TECHNICAL REPORT

The sociodemographic determinants of small area population health in Queensland

Catherine Harper[1]
Peter Baade[2]
Bryan Kennedy[3]

Epidemiology Services Unit (2,3)
Health Information Centre

Central Public Health Unit Network (1)
Brisbane Northside
Public Health Services

Queensland Health

Suggested citation: Harper C, Baade P Kennedy B. The socio-demographic determinants of small area population health in Queensland (June 2002). Queensland Health.

Further copies of this publication are available from:
Manager,
Epidemiology Services Unit
Health Information Centre
Queensland Health
GPO Box 48
Brisbane QLD 4001
Ph (07) 32341881

1. Summary

The provision of appropriate health services is significantly enhanced by knowledge of the health status of the target population. Measurements of area specific health status or health service needs are difficult for small geographical areas. Measures of population health available from routine collections do not always allow for calculation of small area estimates. Furthermore, information obtained through periodic or ad hoc surveys are generally designed to obtain population estimates on either a state or national level. Estimates for smaller geographical areas are generally not possible due to small sample sizes in each of these areas.

This study investigated the application of synthetic estimation techniques for estimating the population mortality rate of small areas in Queensland. In addition this study determines the variation in mortality status in small areas explained by sociodemographic variables.

Poisson regression models were used to synthetically estimate the health of populations in 389 geographic areas in Queensland in 1992-1996. The outcome measures were all cause premature mortality (excluding injury) and mortality due to cardiovascular disease. The explanatory variables were derived from sociodemographic data in the 1996 census.

The regression model for premature all cause mortality (excluding injury) explained 73% of the deviance in area specific mortality, with 42% of the observed values being within the 95% confidence intervals of the predicted value. Similar results were obtained for the cardiovascular disease mortality model. The presence of an Indigenous community in an area inflated mortality counts by a greater amount than any other sociodemographic variable, after the effect of all other variables was removed.

A large component of the variation in area-specific mortality due to all causes excluding injury (0-74 years) and to cardiovascular disease, in small geographic areas in Queensland, can be explained by the differential distribution of census based sociodemographic variables. As the predictive ability of these models were weak, they could not be used to accurately estimate the health status of small geographic areas, suggesting that synthetic estimation techniques utilising sociodemographic variables have limited application for measuring health status for small areas.

2. Introduction

Age and sex are central determinants of health. Australian and international research has affirmed the magnitude of the impact of social and economic factors on health status.[1,2] Specifically, the mortality burden in the Australian population attributable to socioeconomic inequality is large.[3] Indigenous status is a major determinant of health in Australia.[4,5] The health of individuals and populations results from a complex set of interrelationships between individual actions and social, economic, political and cultural processes. While the importance of individual sociodemographic parameters on health outcomes is evident, the extent of the association between these collective parameters and health status has not been the subject of reported study.

Differentials in health status of populations in geographic areas of Queensland and Australia have consistently been reported.[4,5] Concurrently, small area health statistics have become increasingly important, as the gains of a more individualised approach to smaller area population health are recognised.[6] However, the measurement of the health outcomes or health service needs for small geographical areas is generally difficult, primarily due to small population sizes.

Reliable information on the extent of association between sociodemographic variables and health outcome measures would provide an alternative means of assessing the health outcomes for small geographical areas. Synthetic estimation, applying regression analysis techniques, has been widely used in demography and economics, and has been suggested for use in epidemiological assessment of health.[7] This study investigated the application of regression techniques to determine the extent of association between census derived sociodemographic parameters and health outcomes, and for synthetic estimation of the health status of small areas in Queensland.

3. Method

Mortality data were obtained from the Australian Bureau of Statistics (ABS) Causes of Death file. The basic unit of analysis was statistical local area (SLA), which is the smallest geographical area for which mortality data are available. Data were aggregated from 1992 to 1996. However, to account for the numerous boundary changes in the SLAs between 1992 and 1996, some SLAs needed to be aggregated. A total of 389 geographical areas of interest (GAI) were derived from the SLAs, and these GAI covered Queensland.

Two conditions were chosen as the measures of health outcomes: premature mortality due to all causes excluding injury (ICD-9 000-799) for people less than 75 years of age; and mortality due to cardiovascular disease (CVD: ICD-9 410-414). These measures were chosen because the standardised mortality ratios for these conditions were similarly distributed across sociodemographic categories and to assess whether the explanatory variables and parameter estimates differed between general and disease specific mortality. Because injury is caused by external factors and the sociodemographic profile of cases varies markedly from that of all other causes of mortality, injury was excluded from the analysis of mortality due to all causes.

