Online Report: Validation Issues Relating to the Cardiovascular Disease Model Developed in the BODE^3 Programme

Nhung Nghiem, Nick Wilson, Tony Blakely

BODE^3 Programme, University of Otago, Wellington, New Zealand

March 2014

Contents

Figures

Tables

Summary

This Web Appendix presents work on validating the data used in the New Zealand cardiovascular disease (CVD) model we developed for the BODE^3 Programme. The model is largely parameterised with HealthTracker data for consistency, but it was therefore important to first compare HealthTracker estimates (e.g., mortality rates) with other 'official' data, namely NZ Burden of Disease Study (NZBDS) data and Statistics New Zealand (SNZ) data. We also undertook analysis of epidemiological coherence for key parameters for coronary heart disease (CHD) and stroke by using DisModII software.

The results suggest a fairly close match between the HealthTracker mortality data we used (both observed and smoothed) and the mortality data from SNZ. The exceptions were somewhat higher numbers of stroke deaths in older ages using the smoothed data for Maori females and Maori males. However, since the smoothed data produced far more epidemiologically plausible results using DisModII (see below) – we decided to continue to use the smoothed data in the CVD Markov model in TreeAge.

DisModII results for the non-Maori male population (for coronary heart disease, ischaemic stroke, and haemorrhagic stroke) indicated that the use of smoothed input data produced more epidemiologically plausible results than the use of observed (unsmoothed) input data.

DisModII results for the Maori female population (for coronary heart disease, ischaemic stroke, and haemorrhagic stroke) also indicated that the use of smoothed input data was substantially more stable and produced more epidemiologically plausible results than the use of observed (unsmoothed) input data.

Finally we compared the results of the CVD Markov model in TreeAge against currently observed mortality rates, under steady state (i.e. no future time trends in incidence, case fatality and mortality) conditions. From this analysis there seemed to be reasonably good agreement for overall mortality (albeit with some divergence at very old ages i.e., at 90+ years) between our model output and SNZ data. Similarly, there was reasonably good agreement between the model's CVD mortality results and those from the NZBDS.

In summary, these comparisons between data we used and with SNZ data and NZBDS data provide reassurance around our model data inputs, around the use of smoothed data inputs, and around key aspects of the functioning of the CVD Markov model we have built.

Issues addressed in this Appendix, and Objectives

The BODE^3 Programme has rich and varied input data, including HeathTracker, NZ Burden of Disease Study (NZBDS) and other routine (e.g., Ministry of Health and Statistics New Zealand data) from which to choose to build the CVD Markov model. Our preference was to

largely use HealthTracker data, as one consistent source – and indeed, it was the source of much of the NZBDS data as well. With respect to these input variables, three key issues arise: do HeathTracker parameters largely agree with other sources (when available, eg, SNZ all-cause mortality rates); are the data epidemiological coherent?; and are the data stable?

By epidemiological coherence, we mean that in a stable system all of incidence, prevalence, case fatality, remission and mortality should 'agree' through various mathematical relationships.[1] This can be assessed with DisModII. (We note that we do not allow for time trends here; future BODE^3 models probably will.) The key *a priori* threat to coherence was that the HealthTracker data, as good and useful as it is, does not allow an accurate and long look-back period. This would not be a problem for short duration diseases with high 'exit' rates through either remission or case fatality, as (say) with a five year look-back period one would achieve equilibrium by the end of the five years for prevalence, case fatality and remission rates. But consider CHD. If we assume zero remission rates (i.e., a diagnosis for life), then you need a long look-back period to derive accurate prevalence rates. And case fatality rates may not be correct either, as missing prevalent cases (the denominator for the CFR) may have lower CFRs meaning the observed CFR is 'too high'. When this high CFR is inputted to a CVD Markov model that simulates people for decades, one may therefore overshoot the number of CHD deaths. How problematic this is should be judged relative to other (structural) uncertainties in the modelling, and may be 'not that important' or even 'trivial' compared to uncertainty about future incidence rates, the intervention effect, and such like. So perfection in the input data is not strictly necessary; rather parsimony is also an over-riding consideration. Nevertheless, it is possible to at least assess the epidemiological coherence of the input parameters with DisModII, and make a decision as to whether further refinement of input data is required before embarking on other aspects of the model building and parameterisation.

