Spatial inequalities in colorectal and breast cancer survival: premature deaths and associated factors

Susanna M Cramb, Kerrie L Mengersen, Gavin Turrell, Peter D Baade

Abstract

This study examines the influence of cancer stage, distance to treatment facilities and area disadvantage on breast and colorectal cancer spatial survival inequalities. We also estimate the number of premature deaths after adjusting for cancer stage to quantify the impact of spatial survival inequalities. Population-based descriptive study of residents aged <90 years in Queensland, Australia diagnosed with primary invasive breast (25,202 females) or colorectal (14,690 males, 11,700 females) cancers during 1996-2007. Bayesian hierarchical models explored relative survival inequalities across 478 regions. Cancer stage and disadvantage explained the spatial inequalities in breast cancer survival, however spatial inequalities in colorectal cancer survival persisted after adjustment. Of the 6,019 colorectal cancer deaths within 5 years of diagnosis, 470 (8%) were associated with spatial inequalities in non-diagnostic factors, i.e. factors beyond cancer stage at diagnosis. For breast cancers, of 2,412 deaths, 170 (7%) were related to spatial inequalities in non-diagnostic factors. Quantifying premature deaths can increase incentive for action to reduce these spatial inequalities.

Keywords: cancer, premature deaths, relative survival, spatial, Bayesian

Background

Worldwide, breast cancer is the most common cancer in women, while colorectal cancer is the second most commonly diagnosed among women, and third most common among men (Ferlay et al., 2010). In developed nations, including Australia, survival for both these cancers has improved over recent decades (Australian Institute of Health and Welfare and Cancer Australia & Australasian Association of Cancer Registries, 2008), with Australia having one of the highest survival rates in the world (Coleman et al., 2011).

However, the improvement in survival has not been observed equally across all population subgroups. Inequalities for both breast and colorectal cancer survival have been reported by deprivation and differences in health care access (Du et al., 2011; McKenzie et al., 2011). Within Australia, poorer survival has been observed for those in areas of greater socio-economic disadvantage, geographic remoteness and, for rectal cancer, further distance to radiotherapy facilities (Australian Institute of Health and Welfare and Cancer Australia & Australasian Association of Cancer Registries, 2008; Baade et al., 2011b; Cramb et al., 2011).

The quality of patient management can be gauged by survival (Yu et al., 2004). The prognosis for breast and colorectal cancer depends in large part on the stage of disease at diagnosis (Schottenfeld and Fraumeni Jr, 2006), which may vary geographically (Tian et al., 2012; Tian et al., 2011). Beyond that, the outcome depends on other non-diagnostic factors such as treatment, rehabilitation, environmental factors such as area disadvantage, and patient characteristics including comorbidities (Yu et al., 2005a), all of which could potentially contribute to geographical variation in cancer survival. Throughout this paper we use the term “non-diagnostic” to encompass these other factors.

Since only a few population-based cancer registries collect stage information, not many studies have been able to separate the effect of diagnostic from other factors on geographic inequalities in cancer survival on a population basis. In New South Wales (NSW), Australia, it was found that adjusting for stage did not reduce the survival differential for colorectal cancer (Yu et al., 2005a). However, in Italy, stage at diagnosis explained most of the colorectal cancer survival inequalities between Northern and Southern areas, while treatment had a minimal role (Fusco et al., 2010). In England, stage at diagnosis and deprivation were important causes of breast cancer survival inequalities (Davies et al., 2010).

However these previous studies have used relatively large geographical regions, which reduce the ability to measure spatial variation and can limit interpretation because of the greater heterogeneity within those regions. In contrast, inequalities in cancer survival at the small-area level have rarely been examined, typically due to difficulties associated with sparse data in small geographical areas and in accounting for the spatial correlation between neighboring areas (Wakefield and Elliott, 1999). Bayesian hierarchical methods overcome both problems by incorporating information from neighboring areas for each estimate, producing more reliable small-area estimates (Carlin and Xia, 1999).

Spatial survival analysis is an emerging field. Most analyses have focused on cause-specific survival (Henry et al., 2009; Huang et al., 2007; Wan et al., 2012). We chose to instead use Bayesian hierarchical methods to model relative survival (Fairley et al., 2008; Saez et al., 2012), where cancer patient mortality is compared against mortality in the population of similar age, sex and time period. Our focus was on comparing survival up to 5-years after diagnosis.

To quantify the impact of spatial inequalities in cancer survival, previous studies have calculated the number of deaths that could have been prevented within a given timeframe if there was no systematic regional variation in survival (Dickman et al., 1997; Yu et al., 2004). These estimates of avoidable premature deaths provide an objective measure by which to advocate for resource allocation and establish health priorities (Yu et al., 2004).

This study has two aims:

1. To examine the influence of cancer stage at diagnosis, distance to treatment facilities and area-disadvantage on spatial survival inequalities for breast and colorectal cancer, and
2. To estimate the number of premature deaths due to non-diagnostic-related spatial survival inequalities after adjusting for cancer stage at diagnosis.

Methods

Data

Study cohort

Data on colorectal (ICD-O3 C18-C20,C218) and breast (ICD-O3 C50) cancers diagnosed in Queensland during 1996 to 2007 were obtained from the Queensland Cancer Registry (QCR) following approval from Queensland Health (Ethics approval number: HREC/09/QHC/25). Due to small numbers, male breast cancers were excluded from analysis. The QCR is a population-based registry which has been in operation since 1982 (Queensland Cancer Registry, 2010), and covers a population of 4.2 million (in 2007) (Australian Bureau of Statistics, 2008b). Notification of cancer (excluding non-melanoma skin cancer) to the QCR is required by law (Queensland Cancer Registry, 2010). Data quality is high, as evidenced by

the high percentage of cases diagnosed with histological verification (92.1%) and low percentage of cases diagnosed by death certificate only (1.4%) in 2007.

The survival analysis included the first occurrence of a primary colorectal or breast cancer in individuals aged less than 90 years at diagnosis. Cases were excluded if they lacked age or SLA of residence information, were identified at autopsy, notified via death certificate only or had a survival time of less than one day. All cases were followed until 31st December 2007.

Stage at diagnosis

Colorectal cancer stage was extracted from pathology records held by the QCR (Krnjacki et al., 2008) and then classified based on the Dukes staging system (Haq et al., 2009). To increase accuracy (Krnjacki et al., 2008) and reduce problems with sparse data, stage was grouped into three categories: early (localized/non-localized), advanced (regional/distant) and unknown.

The QCR does not collect detailed information about breast cancer stage at diagnosis. However, consistent with recent reports (Baade et al., 2011c; Krnjacki et al., 2008; Youlden et al., 2009), "Early" breast cancer was defined as ≤20 mm diameter with no evidence of lymph node involvement or distant metastases (stage I). Although it was unlikely these cases had metastasized, this could not be established. There was insufficient detail to distinguish between stages II, III or IV, so these were collectively categorized as "Advanced" breast cancers. Cancers diagnosed as a result of metastatic disease were included in this category. The "Unknown" category included those with unknown tumor size or unknown lymph node status if the tumor size was ≤ 20mm.

Geographical location

Statistical Local Areas (SLAs) were used as the region of analysis. Cancer incidence data across all years were mapped to the 2006 SLA boundaries based on suburb and postcode of residence prior to data extraction. In 2006 Queensland had 478 SLAs, which covered the State without gap or overlap, with a median population of 5,810 (range: 7 to 77,523). Based on their SLA of residence, each patient was assigned to a quintile of area disadvantage based on the Australian Bureau of Statistics (ABS) Socioeconomic Indexes for Areas Index of Relative Disadvantage (SEIFA-IRSD) (Australian Bureau of Statistics, 2008a).

