Published in final edited form as:

Cancer Causes Control. 2017 October ; 28(10): 1095–1104. doi:10.1007/s10552-017-0941-8.

A geographic information system-based method for estimating cancer rates in non-census defined geographical areas

Vincent L. Freeman^{1,2,3}, Emma E. Boylan¹, Oksana Pugach^{1,3}, Sara L. McLafferty⁴, Katherine Y. Tossas-Milligan^{1,2}, Karriem S. Watson², and Robert A. Winn²

¹Division of Epidemiology and Biostatistics, University of Illinois at Chicago School of Public Health, 1603 W. Taylor St., Chicago, IL 60612, USA

²University of Illinois at Chicago Cancer Center, University of Illinois Hospital and Health Sciences System, 914 S. Wood St., Chicago, IL 60612, USA

³Institute for Health Research and Policy, University of Illinois School of Public Health, 1747 W. Roosevelt Road, Chicago, IL 60612, USA

⁴Department of Geography and Geographic Information Science, School of Earth, Society, and Environment, University of Illinois at Urbana-Champaign, 605 E. Springfield Ave, Champaign, IL 61820, USA

Abstract

Purpose—To address locally relevant cancer-related health issues, health departments frequently need data beyond that contained in standard census area-based statistics. We describe a geographic information system-based method for calculating age-standardized cancer incidence rates in non-census defined geographical areas using publically available data.

Methods—Aggregated records of cancer cases diagnosed from 2009 through 2013 in each of Chicago's 77 census-defined community areas were obtained from the Illinois State Cancer Registry. Areal interpolation through dasymetric mapping of census blocks was used to redistribute populations and case counts from community areas to Chicago's 50 politically defined aldermanic wards, and ward-level age-standardized 5-year cumulative incidence rates were calculated.

Results—Potential errors in redistributing populations between geographies were limited to <1.5% of the total population, and agreement between our ward population estimates and those from a frequently cited reference set of estimates was high (Pearson correlation $r = 0.99$, mean difference = -4 persons). A map overlay of safety-net primary care clinic locations and ward-level incidence rates for advanced-staged cancers revealed potential pathways for prevention.

Conclusions—Areal interpolation through dasymetric mapping can estimate cancer rates in non-census defined geographies. This can address gaps in local cancer-related health data, inform health resource advocacy, and guide community-centered cancer prevention and control.

Correspondence to: Vincent L. Freeman.

Electronic supplementary material The online version of this article (doi:10.1007/s10552-017-0941-8) contains supplementary material, which is available to authorized users.

Keywords

Local cancer health information; Areal interpolation through dasymetric mapping; Precision public health

Introduction

Precision medicine is an approach to disease treatment and prevention that takes into account individual differences in genes, environment, and lifestyle for each person [1, 2]. “Providing the right treatment to the right patient at the right time” neatly summarizes the mantra of this patient care model [3]. There is now movement among researchers, practitioners, policy makers, and healthcare advocates to extend the precision medicine paradigm to public health. Precision public health can be defined as improving the ability to prevent disease, promote health and reduce health disparities by applying emerging methods and technologies for measuring disease rates, pathogens, exposures, behaviors, and susceptibility in populations, and by developing evidence-based policies and targeted interventions to improve the health of populations [4]. As with the precision medicine initiative [2], the near-term impact of precision public health efforts could be optimized by initially focusing on cancer.

In the United States, cancer incidence and mortality rates for individual cancer sites can vary widely by geographic area. For example, mortality from colorectal cancer has been reported to be higher in the Midwest and South and lower in the West and Northeast regions, with 5-year mortality rates in West Virginia and Mississippi (19.7 and 18.6 per 100,000, respectively, 2009–2013) being nearly twice as those among residents in Utah and Connecticut (10.9 and 11.8 per 100,000, respectively, 2009–2013) [5, 6]. Yet, even similar contrasts at the level of counties can obscure cancer-related health disparities between smaller and more local populations that often do not fit neatly into standard census reporting areas. Moreover, non-census reporting areas, such as state legislative districts, precincts, and aldermanic wards, may be part of critical public health intervention pathways for cancer prevention and control and eliminating disparities. However, public health funding constraints and methodological hurdles make it difficult for local health departments to tailor the collection and reporting of population health data in ways that address locally relevant cancer health issues and inform stakeholders beyond what can be inferred from common census area-based statistics [7, 8]. In situations where public health research is centered on non-census geographic entities, a method of estimating population characteristics for these entities is necessary.

