

36. Parkin DM, Iscovich J. Risk of cancer in migrants and their descendants in Israel: II. Carcinomas and germ-cell tumours. *Int J Cancer*. 1997;70(6):654-660.
37. Abraido-Lanza AF, Dohrenwend BP, Ng-Mak DS, Turner JB. The Latino mortality paradox: a test of the "salmon bias" and healthy migrant hypotheses. *Am J Public Health*. 1999;89(10):1543-1548.
38. Balarajan R, Bulusu L, Adelstein AM, Shukla V. Patterns of mortality among migrants to England and Wales from the Indian subcontinent. *BMJ (Clin Res Ed)*. 1984;289(6453):1185-1187.
39. Balzi D, Geddes M, Brancker A, Parkin DM. Cancer mortality in Italian migrants and their offspring in Canada. *Cancer Causes Control*. 1995; 6(1):68-74.
40. McCredie M, Williams S, Coates M. Cancer mortality in East and Southeast Asian migrants to New South Wales, Australia, 1975-1995. *Brit J Cancer*. 1999;79(7-8):1277-82.
41. McCredie M, Williams S, Coates M. Cancer mortality in migrants from the British Isles and continental Europe to New South Wales, Australia, 1975-1995. *Int J Cancer*. 1999;83(2):179-185.
42. Wild S, McKeigue P. Cross sectional analysis of mortality by country of birth in England and Wales, 1970-92. *BMJ*. 1997;314(7082):705-710.
43. Russell JN, Hitchcock D. *Vital and Health Statistics: The Assessment of Immigration Status in Health Research*. Atlanta, Ga: National Center for Health Statistics; 1999.
44. Merriam-Webster's OnLine Collegiate Dictionary. Merriam-Webster Inc; Springfield, Ma: 2001
45. Shinagawa SM, Kagawa-Singer M, Chen MS, Tsark JU, Palafox NA, Mackura G. Cancer registries and data for Asian Americans and Native Hawaiians and Pacific Islanders: what registrars need to know. *J Reg Mgmt*. 1999; 26(4):128-141.
46. Bhopal R. Spectre of racism in health and health care: lessons from history and the United States. *BMJ*. 1998;316(7149):1970-1973.
47. Sorlie PD, Rogot E, Johnson NJ. Validity of demographic characteristics on the death certificate. *Epidemiology*. 1992;3(2):181-184.
48. National Center for Health Statistics *Comparability of Marital Status, Race, Nativity, and Country of Origin on the Death Certificate and Matching Census Record*. Rockville, Md: National Center for Health Statistics; 1969. NCHS series 2, No. 34.
49. National Center for Health Statistics. *Vital Statistics of the United States, Mortality, Technical Appendix; 1995*. Washington, DC: Public Health Service; 1999.
50. Pablos-Mendez A. Mortality among Hispanics. *JAMA*. 1994;271(16):1237.



## Florida Cancer Registrars Association

... furthering cancer care and research ...

### PROFESSIONAL REVIEW FOR TUMOR REGISTRARS A STUDY GUIDE

Third Edition

EDITOR

Gayle Greer Clutter, RT, CTR

ASSOCIATE EDITORS

Patricia Bentley, CTR • Jean Byers, CTR • April Fritz, CTR

PUBLISHED BY

The Florida Cancer Registrars Association

Cost: \$40.00

#### FCRA VIDEO STUDY COURSES

Cancer Epidemiology

Principles of Biostatistics

Advanced Statistical Processes

Instructor: Tim Aldrich, PhD, MPH, CTR

Video Course Cosponsors:

Duke University Medical Center • NCRA  
Tumor Registrars Association of North Carolina

Includes Workbook & VHS Tape For Each Course  
Cost \$40.00 each or \$100.00 for all three

All costs include Priority Mail Shipping

Contact Pat Bentley at (407) 894-6232

patbentley@cfl.rr.com

Order Forms may be obtained at FCRA Website:

<http://www.fcra.org/members/index.html>

Click on "FCRA Store"

## Original Article

# Assessing and Improving Census Tract Completeness

Francis P. Boscoe, PhD; Christine L. Kielb, MS; Maria J. Schymura, PhD; Tina M. Bolani, BS

**Abstract:** The quality and completeness of address fields is becoming an important issue for central cancer registries because geography is increasingly being incorporated into epidemiologic studies. This paper describes the effort to investigate why 20% of a sample of 500 case reports in the New York State Cancer Registry (NYSCR) could not be successfully assigned a census tract by our automatic matching process. The most common reasons were: 1) addresses missing from the reference file, and 2) addresses with minor errors. In both situations the census tract was obtained with modest additional effort through manual geocoding, increasing the percentage of successful assignments in the sample from 80% to 99.8%. A batch of 85,863 case reports not previously assigned census tracts was also matched against a second reference file. This effort yielded 15,808 additional census tract designations and demonstrates the value of consulting more than one reference file.