Distance to treatment

The distance to the closest radiation facility was calculated by geocoding the location of all radiation facilities in Queensland, and the centroid of each SLA at diagnosis. A custom GIS application was used to calculate the shortest travelling time by road from each SLA centroid to the closest radiation facility by each year to account for increasing coverage of the radiation facilities over time. Radiotherapy facilities are only located in larger cities. By the end of 2007 there were a total of 4 public and 5 private radiotherapy facilities in Queensland. Five (3 public and 2 private) were located in Brisbane, three additional private facilities were located within a 125 km radius of Brisbane, and another public facility in Townsville (1,360 km north of Brisbane).

Distance was classified into three categories based on practical considerations to improve the interpretation of estimates: < 2 hours (return travel within one day), 2-6 hours (one full day of travelling) and > 6 hours (more than one day of travel with overnight accommodation required).

Survival estimates

Unadjusted relative survival estimates were calculated using actuarial (life table) methods. Expected survival was estimated using the Ederer II method (Ederer and Heise, 1959) with the Stata macro strs, based on Queensland life tables generated from mortality data obtained from the ABS. The population mortality was calculated by each SLA, gender and 5-year age group (to ages 90+). Estimates were calculated for two aggregated time periods for greater stability; 1997-2002 and 2003-2007, and then applied to each year within the appropriate time period.

Survival estimates were derived using period analysis, in which survival is calculated using patients alive during the time period of interest (Brenner and Hakulinen, 2009). Since the focus was on estimating survival inequalities up to 5 years after diagnosis, each individual's follow-up time was censored at 5 years after diagnosis.

The expected number of deaths, person time at risk, and deaths in the interval was calculated for each individual, then aggregated over each combination of SLA (1 to 478 areas), follow-up period (1 to 5 years after diagnosis) and covariates consisting of age at diagnosis (0-49, 50-69 and 70-89 years), stage at diagnosis (early, advanced and unknown), distance to treatment facilities (<2 hours, 2 – 6 hours and 6+ hours) and for colorectal cancer, gender (male and female). Since there was an exact concordance between SLA and area disadvantage, aggregating by area disadvantage was not required.

Statistical model

The Bayesian spatial survival model adopted for this analysis assumed the hazards were constant within pre-specified follow-up time intervals, and was based on the model described by Fairley et al (Fairley et al., 2008),

$$d_{kti} \sim \text{Poisson}(\mu_{kti})$$

$$\log(\mu_{kti} - d^*_{kti}) = \log(y_{kti}) + \alpha_t + \mathrm{x}_k\beta + u_i + v_i \qquad [1]$$

where d_{kti}, the observed deaths in the k^{th} stratum, t^{th} follow-up interval and i^{th} SLA follows a Poisson distribution with mean μ_{kti}, $(\mu_{kti} - d^*_{kti})$ is the modeled number of excess deaths with d^*_{kti} representing the expected number of deaths due to other causes, y_{kti} is person-time at risk, α_t is an intercept which varies by time, β_k is the coefficient of the predictor variable vector x (representing broad age at diagnosis groups, distance to treatment, area disadvantage, stage and, for colorectal cancer, gender), u_i is the spatial random effects for the i^{th} SLA and v_i is the unstructured random effects. Non-informative normal distributions were used as priors on the parameters, apart from u_i which was assigned an intrinsic conditional autoregressive (CAR) prior. Refer to the Appendix for further information on the prior distributions.

The effects of age, stage, area disadvantage and distance to treatment were explored by including various combinations in models. The Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002) was used to compare model goodness of fit, with lower values indicating a better model.

The exponential of the parameter estimates are the excess hazard ratios, also called relative excess risks (RERs). The RERs for estimates of the impact of the covariates of interest were calculated as exp(β) from Equation (1), and were in comparison to the baseline level of the

covariate. For the estimates in each SLA the RER was calculated as $\exp(u_i + v_i)$, and provided an estimate of the excess risk of death in that SLA against the Queensland average excess risk of death.

The models were analyzed using Markov chain Monte Carlo (MCMC), via WinBUGS (Imperial College and Medical Research Council, UK), interfaced with Stata (StataCorp, Texas) (Thompson et al., 2006). A burn-in period of 250,000 iterations was discarded, and a further 100,000 iterations monitored (with every 10^{th} iteration kept). A global clustering test (Tango's Maximized Excess Events Test (MEET) (Tango, 2000)) was used to determine if there was significant variation in the RER estimates across the SLAs. This method was preferred over other global clustering tests as it has been shown to effectively identify overall spatial variation across a variety of datasets (Kulldorff et al., 2006).

The 80% credible interval (CrI) was provided for all posterior distributions, as this is considered to provide sufficient coverage (Richardson et al., 2004). For non-Bayesian analyses we used the standard 95% confidence interval (CI).

The probability of a specific estimate being higher than the estimate in the previous category can be used as an alternative to the credible interval when comparing categories. High values (above 80% when expressed as a percentage, consistent with 80% credible intervals) indicate the estimate is likely to be above the former category. This was calculated as the percentage of MCMC iterations for a given estimate that were higher than the preceding stratum-specific estimate, and was provided for RER and the proportion of premature deaths.

Premature deaths

In calculating the number of premature deaths, we set the optimum benchmark for survival to be equal to the 20th centile RER of ranked SLAs, consistent with other published results (Yu et al., 2004). To exclude the diagnostic component of survival, and quantify only the non-diagnostic survival component, we used results from the Bayesian spatial survival model (equation 1) that included stage, along with age and gender. Areas with an RER below the 20th centile were excluded from the calculations of observed deaths (d_{kti}), expected deaths due to other causes (d^*_{kti}) and person-time at risk (y_{kti}) to avoid theoretically increasing their risk of death to the 20th centile.

There were three parameters required to calculate premature deaths resulting from spatial inequalities: observed excess deaths, optimum excess deaths and the spatial fraction.

Observed excess deaths

In relative survival the 'excess deaths' are the deaths considered to be caused by the cancer. Instead of counting a death if the death certificate recorded it as a cancer death (as in cause-specific survival), all deaths among cancer patients are compared against the mortality that would be expected among people of similar age, gender, SLA of residence and broad time period. This prevents bias due to inaccuracies in coding deaths.

The modeled number of excess deaths within five years of diagnosis is calculated as $\mu_{kti} - d^*_{kti}$ (from equation 1). The observed number of excess deaths within five years of diagnosis at each stratum is:

$$D_k = \sum_{t=1}^{5} \sum_{i} (d_{kti} - d^*_{kti}) \quad [2]$$

where d_{kti} is the number of observed deaths from any cause in the k^{th} strata, t^{th} follow-up period, and i^{th} SLA, and d^*_{kti} represents the number of expected deaths due to other causes estimated by applying the population mortality rates to the study cohort.

Optimum excess deaths

The optimum number of excess deaths is the number of deaths that would be observed within five years of diagnosis if there were no inequalities in non-diagnostic factors. This was calculated by multiplying the excess mortality rate at the 20th centile by person-time at risk, separately for each stratum and follow-up interval and then summed over the follow-up intervals:

$$\mathrm{Opt}_k = \sum_{t=1}^{5} (\exp(\alpha_t + x_k\beta + u_{20} + v_{20}) \times y_{kt}) \quad [3]$$

where $u_{20} + v_{20}$ were the random effect values corresponding to the lowest 20th centile of relative excess risk across all the SLAs, and other variables were as described in equation 1.

