke patients not in previous data.
  - Individuals without genetic material
  - Individuals with unusable samples
  - Etc.
- Could these be added?

# Structure

- **We (REGARDS) are flexible with respect to structure**
- **We would propose**
  - **A small working group (2-3 from each study, couple of MS statistician/programmers) from both studies submit the grant**
  - **Those working on the grant would have “first call” on papers.**
  - **It would then be opened to all investigators from both studies**
  - **It would then be opened to other scientists (not associated with either study (with involvement of those from both studies))**

# Timeline?

- **Three year grant**
  - About 9 months to get “expansion data” in place
  - About 9 months for investigators working on the grant to submit papers
  - About 9 months other ISGC/SiGN and REGARDS investigators to submit papers
  - About 9 months for supported open use of data
- **October submission?**



# REGARDS Participants in Non-Stroke Belt CBSAs

CBSA Name	Frequency	CBSA Name	Frequency
Los Angeles-Long Beach-Anaheim, CA	1038	San Diego-Carlsbad, CA	126
New York-Newark-Jersey City, NY-NJ-PA	720	Dayton, OH	118
Chicago-Naperville-Elgin, IL-IN-WI	706	Richmond, VA	113
Detroit-Warren-Dearborn, MI	660	Louisville/Jefferson County, KY-IN	106
Washington-Arlington-Alexandria, DC-VA-MD-WV	527	Indianapolis-Carmel-Anderson, IN	105
San Francisco-Oakland-Hayward, CA	463	Jacksonville, FL	104
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	388	Sacramento--Roseville--Arden-Arcade, CA	92
St. Louis, MO-IL	315	Riverside-San Bernardino-Ontario, CA	89
Cleveland-Elyria, OH	308	Orlando-Kissimmee-Sanford, FL	88
Houston-The Woodlands-Sugar Land, TX	260	Buffalo-Cheektowaga-Niagara Falls, NY	85
Miami-Fort Lauderdale-West Palm Beach, FL	258	Akron, OH	74
Dallas-Fort Worth-Arlington, TX	255	Denver-Aurora-Lakewood, CO	74
Baltimore-Columbia-Towson, MD	252	Oklahoma City, OK	68
Cincinnati, OH-KY-IN	194	Rochester, NY	62
Milwaukee-Waukesha-West Allis, WI	181	Seattle-Tacoma-Bellevue, WA	62
Boston-Cambridge-Newton, MA-NH	149	San Jose-Sunnyvale-Santa Clara, CA	59
Tampa-St. Petersburg-Clearwater, FL	140	Phoenix-Mesa-Scottsdale, AZ	57
Virginia Beach-Norfolk-Newport News, VA-NC	136	Tallahassee, FL	56
Columbus, OH	133	Hartford-West Hartford-East Hartford, CT	55
Minneapolis-St. Paul-Bloomington, MN-WI	131	Toledo, OH	54
Pittsburgh, PA	131	Pensacola-Ferry Pass-Brent, FL	53
Kansas City, MO-KS	129	Tulsa, OK	50

# REGARDS Participants in Stroke Belt or Buckle CBSAs

CBSA Name	Frequency
Atlanta-Sandy Springs-Roswell, GA	981
Columbia, SC	696
Memphis, TN-MS-AR	494
Birmingham-Hoover, AL	464
Charleston-North Charleston, SC	464
New Orleans-Metairie, LA	459
Jackson, MS	416
Baton Rouge, LA	396
Charlotte-Concord-Gastonia, NC-SC	385
Augusta-Richmond County, GA-SC	345
Columbus, GA-AL	254
Myrtle Beach-Conway-North Myrtle Beach, SC-NC	239
Raleigh, NC	231
Greensboro-High Point, NC	230
Shreveport-Bossier City, LA	228
Little Rock-North Little Rock-Conway, AR	222
Macon, GA	217
Savannah, GA	207
Fayetteville, NC	205
Nashville-Davidson--Murfreesboro--Franklin, TN	200
Orangeburg, SC	200
Wilmington, NC	200
Florence, SC	182
Greenville-Anderson-Mauldin, SC	180
Mobile, AL	176
Huntsville, AL	159
Albany, GA	156
Hilton Head Island-Bluffton-Beaufort, SC	154
Winston-Salem, NC	153
Montgomery, AL	152
Lafayette, LA	151
Knoxville, TN	143
Rocky Mount, NC	140

CBSA Name	Frequency
Durham-Chapel Hill, NC	137
Goldsboro, NC	126
Lake Charles, LA	123
Chattanooga, TN-GA	120
Sumter, SC	111
Asheville, NC	104
Greenville, NC	97
Gulfport-Biloxi-Pascagoula, MS	94
New Bern, NC	94
Monroe, LA	93
Warner Robins, GA	86
Pine Bluff, AR	83
Tuscaloosa, AL	81
Spartanburg, SC	80
Georgetown, SC	76
Kinston, NC	69
Jacksonville, NC	68
Alexandria, LA	65
Fayetteville-Springdale-Rogers, AR-MO	64
Wilson, NC	64
Morehead City, NC	63
Anniston-Oxford-Jacksonville, AL	60
Valdosta, GA	60
Bennettsville, SC	59
Roanoke Rapids, NC	56
Brunswick, GA	55
Hammond, LA	52
Hickory-Lenoir-Morganton, NC	52
Meridian, MS	51
Opelousas, LA	51