

**Title:** *Enhancing the Australian National Health Survey Data for Use in a Microsimulation Model*

**Author(s):** Annie Abello, Sharyn Lymer, Laurie Brown, Ben Phillips and Ann Harding

**Institution/Country:** National Centre for Social and Economic Modelling (NATSEM), Canberra

An important component of the economic well-being of low income families in Australia is their access to subsidised drugs through the Australian Pharmaceutical Benefits Scheme (PBS). This paper documents the methodology used to overcome survey data limitations when constructing 'MediSim', a microsimulation model of the PBS. The PBS provides Australians with affordable access to necessary and cost-effective prescription medicines. Federal government expenditure on the PBS in 2004-05 was \$Aus 5.4 billion, or 15 per cent of total Government health expenditure.

There is considerable debate over the sustainability of the PBS, with the scheme being one of the fastest growing sectors within health. MediSim was constructed to estimate current and future use and costs of PBS medicines under different policy settings, and to quantify the distributional effects of policy changes. A necessary step to modelling both the costs and benefits of these medicines is the inclusion of diseases and health conditions in the model's base file. MediSim was built using the most recent Australian National Health Survey (NHS). The 2001 NHS contains the latest person-level information in Australia on long-term health conditions, drug usage for national health priority conditions, and health risk factors.

However, it has a number of limitations as the base data for a microsimulation model, following changes in survey design implemented in the 2001 NHS. These include the incomplete coverage of families (as the survey structure is person-based); lack of data on short-term (ST) health conditions; and lack of data on prescribed drug usage for ST and non-priority long-term (LT) conditions.

A number of steps were taken to overcome these limitations. First, we statistically matched the NHS with another ABS national survey, to create synthetic families and get a complete record for every individual within each family. Family level information is needed to model PBS safety nets. The statistical matching allowed the retention of the health information available on NHS 2001, whilst borrowing the family structure from another survey.

Next, we imputed ST health conditions based on detailed information in the previous (1995) NHS and converted the two-weekly prevalence rates derived from the 1995 survey's two-week recall period to annual figures. ST conditions were imputed according to selected characteristics including age, sex, government concession cardholder status and self-assessed health status. The 1995 prevalence rates were uprated to 2001 levels. To expand the two-week estimate to a full-year estimate, prevalence rates were increased to a bi-monthly rate then the short term conditions were imputed six times in the model dataset until the up-rated prevalence figures were met. To assign ST health conditions across the model dataset with a sufficient level of precision, it was necessary to clone records (i.e. create multiple records of the same person in the dataset). Finally, we imputed annual drug usage for ST and non-priority LT health conditions. These initial estimates were then aligned to administrative data on PBS benefit drugs. The application of statistical matching methods and use of complementary data sets significantly improved the usefulness of the 2001 NHS as a base dataset for MediSim.