

High Degree of Accuracy in Census Tract Assignment within the Metropolitan Detroit Cancer Surveillance System

Cathryn H. Bock, PhD
NAACCR Annual Conference
Detroit, MI
June 7, 2007

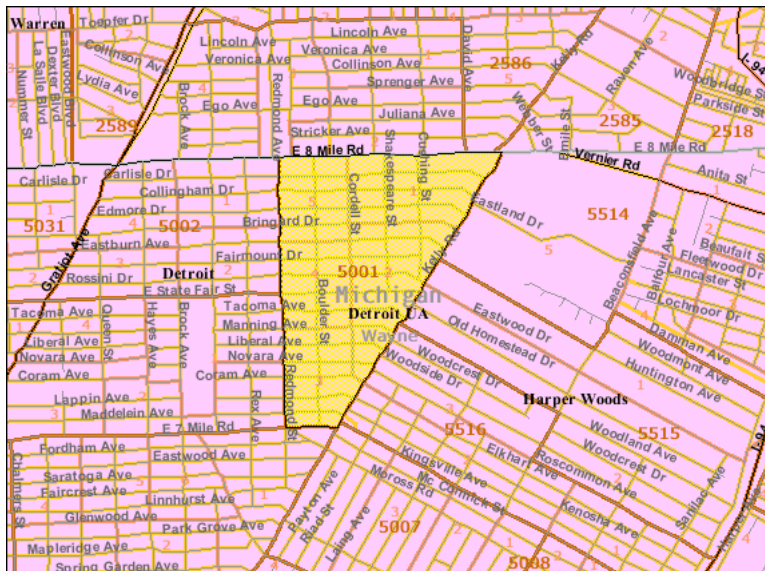
B A R B A R A A N N

KARMANOS

CANCER INSTITUTE

Wayne State University

Background



- Census Tract used as proxy for sociodemographic information
- **“Assessing the Accuracy of Geocoding”**
 - Rapid Response Surveillance Study
 - PI: Kendra Schwartz

MDCSS Census Tract

- Patient's address at time of diagnosis
 - Street number, Street direction, Zip code, County Code
- Matched with 2000 Topographically Integrated Geographic Encoding and Referencing system (TIGER) line files that contain data on census tract
- Non-Matches
 - Editing staff evaluates for address errors
 - Manually assigns census tract by cross-referencing with the 2003-2004 Edition Bresser's Cross-Index Directory

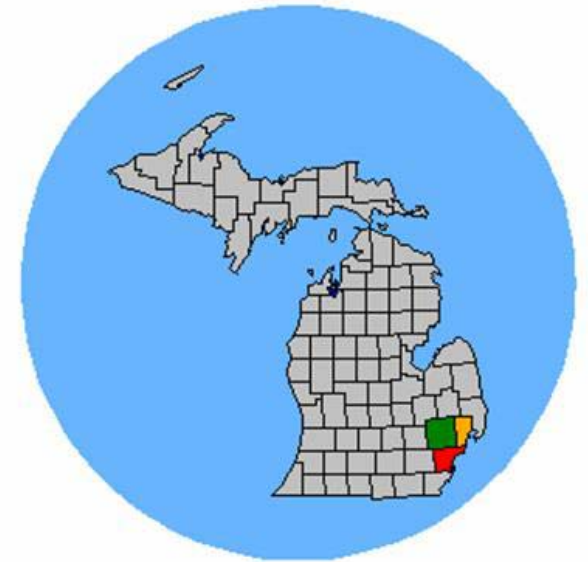
Specific Aims

1. Assess completeness of census tract assignment within the Metropolitan Detroit Cancer Surveillance System (MDCSS)
2. Compare MDCSS-assigned census tract with Michigan Metropolitan Information Center (MIMIC) assigned census tract
3. Compare GPS determined census tract with MDCSS and MIMIC assigned census tract
4. Evaluate predictors of census tract assignment within MDCSS



Study Design

- MDCSS Registry
- All SEER Reportable cases 1998-2002
- Random sample of 750 cases with MDCSS census tract sent to MIMIC
 - 250 per County
- Random sub-sample of 200 cases
 - GPS measurement



MDCSS Census Tract

- 124,798 SEER reportable incident cancer diagnoses 1998-2002 in MDCSS database
 - **124,073 (99.4%)** were assigned a census tract
 - **725 (0.6%)** had no census tract assigned
 - 557 (76.8%) missing information at the city level (only county recorded)
 - 84 (11.6%) PO Box
 - 56 (7.7%) missing information at the street level
 - 28 (3.1%) had complete address information but no census tract assigned.