Poisson regression models were used for each outcome measure (counts of mortality), using the log of the population as an offset variable in a log link function. Predicted number of deaths for each GAI and 95% confidence intervals (CI), and adjusted mortality ratios were determined from the parameter estimates of the final model.

A backward stepwise selection process was used to refine the Poisson models. The same set of explanatory variables was initially used in the Poisson models for both outcome measures. These explanatory variables were the sociodemographic measures (n=41) derived from the 1996 Census, detailed in Appendix 1. In addition, a measure of area-based socioeconomic disadvantage was calculated based on the method used by the ABS, but incorporating additional age-specific variables.[8] Age and sex-specific population counts were obtained from the estimated resident population data from the ABS.

4. Results

The model for GAI-specific mortality due to all causes (excluding injury) for persons aged 0-74 years across Queensland explained 73% of the deviance in the mortality counts. Similarly, the model for CVD explained 60% of the deviance. Both models, particularly for CVD, indicate that the presence of an Indigenous community in the GAI inflated the mortality counts by a greater amount than any other sociodemographic variable, after the effects of all other variables were removed (Table 1). After adjustment for Indigenous populations, rural and remote GAIs were generally associated with lower mortality.

Table 1: Poisson regression model for mortality due to All Causes (excluding injury: 0-74 years) and CVD, Queensland, 1992-1996.

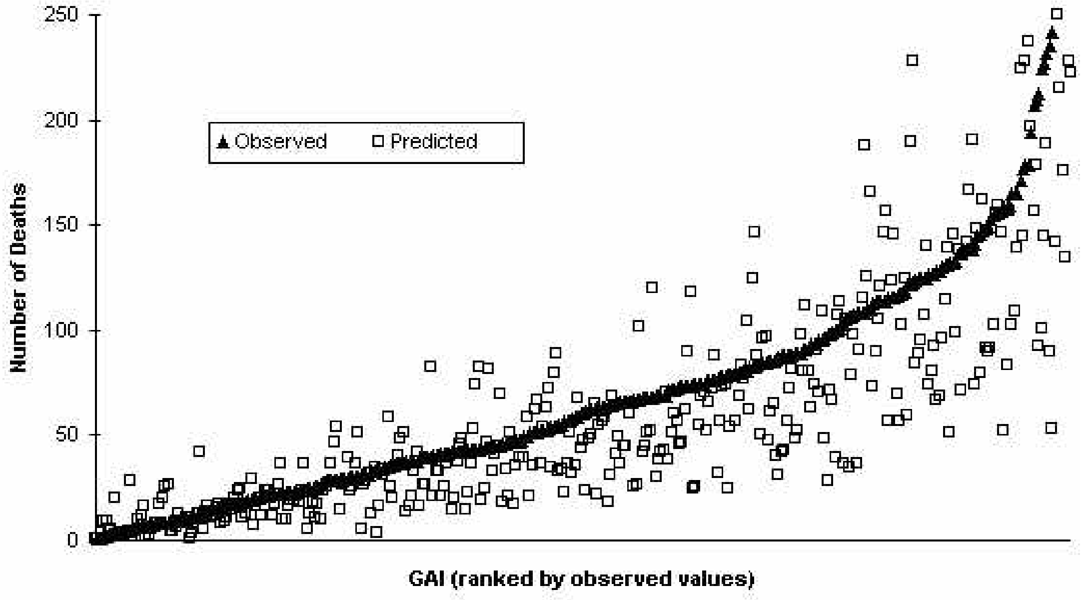
Parameter	Mortality Ratio (95% CI) All causes excluding injury, 0-74years	Mortality Ratio (95% CI) CVD
5% increase		
% females in Labour Force: unemployed	2.09 (1.4 - 3.2)	2.50 (1.2- 5.3)
% persons employed	1.37 (1.2 - 1.5)	1.39 (1.1 - 1.7)
% people 15 years and over who are divorced or separated	1.32 (1.0 - 1.7)	1.69 (1.0 - 2.8)
% Aboriginal or Torres Strait Islander (A&TSI)	1.21 (1.1 - 1.3)	1.30 (1.2 - 1.5)
% persons aged 50-59 years without educational qualifications	1.18 (1.0 - 1.3)	
% persons aged 15-49 years without educational qualifications		1.24 (1.1 - 1.5)
% persons who were non-English speaking	1.13 (1.1 - 1.2)	1.17 (1.0 - 1.3)
% persons in SES decile "9" *	0.93 (0.9 - 1.0)	
% persons in SES quintile "5" *		0.95 (0.9 - 1.0)
% persons with a different address in 1991	0.91 (0.8 - 1.0)	0.83 (0.7 - 1.0)
% persons aged 60 years plus without educational qualifications	0.90 (0.8 - 1.0)	
% persons employed in agriculture/forestry/fishing	0.90 (0.8 - 1.0)	0.81 (0.7 - 0.9)
% persons employed in manufacturing	0.86 (0.8 - 1.0)	0.76 (0.6 - 1.0)
% males employed as tradespersons		0.83 (0.7 - 1.0)
% population aged 45-69 years	0.72 (0.6 - 0.9)	
% population aged 15-24 years	0.64 (0.5 - 0.8)	0.70 (0.5 - 0.9)
% population aged 0-4 years	0.63 (0.4 - 1.0)	
% population aged 25-44 years	0.63 (0.5 - 0.8)	0.79 (0.6 - 1.0)
% population aged 5-14 years	0.52 (0.4 - 0.6)	0.47 (0.3 - 0.6)
Dichotomous		
Indigenous Community in GAI [Yes : No]	3.44 (2.6 - 4.6)	21.07 (9.3 - 47.7)
Capital City GAI [Yes : No]		2.34 (1.6 - 3.5)
Other Urban GAI [Yes : No]	1.28 (1.1 - 1.5)	2.79 (1.9 - 4.2)
Major Rural GAI [Yes : No]	0.53 (0.4 - 0.7)	
Other Remote GAI [Yes : No]	0.45 (0.3 - 0.8)	
Other Rural GAI [Yes : No]	0.30 (0.2 - 0.4)	0.51 (0.3 - 0.8)

