ICTR Data Managers Interest Group

Timothy Shields, MMI tshields@jhsph.edu

Frank Curriero, EHS fcurrier@jhsph.edu

November 26, 2013

Overview - Spatial Analysis

GIS

Spatial Data

Obtain, Collect, Create Images (Satellite, Google Earth) Spatial Data Accuracy Non-spatial data

Maps

Spatial Database/Integration View Patterns Simultaneously Generate/Refine Hypothesis

Spatial Statistics

Quantify Relationships Test Hypothesis Study Design Go Beyond the Map!

Rationale

Identified health outcomes and measured environmental exposures likely vary spatially. Collect and harness this information to improve

- study design
- analysis
- prevention, intervention, control and policy efforts

Overview - Spatial Analysis

GIS

Spatial Data

Obtain, Collect, Create Images (Satellite, Google Earth) Spatial Data Accuracy Non-spatial data

Maps

Spatial Database/Integration View Patterns Simultaneously Generate/Refine Hypothesis

Spatial Statistics

Quantify Relationships Test Hypothesis Study Design Go Beyond the Map!

Data Warehouse •Receive •Compile •Organize •Disseminate

Data Enhancer Integrate Add Context Expand Disseminate

Spatial Data

Wood Basic Science Building Auditorium



•725 N Wolfe St,Baltimore, MD 21205•Wolfe and Madison

Coordinate: •39.298509, -76.590629 •362841 S, 4351118 N

GIS - Mapping Features

Points - - - - -

Lines - -

Polygons - - -

Others layers - -



GIS – Attaching Data To The Map



FID	Shape *	Patitent	Address	City	State	Zip	Age	Sex	Race	Date
0	Point	John Doe1	**** Argonne Dr	Baltimore	MD	21218	33	M	1	5/20/199
1	Point	John Doe2	**** Greenmount Ave	Baltimore	MD	21218	22	M	2	2/3/2004
2	Point	Jane Doe1	**** Harford Rd	Baltimore	MD	21218	34	F	3	4/9/2008
3	Point	John Doe3	**** N Charles St	Baltimore	MD	21218	29	M	1	9/13/200
4	Point	Jane Doe2	**** Harford Rd	Baltimore	MD	21218	54	F	2	6/29/200
5	Point	John Doe4	**** Saint Paul St	Baltimore	MD	21218	23	M	1	1/13/200
6	Point	John Doe5	**** E 25th St	Baltimore	MD	21218	19	M	1	10/15/20
7	Point	Jane Doe3	**** E 25th St	Baltimore	MD	21218	44	F	2	7/31/199
8	Point	Jane Doe4	**** Greenmount Ave	Baltimore	MD	21218	34	F	2	4/15/200
9	Point	John Doe6	**** Harford Rd	Baltimore	MD	21218	27	M	1	6/20/200
10	Point	Jane Doe5	**** Greenmount Ave	Baltimore	MD	21218	36	F	1	11/24/20
11	Point	John Doe7	**** Loch Raven Blvd	Baltimore	MD	21218	27	M	3	9/9/2004
12	Point	Jane Doe6	**** N Howard St	Baltimore	MD	21218	41	F	1	3/26/200
13	Point	John Doe8	**** N Charles St	Baltimore	MD	21218	38	M	2	12/26/20

- Data attached to map features
- Spatial relationships between map layers can be determined
- Questions answered by topology:
 - o What census tract is each case in?
 - How many cases are within each census tract?
 - o Which census tracts are adjacent?
 - What is the distance between patient and clinic, patient and other patient, patient and transport network?
- Data Enrichment

Field sites – Zambia and Zimbabwe, ICEMR project

- International Center of Excellence for Malaria Research:
 - Combined and complimentary studies on:
 - malaria epidemiology, vector biology and parasite genetics/genomics
- 3 sites:

Zambia: Choma, south; Nchelenge, north;