**Key words:** address matching, census tract assignment, geocoding

### Introduction

In recent years, increasing attention has focused on the relationship between place of residence and cancer incidence. Although efforts to determine the link between chronic, low-level exposures to environmental contaminants and cancer have received much publicity,<sup>1,4</sup> perhaps the greatest contribution of the geographic approach has been the identification of areas of increased risk due to socioeconomic factors.<sup>5-11</sup> Since the investigation of such research questions requires complete and accurate information on address at diagnosis, geocoding, defined as the act of assigning additional geographic information to a case report, is an increasingly important aspect of registry practice.

Geocoding applies to multiple spatial scales, from county through latitude and longitude coordinates. The choice of a particular scale to use in an epidemiologic study is motivated by the type of research question being asked, data limitations, and confidentiality concerns. Finer geographic scales are desirable in analyses where it is important to minimize the variability of population characteristics within geographical units, or where a local point source environmental exposure is being investigated. However, address completeness often becomes an issue due to the frequent lack of address information in disease registries. Confidentiality concerns emerge at finer scales as well.

This study focuses on the census tract level of geographic analysis. Census tracts are small, relatively permanent statistical subdivisions of a county that are designed to contain between 2500 and 8000 persons and to be as homogeneous as possible with respect to population characteristics, economic status, and living conditions.<sup>12</sup> The census

tract is a required data element for both the Surveillance, Epidemiology, and End Results (SEER) Program and National Program of Cancer Registries (NPCR).<sup>13</sup> Census tracts confer a number of advantages for analysis: the units are sufficiently small to identify neighborhood-level patterns in urban areas, their relatively uniform population size means they are less biased than larger geographical units with respect to finding patterns in urban versus rural areas, and confidentiality concerns can be addressed if a consistent policy is observed.<sup>14</sup> The Bureau of the Census reports information regarding income, housing, education, and other demographic characteristics at this level, facilitating ecologic analysis.

The main disadvantage of census tracts is the difficulty in attaining sufficient completeness so that this unit of analysis is useful. It is especially difficult to assign census tracts in rural areas where rural route and post office box addresses are in use. For this reason, a study that excludes cases with unknown census tract will be biased toward areas with higher population density.<sup>15</sup> Missing census tracts can be imputed on the basis of ZIP code of residence, though unless this is done carefully, differential bias can still result. Another disadvantage of census tracts is that much of the public is unfamiliar with this unit of analysis, which complicates the communication of findings.

In this paper, the issue of census tract completeness was addressed as part of an ongoing effort to assess the feasibility of census tract level analysis. Specifically, an investigation was undertaken to find out why some incoming case reports fail to have census tract automatically assigned by matching software, and to assess the degree of additional effort required to supply this information.

"Assessing and Improving Census Tract Completeness"

All authors from New York State Cancer Registry, New York State Department of Health, Albany, NY.

Address correspondence to Christine L. Kielb, MS, Research Scientist, Bureau of Environmental and Occupational Epidemiology, New York State Department of Health, 547 River Street, Room 200, Troy, NY 12180. Telephone: 518-402-7976; Fax: 518-402-7959; e-mail: clk03@health.state.ny.us.

Submitted: 11/14/01. Revised: 05/21/02. Accepted: 09/13/02

## Methods

This project consisted of two phases. In the first phase, a random sample of 500 in-state case reports sent to the New York State Cancer Registry (NYSCR) during the 2000 calendar year was selected for analysis. Each case had a date of diagnosis of 1999 or earlier. An analysis was undertaken to find out the proportion of these cases automatically assigned a census tract by matching software, the reasons for nonassignment, and how much and what kinds of effort were required to assign a census tract.