* decile 10 and quintile 5 are the lowest SES levels in the respective schema

For all cause mortality (excluding injury), 42% of the 95% CI of the predicted value included the observed counts, 42% of the predicted intervals were below the observed counts, and 15% were above the observed counts (Figure 1). For CVD the corresponding figures were 42%, 40% and 18% respectively. Thus for both outcome measures, the predictive model tended to underestimate the observed mortality count.

The 20 GAIs with the highest relative difference between observed and predicted mortality were examined (using all cause mortality (excluding injury)) to assess whether there were any GAI-specific biases that might be able to explain the tendency of the models to underestimate mortality. There was no apparent sociodemographic similarity between these 20 GAIs. Of the 20 GAIs with highest observed mortality counts, eight (40%) were within the 95% confidence intervals of the predicted value. These analysis suggest that the model's performance was similar for both the outlier GAIs and the more "typical" GAIs.

Figure 1: Observed and predicted mortality due to All Causes (excluding Injury: 0-74 years), Queensland, 1992-1996, by GAI.*



* Note: the range of the y-axis has been compressed to reflect the majority (98%) of GAI-specific counts.

5. Discussion

Variation in mortality due to all causes (excluding injury: 0-74 years) and CVD in small geographic areas in Queensland, can be largely explained by the differential distribution of sociodemographic variables. Given the complex pathophysiology and long natural history leading to mortality from the majority of causes including CVD, the explanation of up to 73% of the deviance in population mortality by census derived sociodemographic variables at the time of death is noteworthy. The extent that routinely collected sociodemographic variables explain the differences in population mortality between small geographical areas has not previously been reported in the published literature

The set of sociodemographic variables that displayed significant independent associations with mortality due to all causes (excluding injury: 0-74 years) were similar to those for mortality due to CVD. Advancing age was strongly associated with both outcome measures. Mortality was likely to be higher in GAIs that contained an Indigenous community or had higher proportion of Indigenous people. Mortality was likely to be lower if the GAI was in a rural or remote area, after adjustment for Indigenous populations. This latter observation contrasts with the generally higher mortality rates reported for Queensland rural and remote areas,[5,9] and suggests that a substantial component of the reported higher mortality rate may be due to the increased proportion of Indigenous population in these areas.[5] Consistent with considerable literature, this highlights the poor health status of the Indigenous population.

The significant parameters of the predictive model for both all causes (excluding injury) and CVD included both the summary index of socioeconomic disadvantage and up to eight individual socioeconomic variables of education, labour force status and occupation. The significance of these individual socioeconomic status parameters is indicative of the importance of such variables in our understanding of the social and economic processes that produce health inequalities.[10]

Both population mortality models displayed limited predictive ability. Given the degree of accuracy required for population health and health service planning and benchmarking purposes, this suggests that these synthetic estimation techniques may have limited potential for predicting the health status of small populations. Local surveys may be the most accurate source of relevant and reliable information for small geographical areas. However, with the financial and personnel constraints imposed on local health authorities, this may not be practical. Instead, we may need to consider whether small area estimates are necessary for planning. In addition, health service planning based on population health status must recognise that within a geographic area there may be considerable variation in health status of individuals.