Zimbabwe: Mutasa, East





Spatial Data



-Satellite Imagery -OnScreen Digitizing -GPS

Spatial Data



Data

REDCap

Home My Projects 😵 Training Resources 🙆 Help & FAQ 🛛 🖓 Send-It

Listed below are the REDCap projects to which you currently have access. Click the project tille to open the project. Newly created projects begin in Development status are are used to build and design them. When you are ready to begin entering real data in the project, you may move it to Production status are by to designate the project as officially collecting data. When you are finished collecting data or if you wish to stop collection, the project may be set to mactive status are although it may be brought back to Production status at any time when you are ready to begin collecting data again. Also listed is the project type, which designates if the project contains surveys and the advector the protect of the again.

My Projects	Records	Fields	Туре	Status
ICEMR - Surveillance Data	Loading	Loading		\bigcirc
ICEMR - Weather Data	Loading	Loading	10.1	Ø
ICEMR - Household Data	Loading	Loading	1	\bigcirc
ICEMR - Household Member Data	Loading	Loading		\bigcirc

tshields | My Profile | Log out

REDCap Software - Version 5.0.11 - © 2013 Vanderbilt University

1	ICEMRHouseholdMember_DATA_2013-05-24_1412updated														
	F	G	Н	1	J	К	L	М	N	0	Р	Q	R	S	Т
1	hh_numbe	hh_type	today	start_time	worker_id	geopoint	lat	long	elev	error	dob_day	dob_mont	dob_year	age_years	age_m
2	101480	2	4/25/2012	2012-04-2	111	-16.XXXXXXXXXXXXXXX 26.XXXXXXXXXXXXXXX 1102.0 10.0	-16.XXXXX	26.XXXXXX	1102	10	13	5	1967	45	
3	101480														
4															
5	101480														
6	101480	2	4/25/2012	2012-04-2	111	-16.XXXXXXXXXXXXXX 26.XXXXXXXXXXXXXXXX 1102.0 10.0	-16.XXXXX	26.XXXXXX	1102	10	26	3	1971	41	
7	101480														
8															
9	101480														
10															
11	404400	-	4/25/2042	2042.04.2			4.5 200000	25 200000	4400	10			2004		
12	101480	2	4/25/2012	2012-04-2	111	-16.XXXXXXXXXXXXXX 26.XXXXXXXXXXXXXXXX 1102.0 10.0	-16.XXXXX	26.XXXXXX	1102	10	1	6	2001	11	
13	101480														
14															
15	101/190														
17	101480	2	1/25/2012	2012-04-2	111	-16 XXXXXXXXXXXXXXX 26 XXXXXXXXXXXXXXXXXX	-16 YYYYY	26 88888	1102	10	6	11	2005	6	
18	101480	2	. 4/23/2012	2012-04-2	111	-10.777777777777777777777777777777777777	-10./////	20.1111	1102	10	0	11	2005	0	
19	101100														
20															
21	101480														
22	101480	2	4/25/2012	2012-04-2	111	-16.XXXXXXXXXXXXXX 26.XXXXXXXXXXXXXXX 1102.0 10.0	-16.XXXXX	26.XXXXXX	1102	10	12	2	1997	15	
23	101480														