Spatial fraction

The spatial fraction is used to distinguish between deaths influenced by spatially structured factors, and those due to random variation. This parameter estimates the relative contribution of the spatial component in the Bayesian spatial survival model, and is defined as:

$$\psi = \frac{\sigma_{u(m)}}{\sigma_{u(m)} + \sigma_v} \quad [4]$$

where $\sigma_{u(m)}$ is the marginal standard deviation of the spatial component u, and σ_v is the standard deviation of the random component v.

Premature deaths

The number of premature deaths due to spatial inequalities in non-diagnostic factors was defined as the total number of deaths within five years of diagnosis that could be avoided, i.e.:

$$\mathrm{Dprem} = \psi \times \sum_k (\mathrm{D}_k - \mathrm{Opt}_k) \qquad [5]$$

where ψ represents the spatial fraction (see equation 4), D_k represents observed excess deaths in the k^{th} stratum (equation 2), and Opt_k represents the optimum excess number of deaths in the k^{th} stratum (equation 3).

The total number of avoidable premature deaths was calculated as shown in equation 5, and was also calculated by stage, distance and area disadvantage categories. The median values of the 10,000 MCMC iterations were used as the Dprem value, and 80% credible intervals were obtained from the 10$^{\text{th}}$ and 90$^{\text{th}}$ centiles. The premature death percentages were calculated by dividing Dprem by the observed excess deaths ($\sum_k(\mathrm{D}_k)$).

Results

The final study cohort consisted of 25,202 females diagnosed with breast cancer and 26,390 cases of colorectal cancer (14,690 males, 11,700 females) (Table 1), as there were 264 breast cancer cases (1.0%) and 280 colorectal cancer cases (1.0%) that did not meet the inclusion criteria.

Breast cancer

The unadjusted 5-year relative survival from breast cancer by area disadvantage showed that 88% of women in the least disadvantaged quintile were likely to survive at least 5 years, while for the most disadvantaged quintile this decreased to 83% (Table 1). Survival differences by distance to nearest radiation facility were slightly smaller (86% for those living less than 2 hours distance, and 83% for those with at least 6 hours travel time).

Breast cancer survival was greatly impacted by stage at diagnosis. Based on DIC values, the full model containing age, stage, distance and area disadvantage; the model adjusted for age, stage and disadvantage and the model adjusted for age and stage were preferred against the alternatives (Table 2).

After adjustment for all factors in the full model, the oldest age group (ages 70-89), advanced or unknown stage and increasing area disadvantage had higher risk of death. Notably the higher survival observed in the 50-69 compared against the 0-49 age group, as well as the impact of distance from nearest radiotherapy treatment facility were no longer evident.

As indicated in Figure 1, even after adjusting for stage there was still moderate evidence for spatial inequalities (Tango's MEET p=0.042 for the model containing age and stage). After further adjusting for area disadvantage, spatial survival inequalities were attenuated to non-significance (p=0.452).

Between 1998 and 2007 there were 2,850 deaths due to breast cancer among women in Queensland within five years of diagnosis. After removing the SLAs with a risk of death

lower than the 20th centile, there were 2,412 deaths (Table 4). Of these deaths, 170 (7%) were estimated to be avoidable if there were no systematic spatial variation in non-diagnostic survival components, after adjusting for stage and age at diagnosis.

The proportion of premature deaths was high for those residing more than 2 hours distance from a treatment facility, and all area disadvantage quintiles above the least disadvantaged, particularly for the most disadvantaged. By stage, most premature deaths were among those diagnosed at advanced or unknown stage.

Colorectal cancer

The unadjusted 5-year relative survival varied by 8 absolute percentage points between the least disadvantaged (69%) and most disadvantaged (61%) areas (Table 1). There were also survival differences by distance to treatment facilities, with those living within 2 hours travelling time having a better survival than those residing a travelling time distance of at least 6 hours away (65% versus 60%, respectively).

Among the models considered, the best fit to colorectal cancer survival was the full model containing age, sex, stage, distance and area disadvantage as well as the model adjusted for age, sex, stage and area disadvantage based on DIC values (Table 3). After adjusting for stage at diagnosis, survival remained poorer among older patients and as area disadvantage increased. There was also some evidence that survival was poorer among those residing more than 6 hours travelling time to treatment, particularly in comparison to those living within 2-6 hours distance (88% probability that RER is higher).

Mapping the stage-adjusted RER did not substantively alter the observed survival inequalities (Figure 2), with both models (age and gender only; age, stage and gender) having strong evidence of geographical variation (Tango's MEET p=0.001). Further adjusting for area disadvantage and then distance to treatment only slightly reduced the spatial inequalities (p=0.004 and p=0.019, respectively).

There were 7,357 deaths due to colorectal cancer within five years of diagnosis in Queensland during 1998-2007, and 6,019 deaths after removing the SLAs with a risk of death lower than the 20th centile (Table 4). Of these deaths, 470 (8%) deaths would not have occurred if there were no spatial inequalities in the non-diagnostic survival component, after adjusting for stage, age at diagnosis and gender. There was a clear gradient of a greater proportion of premature deaths occurring as distance from radiotherapy treatment facilities increased, and also generally as area disadvantage increased. This contrasted with the consistency of the proportion of premature deaths across cancer stage categories.

The colorectal cancer stage categories differed from those used for breast cancer stage categories. To explore if differences between the cancers may have resulted from different stage groupings, analyses were also run using alternate colorectal stage categories (early: localized; advanced: non-localized/regional/distant; unknown). Results (not shown) were broadly consistent with those reported here.

Discussion

In this population-based study, cancer stage, age group, and disadvantage were important predictors of survival outcomes for people diagnosed with colorectal cancer and women diagnosed with breast cancer. After adjusting for stage and excluding the impact of random

variation, we estimated that 470 (8%) premature deaths due to colorectal cancer and 170 (7%) premature deaths due to breast cancer could be attributed to spatial inequalities in management factors.

Despite fairly similar numbers of incident cases for colorectal cancer and female breast cancer, the lower number of premature deaths attributable to spatial inequalities in management factors for breast cancer is due mainly to the higher survival, but also the strong influence exerted by stage at diagnosis on breast cancer spatial survival inequalities.

A previous study found women living in more remote areas of Queensland were more likely to be diagnosed with advanced breast cancer (Baade et al., 2011c). These diagnostic inequalities have resulted in survival inequalities, as our study showed the lower survival in rural areas was reduced after adjusting for stage. Mammography screening has been shown to be effective in diagnosing tumors early (Australian Institute of Health and Welfare, 2010), and the current low public mammography participation rate of 57% in the target age group (only slightly higher in more regional areas) (Australian Institute of Health and Welfare, 2010) provides considerable scope for the spatial survival inequalities resulting from late diagnosis to be addressed.

Although geographic differences have been demonstrated in the risk of being diagnosed with advanced colon cancer in Queensland (Baade et al., 2011a), we found these spatial differences in stage at diagnosis did not exert an important influence on the spatial survival inequalities. This is consistent with a previous Australian study which found stage at

diagnosis had limited impact on geographical inequalities in colon cancer survival inequalities, but more evidence for rectal cancer (Yu et al., 2005b).

Population-based screening would be expected to reduce the stage at which a cancer is diagnosed. In Queensland, mammography screening was introduced in 1991 and is freely available for women aged 40+ years, with a target age group of 50-69 years. More recently, mobile mammography clinics have been used to overcome the barrier of distance for women in more remote regions, to the extent that women in rural areas now have higher participation rates for public mammography than women in urban areas. Since the screening differential (lower screening in urban areas) is the opposite for the stage differential (less advanced stage in urban areas), it is unlikely that screening patterns can explain this differential stage distribution.

