Shared code

The following code might be useful to other Integrated Data Infrastructure (IDI) users in the Virtual Health Information Network (VHIN).

Address information in the IDI

This piece of SAS code selects the most recently updated address for an individual at a given date. It requires a 2017 annual areas file that can be downloaded from the SNZ website. This method uses all address sources from the address_notification_full table. If an address is not recorded prior to the reference date, the first address update within the 12 months following the reference date is used.

The above method is used in a range of projects and gives a good approximation of address at a given date. However, depending on the specific requirements for your project you may wish to vary the above method by (for example) excluding some address sources, or excluding address updates that occur after the reference date.

Census variables

This is SAS code developed for VHIN Healthier Lives National Science Challenge projects. The code creates a dataset of individuals and their characteristics according to variables derived from the IDI census tables from the 2013 census. It puts descriptive labels and formats on all the retained variables.

Some of these variables have been used previously in the New Zealand Census-Mortality Study and the New Zealand Deprivation Index development.

Variables include:

- Snz_uid, the unique IDI identifier

- Usual residence indicator

- Demographic factors such as sex, urban/rural, family type, partnership status

- Socioeconomic variables such as highest qualification, equivalised income, household crowding according to the Canadian Household Crowding Index

- Ethnicity including total ethnicity (Maori, Pacific and Asian) and prioritised ethnicity

- Smoking status

More information on the names, formats and labels for these variables are listed here.

Earthquake study

This is SAS code developed for VHIN Healthier Lives National Science Challenge project examining the impact of housing damage from the Canterbury earthquakes on the incidence of cardiovascular disease.

The first set of code creates a dataset of individuals who were living in Canterbury and the rest of New Zealand on the 3rd September 2010, and their age, ethnicity, income, meshblock, and area level deprivation. This data is linked to subsequent information on follow-up, including cardiovascular hospital admission or death, and censoring for time outside the country.

Other code available is an example of producing age standardised incidence rates of cardiovascular events by area level housing damage. Finally an example of Poisson regression is provided; it examines the association between housing damage and cardiovascular events.

For more information on this analysis please see this publication in the Lancet Planetary Health.

Costs of cardiovascular disease in NZ

[This is the SAS code](#) from the first stage of the VHIN catalyst project 'Costs of cardiovascular disease in NZ'.
The code creates a file with:
- individual-level health costs from PHO, NMDS, NNPAC, labs and pharmaceuticals
- flags and dates for cardiovascular events
- basic demographic information

Risk factors for congenital malformations in NZ

[This is the SAS code](#) from the VHIN catalyst project 'Risk factors for congenital malformations in NZ'.
The code creates a data file for use in the VHIN 'Congenital Malformations' catalyst project. It includes code for linking babies to parents via birth records and creating flags for pharmaceuticals prescribed during pregnancy.

Getting the denominator right

[This is the first of the SAS code files](#) from the VHIN catalyst project 'Getting the denominator right'.
The code creates a list of the resident population of NZ for a given date, with basic demographic information attached. The methodology is based on that used by Statistics NZ's Census Transformation team (selects people into the population if they have had activity in health, tax, or education data in the last 12 months, then removes people who died or moved overseas).

[This is the second of two SAS code files](#) from the VHIN catalyst project 'Getting the denominator right'.
This code creates a dataset with a range of demographic, geographic, health (mostly cardiovascular) and pharmaceutical variables for everyone in the NZ resident population.

See the page [Identifying the NZ ERP in the IDI](#) for more information about identifying a population in the IDI.

Other

The above code documents are also available on the IDI Wiki and on [Meetadata](#). You can request to access Meetadata by emailing meetadata@snz.govt.nz. You can access Meetadata outside of the IDI using a RealMe login.

Examples of IDI code are also publically accessible from [Statistics New Zealand](#) and on [GitHub - StatisticsNZ/IDI](#) – an IDI code sharing website used by NZ government agencies.

Original by Sheree Gibb, updated by Andrea Teng, June Atkinson and Sheree Gibb 8/12/2017