Spatial Patterns of Disease and Risk Factors using Small Local Rural Datasets

Rachel Whitsed, Ana Horta & Herbert Jelinek
Prevalence of NCDs increasing

Rural Australian population has worse health outcomes compared to the urban population
Diabetes risk factors

- Tobacco and alcohol use
- Lack of physical activity
- Diet
- Obesity and overweight
- Socio-economic disadvantage
Spatial patterns of diabetes

Diabetes map of Australia
National Diabetes Services Scheme
Spatial patterns of diabetes

National scale
• Rural-urban divide

Local scale
• Population characteristics
• Natural environment
• Built environment
Spatial patterns of diabetes

Small rural and regional communities

• National health databases cannot provide data at fine enough spatial scale
• But local databases tend to be very small

Crucial for decision-makers to understand spatial patterns of NCDs and risk factors
Disease cluster analysis

Highlight locations where diseases are clustered

- Looks for areas which are more likely to be near known locations of disease
- May include smoothing
- May incorporate background population
- Other risk factors
Spatial representativeness

• Dataset 515 participants in Albury aged over 45 in a health-screening program
• How representative is the database of the background population?
Generalised Additive Model (GAM)

- Commonly used in health geography for disease hotspot analysis
- **Can also be used to assess database representativeness**
- We used a modified GAM with sensitivity analysis (GAM-SA)
- Areas with relative risk close to 1 are representative of background population
Database in Albury is spatially representative
Generalised Additive Model (GAM)

Commonly used for in health geography for disease hotspot analysis

We used a modified GAM with sensitivity analysis

Areas with high relative risk are potential disease hotspots
Other spatial modelling approaches

Machine learning algorithms commonly used for species distribution modelling or site suitability selection

- **MaxEnt** – uses Maximum Entropy to determine probability of distribution
- **Random Forest** – uses multiple decision trees to determine probability of distribution
MaxEnt and Random Forest

Predict spatial patterns of disease based on spatial patterns of risk factors
Relative importance of risk factors
Relationship of risk factors with disease
MaxEnt results

Probability of diabetes

- High
- Low
Random Forest results

Probability of diabetes

- High
- Low
Comparisons

GAM-SA

MaxEnt

Random Forest
Common patterns

Hotspots of diabetes

Important risk factors

• Age
• Obesity and overweight
• Socio-economic disadvantage
• Terrain ruggedness
• Proximity to parks
Implications for modelling

Hotspot methods

• Can show where unexplained risk factors may be present
• Can show changes in hotspots when different risk factors are included or omitted
• Can smooth spatially, but doesn’t extrapolate
Implications for modelling

Machine learning methods (MaxEnt, Random Forest)

- Can extrapolate to show locations similar to known disease locations
- Can show how each risk factor contributes to the disease
Implications for rural health management

Small local datasets can show spatial patterns of disease and risk that might otherwise be overlooked.

Can be used for evidence-based region-specific population health planning.

Targeted place-based interventions in rural and regional towns.
Thank you

Contact
Rachel Whitsed
Institute for Land, Water and Society
Charles Sturt University
rwhitsed@csu.edu.au
02 6051 9641