Screening is also an unlikely explanation for the colorectal cancer stage differences. In contrast to breast cancer, there is no population-based screening program for colorectal cancer. The Australian bowel cancer screening program (faecal occult blood test) started only in 2006, with gradual implementation for adults aged 50, 55 and 65 years of age (Australian Institute of Health and Welfare, 2009).

We found significant evidence that colorectal and breast cancer patients living in areas of greater socioeconomic deprivation had lower survival than their counterparts living in more affluent areas. This is consistent with other reports (Australian Institute of Health and Welfare and Cancer Australia & Australasian Association of Cancer Registries, 2008), however, as was the case in our study, it is unclear whether these patterns reflect treatment inequalities, or patient characteristics such as obesity, smoking, alcohol consumption and comorbidities rather than area-level characteristics (Frederiksen et al., 2009).

For breast cancer, stage at diagnosis and area-level disadvantage largely explained the survival inequalities, similar to a study in England (Davies et al., 2010). An American study showed much of the socioeconomic inequality in breast cancer survival was explained by stage at diagnosis, initial treatment and race (Yu, 2009). In Western Australia, breast cancer patients in rural (often lower socioeconomic) areas, were less likely to undergo hormone therapy, radiotherapy or be treated by a high-caseload surgeon (Mitchell et al., 2006). These treatment inequalities explained the remaining survival differentials in breast cancer after adjusting for age and tumor characteristics (Mitchell et al., 2006).

In contrast, colorectal cancer survival inequalities only modestly decreased after adjusting for stage, area disadvantage and distance to treatment, suggesting additional factors are important. Colorectal cancer survival is influenced by treatment factors such as the type of surgery provided, hospital caseload and specialist expertise at the treating institution (Yu et al., 2005b). Treatment inequalities may not have been adequately captured, as an American study found treatment disparities only slightly reduced after adjusting for socio-demographic characteristics and the availability of specialist oncology services (Haas et al., 2011). Further investigation to identify factors influencing colorectal cancer spatial inequalities is important.

The higher proportions of premature deaths due to spatial inequalities in non-diagnostic factors among the more disadvantaged and distant regions further suggest treatment inequalities. Enabling more remote patients to access the same level of treatment and care as urban patients is extremely challenging in the Australian environment of very large distances. Nonetheless, quantifying the impact of these inequalities will encourage efforts to identify ways to reduce these inequalities, leading to substantial public health gains.

Our analysis was based on 478 geographical regions, finding that 7.1% and 7.8% of breast and colorectal cancer premature deaths respectively were due to spatial inequalities in non-diagnostic factors. Other studies have tended to use a smaller number of geographical regions. One study examined 81 regions across four Nordic countries finding 2.6% of breast cancer deaths and 5.0% of colon deaths were considered 'savable' if there was no regional variation (Dickman et al., 1997). A study in NSW, Australia, after adjusting for broad stage categories calculated a similar estimate across 25 regions as 4.4% of breast and 6.9% of colon cancer deaths within 5 years of diagnosis (Yu et al., 2004). Our higher percentages could reflect the increased potential to detect variability using small geographical areas, or greater variability in survival outcomes within Queensland.

Strengths of this study include the high quality, population-based coverage of the Queensland Cancer Registry, the ability to adjust for stage at diagnosis, the benefits of analyzing small-area data using Bayesian hierarchical models and the use of period analysis to provide more up-to-date survival estimates.

Limitations include the lack of data on individual socioeconomic characteristics and comorbidities, the use of broad rather than clinical stage categories, the substantial proportion of cancers with unknown stage at diagnosis, the lack of treatment information and the relatively small number of covariates included in the models.

In conclusion, although earlier cancer diagnosis would decrease survival inequalities for breast cancer patients in rural areas, there remain an important number of premature deaths for breast and colorectal cancer that could be avoided by removing spatial inequalities in non-

diagnostic factors. Despite a freely available public health service, spatial variation in treatment utilization is likely to play an important role, although other environment or patient-factors may also be contributing. Identifying the precise non-diagnostic factors that cause these premature deaths will not be easy, but unless quantitative data such as these are disseminated, there will be little incentive on the part of researchers and health providers to investigate, develop and implement the necessary interventions.

Acknowledgements

Peter Baade was supported by a National Health and Medical Research Council (NHMRC) Career Development Fellowship (ID 1005334). Kerrie Mengersen acknowledges support from the CRC for Spatial Information. Gavin Turrell was supported by an NHMRC Senior Research Fellowship (ID 1003710). The extraction of colorectal cancer stage was supported by an NHMRC grant (ID 561700). The authors would like to thank Nicole White, Queensland University of Technology, for helpful methodological discussions, and Joanne Aitken, Cancer Council Queensland, for her review of a previous draft of this manuscript. The views expressed in this paper are those of the authors and not of any funding body.

References

Aitkin, M., Liu, C.C., Chadwick, T., 2009. Bayesian model comparison and model averaging for small-area estimation. The Annals of Applied Statistics 3, 199-221.

Australian Bureau of Statistics, 2008a. Information Paper: An Introduction to Socio-Economic Indexes for Areas (SEIFA), 2006, ABS Cat. no. 2039.0. ABS, Canberra.