In this paper, we describe a geographic information systems (GIS)-based method for calculating cancer incidence rates in non-census reporting areas using publically available data sources for the City of Chicago, IL. Areal interpolation is the process by which data collected from one set of zonal units (termed “source” zones) can be estimated for another zonal division of the same space (termed “target” zones) that shares few or no boundaries with the former [9]. We selected Chicago’s 50 aldermanic wards as the non-census reporting geographical areas of interest for several reasons. In Chicago, aldermen are individuals

elected by their constituency to represent the needs of the ward during a four-year term. Aldermen also comprise the City Council, which constitutes the legislative branch of the City of Chicago. Their membership in the Chicago City Council grants them the authority to regulate for the protection of public health, safety, morals, and welfare among others [10]. Thus, constituents expect their ward alderman to engage in public health policy and resource advocacy on their behalf.

The source zones in our application are Chicago's 77 community areas which are standard geographic reporting areas for cancer data. Community areas are neighborhood divisions and clusters created in the 1920s by the Social Science Research Committee at the University of Chicago, and they are officially recognized by the City of Chicago for some urban planning purposes (e.g., transportation infrastructure and recreational attractions) and for describing and tracking health indicators [11, 12]. Census, health, and socioeconomic data are readily available for these community areas. The areal boundaries for community areas and aldermanic wards do not overlap, making it difficult to transfer cancer data from one set of zones to the other. Our objective is to estimate age-adjusted cancer incidence rates for each ward based on cancer data for community areas and population data for census tracts and blocks.

Materials and methods

Data sources

Cancer case records were obtained from the Illinois State Cancer Registry (ISCR) (<http://www.idph.state.il.us/cancer/statistics.htm>). ISCR is the only population-based source for cancer incidence information in Illinois. Cancer cases are collected through mandated reporting by hospitals, ambulatory surgical treatment centers, non-hospital affiliated radiation therapy treatment centers, independent pathology labs, dermatologists, and through the voluntary exchange of cancer patient data with other states. For this study, we used the Chicago Community Area subfile (CA0413.dat) from the ISCR public use dataset of cancer incidence among Illinois residents diagnosed from 1986 through 2013 [13]. The ISCR public use dataset is also available at the state, county, and ZIP code levels, and for Chicago and Cook County suburbs. We selected the Community Area subfile for this analysis because it was the only product limited to Chicago residents. ZIP codes are not perfectly coincident with the city boundary. Therefore, using them would result in estimation of the total case count, in addition to case locations within the study area. Community areas also provided better compatibility with the other data sources used in this analysis. In most Census years, tracts nest within community areas. In 2010, ISCR published a list of exceptions to this rule that they used to geocode cases. An alternative geography, the Census ZIP code tabulation area (ZCTA), was not interchangeable with ZIP code within some inhabited areas of Chicago. The public use dataset contains individual-level records of incident cancer grouped in 5-year intervals. The Chicago Community Area subfile contains incidence case counts within each of the 77 community areas from 2004 to 2008 and 2009 to 2013. Fully 99.5% of all cases were geocoded to the address level before assignment to a community area, and no community area contained less than 93% of cases geocoded to the address level.

Population counts for the city of Chicago at the census block level were obtained from the 2010 Decennial Census Summary File 1, Table P12, Total Population Sex by Age [14]. Census TIGER/Line shapefiles of the associated Census tract and block geography were obtained from the US Census Bureau's MAF/TIGER Database [15]. A shapefile of the Chicago community area and ward boundaries as of May 2015 was obtained from the City of Chicago Data Portal [16, 17].

Cases

Cases were included if they were diagnosed with invasive cancer between 2009 and 2013 ($n = 56,088$). Variables included community area, diagnosis year grouped in 5-year period, Surveillance, Epidemiology, End Results (SEER) Program cancer site group, age group at diagnosis, sex, race group, SEER summary disease stage, Hispanic ethnicity, and geocode precision. Reported age groups were <5, 5–14, 15–34, 35–44, 45–54, 55–64, and 65 years. Race codes included white, black, other, and unknown.

