

NAACCR 2006 Conference
NAACCR SHOWCASE
Great Circle Distance Calculator

Chris Johnson
NAACCR GIS Committee
Cancer Data Registry of Idaho



Great Circle Distance Calculator

- What is it?
- How do I use it?
- What can I do with it?

(in 10 minutes or less)



NAACCR

Great Circle Distance

What is it?

“The great circle distance is the shortest distance between any two points on the surface of the Earth measured along a path on the surface of the Earth.”

“The distances are ‘as the crow flies’ if a crow could fly at sea level.”

“The shape of the Earth closely resembles a flattened spheroid with extreme values for the radius of curvature of 6,336 km at the equator and 6,399 km at the poles.”

“Using a sphere with a radius of 6,367 km results in an error of up to about 0.5%.”

■ Huh?

The logo for NAACCR (North American Association of Certified Crime Reporters) features the acronym in a bold, blue, serif font. A red swoosh underline is positioned beneath the letters 'A', 'C', 'C', and 'R'.

Great Circle Distance

What is it?



Great Circle Distance Calculator

What is it?

- A SAS program that calculates the great circle distance between the locations of cases at the time of diagnosis and the locations of treatment facilities.
- Case locations are taken from NAACCR items 2352 (latitude) and 2354 (longitude) in a NAACCR v10 or v11 record layout file.
 - ◆ The program can use either source (unconsolidated) or consolidated case records as input.
- A second input file (that you provide) contains facility IDs, latitude, and longitude.

Great Circle Distance Calculator

How do I use it?

GIS Committee - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.naacr.org/gis

Getting Started Latest Headlines

NAACCR
The North American Association of Central Cancer Registries

Site Map Questions Contact NAACCR Request Publications

Search

Home Page

About NAACCR

- About NAACCR, Inc.
- Annual Conference
- Annual Conference Scholarship Program
- Board
- Committees
- Executive Office
- Gifts, Donations & Bequest Program
- Goals and Objectives
- Governance
- Institutional Review Board (IRB)
- Memberships
- Photo Album
- Research Program
- Sponsoring Organizations
- Town Meetings

NAACCR Membership

Registration Standards

Education & Training

Registry Certification

Statistics & Reports

Newsletters & Calendar

CINA + Online

Geographic Information Systems Committee

The GIS *Ad Hoc* Committee was formed to address the appropriate uses of geographic information systems (GIS) in cancer registry practice. The committee will work towards fulfilling the recommendations and encouraging GIS education and research in the areas outlined on pages 53-55 in the October 2002 NAACCR publication, "Using Geographic Information Systems Technology in the Collection, Analysis and Presentation of Cancer Registry Data: A Handbook of Basic Practices".

This committee's web page can also be accessed using the address: <http://www.naacr.org/GIS>

- [GIS Committee 2004-05 Objectives](#) (PDF)
- [GIS Committee 2005-06 Objectives](#) (Word)
- [GIS Committee Roster](#) (Word)

Publications

- [Using Geographic Information Systems Technology in the Collection, Analysis, and Presentation of Cancer Registry Data: A Handbook of Basic Practices](#) (PDF)
- [Review of Cluster Analysis Software](#) (PDF)

Education

- [Learn about GIS-related web sites from this descriptive list](#) (PDF)
- [Learn about Geocoding](#)
- [Learn about Cartography](#)
- [Learn about Map Design](#)

GIS Survey

- This survey is meant to assess the GIS capabilities and training needs of the NAACCR membership registries. The survey results will be used to identify areas where GIS tools or resources may be needed and encourage NAACCR to promote development of education and training activities, and resources that could be pooled or shared by its member registries.
- [GIS Survey](#) (Word)

