Cancer Epidemiology 43 (2016) 56–62

Contents lists available at ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.cancerepidemiology.net

Impact of geographic area level on measuring socioeconomic disparities in cancer survival in New South Wales, Australia: A period analysis

Julia F. Stanbury[a,b,*], Peter D. Baade[c,d], Yan Yu[b], Xue Qin Yu[a,b]

[a] Sydney School of Public Health, The University of Sydney, Sydney, Australia
[b] Cancer Research Division, Cancer Council New South Wales, Sydney, Australia
[c] Cancer Research Centre, Cancer Council Queensland, Brisbane, Australia
[d] School of Public Health and Social Work, Queensland University of Technology, Brisbane, Australia

ARTICLE INFO

Article history:
Received 24 September 2015
Received in revised form 12 May 2016
Accepted 2 June 2016
Available online 6 July 2016

Keywords:
Cancer
Survival analysis
Socioeconomic variation
Disparity
Geographic
Area level

ABSTRACT

Background: Area-based socioeconomic measures are widely used in health research. In theory, the larger the area used the more individual misclassification is introduced, thus biasing the association between such area level measures and health outcomes. In this study, we examined the socioeconomic disparities in cancer survival using two geographic area-based measures to see if the size of the area matters.
Methods: We used population-based cancer registry data for patients diagnosed with one of 10 major cancers in New South Wales (NSW), Australia during 2004–2008. Patients were assigned index measures of socioeconomic status (SES) based on two area-level units, census Collection District (CD) and Local Government Area (LGA) of their address at diagnosis. Five-year relative survival was estimated using the period approach for patients alive during 2004–2008, for each socioeconomic quintile at each area-level for each cancer. Poisson-regression modelling was used to adjust for socioeconomic quintile, sex, age-group at diagnosis and disease stage at diagnosis. The relative excess risk of death (RER) by socioeconomic quintile derived from this modelling was compared between area-units.
Results: We found extensive disagreement in SES classification between CD and LGA levels across all socioeconomic quintiles, particularly for more disadvantaged groups. In general, more disadvantaged patients had significantly lower survival than the least disadvantaged group for both CD and LGA classifications. The socioeconomic survival disparities detected by CD classification were larger than those detected by LGA. Adjusted RER estimates by SES were similar for most cancers when measured at both area levels.
Conclusions: We found that classifying patient SES by the widely used Australian geographic unit LGA results in underestimation of survival disparities for several cancers compared to when SES is classified at the geographically smaller CD level. Despite this, our RER of death estimates derived from these survival estimates were generally similar for both CD and LGA level analyses, suggesting that LGAs remain a valuable spatial unit for use in Australian health and social research, though the potential for misclassification must be considered when interpreting research. While data confidentiality concerns increase with the level of geographical precision, the use of smaller area-level health and census data in the future, with appropriate allowance for confidentiality

1. Introduction

Many published studies which report socioeconomic disparities in cancer survival use area-based measures of socioeconomic status (SES) [1,2]. Individual-level demographic data is preferable and most accurate, but often very difficult to obtain in population-based studies. Instead, these "ecological" studies use census-derived area-based measures of SES to classify patients based on characteristics of the aggregate population of the area in which

Abbreviations: SES, Socioeconomic status; NSW, New South Wales; RER, Relative excess risk; LGA, Local Government Area; CD, Census District; ABS, Australian Bureau of Statistics.

* Corresponding author at: Cancer Research Division, Cancer Council New South Wales, P.O. Box 572, Kings Cross, NSW 1340, Australia.
E-mail addresses: julias@nswcc.org.au, jsta6827@uni.sydney.edu.au, juliastanbury@gmail.com (J.F. Stanbury), PeterBaade@cancerqld.org.au (P.D. Baade), yu1985yan@gmail.com (Y. Yu), xueqiny@nswcc.org.au (X.Q. Yu).

http://dx.doi.org/10.1016/j.canep.2016.06.001
1877-7821/

they live. Misclassification of individuals may result depending on the extent of variation within the population of a specific area [3]. Small spatial areas are known to represent more socioeconomically homogeneous populations compared to larger areas, primarily due to their smaller resident population and so their socioeconomic index values are more likely to accurately represent the characteristics of that population [4,5].

We previously reported socioeconomic disparities in cancer survival in New South Wales (NSW), Australia using Local Government Areas (LGA) to classify cases by SES (*Submitted paper BMC Cancer 2015*). LGAs are a valuable and widely used spatial unit in Australian health and social research, since data are readily available at this level. Compared to other spatial units in Australia, LGAs are considered to be 'relatively' small. However the use of LGAs in ecological studies has been criticised due to the inherent population heterogeneity within each LGA introducing potential misclassification of individuals [6]. It is unknown to what extent this misclassification may occur and what impact it may have on research results. Similar misclassification effects have been observed in previous studies where area-based geographic units have been used [3,5]. Consequently, the true disparities in cancer survival in NSW may vary from those previously reported.