Selection of census geography

For most tracts within Chicago, the community area number is given by the first two digits of the tract ID. However, boundary changes and new tracts introduced in the 2010 Census created multiple exceptions to this rule. In addition, at least one tract in 2010 was split between the city of Chicago and unincorporated Cook County.

New tracts and tracts intersected by city or community area boundaries were assigned based on the documentation for the ISCR public use data file [13]. The documentation includes a table describing how new tracts or, if necessary, blocks, were assigned to community areas. No blocks had their populated area split by a city or community area boundary. Since these tracts also define the reporting area in the public use dataset, other tracts were excluded. Tracts included in the public use dataset were not allowed to be split by the city boundary during the following steps.

Areal interpolation of the at-risk population in each aldermanic ward

The method below describes a two-step procedure for constructing population characteristics of the 50 Chicago Aldermanic Wards by aggregating population characteristics at the U.S. census tract level. Tracts are an appealing unit of analysis for this task because they nest within other census geography and are smaller than many non-census reporting areas of interest, including wards. They also offer good availability of population characteristics such as marital status and language that are of interest to disparities researchers but that are suppressed in blocks due to their smaller population [18]. Once a population density weight is built for tract/ward piece, it can be reused to estimate any tract-level population characteristic.

A U.S. census tract is a smaller geographic unit than a Chicago aldermanic ward. However, the straightforward summing up of the population of U.S. census tracts was not reasonable due to lack of agreement between ward and tract boundaries, which resulted in a high percentage of census tracts being located in more than one aldermanic ward. The role of each areal unit in the final analysis is summarized in Supplemental Table A.

Step 1—estimation of population density weights

Blocks-to-ward matching

The smallest geographic area provided by the U.S. Census is a block. There is a clear correspondence between block and census tract geographic entities, such that multiple blocks are uniquely contained in one census tract. Therefore, to minimize approximation error while aggregating census tracts to the ward level, we used blocks-to-ward matching. ArcMap 10.3.1 (Esri, Redlands, CA, USA) was used to combine the block and ward shapefiles. The intersect function was used to output a single map of Chicago with a boundary for each block/ward combination. Block/ward combinations whose areas made up less than 1% of the total block area were assumed to be the result of spatial error and removed from further analysis. Under the assumption of a homogeneous population distribution within a census block, the area-weighted population of the block/ward combination was estimated as follows:

$$\hat{P}_{B/W} = \frac{A_{B/W}}{A_B} P_B, \quad (1)$$

where $\hat{P}_{B/W}$ is the estimated population of the block/ward combination; P_B is the population of the census block; $A_{B/W}$ is the area of the block/ward combination; and A_B is the block area. The fact that blocks are small in size (the median size of populated blocks in Chicago is 0.01 km².) means that errors arising from the assumption that population is uniformly distributed within blocks will be minimized.

Block-to-within-ward census tract matching

The estimated population of census tracts within the ward was calculated from the area-weighted population of constituent block/ward combinations as follows:

$$\hat{P}_{T/W} = \sum_{\text{within tract/ward}} \hat{P}_{B/W} = \sum_{\text{within tract/ward}} \frac{A_{B/W}}{A_B} P_B, \quad (2)$$

where $\hat{P}_{T/W}$ is the estimated population of the tract/ward combination, and summation is over the blocks contained in the appropriate tract/ward combination. This method takes advantage of the fact that blocks are smaller than tracts and less likely to be split by ward boundaries. It also avoids assigning population to non-residential areas of tracts. This is essentially the binary dasymetric method for interpolating population data between incompatible spatial zones in which unpopulated areas were excluded from the calculations [9].

Population weight of each tract/ward combination

The population weight of each tract/ward combination was estimated as the ratio of tract/ward combination population to total tract population (P_T):

$$\hat{w}_{T/W} = \frac{\hat{P}_{T/W}}{P_T}. \quad (3)$$

The SQL procedure in SAS, v. 9.4 (SAS Institute Inc., Cary, NC, USA) was used to join the relationship file to population counts and perform all data manipulation. The resulting file was a crosswalk containing tract ID, ward number, and estimated population weight of each tract/ward combination. The estimated population weight was used to obtain other tract-level population characteristics at the ward level.