Distance Calculator

- This SAS code calculates the great circle distance between the locations of cases at the time of diagnosis and the locations of treatment facilities. Case locations are taken from NAACCR items 2352 (latitude) and 2354 (longitude) in a NAACCR v10 or v11 record layout file. The program can use either source (unconsolidated) or consolidated case records as input. A second input file contains facility IDs, latitude, and longitude.
- [Distance Calculator SAS code](#) (Text)
- [Distance Calculator documentation](#) (Word)
- [Distance Calculator - How to add coordinates to an GIS attribute table](#) (PDF)

http://www.naacr.org/filesystem/pdf/GIS%20handbook%206-3-03.pdf

Great Circle Distance Calculator

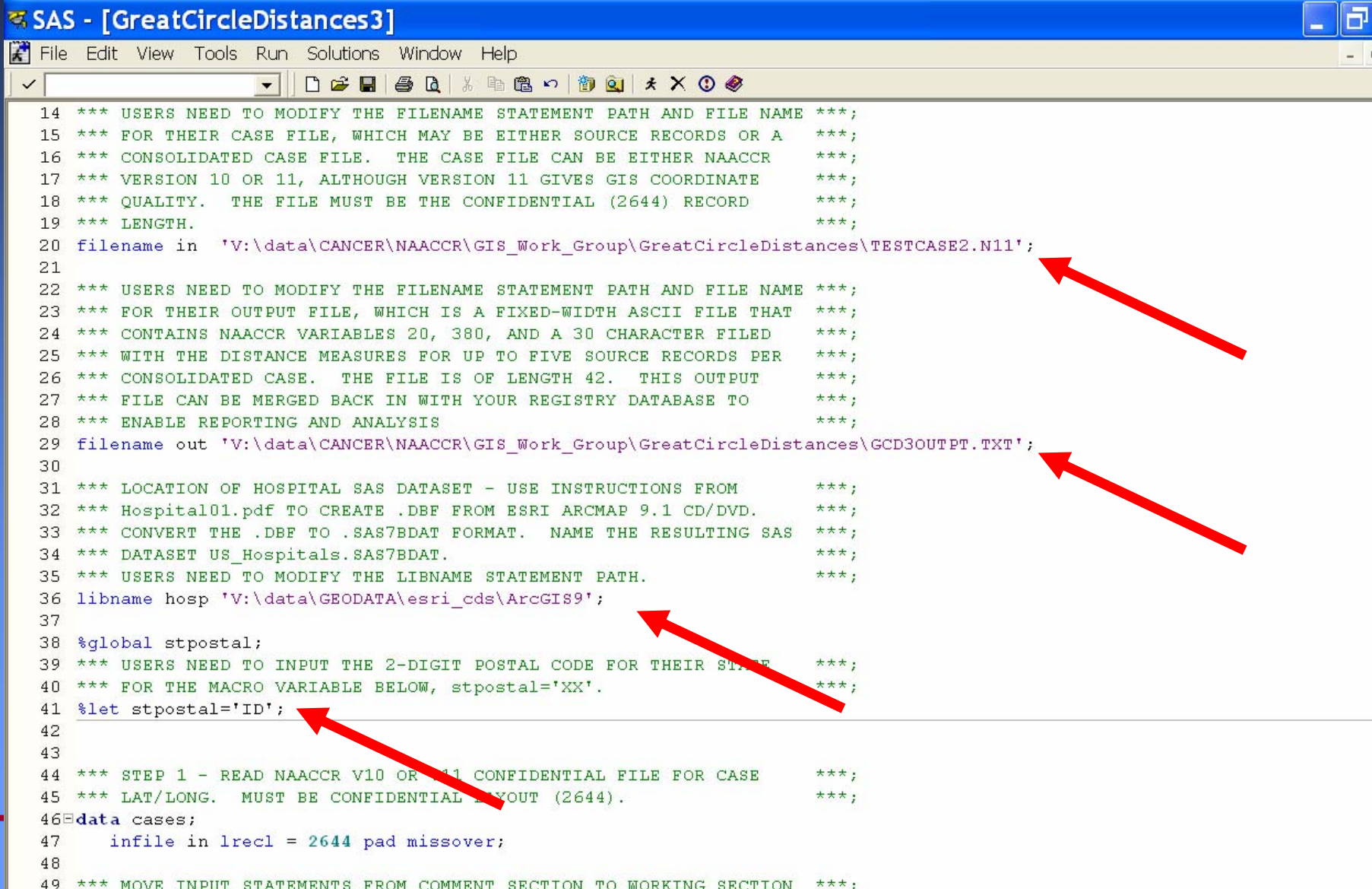
How do I use it?

