Using Census Data in Child Health Research

Hoan Nguyen, Peter Jacoby
Telethon Institute for Child Health Research

Census Analysis Conference - Canberra 2006
A Presentation in Two Parts

• Maternal and Child Health Research Database (MCHRDB) and Census Data (Hoan)

• Case Study (Peter)
Background of Maternal & Child Health Research Database (MCHRDB)

- 1980 Establishment of the midwives registry and birth defects registry
- MCHRDB was established in the mid 1980’s
- In 1984 and 1988 data from the midwives register, registrar general births and deaths and morbidity data were included in the database.
- 1990’s and 2000’s ongoing projects funded by National Health and Medical Research Council (NHMRC) grants and other sources
- Currently funded by an NHMRC Program Grant
- Purpose of the database is to provide child and maternal health data to support health research, health planning and policy development
Who are the users of the Database?

- Program Grant Researchers
- Other researchers within and outside the Institute
- Students - MPH, PhD and Postdoc
Census Products Used at ICHR

- CDATA 1986 (with Supermap)
- CDATA 1991, 1996 and 2001 (with MapInfo)
- Add-on Datapaks:
  - Socio Economic Indexes For Areas (SEIFA) – Profiling the Population
  - Usual Residents Profile
  - Indigenous Profile

- Urban Index of Relative Socio-Economic Advantage
- Rural Index of Relative Socio-Economic Advantage
- Index of Relative Socio-Economic Disadvantage
- Index of Economic Resources
- Index of Education and Occupation
SEIFA 2001

- Index of Advantage/ Disadvantage
- Index of Disadvantage
- Index of Economic Resources
- Index of Education and Occupation
FIGURE 1
CAUSAL MODEL RELATING GEOSOCIAL AND PSYCHOSOCIAL FACTORS TO HEALTH AND QUALITY OF LIFE

So, what did this project want to know?

- Location (urban versus rural)
- Health service provision level available
- SEIFA indexes that reflect the proportion of poverty of the area and the family environment where the child belongs
How do we match the addresses?

- Use the in-house application programs
- Select tools for manually checking unlinked records
- Cross-check with other projects
Using Census Data
Example 1

Association of sociodemographic characteristics of children with intellectual disability in Western Australia

Helen Leonard, Beverly Petterson, Nick De Klerk, Stephen R. Zubrick, Emma Glasson, Richard Sanders, Carol Bower

Social Science & Medicine 60 (2005) 1499-1513
Using Census Data

Example 1

Aims

• Investigating sociodemographic correlates of intellectual disability (ID) of unknown cause in WA born children

Methods

• Maternal sociodemographic characteristic of 2871 ID children born between 1983 and 1992 were compared with those of 236,964 children without ID

• SEIFA based on the census district of mother’s residence were included in analysis
Mild-moderate ID by
Index of relative social disadvantage

Odds Ratio

least disadvantaged  →  most disadvantaged
Mild-moderate ID
by index of education and occupation

Odds Ratio

least disadvantaged  most disadvantaged
Using Census Data

Example 2

Independent effects of socioeconomic status and place of residence on the incidence of childhood type 1 diabetes in Western Australia

Aveni Haynes, Max K Bulsara, Carol Bower, Jim P Codde, Timothy W Jones and Elizabeth A Davis

Pediatric Diabetes (2006: 7) 94 - 100
Aims

• To analyse the incidence of type 1 diabetes in 0-14 year olds from 1985 to 2002 by socioeconomic status and for regional variation

Methods

• Residential postcode at the time of diagnosis was used to categorize cases into urban, rural or remote areas

• Index of Disadvantage was used to analyse the incidences
Difference in incidence of type 1 diabetes by area (1985-2002)

Mean age standardised incidence (per 100,000)

- Metropolitan: 18.1
- Rural: 14.3
- Remote: 8.0

Haynes et al (Ped Diabetes, 2006)
Difference in incidence of type 1 diabetes by socioeconomic status (1985 - 2002)

Group 5 v Group 1: IRR 1.56 (1.29 - 1.88), $p < 0.001$

Haynes et al (Ped Diabetes, 2006)
Difference in incidence of type 1 diabetes by socioeconomic status - metropolitan area only

Group 5 v Group 1: IRR 1.54 (1.25 - 1.91), p < 0.001

Haynes et al (Ped Diabetes, 2006)
Using Census Data
Example 2

Selected Results

- Significant difference in incidence of type 1 diabetes between metropolitan, rural and remote areas of Western Australia
- Highest incidence found in the metropolitan area
- Incidence of type 1 diabetes increased with higher socioeconomic status
Using Census Data in Child Health Research

Hoan Nguyen, Peter Jacoby
Telethon Institute for Child Health Research

Census Analysis Conference - Canberra 2006
Case Study
Disease Clustering

Is there geographical clustering of cases?

Can this be accounted for by socio-demographic variations?

Investigate environmental factors
- Air pollution (roads, industry etc)
- EM Radiation (power lines, mobile phone towers etc)
- UV Radiation
- Landfill sites
- Agriculture (pesticide spraying etc)
- Etc

Clustering in space AND time
All Childhood Cancers in Western Australia 1983-2002

Detection of space/time clusters at the census collection district (CD) level

Initial Data:

- Date of diagnosis for each case (N=879)
- CD of residence for each case
- Total child population (<15) for each CD from 2001 census
Western Australia
2001 Census
4408 Collection Districts (CDs)
Clustering:

A more heterogeneous and clumped distribution of disease cases than would be expected from variations in the population density and chance fluctuations (Alexander and Cuzick 1992).

**SaTScan Software** (Kulldorff 2005)

- Uses spatial scan statistic to detect spatial or space-time clusters of disease and assess whether they are statistically significant
- Free download from [http://www.satscan.org](http://www.satscan.org)
Primary Cluster (1985-93)
p=0.004
Assumption made: child population has remained constant during study period at 2001 levels

Include population data for each CD from 1991 census in analysis
(SaTScan interpolates between dates)

- **PROBLEM**
  CD boundaries change from one census to the next

- **SOLUTION**
  Use area-weighting in ArcGIS to apportion 1991 child population on to 2001 CD boundaries
Area weighting

Population \((P_i)\)  
Area \((A_i)\)

Source file

Target file

Area \((A_k)\)

Union

Unique polygons \((\text{source} \times \text{target})\)

Calculate population \((A_{ik}/A_i)*P_i\)

Dissolve on target file boundaries

Population by target area \((P_k)\)
Old Boundaries

New Boundaries

Union

Area Weighting

Pop_{U4} = Pop_{14} \times \frac{\text{Area}_{U4}}{\text{Area}_{14}}

etc

Dissolve

Pop_{21} = Pop_{U1}
Pop_{22} = Pop_{U2} + Pop_{U3}
Pop_{23} = Pop_{U4}
Pop_{24} = Pop_{U5}
South-Western Australia - Disease clusters after including population changes
Conclusions

- No evidence of space/time clustering of childhood cancers in WA
  BUT.......power of study is limited by sample size

- Failure to account for population changes can lead to dangerously misleading conclusions
Acknowledgements

We would like to thank the following ICHR staff for their contributions:

- Stephen Zubrick
- Helen Leonard
- Aveni Haynes
- Peter Cosgrove
- Liz Milne
- Tim Threlfall (WA Dept. of Health)
- Nick De Klerk

Also, as CDATA users, we would like to thank the Census Technical team who have done a great job in supporting and helping with CDATA and Mapinfo issues for the past 3 censuses.

We’re hoping that ABS will maintain this service for future censuses.