Step 2—estimation of ward populations

The estimated aldermanic ward population $\hat{P}_W$ was determined by summing over all tracts within each ward using the estimated populations of the ward/tract combinations:

$$\hat{P}_W = \sum_{\text{within ward}} \hat{P}_{T/W} = \sum_{\text{within ward}} \sum_{\text{within tract/ward}} \hat{P}_{B/W} = \sum_{\text{within ward}} \sum_{\text{within tract/ward}} \frac{A_{B/W}}{A_B} P_B. \quad (4)$$

The procedure described above was repeated to estimate the population by age group of each ward.

Due to block splitting and changes in official ward boundaries, there is no gold standard source for ward population totals. A frequently cited set of ward population estimates produced by Elliott Ramos for WBEZ Chicago was generated by joining census block centroids to wards [19]. Agreement with the estimates from this block centroid assignment-based method was evaluated in RStudio using Bland–Altman plots and Pearson correlation coefficients [20]. The BlandAltmanLeh package for R was used to produce the plots [21].

Attributing cancer case counts to aldermanic wards using block population weights

Next, we extended the method used to estimate the ward population denominators to estimate the corresponding numerators from the aggregated case count data. To attribute case counts to their respective aldermanic wards, we assumed that the distribution of cancer cases within community areas was non-uniform, and that the non-uniform distribution of cases reflected the non-uniform distribution of the underlying at-risk population within the selected census geographies. Therefore, we created a population weight reflecting the proportion of a community area's total population living in each block or block/ward combination. This population weight equals the share of total community area cases assigned to each area.

Block or block/ward combination case counts were summed up to the ward level as described in steps 1 and 2 above. The procedure was repeated restricted to advanced-stage

cases ($n = 27,231$, 48.5% of all cases), defined as cases with regional or distant-stage disease.

Calculation of age-standardized ward cancer incidence rates and 95% confidence intervals

The STD RATE procedure, SAS 9.4, was used to calculate age-standardized overall and advanced-stage cancer risk by ward in Chicago. Incidence was directly standardized to the total US population in 2010 to facilitate comparisons with state and national data, and 5-year (2009–2013) cumulative incidence rates within each ward were calculated. The STD RATE procedure generates 95% confidence intervals calculated using the Tiwari method and the Fay and Feuer method, with the latter tending to yield slightly more conservative (i.e., wider) intervals [22, 23]. However, it must be noted that the interpolation described above is a deterministic estimation method, as opposed to a probabilistic estimation method, for which no estimate of the errors is available. Consequently, we report only the more conservative 95% confidence intervals from the Fay and Feuer method.

Results

Lack of agreement between community area and ward boundaries

Figure 1 shows the Chicago community areas with an overlay of the aldermanic ward boundaries. The overlay reveals substantial lack of agreement between community area boundaries. This illustrates the degree of incompatibility between the spatial zones and further justifies interpolating the ward rates directly rather than inferring them from the publically available community area rates.

Using census blocks instead of tracts to redistribute populations from community areas to wards minimized approximation errors

Table 1 summarizes the number of census tracts and blocks split by ward boundaries and the size of the affected population. Over 40% of census tracts were split by a ward boundary, affecting 46% of the Chicago population, meaning that the ward location for 46% of the total population had to be assigned using area-based population density weights. In contrast, only 1.1% of census blocks were split by a ward boundary, affecting only 1.4% of the Chicago population. Thus, using census blocks minimized the need to apply area-based population density weights to assign community area populations to the correct ward, and in doing so, helped minimize potential approximation errors during interpolation. Agreement with the set of ward population estimates produced for WBEZ Chicago that utilized the census block centroid assignment method [19] was high (Pearson correlation $r = 0.99$, average difference [95% confidence interval] = -4 persons [-39 to 31], Supplemental Fig. A), with the total Chicago population estimates differing by only 205 (.008%).