Agreement with MIMIC

- 750 with MDCSS census tract sent to MIMIC
 - 671 (89.5%) MIMIC census tract
 - **Census tract: 95.2% agreement**
 - **County: 99.9% agreement**
 - 28 with complete address but no MDCSS census tract also sent to MIMIC
 - 7 were assigned census tract by MIMIC

GPS agreement

- 200 cases sampled from the 750 that were sent to MIMIC
 - GPS readings obtained, census tract assigned
- GPS and MDCSS
 - Census tract: 94.3% agreement
 - County: 100% agreement
- GPS and MIMIC
 - Census tract: 97% agreement
 - County: 100% agreement
- 20 cases randomly sampled from the 200
 - GPS readings provided 100% agreement with first reading



Characteristics of cases missing census tract information

Census Tract Status by Age

Age Group	Known		Unknown	
0-9	474	0.4%	4	0.6%
10-19	989	0.8%	32	4.4%
20-29	4896	4.0%	136	18.8%
30-39	6557	5.3%	115	15.9%
40-49	12382	10.0%	72	9.9%
50-59	21002	16.9%	90	12.4%
60-69	26555	21.4%	85	11.7%
70-79	33113	26.7%	97	13.4%
80-89	15899	12.8%	57	7.9%
≥ 90	2201	1.8%	1	0.1%
Unknown	5	0.0%	36	5.0%

Results: Ages 18-79

- Age
- Gender
- Race
- Ethnicity
- Reporting Source
- Histology
- Stage
- Site

Mean Age: Ages 18-79

	Known Census Tract n=104,974	Unknown Census Tract n=608
Mean Age (years)	59	45
	p<0.0001	

Gender: Ages 18-79

Gender	Known Census Tract n=104,974		Unknown Census Tract n=608	
Male	49,668	47.3%	218	35.9%
Female	55,303	52.7	389	64.0%
Unknown	3	0.0%	1	0.2%
	P<0.0001			

Race: Ages 18-79

Race	Known Census Tract n=104,974		Unknown Census Tract n=608	
White	79,460	75.7%	297	48.9%
Black	23,040	22.0%	53	8.7%
Other	1,117	1.1%	4	0.7
Unknown	1,357	1.3%	1254	41.8%
	P<0.0001			

Ethnicity: Ages 18-79

Ethnicity	Known Census Tract n=104,974		Unknown Census Tract n=608	
Non-Hispanic	101,915	97.1%	340	55.9%
Hispanic	1,532	1.5%	9	1.5%
Unknown	1,527	1.5%	259	42.6%
	P<0.0001			

Reporting Source: Ages 18-79

Reporting Source	Known Census Tract N=104,974		Unknown Census Tract n=608	
Hospital/Clinic	97,849	93.2%	137	22.5%
Laboratory	3,615	3.4%	342	56.3%
Physician	2,914	2.8%	128	21.1%
Autopsy/Death Certificate	596	0.6%	1	0.2%

p<0.0001

Histology: Ages 18-79

Histology	Known Census Tract n=104,974		Unknown Census Tract n=608	
In Situ	17,616	16.8%	334	54.9%
Malignant	87,358	83.2%	274	45.1%
	p<0.0001			

Stage: Ages 18-79

Stage	Known Census Tract n=104,974		Unknown Census Tract n=608	
In Situ	17,642	16.8%	334	54.9%
Local	42,710	40.7%	144	23.7%
Regional	19,112	18.2%	20	3.3%
Distant	20,265	19.3%	25	4.1%
Unstaged	5,245	5.0%	85	14.0%
	p<0.0001			

Site: Ages 18-79

Cancer Site	Known Census Tract n=104,974		Unknown Census Tract n=608	
Cervix Uteri	8,657	8.3%	268	44.1%
Other	96,317	91.7%	340	55.9%
	P<0.0001			

Predictors of coded census tract: Ages 18-79

Variable	Odds Ratio	95% CI	p-value
Age	1.01	1.007-1.02	0.0002
Sex	1.38	1.06-1.79	0.02
Lab/Physician Report	0.05	0.04-0.07	<0.0001
Stage: Distant*	3.62	2.09-6.30	0.001
Stage: In Situ*	1.32	0.91-1.92	0.003
Stage: Local*	1.95	1.36-2.79	0.91
Stage: Regional*	2.8	1.61-4.98	0.05

* vs. Unstaged

Results: Ages 18-79

- Mean age is younger (45 vs. 59 years)
- Source of data – more likely to be by laboratory report or physician
- Greater proportion of in situ cancer
- Greater proportion of cervical cancer

Study Considerations

- Since 2002, MDCSS has moved to SEER*DMS database, with an improved built-in census tract assignment algorithm
 - More recent cancer cases are more likely to have a census tract assigned
 - Census tracts are assigned more quickly
- Unlike most other registries, MDCSS collects incident in situ cervical cancers
 - does not follow up these cases
 - these cases are less likely to have complete address information

Study Limitations

- Results may not apply to other registry populations with different urban/rural distributions
- Results are dependent on the correctness of latest TIGER line files, which varies by geographic location
- Results could vary in regions undergoing growth in new housing development
- Results are dependent on registry-specific census tract assignment methods

Study Strengths

- Large sample size of registry cancer cases
- Centered around Census 2000
- Racial and Ethnic diversity

Conclusions

- Census tract assignment by MDCSS has a high degree of accuracy, particularly among adults $>$ age 40 and malignant cases.
- Missing census tract is often attributable to the data source being a laboratory or physician report.



Acknowledgments

- Kendra Schwartz – PI
- Doug Towns – MIMIC
- Eric Pietsch – Research Assistant
- Khairul Islam – Biostatistician
- MDCSS Registry Staff