1. Assumes you have geocoded your cancer cases
 - Residential address at diagnosis
2. Read the instructions on the NAACCR website (www.naaccr.org/GIS)
3. Export NAACCR V10 or V11 Confidential layout (with case latitude and longitude)
4. Create file of Reporting Hospital (NAACCR item 540) locations
5. Run the GreatCircleDistances2.sas program

Great Circle Distance Calculator

How do I use it?

```
SAS - [GreatCircleDistances3]
File Edit View Tools Run Solutions Window Help
14 *** USERS NEED TO MODIFY THE FILENAME STATEMENT PATH AND FILE NAME ***;
15 *** FOR THEIR CASE FILE, WHICH MAY BE EITHER SOURCE RECORDS OR A ***;
16 *** CONSOLIDATED CASE FILE. THE CASE FILE CAN BE EITHER NAACCR ***;
17 *** VERSION 10 OR 11, ALTHOUGH VERSION 11 GIVES GIS COORDINATE ***;
18 *** QUALITY. THE FILE MUST BE THE CONFIDENTIAL (2644) RECORD ***;
19 *** LENGTH. ***;
20 filename in 'V:\data\CANCER\NAACCR\GIS_Work_Group\GreatCircleDistances\TESTCASE2.N11';
21
22 *** USERS NEED TO MODIFY THE FILENAME STATEMENT PATH AND FILE NAME ***;
23 *** FOR THEIR OUTPUT FILE, WHICH IS A FIXED-WIDTH ASCII FILE THAT ***;
24 *** CONTAINS NAACCR VARIABLES 20, 380, AND A 30 CHARACTER FILED ***;
25 *** WITH THE DISTANCE MEASURES FOR UP TO FIVE SOURCE RECORDS PER ***;
26 *** CONSOLIDATED CASE. THE FILE IS OF LENGTH 42. THIS OUTPUT ***;
27 *** FILE CAN BE MERGED BACK IN WITH YOUR REGISTRY DATABASE TO ***;
28 *** ENABLE REPORTING AND ANALYSIS ***;
29 filename out 'V:\data\CANCER\NAACCR\GIS_Work_Group\GreatCircleDistances\GCD3OUTPT.TXT';
30
31 *** LOCATION OF HOSPITAL SAS DATASET - USE INSTRUCTIONS FROM ***;
32 *** Hospital01.pdf TO CREATE .DBF FROM ESRI ARCMAP 9.1 CD/DVD. ***;
33 *** CONVERT THE .DBF TO .SAS7BDAT FORMAT. NAME THE RESULTING SAS ***;
34 *** DATASET US_Hospitals.SAS7BDAT. ***;
35 *** USERS NEED TO MODIFY THE LIBNAME STATEMENT PATH. ***;
36 libname hosp 'V:\data\GEODATA\esri_cds\ArcGIS9';
37
38 %global stpostal;
39 *** USERS NEED TO INPUT THE 2-DIGIT POSTAL CODE FOR THEIR STATE ***;
40 *** FOR THE MACRO VARIABLE BELOW, stpostal='XX'. ***;
41 %let stpostal='ID';
42
43
44 *** STEP 1 - READ NAACCR V10 OR V11 CONFIDENTIAL FILE FOR CASE ***;
45 *** LAT/LONG. MUST BE CONFIDENTIAL (2644) OUTPUT (2644). ***;
46 data cases;
47     infile in lrecl = 2644 pad missover;
48
49 *** MOVE INPUT STATEMENTS FROM COMMENT SECTION TO WORKING SECTION ***;
```

A screenshot of the SAS software interface. The title bar reads "SAS - [GreatCircleDistances3]". The menu bar includes "File", "Edit", "View", "Tools", "Run", "Solutions", "Window", and "Help". The main window displays a SAS program with line numbers 14 through 49. Four red arrows point to specific lines of code: line 20, line 29, line 36, and line 41. The code contains several comments and SAS statements for file handling and data processing.

Great Circle Distance Calculator

How do I use it?