By stability, we mean that the sample sizes from whence the parameters are sourced are large enough that large random error is not problematic in the observed parameters by sex, age, and ethnic groups. This is particularly an issue for older Māori. Assuming that epidemiological coherence is first assured, there are several possible approaches to instability:

- Just use the observed data as inputs to the CVD Markov model, accept that there will be 'unders and overs' across various strata or socio-demographic groups, but assume for all ages combined (say) that these unders and overs largely cancel or net out.
- Smooth or stabilise data before inputting it to the CVD disease model. There are a multitude of ways to smooth. Two we consider in this document are: regression-based smoothing on incident rate, prevalent rate, case fatality rate within 28 days after the first ever CVD event and case fatality rate after 28 days, disease specific mortality rate and background mortality rate, by assuming that relative differences by sex, age and ethnicity apply (i.e. a main effects log-link regression from which we use the predicted values by socio-demographic strata); actually using DisModII outputs – which are necessarily smoothed.

The second major issue, and highly dependent on the above, is assessing the validity of the CVD Markov model.[2] The aspect of validity addressed in this Appendix is external validity, and more specifically calibrating the model against the observed input data. That is, ensuring that the outputs (e.g. death rates, prevalence, etc) from the CVD Markov model under steady state settings (e.g. no future change in incidence and case fatality rates, no change in background mortality rates) with no intervention applied agree with that observed in New Zealand currently. Thus, it also acts as a 'total' assessment on all structural assumptions in the model, input parameters (especially those about epidemiological coherence above).

Thus, this Appendix has four objectives:

1. To assess agreement of HeathTracker mortality estimates with other Ministry and SNZ data.
2. To assess the epidemiological coherence of the preferred input parameters from HealthTracker, using DisModII.
3. To assess whether observed or smoothed input parameter data-sets should be used in the CVD Markov model.
4. To calibrate the CVD Markov model against currently observed mortality rates, under steady state (i.e. no future time trends in incidence, case fatality and mortality) conditions.

By way of context, the first section of this Appendix gives the structure of the final CVD model – it helps to know what structure we are aiming for. Then the remaining sections address the four objectives above in turn.

Data dictionary

This section describes how key epidemiological parameters inputs in the CVD model were generated and calculated from Health Tracker and other data sources (eg, Statistics New Zealand). All the epidemiological input rates used in the CVD modelling were smoothed using Poisson regression in SAS 9.3. All the negative rates, if any, were dealt with in SAS.

Table 1 Definition of key epidemiological parameters used in the CVD model and how they were generated from Health Tracker data

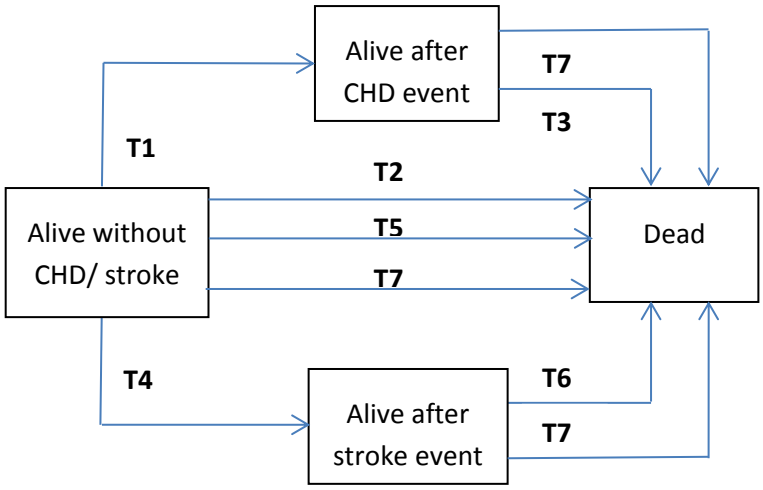
Key epidemiological parameters	Definitions
Input variables (that were actually put into TreeAge)	
Incident rate for people who survive at least 28 days after their first-ever CVD event.	We used the Health Tracker database from 2001 to 2010 to extract incident cases, prevalent cases, and death counts in order to calculate epidemiological data inputs. We used a look-back period of 5 years (2001-2006) to identify a "first-ever CVD event" (eg, a non-fatal heart attack, stroke, or being started on CVD medicines after a diagnosis of CVD).