Valid application of synthetic estimation techniques assumes strong and stable associations between the explanatory variables and the outcome variable of interest,[7] and similarity between the outcome variable in the calibration and target populations.[11] The authors felt that neither of these assumptions were violated in this study. The expectation that synthetic estimation techniques could be used to predict small area health status with a high degree of accuracy was an extension of prior application of these techniques for larger populations and probably ambitious in light of the multifaceted aetiology, pathophysiology and long natural history of many diseases and the complex nature of populations.

Sociodemographic variables explain the majority of variation in mortality between small geographic areas in Queensland. However, the predictive ability of synthetic estimation of mortality is weak. Appropriate application of synthetic estimation techniques is dependent upon the knowledge of the relationships driving the outcome, the required accuracy of the estimates, availability of empirical data and the purpose of health status assessment.

Acknowledgments
The authors would like to acknowledge the valued advice, support and assistance provided by David Firman, Christine McClintock and Roscoe Taylor.

References

1. World Health Organisation. *Social determinants of health. The solid facts*. Copenganen: WHO, 1998.
2. Blane D. Commentary: Socioeconomic health differentials. *Int J Epidem* 2001; 30: 292-3.
3. Turrell G, Mathers C. Socioeconomic inequalities in all-cause and specific cause mortality in Australia: 1985-1987 and 1995-1997. *Int J Epidem* 2000; 29: 231-239.
4. Glover J, Tennant S. A Social Health Atlas of Australia (2nd ed.). Adelaide: Public Health Information development Unit, University of Adelaide, 1999.
5. Queensland Health. *Health Indicators for Queensland*. Public Health Services, Queensland Health. Brisbane 2001.
6. Taylor R. Small area population disease burden. *Aust NZ J Public Health* 2001; 25: 289-93.
7. Spasoff RA, Strike CJ, Nair RC, et al. Small group estimation for public health. *Can J Public Health* 1996; 87: 130-4.
8. Australian Bureau of Statistics. *Information Paper: 1996 Census of Population and Housing - Socio-economic indexes for areas*. Canberra: Government Printer, 1998.
9. Queensland Health. *1997/1998 Annual Report*. Brisbane: Queensland Health, 1998.
10. McCraken K. Into a SEIFA SES cul-de-sac? *Aust NZ J Public Health* 2001; 25: 305-6.
11. MacKenzie EJ, Shapiro S, Yaffe R. The utility of synthetic and regression estimation: techniques for local planning. *Med Care* 1985; 23: 1-13.

Appendix 1: GAI-specific sociodemographic variables included in regression models.

Sex
% of population who were male
% of population who were female

Age
% of population less than 5 years old
% of population aged 5 to 14 years
% of population aged 15 to 24 years
% of population aged 25 to 44 years
% of population aged 45 to 69 years
% of population aged 70 years and older

Ethnicity
% of population who were Aboriginal or Torres Strait Islander
Whether there was a DOGIT community in the GAI

Housing status
% of population aged 40 years and older renting from non – government authority
% of population renting from government authority
% of population renting from non - government authority
% of households with one or no bedrooms
% of population with a different address in 1996 compared to 1991

Employment
% of males unemployed
% of females unemployed
% of population employed
% of males employed as tradepersons
% of population employed as tradepersons
% of population employed in agriculture, forestry or fishing industries
% of population employed in the mining industry
% of population employed in the manufacturing industry
% of population employed in the construction industry
% of population employed in transport or storage industries

Families
% of one parent families with dependent children

Rurality (indicator variables)
Capital City area
Other Metropolitan area
Major Rural area
Other Rural area
Major Remote area
Other Remote area

SEIFA (where quintile 1 and decile 1 represent the population in the least socioeconomically disadvantaged group)
% of population living in SEIFA quintile 1
% of population living in SEIFA quintile 5
% of population living in SEIFA decile 1
% of population living in SEIFA decile 2
% of population living in SEIFA decile 9
% of population living in SEIFA decile 10

Educational Qualifications
% persons aged 15 - 49 with no qualifications
% of persons aged 50 - 59 with no qualifications
% of population aged 60 years and older with no qualifications

Marital status
% of population aged 15 years and older who were divorced or separated

Income
% of two - parent families with one or more dependent children and income less than $26,000
% of families with income less than $20,800

Language
% of populations who were non - English speaking

Access to motor vehicle
% of households with no motor vehicle