SAS - [Results Viewer - V:\data\IHA\SAL_SURV\2006\gensalsurv\Results\sashtml1.htm]

File Edit View Go Tools Solutions Window Help

Results

Results

- Freq: Great Ci
- Table N10_
- Freq: Great Ci
- Table N10_
- One-Way
- One-W
- One-W

Great Circle Distance V2

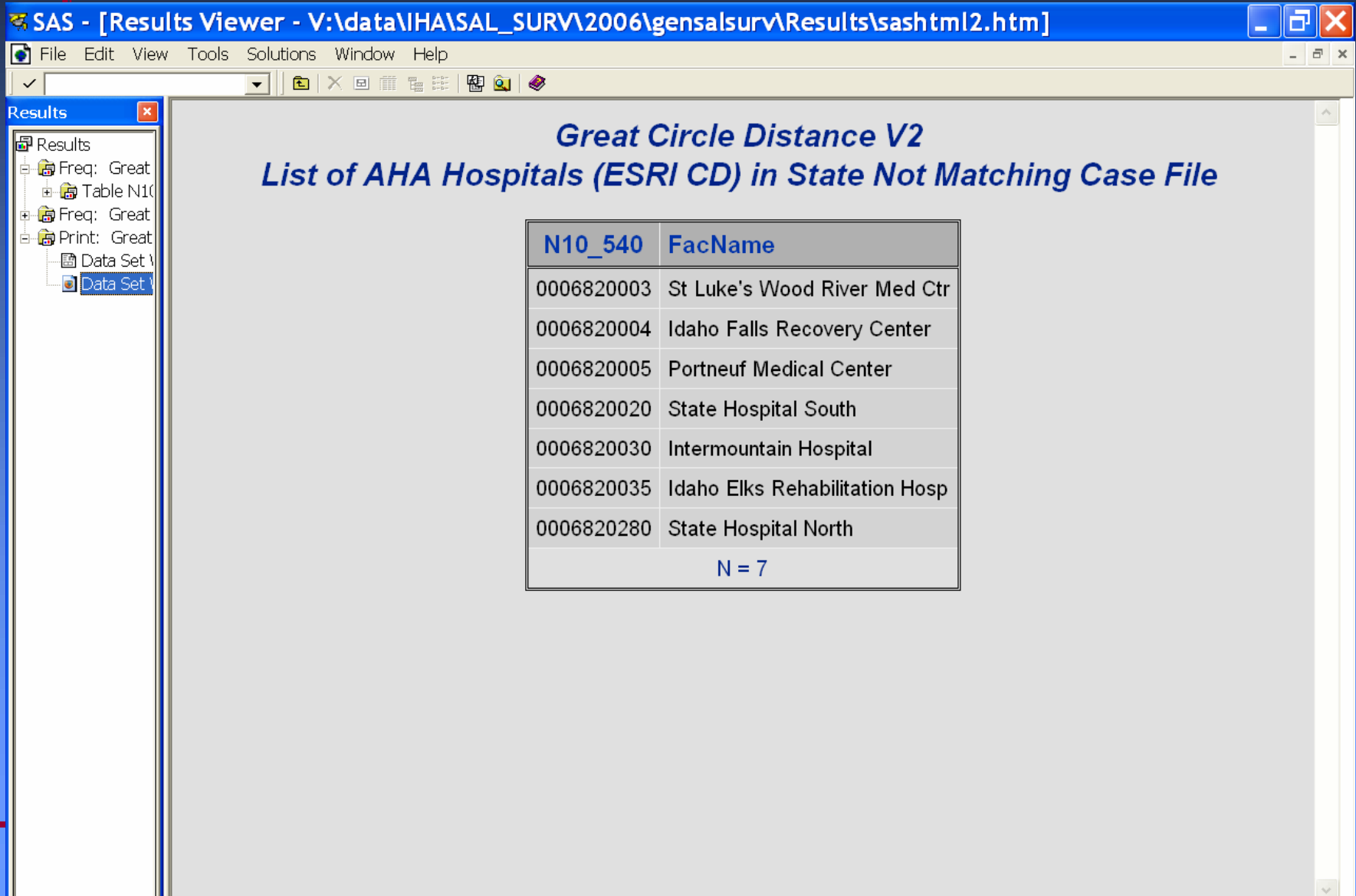
List of Reporting Hospital IDs Not Matching ESRI File

The FREQ Procedure

Reporting Hospital				
N10_540	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0006000000	19	0.07	19	0.07
0006666666	15988	61.59	16007	61.66
0006820123	1	0.00	16008	61.66
0006820157	134	0.52	16142	62.18
0006820235	114	0.44	16256	62.62
0006820290	3906	15.05	20162	77.67
0006820300	1548	5.96	21710	83.63
0006820364	2	0.01	21712	83.64
0006820375	604	2.33	22316	85.96
0006820385	2054	7.91	24370	93.88
0006999999	1589	6.12	25959	100.00
0099999999	1	0.00	25960	100.00

Great Circle Distance Calculator

How do I use it?