The process of census tract assignment included the following procedure. Case reports were run through matching software, AutoMatch version 4.1 and AutoStan version 4.6d, using Dynamap version 10.3 as the reference file.<sup>16-18</sup> Both the software and reference file are considered state-of-the-industry.<sup>19</sup> Case reports that failed to have a census tract assigned were forwarded to the geocoding staff for review, who assigned them to one of 4 categories. The first category included case reports with an address that matched a previously existing, census-tracked address in the registry. In such instances, the case report was simply linked to the previously existing address. The second category included case reports where the address was complete and correct according to the United States Post Office (USPS) ZIP+4 Web site,<sup>20</sup> but simply missing from the reference file. For these case reports, the geocoding staff attempted to assign census tract manually using the Bureau of the Census' address-census tract crosswalk file.<sup>21, 22</sup> The third category included case reports with minor address errors such as misspellings or transpositions. These case reports were corrected and sent back for reprocessing. The fourth category included case reports in which the address was either incomplete or not similar to any address recognized by the USPS, as well as case reports with correct addresses that could not be assigned a census tract using the address-census tract crosswalk file. The geocoding staff attempted to assign census tract to these case reports by other available resources, including Department of Motor Vehicles (DMV) records, Real Property Data, and commercial people-finding Web sites. Geocoding staff are typically able to complete about 20 to 25 cases per hour.

In the second phase of the project, the Bureau of the Census' address-census tract crosswalk file was used to match case reports in the registry with reporting source dates from 1996 to 1999 which had not previously been

assigned a census tract. To facilitate the matching process, a SAS program was first written to remove extraneous characters from both the registry and census address fields, and to provide several different versions of each address in the registry file, reflecting common alternate spellings and abbreviations. The actual matching process consisted of two steps. First, all case reports whose ZIP code and street name matched the census file, and whose street number fell within the number range specified in the census file were isolated. In the second step, census tracts were assigned to case reports that could not be matched to the census file in the first step because they had missing or out-of-range street numbers, but involved streets that fell entirely within a single census tract.

**Table 1.** Outcomes of Census Tract Assignment Process in the Sample of Case Reports

Type of Outcome	N	%
Automatic assignment	400	80.0
Case matching	3	0.6
USPS/Census Crosswalk		
Address completely correct	39	7.8
Address nearly correct	26	5.2
Additional sources	31	6.2
Total Assigned	499	99.8
Census Tract Unassigned	1	0.2
Total	500	100.0

## Results

It was possible to assign a census tract to 499 (99.8%) of the 500 cases in the sample. A breakdown by category is given in Table 1. The 26 cases whose addresses were nearly correct displayed a range of problems. By far the most common was an incorrect ZIP code in a city with multiple ZIP codes, accounting for half of such case reports (Table 2). The 32 most difficult cases consisted largely of post office boxes, rural routes, and recent construction (Table 3). All but 1 of these could be assigned a census tract by consulting additional reference sources, shown in Table 4 in approximate order of usefulness.

In the second phase of the project, 85,863 case reports were matched against the Bureau of the Census crosswalk file. Of these previously unassigned case reports, 15,808, or 18.4%, were assigned a census tract: 12,106 in the first step, and 3702 in the second step. These were added to the registry in batch mode.

**Table 2.** Range of Problems in Nearly Correct Addresses

Problem	Example	N	%
Wrong ZIP Code in a multi-ZIP city	12202 instead of 12203	13	2.6
Slight misspelling of a street	Mrple Avenue instead of Maple Avenue	7	1.4
Extra or garbage characters in address field	MapleHAvenue instead of Maple Avenue	2	0.4
An address that changed between the time the address was first collected and the time the case was processed	Part of 10707 was reassigned to 10709 in 1999	2	0.4
Facility name provided	Shady Grove Nursing Center	1	0.2
Address information out of sequence	Maple Avenue 8 instead of 8 Maple Avenue	1	0.2
Total		26	5.2

## Methods

This project consisted of two phases. In the first phase, a random sample of 500 in-state case reports sent to the New York State Cancer Registry (NYSCR) during the 2000 calendar year was selected for analysis. Each case had a date of diagnosis of 1999 or earlier. An analysis was undertaken to find out the proportion of these cases automatically assigned a census tract by matching software, the reasons for nonassignment, and how much and what kinds of effort were required to assign a census tract.