Key epidemiological parameters	Definitions
	An incident case in a particular year, say 2007, was identified if the person had a CVD event in that year and hadn't had any CVD event between 2001 and 2006. Only incident cases in 2010 were used to calculate incident rates to put into TreeAge. Incident 28 days survivors refer to people who had a first-ever CVD event and survived for at least 28 days after that event. The incidence rate for people who survive at least 28 days after their first-ever CVD event was calculated by dividing the number of 28 days survivor incident cases by the total healthy population (by age, sex & ethnicity).
Case fatality rate (CFR) pre-28 days	Case fatality rates pre-28 days were calculated by dividing all CVD deaths from people who had a first-ever CVD event and died in under 28 days after that event, by total CVD incident cases.
Case fatality rate post-28 days	Case fatality rates post-28 days were calculated by dividing the difference between deaths observed from 28-day-incident-survivors and deaths expected (from the non-CVD background mortality rate – see below), by person-years lived among those with a past CVD event. Of note is that we suspect that due to the constraints on the look-back period with HT data (leading to an under-estimate in the number of CVD prevalent cases), the estimated CFR will tend to be over-estimated by our methods. However, this is something that can potentially be addressed (eg, via DisModII outputs).
Non-CVD background mortality rate	Non-CVD mortality rates were calculated by dividing the total deaths (excluding CVD deaths) by the total New Zealand population in 2010. The death counts were estimated from Health Tracker, and the New Zealand population data were adopted from Statistics New Zealand.
Prevalent cases (CVD)	Prevalent cases were defined as all people with a reported CVD event from 2001 to 2010 (but incident cases arising in each year were excluded from prevalent counts for that year, but included in subsequent years). We recognise that this will be an underestimate for two reasons: (i) Individuals could have had a CVD much earlier (even decades) and still be alive; (ii) Some individuals will have had CVD events eg, a "silent myocardial infarction" that have never been diagnosed. We consider this issue further when using DisModII outputs. We considered prevalent cases for coronary heart disease (CHD) and the different forms of stroke.
Other related input variables (that were not put into TreeAge)	
Person-years	Person-time (person-years) was calculated by summing up all the days that people lived from the date that they were diagnosed with various CVD conditions until they died or until the end of the period. The starting date was 1 January 2007, and the ending date was 31 December 2010. So a person could have a maximum of four years of "person-time". The reason we used the period of 2007-2010 to calculate CFR was that the TreeAge model involves a one year step and we applied the same CFR for all the 28 day survivors until they died. It would be ideal to use a longer period but we only had cost data from 2007 and mortality data up to 2010.
Deaths observed	Deaths observed from 28-day-incident-survivors were for all causes of death, based on death certificates for the period of 2007-2010.
Deaths expected	Deaths expected were all causes background mortality rate multiplying with CVD person-years for the period of 2007-2010.

Model structure – CVD model for New Zealand

For the purposes of making this Appendix on validation issues more self-contained, we repeat some material of the basics of the CVD model below. But for further detail please see the "Main Technical Appendix" on the CVD model.

Figure 1 Structure of the Markov model for the CVD model built in TreeAge (simplified and not showing the different forms of stroke)



Transitions:
T1: Incident CHD, non-fatal in 1st year
T2: Incident CHD, fatal in 1st year
T3: CHD case fatality in CHD survivors
T4: Incident stroke, non-fatal in 1st year
T5: Incident stroke, fatal in 1st year
T6: Stroke case fatality in stroke survivors
T7: Non-CHD and non-stroke mortality

Table 2 Details around the transition probabilities in the CVD Markov model: mathematical formula and descriptions