The screenshot shows the SAS Results Viewer interface. The title bar indicates the file path: V:\data\IHA\SAL_SURV\2006\gensalsurv\Results\sashtml2.htm. The main content area displays the following text:

Great Circle Distance V2
List of AHA Hospitals (ESRI CD) in State Not Matching Case File

N10_540	FacName
0006820003	St Luke's Wood River Med Ctr
0006820004	Idaho Falls Recovery Center
0006820005	Portneuf Medical Center
0006820020	State Hospital South
0006820030	Intermountain Hospital
0006820035	Idaho Elks Rehabilitation Hosp
0006820280	State Hospital North

N = 7

Great Circle Distance Calculator

How do I use it?

SAS - [Results Viewer - V:\data\IHA\SAL_SURV\2006\gensalsurv\Results\sashtml6.htm]

File Edit View Go Tools Solutions Window Help



Great Circle Distance V2

GC Distances in Statute Miles from Cases to Hospitals

Overall



Hospital Results



FacName	NumberCases	MeanDistance	MinDistance	MaxDistance
Overall	66253	22.99	0.00	499.51
Hospital 1	78	16.16	0.02	403.60
Hospital 2	281	23.90	0.01	405.41
Hospital 3	558	11.11	0.03	276.09
Hospital 4	898	22.77	0.01	465.58
Hospital 5	25	39.50	0.33	276.12
Hospital 6	59	13.01	0.09	65.59
Hospital 7	6	96.64	32.12	239.51
Hospital 8	962	11.95	0.03	436.29
Hospital 9	101	16.98	0.33	196.92
Hospital 10	5596	20.73	0.00	420.86
Hospital 11	213	27.70	0.08	165.74
Hospital 12	58	8.03	0.15	64.98

(Masked Names)

Great Circle Distance Calculator

What can I do with it?

- The SAS program creates a record-level dataset with up to five distance measures per case
 - ◆ one each for up to 5 source records associated with every case
 - ◆ and a fixed-width ASCII file for importing into your cancer registry database
- These can be used for studying the relationships between distance to treatment facility and other variables (type of treatment, survival...)

Great Circle Distance Calculator

What can I do with it?

- Recent studies using distances from addresses at diagnosis to treatment facilities include:
 - ◆ Treatment of local breast carcinoma in Florida: the role of the distance to radiation therapy facilities. Voti L et al. *Cancer*. 2006 Jan 1;106(1):201-7.
 - ◆ Geographic access to health care for rural Medicare beneficiaries. Chan L et al. *J Rural Health*. 2006 Spring;22(2):140-6.
 - ◆ Impact of patient distance to radiation therapy on mastectomy use in early-stage breast cancer patients. Schroen AT et al. *J Clin Oncol*. 2005 Oct 1;23(28):7074-80.
 - ◆ Relationship Between Travel Distance and Utilization of Breast Cancer Treatment in Rural Northern Michigan. Meden T et al. *JAMA*. 2002;287:111.

Great Circle Distance Calculator

Caveats & Limitations

- Current version requires access to ESRI ArcGIS data - on either the Data and Maps DVD or United States CD 2.
 - ◆ Instructions for converting from the ESRI CD data to a .dbf file with latitude and longitude for each hospital are in Hospital01.pdf. These cover ArcGIS and MapInfo users.
 - ◆ The ESRI hospital layer is from the American Hospital Association, may not include all hospitals, is not current, and may not reflect the location of a facility at the time the case was diagnosed and/or treated.
 - ◆ No redistribution rights for the ESRI US “Hospitals” data are granted by the vendor.
- Hey, what about Canada?
 - ◆ Need a file of Canadian treatment facility locations to implement.

Great Circle Distance Calculator Caveats & Limitations

- All the usual limitations associated with case geocoding, including varying accuracy in geocoding rural versus urban addresses.
- Different registries may have different rules for how source records are consolidated.
- Great circle distance is not driving distance.
 - ◆ (but it is better than linear distance)



Great Circle Distance Calculator

- What it is
- How I use it
- What I can do with it

(in 10 minutes or less)



- Questions??

NAACCR