Median total cancer incidence rates were similar, but there were incidence rate patterns at the aldermanic level that were not apparent at the community area level

Figure 2 shows 5-year age-adjusted cumulative incidence rates expressed as quartiles for all cancers combined across Chicago community areas (panel A) and across the politically designated aldermanic wards (panel B). The median rate at the community area level and aldermanic ward level was essentially the same (254.1 per 10,000 [95% CI 239.8–263.0] and

249.2 per 10,000 [95% CI 233.9–263.6], respectively). This suggests that case counts and source populations were successfully interpolated between the two geographies; their rate ranges may differ, but their medians should be similar. Both maps also show a disproportionate burden of incident cancers on the south and far west sides of Chicago, which are predominantly African-American and/or socioeconomically disadvantaged regions of the city. Nevertheless, there are incidence rate patterns at the aldermanic ward level that are not evident at the community area level. For example, the aldermanic ward with the highest incidence rate was ward 32 (Supplemental Table B). However, it is derived from six community areas that have consistently lower incidence rates (Supplemental Fig. B and Table B, community areas 5, 6, 7, 21, 22, and 24). Thus, the interpolation procedure appears to have located a population that is shared by the six community areas but whose rate ranks among the highest in the city.

Relating ward-level cancer incidence rates to resources for cancer prevention and control

We examined the geographic distribution of ward-level cancer incidence rates and locations of safety-net clinics, limiting the analysis to advanced-staged cancers and facilities offering primary care. Figure 3 shows 5-year age-adjusted cumulative incidence rates for advanced-stage cancers at the ward level and locations of the 103 federally qualified health centers (FQHCs) in Chicago that offer adult primary care services. As with total incidence, the wards with the highest rates of advanced-staged cancer tend to be located on the south and far west regions of the city. The map also shows FQHCs in nearly all of these higher-risk wards, drawing attention to potential intervention pathways for improved cancer prevention and control that could involve residents, the alderman, a local health care network, and health providers.

Discussion

Technical limitations are a frequently cited barrier to the development and dissemination of local health information [8, 12, 24–26]. For example, much of the data that are available are reported for geographic areas that do not provide the local-level information needed for public health intervention. Changing local boundaries, as is frequently the case with political and administrative areas, complicates measuring and tracking relevant characteristics of the underlying population. Small numbers issues and the timeliness of reporting are additional technical challenges. For example, cancer incidence rates in small local populations can fluctuate greatly from year-to-year, and the lag-time between a cancer diagnosis and when information about it is available to local health departments can be up to two years or more [27]. Such factors impede both local cancer-related health surveillance and evaluation of population-based intervention programs.

In this paper, we have demonstrated how a widely used areal interpolation method can be applied to estimate cancer rates within non-census defined, but locally relevant, geographical areas. The method has several advantages. It is (1) easy to implement using GIS software, (2) uses existing census-defined area data, and (3) can be customized to specific cancer-related health surveillance and reporting needs of local health departments and the communities they serve. The method is applicable when estimates of population at risk are

needed for non-census reporting areas, and the actual case locations are not known, which are situations frequently encountered in public health practice. Moreover, the method can accommodate situations where local area boundaries change over time due to environmental, sociodemographical, political, and/or other reasons. Thus, it can offer greater flexibility and precision in what local health information is reported and for whom.

Areal interpolation methods have been extended in many ways to provide even more accurate estimates of population and health statistics for alternative geographies, and these methods have great potential to inform public health intervention. For example, areal interpolation methods can incorporate ancillary data on roads, vegetation, housing, and other characteristics to better estimate geographic variation in population and related socioeconomic characteristics [28–30]. In distributing cancer cases to blocks, one can take into account demographic characteristics such as age, gender, and race to reflect fine-grained spatial and social variation in population at risk [31, 32]. Similarly, methods of small area estimation can and are being used to produce reliable health indicator estimates based on individual demographic factors and local environmental characteristics [33]. Furthermore, researchers have applied Monte Carlo methods to assess the accuracy of cancer rates estimated via areal interpolation [32, 34].