The process of census tract assignment included the following procedure. Case reports were run through matching software, AutoMatch version 4.1 and AutoStan version 4.6d, using Dynamap version 10.3 as the reference file.<sup>16-18</sup> Both the software and reference file are considered state-of-the-industry.<sup>19</sup> Case reports that failed to have a census tract assigned were forwarded to the geocoding staff for review, who assigned them to one of 4 categories. The first category included case reports with an address that matched a previously existing, census-tracked address in the registry. In such instances, the case report was simply linked to the previously existing address. The second category included case reports where the address was complete and correct according to the United States Post Office (USPS) ZIP+4 Web site,<sup>20</sup> but simply missing from the reference file. For these case reports, the geocoding staff attempted to assign census tract manually using the Bureau of the Census' address-census tract crosswalk file.<sup>21, 22</sup> The third category included case reports with minor address errors such as misspellings or transpositions. These case reports were corrected and sent back for reprocessing. The fourth category included case reports in which the address was either incomplete or not similar to any address recognized by the USPS, as well as case reports with correct addresses that could not be assigned a census tract using the address-census tract crosswalk file. The geocoding staff attempted to assign census tract to these case reports by other available resources, including Department of Motor Vehicles (DMV) records, Real Property Data, and commercial people-finding Web sites. Geocoding staff are typically able to complete about 20 to 25 cases per hour.

In the second phase of the project, the Bureau of the Census' address-census tract crosswalk file was used to match case reports in the registry with reporting source dates from 1996 to 1999 which had not previously been

assigned a census tract. To facilitate the matching process, a SAS program was first written to remove extraneous characters from both the registry and census address fields, and to provide several different versions of each address in the registry file, reflecting common alternate spellings and abbreviations. The actual matching process consisted of two steps. First, all case reports whose ZIP code and street name matched the census file, and whose street number fell within the number range specified in the census file were isolated. In the second step, census tracts were assigned to case reports that could not be matched to the census file in the first step because they had missing or out-of-range street numbers, but involved streets that fell entirely within a single census tract.

**Table 1.** Outcomes of Census Tract Assignment Process in the Sample of Case Reports

Type of Outcome	N	%
Automatic assignment	400	80.0
Case matching	3	0.6
USPS/Census Crosswalk		
Address completely correct	39	7.8
Address nearly correct	26	5.2
Additional sources	31	6.2
Total Assigned	499	99.8
Census Tract Unassigned	1	0.2
Total	500	100.0

## Results

It was possible to assign a census tract to 499 (99.8%) of the 500 cases in the sample. A breakdown by category is given in Table 1. The 26 cases whose addresses were nearly correct displayed a range of problems. By far the most common was an incorrect ZIP code in a city with multiple ZIP codes, accounting for half of such case reports (Table 2). The 32 most difficult cases consisted largely of post office boxes, rural routes, and recent construction (Table 3). All but 1 of these could be assigned a census tract by consulting additional reference sources, shown in Table 4 in approximate order of usefulness.

In the second phase of the project, 85,863 case reports were matched against the Bureau of the Census crosswalk file. Of these previously unassigned case reports, 15,808, or 18.4%, were assigned a census tract: 12,106 in the first step, and 3702 in the second step. These were added to the registry in batch mode.

**Table 2.** Range of Problems in Nearly Correct Addresses

Problem	Example	N	%
Wrong ZIP Code in a multi-ZIP city	12202 instead of 12203	13	2.6
Slight misspelling of a street	Mrple Avenue instead of Maple Avenue	7	1.4
Extra or garbage characters in address field	MapleHAvenue instead of Maple Avenue	2	0.4
An address that changed between the time the address was first collected and the time the case was processed	Part of 10707 was reassigned to 10709 in 1999	2	0.4
Facility name provided	Shady Grove Nursing Center	1	0.2
Address information out of sequence	Maple Avenue 8 instead of 8 Maple Avenue	1	0.2
Total		26	5.2

**Table 3.** Problems that Make Addresses Difficult to Geocode

Problem	Example	Census Tract locatable	Percent	Census Tract not locatable
Address correct but unknown to Bureau of the Census	Any street built in the past year	10	2.0	1
Post Office Box in Rural Area	PO Box 777 Accord, NY 12404	8	1.6	0
Rural Route address	RR 1 Box 555 Accord, NY 12404	3	0.6	0
House number missing or out of range	118 Maple Avenue (1-99 valid range)	4	0.8	0
Street address is given, but case falls within an area where only PO Box and/or RR addresses are officially recognized	Turkey Hill Road Accord, NY 12404	2	0.4	0
Street does not exist and is not similar to any named street	Manitoba Street Albany, NY 12207	2	0.4	0
Post Office Box in an urban area	PO Box 325 Albany, NY 12207	1	0.2	0
No information provided	Unknown Bronx, NY 99999	1	0.2	0
Totals		31	6.2	1

## Discussion

In this study, 80% of a sample of case reports were able to have census tracts assigned to them through the registry's automatic matching process. 19.8% had census tract assigned by the geocoding staff, at the rate of 20 to 25 cases per hour. In addition, many case reports in the NYSCR that were not assigned a census tract through the first matching process were assignable using an alternate reference file. The results of this study give much cause for optimism regarding the ability to attain very high completeness in the census tract field. For example, at the rate of 20 to 25 cases per hour, and assuming similar incoming data quality, a registry in a state the size of New Jersey or Ohio could potentially achieve census tract completeness through the efforts of a single staff person.

