

Synthetic predictions

The synthetic predictions of the prevalence of chronic disease and associated risk factors have been produced for a majority¹ of Statistical Local Areas (SLAs) in Australia, using modelled survey data collected in the 2007-08 ABS National Health Survey (NHS) and known characteristics of the area.

A synthetic prediction can be interpreted as the likely value for a 'typical' area with those characteristics: the SLA is the area level of interest for this project (where SLAs had small populations they were grouped to larger areas). This work was undertaken by the Australian Bureau of Statistics (ABS), as they hold the NHS unit record files on which the model is based: the small area data were compiled by PHIDU.

The approach used is to undertake an analysis of the survey data for Australia to identify associations in the NHS data between the variables that we wish to predict at the area level (e.g., prevalence of chronic conditions and risk factors) and the data we have at the area level (e.g., socioeconomic status, use of health services). The relationship between these variables for which we have area level data (the predictors) and the reporting of chronic conditions in the NHS is also a part of the model that is developed by the ABS. For example, such associations might be between the number of people reporting specified chronic conditions in the NHS and:

- the number of visits to a general medical practitioner,
- the proportion of the population receiving a pension or benefit,
- the proportion of the population with a profound or severe disability and
- socioeconomic status (as indicated by Census data).

The results of the modelling exercise are then applied to the SLA counts of the predictors. The prediction is, effectively, the likely value for a typical area with those characteristics. The raw numbers were then age standardised, to control for the effects of differences in the age profiles of areas.

The numbers are estimates for an area, not measured events as are death statistics: they should be used as indicators of likely levels of a condition or risk factor in an area.

| | Outcome | | Predictor | | Relationship |
|-------|----------|---|-----------|---|--------------|
| NHS | <u>A</u> | x | <u>B</u> | = | <u>C</u> |
| Local | E | = | <u>D</u> | x | <u>C</u> |
| --- | | | | | |

A = Variables in NHS to be predicted at small area level – outcome variables

B = Variables in NHS (and available at local area level (D)) – predictor variables

C = relationships between A and B found in NHS data (from modelling NHS data) for predictor variables

D = Variables at local area level also in NHS – predictor variables

E = Synthetic estimates

Note: Underlining indicates known data (for C, is a known (calculated) relationship)