Australian Bureau of Statistics, 2008b. Population by Age and Sex, Regions of Australia, 2007, Cat. no. 3235.0. ABS, Canberra.
Australian Institute of Health and Welfare, 2009. National Bowel Cancer Screening Program: annual monitoring report 2009, Cat. no. CAN 45. AIHW, Canberra.
Australian Institute of Health and Welfare, 2010. BreastScreen Australia monitoring report 2006–2007 and 2007–2008, Cancer series no. 55. Cat. no. CAN 51. AIHW, Canberra.
Australian Institute of Health and Welfare, Cancer Australia & Australasian Association of Cancer Registries, 2008. Cancer survival and prevalence in Australia: cancers diagnosed from 1982 to 2004, Cancer Series no. 42, Cat. no. CAN 38. AIHW, Canberra.
Baade, P.D., Dasgupta, P., Aitken, J., Turrell, G., 2011a. Geographic remoteness and risk of advanced colorectal cancer at diagnosis in Queensland: a multilevel study. Br J Cancer.
Baade, P.D., Dasgupta, P., Aitken, J.F., Turrell, G., 2011b. Distance to the closest radiotherapy facility and survival after a diagnosis of rectal cancer in Queensland. Med J Aust 195, 350-354.
Baade, P.D., Turrell, G., Aitken, J.F., 2011c. Geographic remoteness, area-level socio-economic disadvantage and advanced breast cancer: a cross-sectional, multilevel study. J Epidemiol Community Health.
Besag, J., York, J., Mollie, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Statist Math 43, 1-59.
Brenner, H., Hakulinen, T., 2009. Up-to-date cancer survival: period analysis and beyond. Int J Cancer 124, 1384-1390.
Carlin, B.P., Xia, H., 1999. Assessing environmental justice using Bayesian hierarchical models: two case studies. J Expo Anal Environ Epidemiol 9, 66-78.
Coleman, M.P., Forman, D., Bryant, H., Butler, J., Rachet, B., Maringe, C., Nur, U., Tracey, E., Coory, M., Hatcher, J., McGahan, C.E., Turner, D., Marrett, L., Gjerstorff, M.L., Johannesen, T.B., Adolfsson, J., Lambe, M., Lawrence, G., Meechan, D., Morris, E.J., Middleton, R., Steward, J., Richards, M.A., 2011. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data. Lancet 377, 127-138.
Cramb, S.M., Mengersen, K.L., Baade, P.D., 2011. Atlas of Cancer in Queensland: geographical variation in incidence and survival, 1998 to 2007. Viertel Centre for Research in Cancer Control, Cancer Council Queensland, Brisbane.
Davies, E.A., Linklater, K.M., Coupland, V.H., Renshaw, C., Toy, J., Park, R., Petit, J., Housden, C., Moller, H., 2010. Investigation of low 5-year relative survival for breast cancer in a London cancer network. Br J Cancer 103, 1076-1080.
Dickman, P.W., Gibberd, R.W., Hakulinen, T., 1997. Estimating potential savings in cancer deaths by eliminating regional and social class variation in cancer survival in the Nordic countries. J Epidemiol Community Health 51, 289-298.
Du, X.L., Lin, C.C., Johnson, N.J., Altekruse, S., 2011. Effects of individual-level socioeconomic factors on racial disparities in cancer treatment and survival: Findings from the National Longitudinal Mortality Study, 1979-2003. Cancer.
Eberly, L.E., Carlin, B.P., 2000. Identifiability and convergence issues for Markov chain Monte Carlo fitting of spatial models. Statistics in Medicine 19, 2279-2294.
Ederer, F., Heise, H., 1959. Instructions to IBM 650 Programmers in Processing Survival Computations. Methodological note No. 10. End Results Evaluation Section, National Cancer Institute, Bethesda MD.
Fairley, L., Forman, D., West, R., Manda, S., 2008. Spatial variation in prostate cancer survival in the Northern and Yorkshire region of England using Bayesian relative survival smoothing. Br J Cancer 99, 1786-1793.
Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C., Parkin, D.M., 2010. GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide. International Agency for Research on Cancer, Lyon, France.
Frederiksen, B.L., Osler, M., Harling, H., Ladelund, S., Jorgensen, T., 2009. Do patient characteristics, disease, or treatment explain social inequality in survival from colorectal cancer? Soc Sci Med 69, 1107-1115.
Fusco, M., Pezzi, A., Benatti, P., Roncucci, L., Chiodini, P., Di Maio, G., Di Napoli, R., de Leon, M.P., 2010. Clinical features and colorectal cancer survival: an attempt to explain differences between two different Italian regions. Eur J Cancer 46, 142-149.
Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, in: Bernardo, J.M., Berger, J., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics 4. Oxford University Press, Oxford, pp. 169-193.
Haas, J.S., Brawarsky, P., Iyer, A., Fitzmaurice, G.M., Neville, B.A., Earle, C., 2011. Association of area sociodemographic characteristics and capacity for treatment with disparities in colorectal cancer care and mortality. Cancer 117, 4267-4276.
Haq, A.I., Schneeweiss, J., Kalsi, V., Arya, M., 2009. The Dukes staging system: a cornerstone in the clinical management of colorectal cancer. Lancet Oncol 10, 1128.

Henry, K.A., Niu, X., Boscoe, F.P., 2009. Geographic disparities in colorectal cancer survival. Int J Health Geogr 8, 48.
Huang, L., Kulldorff, M., Gregorio, D., 2007. A spatial scan statistic for survival data. Biometrics 63, 109-118.
Krnjacki, L.J., Baade, P.D., Lynch, B.M., Aitken, J.F., 2008. Reliability of collecting colorectal cancer stage information from pathology reports and general practitioners in Queensland. Aust N Z J Public Health 32, 378-382.
Kulldorff, M., Song, C., Gregorio, D., Samociuk, H., DeChello, L., 2006. Cancer map patterns: are they random or not? Am J Prev Med 30, S37-49.
Leroux, B.G., Lei, X., Breslow, N., 2000. Estimation of disease rates in small areas: a new mixed model for spatial dependence, in: Halloran, M.E., Berry, D. (Eds.), Statistical models in epidemiology, the environment and clinical trials. Springer, New York, pp. 135-178.
McKenzie, F., Ellison-Loschmann, L., Jeffreys, M., 2011. Investigating reasons for ethnic inequalities in breast cancer survival in New Zealand. Ethn Health, 1-15.
Mitchell, K.J., Fritschi, L., Reid, A., McEvoy, S.P., Ingram, D.M., Jamrozik, K., Clayforth, C., Byrne, M.J., 2006. Rural-urban differences in the presentation, management and survival of breast cancer in Western Australia. Breast 15, 769-776.
Queensland Cancer Registry, 2010. Cancer in Queensland: Incidence, Mortality, Survival and Prevalence,1982 to 2007. QCR, Cancer Council Queensland and Queensland Health, Brisbane.
Richardson, S., Thomson, A., Best, N., Elliott, P., 2004. Interpreting posterior relative risk estimates in disease-mapping studies. Environ Health Perspect 112, 1016-1025.
Saez, M., Barceló, M.A., Martos, C., Saurina, C., Marcos-Gragera, R., Renart, G., Ocaña-Riola, R., Feja, C., Alcalá, T., 2012. Spatial variability in relative survival from female breast cancer. Journal of the Royal Statistical Society: Series A (Statistics in Society) 175, 107-134.
Schottenfeld, D., Fraumeni Jr, J.F., 2006. Cancer epidemiology and prevention. Oxford University Press, New York.
Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. J R Statist Soc B 64, 583-639.
Tango, T., 2000. A test for spatial disease clustering adjusted for multiple testing. Stat Med 19, 191-204.
Thompson, J., Palmer, T., Moreno, S., 2006. Bayesian analysis in Stata using WinBUGS. Stata J 6, 530-549.
Tian, N., Goovaerts, P., Zhan, F.B., Chow, T.E., Wilson, J.G., 2012. Identifying Risk Factors for Disparities in Breast Cancer Mortality among African-American and Hispanic Women. Womens Health Issues.
Tian, N., Wilson, J.G., Zhan, F.B., 2011. Spatial association of racial/ethnic disparities between late-stage diagnosis and mortality for female breast cancer: where to intervene? Int J Health Geogr 10, 24.
Wakefield, J., Elliott, P., 1999. Issues in the statistical analysis of small area health data. Stat Med 18, 2377-2399.
Wan, N., Zhan, F.B., Lu, Y., Tiefenbacher, J.P., 2012. Access to healthcare and disparities in colorectal cancer survival in Texas. Health Place 18, 321-329.
Youlden, D.R., Cramb, S.M., Baade, P.D., 2009. Current status of female breast cancer in Queensland: 1982 to 2006. Viertel Centre for Research in Cancer Control, Cancer Council Queensland, Brisbane.
Yu, X.Q., 2009. Socioeconomic disparities in breast cancer survival: relation to stage at diagnosis, treatment and race. BMC Cancer 9, 364.
Yu, X.Q., O'Connell D, L., Gibberd, R.W., Armstrong, B.K., 2005a. A population-based study from New South Wales, Australia 1996-2001: area variation in survival from colorectal cancer. Eur J Cancer 41, 2715-2721.
Yu, X.Q., O'Connell, D.L., Gibberd, R.W., Armstrong, B.K., 2005b. A population-based study from New South Wales, Australia 1996-2001: area variation in survival from colorectal cancer. Eur J Cancer 41, 2715-2721.
Yu, X.Q., O'Connell, D.L., Gibberd, R.W., Smith, D.P., Dickman, P.W., Armstrong, B.K., 2004. Estimating regional variation in cancer survival: a tool for improving cancer care. Cancer Causes Control 15, 611-618.

Appendix: Prior distributions

Prior distributions for α and β were diffuse normal distributions with mean 0 and variance 1.0 x 10^6. An intrinsic CAR prior for u_i was employed to describe local spatial dependence across the SLAs, specified as:

$$u_i|u_j \sim N\left(\frac{\sum_j \omega_{ij} u_j}{\sum_j \omega_{ij}}, \sigma_u^2\right)$$

where ω_{ij}= 1 if *i, j* are adjacent, and 0 otherwise (Besag et al., 1991). As usual, the unstructured residual form was modeled with a normal prior, $v_i \sim N(0, \sigma_v^2)$.