Transition probability	Formulae used in the TreeAge model	Mathematical formula and additional descriptions
T1: Incident CHD, non-fatal in 1st year	**Formula for Partial Null:** paCHD= (IncUCHD*RRCHDvin-5/12*CFR[age+iDat;7]*IncUCHD*RRCHDvin*If (trendIdx<15;exp(-CfrCHDtr *(trendIdx+5));exp(-CfrCHDtr*20)))*If (trendIdx<15;exp(-IncCHDtr*(trendIdx+5));exp(-IncCHDtr*20)) **Formula for Intervention:** paCHDt=paCHD*(RRCHD_T*RRCHDsalt)	**Partial Null:** $T1 = \left(Inc28_{a,s,e} - \frac{5}{12} \times CFRpost28_{a,s,e} \times Inc28_{a,s,e}\right) \times CFRTrend \times IncTrend$ *Inc28*: 28 days survivor incident rate (see the formula below). *CFRpost28*: case fatality rate for people who survive at least 28 days after the first-ever CHD event (see the definition in Table 1). *CFR/IncTrend*: time trend for case fatality rate (CFR)/incident rate. *a*: age; *s*: sex; *e*: ethnicity. These subscripts applied for all key epidemiological parameters, eg, mortality rates. The value of 5/12 was used in the *CFRpost28* formula since we used a half-cycle correction for the model (so it was 6/12 months) and took into account 28 days for *CFRpre28*, which was 1/12 months. **Intervention:** $T1 = \left(Inc28_{a,s,e} - \frac{5}{12} \times CFRpost28_{a,s,e} \times Inc28_{a,s,e}\right) \times CFRTrend \times IncTrend \times RRCHD'_{sodium}$ $RRCHD'_{sodium}$: new relative risk for CHD after an intervention (see formula below) as a result of decreasing dietary sodium intake.
T2: Incident CHD, fatal	**Partial Null:** pdCHD= (CFR[age+iDat;6]*If (trendIdx<15;exp(-CfrCHDtr *(trendIdx+5));exp(-	**Partial Null:**

Transition probability	Formulae used in the TreeAge model	Mathematical formula and additional descriptions
in 1st year	CfrCHDtr*20))/(1-CFR[age+iDat;6]*If (trendIdx<15;exp(-CfrCHDtr *(trendIdx+5));exp(-CfrCHDtr*20)))+5/12*CFR[age+iDat;7]*If (trendIdx<15;exp(-CfrCHDtr *(trendIdx+5));exp(-CfrCHDtr*20)))*IncUCHD*RRCHDvin*If (trendIdx<15;exp(-IncCHDtr*(trendIdx+5));exp(-IncCHDtr*20)) **Intervention:** pdCHDt= pdCHD*RRCHD_T*RRCHDsalt	$T2 = \left(\frac{CFRpre28 \times Trend}{1 - CFRpre28 \times Trend} + \frac{5}{12} CFRpost28 \times Trend\right) \times (Inc28 \times Trend)$ *CFRpre28*: case fatality rate for people who died within 28 days after the first ever CHD event (see the definition in Table 1). *Trend*: a time trend for that variable. **Intervention:** See formula for *T1* intervention as above.
T3: CHD case fatality in CHD survivors	**Partial Null & Intervention:** pdpCHD= CFR[age+iDat;7]*If (trendIdx<15;exp(-CfrCHDtr *(trendIdx+5));exp(-CfrCHDtr*20))	$T3 = CFRpost28 \times Trend$
T4: Incident stroke, non-fatal in 1st year	**Partial Null:** IS: paS= ((IncUIS*RRSTRvin)-5/12*CFR[age+iDat;4]*(IncUIS*RRSTRvin)*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20)))*If (trendIdx<15;exp(-IncSTRtr *(trendIdx+5));exp(-IncSTRtr*20)) HS: paHS= ((IncUHS*RRSTRvin)-5/12*CFR[age+iDat;8]*(IncUHS*RRSTRvin)*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20)))*If (trendIdx<15;exp(-IncSTRtr *(trendIdx+5));exp(-IncSTRtr*20)) **Intervention:** IS: paSt= ((IncUIS*RRSTRvin*(RRIS_T*RRSsalt))-5/12*CFR[age+iDat;4]*(IncUIS*RRSTRvin*(RRIS_T*RRSsalt))* If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20)))* If (trendIdx<15;exp(-IncSTRtr *(trendIdx+5));exp(-IncSTRtr*20)) HS: paHSt= ((IncUHS*RRSTRvin*(RRHS_T*RRSsalt))-5/12*CFR[age+iDat;8]*(IncUHS*RRSTRvin*(RRHS_T*RRSsalt))* If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20)))* If (trendIdx<15;exp(-IncSTRtr *(trendIdx+5));exp(-IncSTRtr*20))	See formula for *T1* (albeit adapted for stroke). We considered multiple forms of stroke (ischaemic, haemorrhagic).