The main advantage of the methods employed here compared to these other spatial disaggregation approaches is their ease of implementation and limited data requirements. Our population density-based method relies on widely available census data for different geographies, and it is straightforward to implement within GIS. Analysts with basic knowledge in GIS should be able to apply the method effectively to create cancer maps based on alternative geographies. Although alternative approaches include more sophisticated modeling of small area variation in cancer risk, those approaches require detailed data, software, and statistical or epidemiological expertise that may not be readily available in local public health departments.

In this study, the use of census blocks to redistribute the populations of the community areas to the ward geography limited potential redistribution errors to less than 1.5% of the total population. The resulting ward population estimates (the ‘denominator’) also agreed with a commonly referenced set of estimates within approximately one-hundredth of a percent on average. The accuracy of the ward assignments for the cancer case counts (the ‘numerator’) is more difficult to assess, because the counts originated as aggregated data. However, we assumed that cancer cases would be non-uniformly distributed within community areas as a function of the non-uniform distribution of the underlying populations within census blocks and tracts, and weighted the ward assignments of case counts accordingly; both of which are reasonable strategies. Also, aggregated count data are commonly encountered in public health practice, and tools that help practitioners extract as much information that is validly possible while protecting privacy would be welcomed.

Public health implications

Intervention pathways

Areal interpolation may not always lead to a different understanding of geographical disparities in disease burden, perhaps especially in large urban areas characterized by high levels of residential segregation by race/ethnicity and class, geographical disparities in potential access to health services, and other socio-structural determinants of health. However, the process may reveal a different and potentially more effective public health intervention pathway. For example, the geographical distribution of incident cancers across Chicago community areas was similar to that across aldermanic wards, with both showing a disproportionate share of cases on the south and west regions of the city. To address the regional disparities, it may be more efficacious to engage the aldermen of the affected wards, since they can, and are expected to, advocate for healthcare infrastructure and services for their constituents. Also, even though nearly all of the wards with the highest cancer rates contain FQHCs offering primary care, the fact remains that rates are disproportionately high in these wards. Thus, the descriptive data derived from this areal interpolation method can not only drive health resource advocacy but help guide health system services planning as well.

Another example illustrating the public health practice implications of this method comes from the south side of Chicago. Community area 47 has a very high cancer incidence rate but is absorbed into a large area of moderately high incidence after reallocation to aldermanic wards (Supplemental Fig. B). This is because the 2010 population of community area 47 was 2,916 while the estimated 2010 population of the 8th ward, in which it is contained, was 51,684. While community areas have the advantage of stable boundaries over time, they vary widely in both geographic area and total population. The wards, with an estimated mean population of 53,908, would be expected to have more stable rates than many community areas, because they are required by law to have similar total populations. Since the wards are politically defined, case counts at the ward level may be more comparable to each other demographically than those at the community area level. In this case, both the community area and ward map provide valuable information to local stakeholders. The overall cancer incidence rate in the 8th ward is moderately high compared to the city of Chicago and broadly similar to its south side neighbors. If cancer prevention and control is a priority for 8th ward residents, interventions may be more effective if they are focused on residents of community area 47.

In conclusion, areal interpolation through dasymetric mapping of census blocks can be used to estimate age-standardized cancer rates geographies that the user deems to be locally relevant but are not census-defined. For a local health department, having the capacity to develop descriptive data that pinpoints susceptible populations and communities affected by cancer-related health disparities would facilitate cancer prevention and control interventions. Privacy restrictions and concerns may ultimately limit the precision to which locally important health information can be reported [35]. Nonetheless, this method addresses an important technical barrier to flexible and responsive reporting of cancer risk in local communities and populations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Richard E. Barrett, PhD of the Department of Sociology at the University of Illinois at Chicago for his substantive knowledge, and Tiefu Shen, PhD of the Division of Epidemiologic study, Illinois Department of Public Health, for his comments and technical assistance.

Funding This research was supported by a research project grant from the American Cancer Society (117534-RSGT-09-286-01-CPHPS) (VLF) to the University of Illinois at Chicago and the University of Illinois at Chicago Cancer Center and (VLF).