The variances σ_u^2 and σ_v^2 influence the relative weight given to describing residual variation through spatial correlation between the estimates or some other (random) source. Since this is often unknown, it is typical to place priors on these terms. Commonly the precision (the inverse of the variance) is described as a gamma distribution. To examine the impact of the selection of distributions for the precision parameters ($\tau_u$and τ_v) on the random effect components (u_i and v_i), sensitivity analyses were conducted by comparing three combinations of gamma distributions (Γ) on the precision and 2 combinations of uniform distributions (Unif) on the standard deviation:

1. $\tau_u \sim \Gamma(0.1,100)$, $\tau_v \sim \Gamma(0.1,100)$
2. $\tau_u \sim \Gamma(0.5,1000)$, $\tau_v \sim \Gamma(0.5,1000)$
3. $\tau_u \sim \Gamma(0.1,10)$, $\tau_v \sim \Gamma(0.001,1000)$
4. $\sigma_u \sim \text{Unif}(0,10)$, $\sigma_v \sim \text{Unif}(0,10)$
5. $\sigma_u \sim \text{Unif}(0,1000)$, $\sigma_v \sim \text{Unif}(0,1000)$

These gamma distributions have means and variances on the precisions of (10, 1000); (500, 500000); and for the third option, τ_u has (1,10), while τ_v has (1,1000), respectively. The uniform distributions have means and variances on the standard deviations of (5, 8.3) and (500, 83333.3).

The priors were evaluated and compared on the basis of summary measures of the posterior distribution of the relative excess risk values, DIC values (Spiegelhalter et al., 2002), cumulative distribution function plots of the deviance (Aitkin et al., 2009), and convergence

diagnostics including trace and density plots as well as the Geweke diagnostic (Geweke, 1992). On the basis of these, the first option was selected for this study.

The model formulation employed for this study has been criticized for potential lack of identifiability of the individual *u* and *v* components (Eberly and Carlin, 2000). Despite this, it is recommended to use the spatial fraction (ψ) to determine the relative spatial and random effects (Eberly and Carlin, 2000). Leroux et al (2000) proposed an alternative model, where only one parameter is included for the spatial/random effects, but the prior on this term acts as a mixture distribution incorporating a spatial smoothing parameter λ, where λ provides the spatial proportion, similar to our spatial fraction (ψ). We found this approach was not feasible in this study, as the posterior distribution on λ failed to converge.

Figure 1: Relative excess risk of death from breast cancer among females in Queensland, 1998-2007. A) after adjusting for age; B) after adjusting for age and cancer stage at diagnosis; C) after adjusting for age, stage and area disadvantage; D) after adjusting for age, stage, distance to treatment and area disadvantage.
Note: The Queensland average RER=1.0

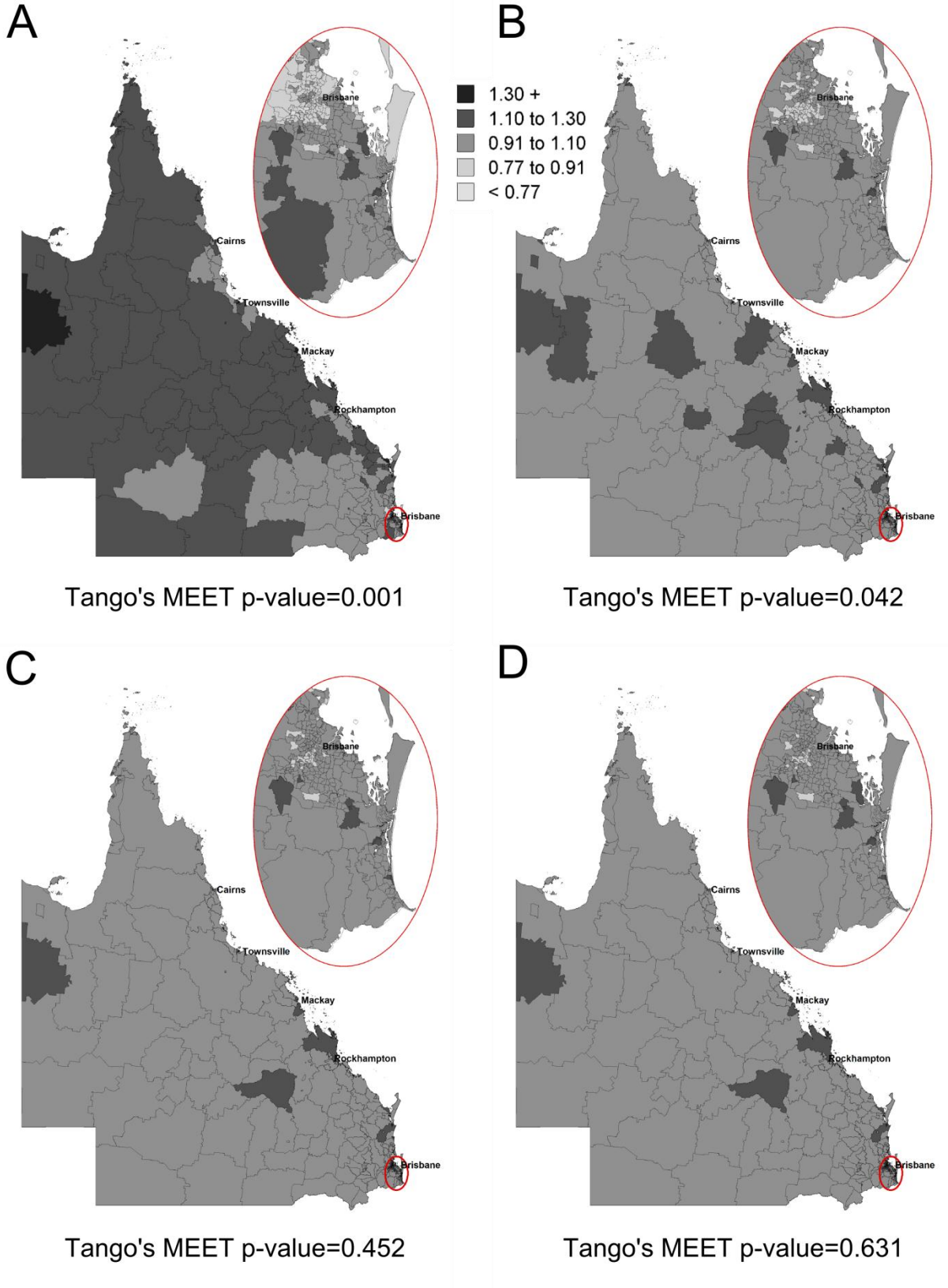

Figure 2: Relative excess risk of death from colorectal cancer among persons in Queensland, 1998-2007. A) after adjusting for age and gender; B) after adjusting for age, gender and cancer stage at diagnosis; C) after adjusting for age, gender, stage and area disadvantage; D) after adjusting for age, gender, stage, distance to treatment and area disadvantage.
Note: The Queensland average RER=1.0