Transition probability	Formulae used in the TreeAge model	Mathematical formula and additional descriptions
T5: Incident stroke, fatal in 1st year	**Partial Null:** IS: pdIS= (CFR[age+iDat;5]*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20))/(1-CFR[age+iDat;5]*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20)))+5/12*CFR[age+iDat;4]*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20)))*IncUIS*RRSTRvin*If (trendIdx<15;exp(-IncSTRtr *(trendIdx+5));exp(-IncSTRtr*20)) HS: pdHS= (CFR[age+iDat;3]*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20))/(1-CFR[age+iDat;3]*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20)))+5/12*CFR[age+iDat;8]*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20)))*IncUHS*RRSTRvin*If (trendIdx<15;exp(-IncSTRtr *(trendIdx+5));exp(-IncSTRtr*20)) **Intervention:** IS: pdISt= pdIS*RRIS_T*RRSsalt HS: pdHSt= pdHS*RRHS_T*RRSsalt	See formula for *T2* (albeit adapted for stroke).
T6: Stroke case fatality in stroke survivors	**Partial Null & Intervention:** IS: pdpIS= CFR[age+iDat;4]*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20)) HS: pdpHS= CFR[age+iDat;8]*If (trendIdx<15;exp(-CfrSTRtr *(trendIdx+5));exp(-CfrSTRtr*20))	See formula for *T3* (albeit adapted for stroke).
T7: Non-CHD and non-stroke mortality	**Partial Null & Intervention:** pdO= CFR[if(IDinequality+iDateth=1;age+iDat+600;age+iDat);2]* If (trendIdx<15;exp(-mortNonCVDtr *(trendIdx+5));exp(-mortNonCVDtr*20))	Direct input into the model: non-CVD background mortality rate by age, sex & ethnicity. (The formula in TreeAge took into account scenario analysis).
IncUCHD: Inc28	(IncCHD[age+iDat;1]/(1-IncCHD[age+iDat;2]+IncCHD[age+iDat;2]*exp(RRCHDstr_G)))*IncCHD[age+iDat;3]	Direct input into the model: incident rate by age, sex & ethnicity for people who survived at least 28 days after the first CVD event. (The formula in TreeAge was built based on the original ACE-Prevention (Australia) CVD model in an Excel spreadsheet but two variables in columns 2&3 of the "IncCHD" table were no longer used).

Transition probability	Formulae used in the TreeAge model	Mathematical formula and additional descriptions
CfrCHDtr/CfrSTRtr	2%	Time trend in case fatality rate for CHD and all forms of stroke (see elsewhere for our justification of this future time trend)
IncSTRtr/IncCHDtr	2%	Time trend in incident rate for CHD and all forms of stroke (see elsewhere for our justification of this future time trend)
RRCHDvin	1	Relative risk CHD for different age groups: a reserved variable for future CVD modelling.
RRCHD_T	1	A reserved relative risk variable for future CVD modelling
$RRCHD'_{sodiu}$ (RRCHDsalt/ RRSsalt)	dBP*effCHDsU_var*100+1	New relative risk for CHD after an intervention as a result of decreasing dietary sodium intake (with a pre-intervention relative risk equal to 1.0): $$RRCHD'_{sodium} = \Delta BP \times EffectSize \times 100 + 1$$ ΔBP*:* percentage change in systolic blood pressure after an intervention (see formula below). *Effect size:* percentage change in CHD incident rate with a one percent change in systolic blood pressure.
dBP	(SBPm[age+iDat;1+iCounselling]*(dNAm/100)/SBPm[age+iDat;2])	Percentage change in systolic blood pressure after an intervention regarding changing dietary sodium intake: $$\Delta BP = \frac{\Delta BP_{\mu} \times \frac{\Delta Na}{100}}{SBP}$$ ΔBP_{μ} : the absolute change in SBP (mmHg) by age group for each 100mmol/24h change in dietary sodium intake. ΔNa: the absolute reduction in dietary sodium intake (mmol/24hours) by sex. SBP: systolic blood pressure (mmHg) by age, sex & ethnicity before an intervention.