We acknowledge that the NYSCR is in an advantageous position with respect to both staff resources and the quality of geocoded information on incoming case reports. As a large registry processing over 88,000 case reports per year, NYSCR is fortunate to have 5 full-time geocoding clerks as well as a research scientist with a geography background. New York's e-911 initiative, designed to improve the ability of emergency vehicles to reach rural addresses by naming all roads and numbering all properties, has eliminated most of the rural route and post office box addresses that make geocoding difficult. As of 2002, only a small handful of counties have not completed their e-911 readdressing, and rural route addresses are expected to be obsolete in New York State by 2003.

Still, all state cancer registries can benefit from the findings in this study, since even modest efforts have the potential to yield large returns. Simply examining the case reports that failed to match can reveal systematic problems that can be readily corrected. For instance, in the course of choosing

the random sample of 500 cases, a systematic problem with addresses in the borough of Queens was identified. Queens is unique in that hyphenated house numbers are permissible, such as 162-71 79<sup>th</sup> Avenue. (This is not an actual address.) A certain number of incoming case reports omit the hyphen, but retain the space, as in 162 71 79<sup>th</sup> Avenue. These were being standardized as 162 71<sup>st</sup> Avenue, and not being geocoded. Fortunately, none of the numbered streets in Queens have 3 digit house numbers, so no incorrect matches were accidentally made.

Perhaps the most intriguing finding is that geocoding quality is more a function of the quality of the reference file than the quality of the incoming case reports. It was possible to assign census tract to 18.4% of a large sample of reports that were not matched using the primary reference file. A combination of reference files, therefore, yields far more utility than any single file considered on its own. As might be expected, the addresses missed by a commercial reference file include many new or recent addresses, but they also include some addresses dating back a century or more. The reasons for this are manifold, but it is important to remember that commercial data companies tailor their products to the needs of their clients, who include many retailers and developers. The quality of a commercial reference file in a particular geographic area may thus be proportional to the level of economic activity in that area. Cancer incidence, of course, does not necessarily follow any such pattern.

It may have been possible to improve results further if automated matching criteria had been relaxed. Taking the necessary steps to allow MapleHStreet to be read as Maple Street would reduce the amount of manual effort required further downstream in the geocoding process. Given the kinds of fine distinctions that have to be made, such as between 119<sup>th</sup> Street, 119<sup>th</sup> Road, 119<sup>th</sup> Avenue, and 119<sup>th</sup>

Drive in Queens, the more stringent matching criteria are preferred. There is a certain art to geocoding that requires a human presence to follow promising common sense leads and discriminate between valid and invalid addresses.

### Summary

In this paper, the degree to which a sample of incoming case reports could be automatically assigned a census tract, the reasons for non-assignment, and the amount of additional effort required for assignment were assessed. Automatic matching assigned census tracts to 80% of cases; an additional 13.4% were assigned by consulting up to two Web sites; 6.2% more were assigned by consulting one or more additional resources. The primary reasons for non-assignment through automatic matching were minor errors in case report addresses and addresses missing from the reference file. In addition, 85,863 previously unassigned case reports were matched against a second reference file, and over 18% of them were assigned. This exercise demonstrates the value of using different methods to geocode case reports and the potential for many population-based cancer registries to attain a high proportion of completeness in the census tract field.