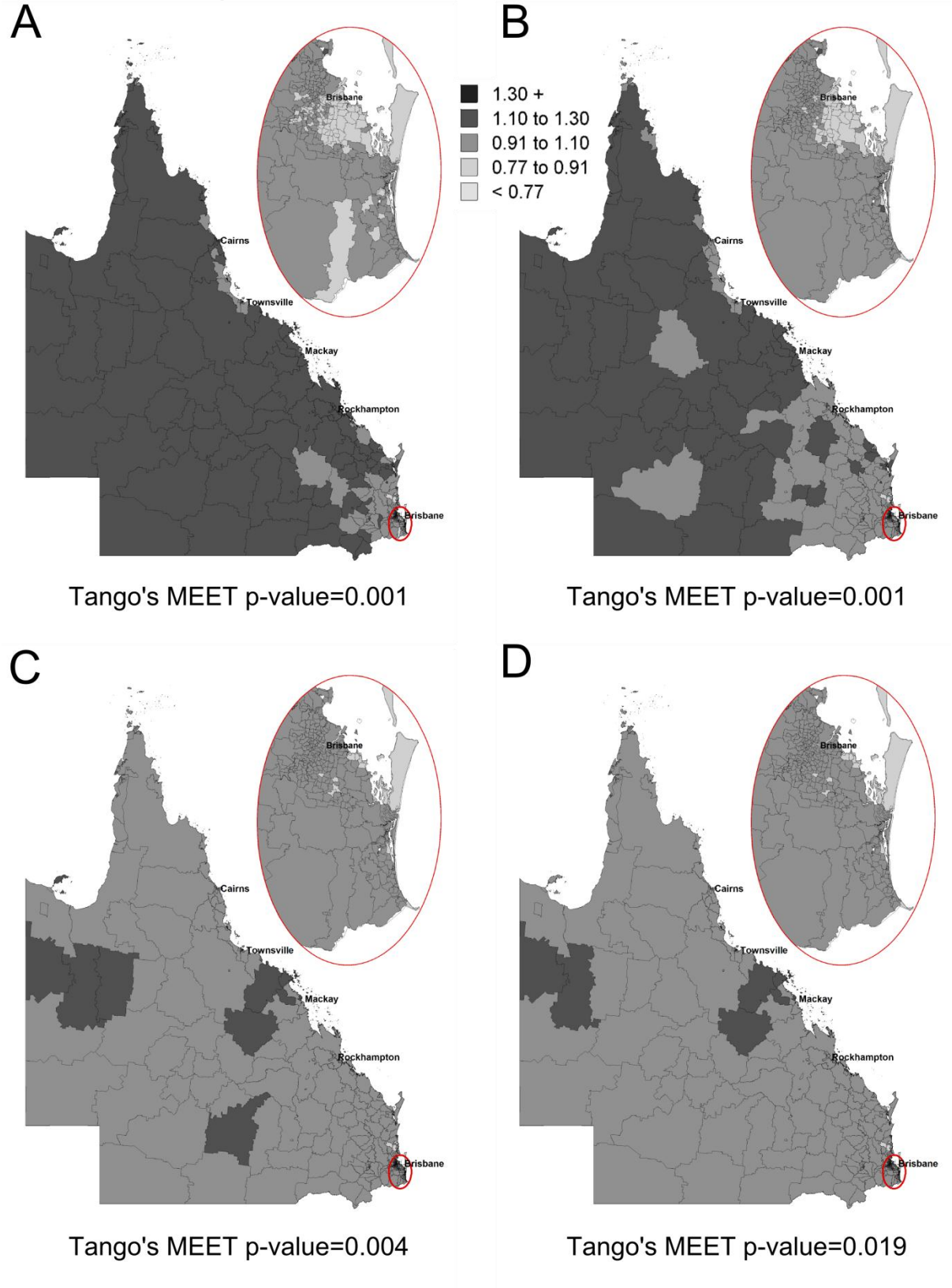

Table 1: Demographic characteristics of the study population and 5-year survival estimates

Variable	N	Colorectal cancer 5-year relative survival [95% CI]	p-value	N	Breast cancer 5-year relative survival [95% CI]	p-value
Sex						
Males	14,690	61.2 [60.2, 62.3]				
Females	11,700	67.1 [65.9, 68.3]	<0.001	25,202	85.2 [84.5, 85.8]	
Age group						
0-49	2,067	68.9 [66.5, 71.2]		6,517	85.8 [84.7, 86.7]	
50-69	11,525	66.5 [65.4, 67.5]		12,622	87.4 [86.7, 88.1]	
70-89	12,798	60.4 [59.1, 61.7]	<0.001	6,063	79.4 [77.7, 81.1]	<0.001
Stage						
Early	12,299	84.6 [83.6, 85.7]		11,505	94.8 [94.1, 95.4]	
Advanced	9,672	41.1 [39.8, 42.4]		10,707	80.2 [79.2, 81.2]	
Unknown	4,419	54.0 [52.1, 56.0]	<0.001	2,990	56.0 [53.5, 58.5]	<0.001
Distance						
< 2 hours	19,865	64.9 [64.0, 65.8]		19,490	85.7 [85.0, 86.4]	
2 – 6 hours	4,554	60.6 [58.7, 62.5]		3,978	83.2 [81.5, 84.7]	
6+ hours	1,971	59.7 [56.8, 62.5]	<0.001	1,734	83.2 [80.8, 85.5]	0.001
Area disadvantage						
Least disadvantaged	3,664	69.0 [66.9, 71.1]		4,098	88.4 [87.0, 89.8]	
Less disadvantaged	5,908	64.4 [62.7, 66.1]		5,829	85.4 [84.1, 86.6]	
Middle	6,587	64.2 [62.6, 65.8]		6,111	85.0 [83.8, 86.2]	
More disadvantaged	6,834	61.7 [60.1, 63.2]		6,189	84.1 [82.8, 85.3]	
Most disadvantaged	3,397	60.5 [58.2, 62.7]	<0.001	2,975	82.6 [80.7, 84.4]	<0.001

Note: p-values calculated using log-rank test for equality of survivor functions.

Table 2: Covariate fixed effects of Relative Excess Risk of death (RER) estimates (80% credible interval) for breast cancer, females

	Including age	Including age and stage	Including age and distance	Including age and area disadvantage	Including age, distance and area disadvantage	Including age, stage and area disadvantage	Including age, stage, distance and area disadvantage	From the fully adjusted model: Probability RER above preceding category (%)
Fixed effects								
Age group								
0-49	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
50-69	0.89 (0.84, 0.95)	1.02 (0.95,1.08)	0.89 (0.84, 0.95)	0.89 (0.83,0.94)	0.89 (0.83,0.94)	1.01 (0.95,1.08)	1.01 (0.95,1.08)	58.5
70-89	1.48 (1.37, 1.59)	1.62 (1.51,1.74)	1.48 (1.37, 1.60)	1.46 (1.35,1.57)	1.46 (1.35,1.58)	1.61 (1.49,1.73)	1.61 (1.49,1.73)	100.0
Stage								
Early		1.00				1.00	1.00	
Advanced		3.34 (3.10,3.61)				3.33 (3.10,3.60)	3.33 (3.08,3.60)	100.0
Unknown		7.93 (7.30,8.65)				7.89 (7.25,8.58)	7.87 (7.23,8.58)	100.0
Distance								
< 2 hours			1.00		1.00		1.00	
2-6 hours			1.07 (0.94, 1.20)		1.03 (0.91,1.15)		1.02 (0.91,1.13)	58.4
6+ hours			1.14 (0.98, 1.31)		1.10 (0.95,1.26)		1.03 (0.90,1.18)	53.9
Area disadvantage								
Least disadvantaged				1.00	1.00	1.00	1.00	
Less disadvantaged				1.28 (1.16,1.43)	1.29 (1.16,1.43)	1.19 (1.08,1.32)	1.19 (1.08,1.32)	98.8
Middle				1.27 (1.14,1.42)	1.27(1.14,1.42)	1.17 (1.06,1.31)	1.17 (1.05,1.30)	38.8
More disadvantaged				1.35 (1.20, 1.51)	1.34 (1.20,1.50)	1.23 (1.10,1. 37)	1.21 (1.09,1. 36)	71.7
Most disadvantaged				1.51 (1.33,1.72)	1.50 (1.32,1.71)	1.34 (1.19,1.51)	1.32 (1.17,1.49)	84.6
Spatial fraction (80% CrI)	0.57 (0.31,0.80)	0.42 (0.19,0.73)	0.46 (0.19,0.76)	0.37 (0.16,0.68)	0.31 (0.12,0.61)	0.29 (0.11,0.54)	0.29 (0.13,0.57)	
DIC	18333.7	17214.1	18337.7	18326.9	18329.6	17212.2	17215.3	
pD	64.2	59.7	63.4	60.6	64.2	61.3	61.9	

Notes: DIC = Deviance Information Criterion. Smaller values signify a better model fit if the difference is at least 5.
pD represents the effective number of parameters in the model. Larger values indicate estimates have undergone less smoothing.