References

1. Desmond-Hellman, S., Sawyers, CL. Committee on a framework for development a new taxonomy of disease; National Research Council. Why now?. In: National Research Council. , editor. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. National Academies Press; Washington, DC: 2011. p. 21-39.
2. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015; 372:793–795. [PubMed: 25635347]
3. [Accessed 14 Jan 2017] <https://www.whitehouse.gov/the-press-office/2015/01/30/remarks-president-precision-medicine>
4. Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. *Am J Prev Med*. 2016; 50:398–401. [PubMed: 26547538]
5. Naishadham D, Lansdorp-Vogelaar I, Siegel R, Cokkinides V, Jemal A. State disparities in colorectal cancer mortality patterns in the United States. *Cancer Epidemiol Biomark Prev*. 2011; 20:1296–1302.
6. U.S. Cancer Statistics Working Group. United States cancer statistics: 1999–2013 incidence and mortality web-based report. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute: Atlanta: 2016. www.cdc.gov/uscs
7. Himmelstein DU, Woolhandler S. Public health's falling share of US health spending. *Am J Public Health*. 2016; 106:56–57. [PubMed: 26562115]
8. Simon PA, Wold CM, Cousineau MR, Fielding JE. Meeting the data needs of a local health department: the Los Angeles County Health Survey. *Am J Public Health*. 2001; 91:1950–1952. [PubMed: 11726372]
9. Fisher PF, Langford M. Modeling sensitivity to accuracy in classified imagery: a study of areal interpolation by dasymetric mapping. *Prof Geogr*. 1996; 48:299–309.
10. City of Chicago City Council. [Accessed 29 Sept 2016] Description of City of Chicago aldermanic roles. 2016. <http://www.cityofchicago.org/city/en/about/council.html>
11. Chicago Fact Book Consortium. Local community fact book: Chicago metropolitan area, 1990. Academy Chicago Publishers; Chicago: 1995.
12. Shah AM, Whitman S, Silva A. Variations in the health conditions of 6 Chicago community areas: a case for local-level data. *Am J Public Health*. 2006; 96:1485–1491. [PubMed: 16809605]
13. Illinois Department of Public Health. [Accessed 23 Jan 2017] Illinois State Cancer Registry, public data set v23, 1986–2013. data as of November 2015. <http://www.idph.state.il.us/cancer/16/READMEv23ChicagoCA.pdf>
14. [Accessed 23 Jan 2017] U.S. Decennial Census Summary File 1, Table P12, Total Population Sex by Age. http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_00_SF1_P012&prodType=table
15. [Accessed 23 Jan 2017] TIGER/Line tracts and blocks. <https://www.census.gov/geo/maps-data/data/tiger-line.html>