### References

- Morris MS, Knorr RS. Adult leukemia and proximity-based surrogates for exposure to Pilgrim plant's nuclear emissions. *Arch Environ Health*. 1996;5(4):266-74.
- Pless-Mullooli T, Phillimore P, Moffatt S, et al. Lung cancer, proximity to industry, and poverty in northeast England. *Environ Health Perspect*. 1998;106(4):189-96.
- Michelozzi P, Fusco D, Forastiere F, Ancono C, Dell'Orco V, Perucci CA. Small area study of mortality among people living near multiple sources of air pollution. *Occup Environ Med*. 1998;55(9):611-15.
- Zaridze DG, Li N, Men T, Duffy SW. Childhood cancer incidence in relation to distance from the former nuclear testing site in Semipalatinsk, Kazakhstan. *Int J Cancer*. 1994;59(4):471-75.
- Liu L, Cozen W, Bernstein L, Ross RK, Deapen D. Changing relationship between socioeconomic status and prostate cancer incidence. *J Natl Cancer Inst*. 2001;93:705-9.
- Krieger N, Quesenbery C, Peng T, et al. Social class, race/ethnicity, and incidence of breast, cervix, colon, lung, and prostate cancer among Asian, black, Hispanic, and white residents of the San Francisco Bay Area, 1988-92 (United States). *Cancer Causes Control*. 1999;10:525-37.
- Boyd C, Zhang-Salomons JY, Groome PA, Mackillop WJ. Associations between community income and cancer survival in Ontario, Canada, and the United States. *J Clin Oncol*. 1999;17(7):2244-55.
- Liu L, Deapen D, Bernstein L. Socioeconomic status and cancers of the female breast and reproductive organs: a comparison across racial/ethnic populations in Los Angeles County, California (United States). *Cancer Causes Control*. 1998;9:369-80.
- Gorey KM, Vena JE. The association of near poverty status with cancer incidence among black and white adults. *J Comm Health*. 1995;20:359-66.
- Gorey KM, Vena JE. Cancer differentials among US blacks and whites: quantitative estimates of socioeconomic-related risks. *J Natl Med Assoc*. 1994;86(3):209-15.
- Baquet CR, Horm JW, Gibbs T, Greenwald P. Socioeconomic factors and cancer incidence among blacks and whites. *J Natl Cancer Inst*. 1991;83:551-7.
- United States Census Bureau. Census Tracts and Block Numbering Areas. Available at: [http://www.census.gov/geo/www/cen\\_tract.html](http://www.census.gov/geo/www/cen_tract.html). Accessed July 16, 2001.
- Hultstrom D. *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary Version 9.1, Sixth Ed.* Springfield, Ill: North American Association of Central Cancer Registries, March 2001;106.
- McLaughlin CC. Confidentiality protection in publicly released central cancer registry data. *J Reg Mgmt*. in press.
- Sheehan T J, Gershman ST, MacDougall LA, et al. Geographic assessment of breast cancer screening by towns, zip codes, and census tracts. *J Public Health Manage Pract*. 2000;6(6):48-57.
- Automatch – [Automated Record Linkage System]. Version 4.1c. Kennebunk, Me MatchWare Technologies Inc, Copyright 1990-1995.
- Autostan – [Automated Standardizer]. Version 4.6d. Kennebunk, Me: MatchWare Technologies Inc, Copyright 1992-1997.
- Dynamap/2000. [database on CD-ROM] Version 10.3. Hanover, NH: Geographic Data Technologies Inc; 2000.
- Johnson, SD Address matching with commercial spatial data. *Business Geographics*. March 1998;24-36.
- ZIP+4 Lookup. [database online] United States Post Office. Available at: [http://www.usps.gov/ncsc/lookups/lookup\\_zip+4.html](http://www.usps.gov/ncsc/lookups/lookup_zip+4.html).
- The Census Tract Street Locator. [database online] United States Census Bureau. Available at: <http://tier2.census.gov/ctsl/ctsl.htm>.
- TIGER/Census Tract Street Index. [database on CD-ROM] Version 4. United States Census Bureau; 2001.

**Table 4.** Geocoding Resources

Source	Description	Location
New York State Department of Motor Vehicles database	Usually contains a complete physical address; useful in situations where the registry receives an incomplete or Post Office Box address.	Not publicly accessible
Freeality Internet Search	Contains links to 11 different people-search engines. This is occasionally useful since not everyone has a driver's license or owns property.	<a href="http://www.freeality.com/find.htm">http://www.freeality.com/find.htm</a>
Eagle Coder	Given an address, retrieves census tract, block group, municipality, latitude and longitude; also generates a map	<a href="http://www.geocode.com/eagle.html-ssi">http://www.geocode.com/eagle.html-ssi</a>
Ancestry.com, Roots Web	These sites search the Social Security death index and return the ZIP code of the last known place of residence.	<a href="http://www.ancestry.com">http://www.ancestry.com</a> <a href="http://www.rootsweb.org">http://www.rootsweb.org</a>
New York State Real Property Data	Contains a complete physical address	
Telephone books	Useful for older cases, since the NYSCR has a fairly complete collection dating to the 1980s	Various publishers
Paper maps	Can include new and proposed streets. Older maps are useful for older cases where a street has been renamed, for example.	Various publishers