Table 3: Covariate fixed effects of Relative Excess Risk of death (RER) estimates (80% credible interval) for colorectal cancer, persons

	Including age and sex	Including age, gender and stage	Including age, gender and distance	Including age, gender and area disadvantage	Including age, gender, distance and area disadvantage	Including age, stage and area disadvantage	Including age, gender, stage, distance and area disadvantage	From the fully adjusted model: Probability RER above preceding category (%)
Fixed effects								
Sex								
Males	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
Females	0.94 (0.91, 0.97)	0.96 (0.93, 1.00)	0.94 (0.91, 0.98)	0.95 (0.91, 0.98)	0.95 (0.91, 0.98)	0.96 (0.93, 1.00)	0.97 (0.93, 1.00)	8.5
Age group								
0-49	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
50-69	1.10 (1.03,1.17)	1.16 (1.09,1.23)	1.09 (1.03,1.16)	1.10 (1.03,1.17)	1.09 (1.03,1.16)	1.15 (1.08,1.22)	1.16 (1.09,1.23)	99.8
70-89	1.49 (1.41, 1.59)	1.72 (1.61,1.83)	1.49 (1.40, 1.58)	1.49 (1.40, 1.59)	1.48 (1.39, 1.58)	1.70 (1.60,1.81)	1.71 (1.61,1.82)	100.0
Stage								
Early		1.00				1.00	1.00	
Advanced		5.45 (5.21, 5.71)				5.44 (5.20, 5.70)	5.45 (5.21, 5.71)	100.0
Unknown		4.19 (3.96, 4.41)				4.17 (3.96, 4.41)	4.18 (3.96, 4.41)	0.0
Distance								
< 2 hours			1.00		1.00		1.00	
2-6 hours			1.06 (0.98, 1.15)		1.05 (0.97, 1.12)		0.99 (0.91,1.06)	42.1
6+ hours			1.09 (0.99, 1.20)		1.08 (0.99, 1.17)		1.07 (0.97,1.16)	87.6
Area disadvantage								
Least disadvantaged				1.00	1.00	1.00	1.00	
Less disadvantaged				1.15 (1.08, 1.23)	1.16 (1.08, 1.24)	1.12 (1.05,1.20)	1.13 (1.05,1.20)	99.2
Middle				1.16 (1.09, 1.24)	1.17 (1.09, 1.25)	1.14 (1.07, 1.22)	1.15 (1.07, 1.23)	65.1
More disadvantaged				1.23 (1.15, 1.31)	1.23 (1.15, 1.32)	1.20 (1.12, 1.28)	1.21 (1.13, 1.29)	89.1
Most disadvantaged				1.29 (1.20, 1.39)	1.29 (1.19, 1.39)	1.26 (1.17, 1.36)	1.27 (1.17, 1.37)	83.8
Spatial fraction (80% CrI)	0.69 (0.50,0.83)	0.62 (0.43,0.79)	0.60 (0.34,0.81)	0.57 (0.38,0.74)	0.45 (0.24,0.68)	0.52 (0.32,0.71)	0.49 (0.25,0.72)	

DIC	34660.4	31432.1	34664.6	34651.4	34654.5	31423.4	31425.0
pD	68.6	65.5	65.0	65.0	63.0	63.9	65.5

Notes: DIC = Deviance Information Criterion. Smaller values signify a better model fit if the difference is at least 5.
pD represents the effective number of parameters in the model. Larger values indicate estimates have undergone less smoothing.

Table 4: Premature deaths due to non-diagnostic spatial inequalities after adjusting for age, gender and stage at diagnosis among SLAs with a Relative Excess Risk of death above the 20th centile by stage, distance and area disadvantage categories

	Observed excess deaths	Optimum excess deaths	Premature deaths due to non-diagnostic spatial inequalities (random variation excluded)				Probability % premature deaths above preceding category (%)
			N	[80% CrI]	%	[80% CrI]	
Colorectal cancer							
Total	*6019*	*5236*	*470*	*[321, 637]*	*7.8*	*[5.3, 10.6]*	
Stage							
Early	1152	1014	81	[45, 125]	7.0	[3.9, 10.8]	
Advanced	3571	3099	282	[190, 388]	7.9	[5.3, 10.8]	64.5
Unknown	1297	1121	104	[65, 149]	8.0	[5, 11.5]	53.1
Distance							
< 2 hours	4066	3637	257	[167, 359]	6.3	[4.1, 8.8]	
2 - 6 hours	1356	1127	136	[95, 181]	10.1	[7.1, 13.3]	100.0
6+ hours	619	489	78	[55, 102]	12.8	[9.1, 16.7]	100.0
Area disadvantage							
Least disadvantaged	606	623	0	[0, 4]	0.0	[0.0, 0.6]	
Less disadvantaged	1218	1088	77	[48, 109]	6.4	[4.0, 9.1]	100.0
Middle	1413	1245	100	[66, 139]	7.1	[4.6, 9.9]	72.4
More disadvantaged	1863	1552	185	[129, 247]	10.0	[7.1, 13.1]	99.9
Most disadvantaged	937	739	120	[85, 157]	12.9	[9.0, 16.8]	100.0
Breast cancer							
Total	*2412*	*1975*	*170*	*[86, 307]*	*7.1*	*[3.6, 12.7]*	
Stage							
Early	414	375	14	[1, 35]	3.4	[0.3, 8.2]	
Advanced	1339	1069	106	[53, 190]	7.9	[3.9, 14.2]	96.9
Unknown	659	530	50	[24, 92]	7.6	[3.6, 13.9]	43.5
Distance							
< 2 hours	1731	1453	106	[53, 197]	6.2	[3.1, 11.4]	
2 - 6 hours	469	356	45	[22, 77]	9.6	[4.8, 16.6]	100.0
6+ hours	221	170	19	[9, 34]	8.9	[4.4, 15.6]	21.4
Area disadvantage							
Least disadvantaged	260	255	1	[0, 10]	0.6	[0.0, 4.0]	
Less disadvantaged	546	444	39	[19, 71]	7.3	[3.6, 13.0]	100.0
Middle	598	492	40	[20, 73]	6.8	[3.4, 12.4]	35.1
More disadvantaged	665	530	52	[26, 94]	7.9	[4.0, 14.0]	87.2
Most disadvantaged	354	258	38	[19, 66]	10.9	[5.3, 18.7]	99.8

Notes: Due to the methodology employed and/or rounding, numbers may not sum to the total.
Negative numbers of avoidable premature deaths were capped at zero.
CrI=credible interval.