16. Chicago Ward Boundaries. [Accessed 29 Sept 2016] 2016. <https://data.cityofchicago.org/Facilities-Geographic-Boundaries/Boundaries-Wards-2015-/sp34-6z76>
17. Chicago Community Area Boundaries. [Accessed 23 Jan 2017] 2017. <https://data.cityofchicago.org/Facilities-Geographic-Boundaries/Boundaries-Community-Areas-current-/cauq-8yn6>
18. United States Census Bureau. [Accessed 3 Feb 2017] American Community Survey, Data Suppression Practices, September 27, 2016. http://www2.census.gov/programs-surveys/acs/tech_docs/data_suppression/ACSO_Data_Suppression.pdf
19. Ramos, E. Interactive: City Council-approved Chicago Ward Map [Internet]. WBEZ Chicago; 2017. <https://www.wbez.org/shows/wbez-news/interactive-city-council-approved-chicago-ward-map/52630c78-06ad-4e98-8aab-05bd3ca72e24> [Accessed 21 Mar 2017]
20. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 327:307–310.
21. Lehnert, B. [Accessed 20 Mar 2017] Package “BlandAltmanLeh” [Internet]. 2017. 2015. <https://cran.r-project.org/web/packages/BlandAltmanLeh>
22. Tiwari RC, Clegg LX, Zou Z. Efficient interval estimation for age-adjusted cancer rates. *Stat Methods Med Res*. 2006; 15:547–569. [PubMed: 17260923]
23. Fay MP, Feuer EJ. Confidence intervals for directly standardized rates: method based on the gamma distribution. *Stat Med*. 1997; 16:791–801. [PubMed: 9131766]
24. Luck J, Chang C, Brown ER, Lumpkin J. Using local health information to promote public health. *Health Aff*. 2006; 25:979–991.
25. IOM (Institute of Medicine). For the public’s health: the role of measurement in action and accountability. The National Academies Press; Washington, DC: 2011.
26. Castucci BC, Rhoades EK, Leider JP, Hearne S. What gets measured gets done: an assessment of local data uses and needs in large urban health departments. *J Public Health Manag Pract*. 2015; 21(1 Supp):S38–S48. [PubMed: 25423055]
27. Clegg LX, Feuer EJ, Midthune DN, Fay MP, Hankey BF. Impact of reporting delay and reporting error on cancer incidence rates and trends. *J Natl Cancer Inst*. 2002; 94:1537–1545. [PubMed: 12381706]
28. Holt JB, Lo CP, Hodler TW. Dasymetric estimation of population density and areal interpolation of census data. *Cartogr Geogr Inf Sci*. 2004; 31:103–121.
29. Maantay JA, Maroko AR, Herrmann C. Mapping population distribution in the urban environment: the cadastral-based expert dasymetric system (CEDS). *Cartogr Geogr Inf Sci*. 2007; 34:77–102.
30. Reibel M. Geographic Information Systems and spatial data processing in demography: a Review. *Popul Res Policy Rev*. 2007; 26:601–618.
31. Curriero FC, Kulldorff M, Boscoe FP, Klassen AC. Using imputation to provide location information for nongeocoded addresses. *PLoS ONE*. 2010; 5:e8998. [PubMed: 20161766]
32. Luo L, McLafferty S, Wang F. Analyzing spatial aggregation error in statistical models of late-stage cancer risk: a Monte Carlo simulation approach. *Int J Health Geogr*. 2010; 9:51. [PubMed: 20959015]
33. Zhang X, Holt JB, Yun S, Lu H, Greenlund KJ, Croft JB. Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the Behavioral Risk Factor Surveillance System. *Am J Epidemiol*. 2015; 182(2):127–137. [PubMed: 25957312]
34. Henry K, Boscoe FP. Estimating the accuracy of geographical imputation. *Int J Health Geogr*. 2008; 7:3. [PubMed: 18215308]
35. Hodge JG Jr. Health information privacy and public health. *J Law Med Ethics*. 2003; 31:663–671. [PubMed: 14968668]

Fig. 1. Chicago Community Area Boundaries versus Aldermanic Ward Boundaries. Population range (median) for community areas and aldermanic wards was 2,876 to 98,514 (31,028) and 51,428 to 56,176 (53,940), respectively

Fig. 2.

5-year age-adjusted cumulative incidence rates for all cancers combined by Chicago community area (**a**) and by aldermanic ward (**b**), 2009–2013. Rates are age-adjusted to the U.S. 2010 standard population, and the incidence rate ranges correspond to first (lowest), second, third, and fourth incidence rate quartiles for community areas. The community area rate quartiles were applied to aldermanic wards to facilitate comparisons between the two maps. The ranges for the aldermanic ward incidence rate quartiles were 189.1 to 234.8, 234.9 to 249.3, 249.4 to 268.0, and 268.1 to 295.9 per 10,000 for the first through fourth quartiles, respectively. The median rate for community areas and aldermanic wards was 254.1 per 10,000 (95% CI 239.8, 263.0) and 249.2 per 10,000 (95% CI 233.9, 263.6), respectively

Fig. 3.

5-year age-adjusted cumulative incidence rates for advanced-stage cancers by Chicago aldermanic ward, 2009–2013, and locations of federally qualified health centers offering adult primary care services ($n = 103$). Rates are age-adjusted to the U.S. 2010 standard population, and the incidence rate ranges correspond to first (lowest), second, third, and fourth incidence rate quartiles, respectively

Table 1

Number of census tracts and census blocks split by ward boundaries and affected population sizes

Census area ^a	<i>n</i>	Number of areas split by ward boundaries	Affected population ^b	Average number affected per census area
Tracts	794	343 (43.2%)	1,248,117 (46.3%)	3,639
Blocks	36,875	405 (1.1%)	36,875 (1.4%)	91

^aUninhabited areas were excluded^bThe total interpolated population of the City of Chicago in 2010 was 2,695,393