Cancer incidence data from the NSW Central Cancer Registry has recently become available at the smaller area unit of census Collection District (CD), the smallest area unit for which a measure of SES is available [7]. Comparing analyses of LGA and CD geocoded data will be able to more accurately detect and identify the extent to which cancer cases may be misclassified according to SES when investigating cancer survival disparities. To date, few studies have used population-based data to compare cancer survival disparities between area level measures [4,5,8]. This study aims to compare the area units of CD and LGA to quantify the extent to which cancer patients could be misclassified by SES between these two area-levels and the impact of such misclassification on estimating socioeconomic disparity in patient survival, with specific reference to cancer survival data in NSW in 2004–2008.

2. Methods

Data were obtained from the population-based NSW Central Cancer Registry for all patients aged 15–89 at diagnosis of a first primary cancer between January 1999 and December 2008 that were prevalent cases between 2004 and 2008. Notification of a cancer diagnosis to the Registry is mandatory in NSW since 1972. We chose ten cancers for analysis as defined by International Classification of Diseases for Oncology 3rd Edition [9] codes (see Table 2). These cancers were chosen based on their high incidence and large contribution to population mortality. Cases were linked to records from the NSW State Registry of Births, Deaths and Marriages and the National Death Index and followed up to 31 December 2008 for survival status. Cases were excluded if notified to the registry by death certificate only or first identified at post-mortem.

LGAs in NSW range from small urban areas with large populations to extremely large rural areas with small populations. In 2001 there were 175 LGAs in NSW, each with an average population of 35,954 (IQR: 4713–43,809) [Australian Bureau of Statistics (ABS) online data 2001]. Comparatively, CDs were the smallest area units used by the ABS at the time of the study period (2004–2008) [10], and represent a more socioeconomically homogenous population than LGAs [7]. In 2001, NSW contained 11,510 CDs, each containing about 200 'dwellings' or an average population of 547 residents (IQR: 369–696) [ABS data 2001].

Individual SES was measured by the 2001 ABS Index of Education and Occupation score where a score indicates a relatively high level of educational attainment and skilled employment [10]. This index also allows us to maintain comparability with previous studies of SES and cancer survival in NSW [11]. Two versions of this measure were used – the first aggregated by CD and the second LGA. Residential address information collected by the Registry at diagnosis was used to assign cases to their CD and LGA, and corresponding SES quintiles. Cases were excluded from analysis if they had insufficient information to assign a CD or LGA or if index scores were not available.

Stage of disease at diagnosis was based on pathology reports and statutory notifications by hospitals, coded using a modified summary classification: localised (stage I), regional (a combination of stages II and III), distant (stage IV) and unknown stage.

2.1. Statistical analysis

Relative survival was used to estimate net survival in this study, which is the ratio of the observed proportion of people surviving 5 years in a group of cancer patients, to the expected proportion of people who would have survived in a comparable group (same age and sex distribution); in this case the general population. Observed survival for each case was calculated from the month of diagnosis to the month of death or censoring (31 December 2008) using life-table methods and relative survival was calculated using the Pohar-Perme approach [12]. We constructed state-wide life tables by sex for each calendar year 2004–2008 using NSW all-cause mortality data and corresponding population data.

We used the same analysis strategy for both CD and LGA classification and then compared the two sets of results. Five-year relative survival by SES quintile was calculated for each cancer using the period approach, which focuses on the survival observed among a specified cohort of patients during a recent time interval (in this case, 2004–2008) and thus provides a more recent estimate of patient survival [13]. Using this method, the survival calculated is a product of the short-term survival experience of patients diagnosed more recently, combined with the longer-term survival experience of patients diagnosed earlier in the study period.

We then investigated the effect of SES on survival time for each cancer using multivariate models to adjust for potentially confounding variables. We used a Poisson-regression model to calculate the relative excess risk (RER) of death due to cancer, after controlling for the other factors included in the model [14]. The RER is the ratio of excess risk of death in a particular SES quintile compared to that of the reference (least disadvantaged) SES group, after controlling for the other factors. In this model, the main-effect variables were SES quintile, age group at diagnosis (<50 years, 50–59 years, 60–69 years, 70–79 years, 80–89 years), sex, stage of disease at diagnosis and year of follow-up (1–5 years), with the natural logarithm of the population size as the offset. The estimated coefficients and standard errors from the Poisson model were used to calculate ninety-five percent confidence intervals (CIs) for the RERs.

All significance tests with p-value <0.05 were taken to indicate statistical significance. All statistical analysis was completed using STATA v13.1 software (StataCorp LP: College Station, TX).

3. Results

A total of 236 690 cases were diagnosed with one of the ten cancers between 1999 and 2008 that were identified from the Registry. 944 cases (0.4%) were excluded due to registry notification by death certificate only or first identified post-mortem. A further 1879 cases were excluded due to missing socioeconomic data. The final cohort used for analyses contained 176 322 cases who were prevalent (alive) at some time during 2004–2008.