GIS – Mapping Data



Mutasa **RDT 0-5 Years**

GIS – Data Enrichment

RDT Summary



Formulas Data

Rovi

Page Lavout

- Integrate data layers

- Spatially link data
- Explore

-	Н		n	0	th	e	SE	25
		' y	Μ				30	50

House_Master_DistRivCat [Compatibility Mode] - Microsoft Excel

Legend
2007-200
Househol

A STATE			1												
5	ŀ	E24	- (0	f _æ											
	Δ Α	В	С	E	F	Q	R	S	Т	U	V	W	х	Y	Z
-	1 ID	utmx	utmy	2007Pos	2008Pos	Dist_RivCat1	Dist_RivCat2	Dist_RivCat3	Dist_RivCat4	Dist_RivCat5	DEM	DEMSLP	ASPECT_DEM	Aspect_Descr	Risk_RivDEM
-	2 E1036	50XXXX	8194633	1		1152.6	2101.1	1465.1	6086.2	17355.3	1081	0.000	123.7	Southeast	0.596
	3 E1084	50XXXX	8192986	1		90.0	1113.2	450.0	4306.9	19282.1	1066	0.022	26.6	Northeast	0.662
	4 E1126	50XXXX	8199371	1		900.0	1973.9	5407.5	5877.6	17688.8	1058	0.000	346.0	North	0.666
Ĺ	5 E1145	50XXXX	8192439	1		1892.1	524.8	254.6	1998.3	21534.3	1058	0.031	80.0	East	0.692
_	6 E1194	50XXXX	8196747	1		484.7	1083.7	3784.3	3245.0	20633.4	1056	0.011	78.7	East	0.666
	7 E1268	51XXXX	8193021	1		1980.0	1731.2	324.5	569.2	22832.4	1062	0.067	106.4	East	0.662
	8 E129	50XXXX	8180603	1		254.6	2291.0	360.0	7486.8	22981.0	1112	0.008	33.7	Northeast	0.493
	9 E1294	51XXXX	8195096	1		1361.9	1820.1	2251.8	1083.7	22630.1	1059	0.040	77.5	East	0.634
	10 E1302	51XXXX	8200177	1		360.0	763.7	6091.5	2707.5	19346.9	1040	0.011	135.0	Southeast	0.397
	11 E1397	51XXXX	8200327	1		0.0	547.4	6120.7	2377.8	19235.0	1040	0.022	135.0	Southeast	0.397
	12 E1421	51XXXX	8198399	1		891.0	484.7	6016.6	1049.6	21324.1	1044	0.016	71.6	East	0.397
	13 E1450	51XXXX	8196186	1		956.7	649.0	4583.8	810.0	23616.7	1046	0.011	0.0	North	0.696
	14 E1769	51XXXX	8184754	1		1642.3	829.8	3798.2	2795.8	29346.8	1097	0.023	180.0	South	0.532
	15 E1803	51XXXX	8187373	1		180.0	1932.4	5152.1	4306.9	28990.9	1095	0.044	11.3	North	0.532
	16 E1812	51XXXX	8186600	1		1053.4	1260.0	5490.0	4680.0	29789.3	1102	0.039	168.7	South	0.498
	17 E1855	51XXXX	8187146	1		926.6	1820.1	5959.7	5158.3	29906.2	1105	0.035	351.9	North	0.221
	18 E194	50XXXX	8177580	1		655.2	685.4	3331.2	6292.9	27003.8	1131	0.031	49.8	Northeast	0.394
	19 E2019	51XXXX	8184894	1		1719.4	815.0	8580.3	6766.2	33370.7	1111	0.016	188.1	South	0.198
	20 E2194	50XXXX	8189713	1		90.0	509.1	2737.2	3646.9	20611.8	1075	0.016	60.9	Northeast	0.601

Acroba

Spatial Statistics



Moss et al. Malaria Journal 2011, 10:163 http://www.malariajournal.com/content/10/1/163

- Risk map estimation
- Assess clustering
- Detect clusters
- Identify Env factors for modeling spatial variation
- Assess trends over space & time

New Trends – More Data

Increase in gps-enabled devices :

 a) Handhelds
 b) Tablets (Electronic Forms!!)
 c) Smartphones

Social media – twitter, yelp, crowdsourcing, Waze, ...
Cloud based systems

3. Cloud-based systems

New Trends – Improved Software

- Trajectory analysis better software for studying space across time
- 2. Overall improvement of combination of space/time tools

Overall Improvement of Data Flow Into GIS



Trajectory Data Analyses for Pedestrian Space-time Activity Study Feng Qi¹, Fei Du² ¹School of Environmental and Life Sciences, Kean University, ²Department of Geography, University of Wisconsin-

Madison