Objective 1: Considering Agreement Between HT Mortality Estimates and Official Data Sources

Our input data was from HealthTracker (HT) for these reasons:

- The NZ Burden of Disease Study (NZBDS) data did not have incidence rates for CVD
- The BDS did not have case fatality rates for the time periods of relevance to our preferred model design (pre and post 28 days of a CVD event)
- To maximise internal consistency it was considered best to derive background mortality rates, non-CVD mortality rates and CVD mortality rates, all from the same data source (HT)

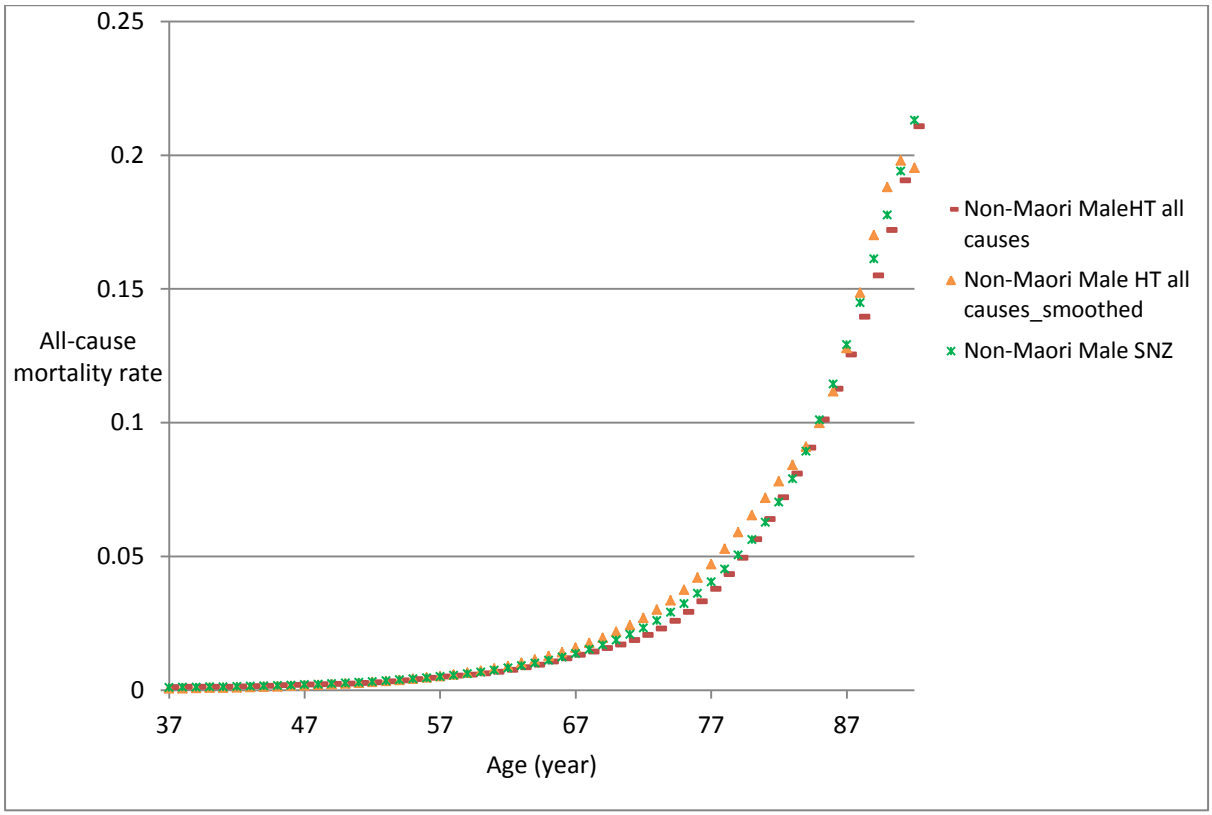
We present results for the non-Māori male population (starting in the 35-39 age group [mid-point 37 years]) – as this is the socio-demographic group with the largest numerators and denominators. We also present results for Māori females as this is the socio-demographic group in the model with the lowest CVD mortality rates.

Comparisons for all-cause mortality rates between Health Tracker and Statistics New Zealand

As per the figures below, there is a fairly good match between both the observed (unsmoothed) and smoothed HT data and SNZ data for all-cause mortality in non-Māori males. The observed data from HT is also generally slightly closer to the SNZ data, than the smoothed data.

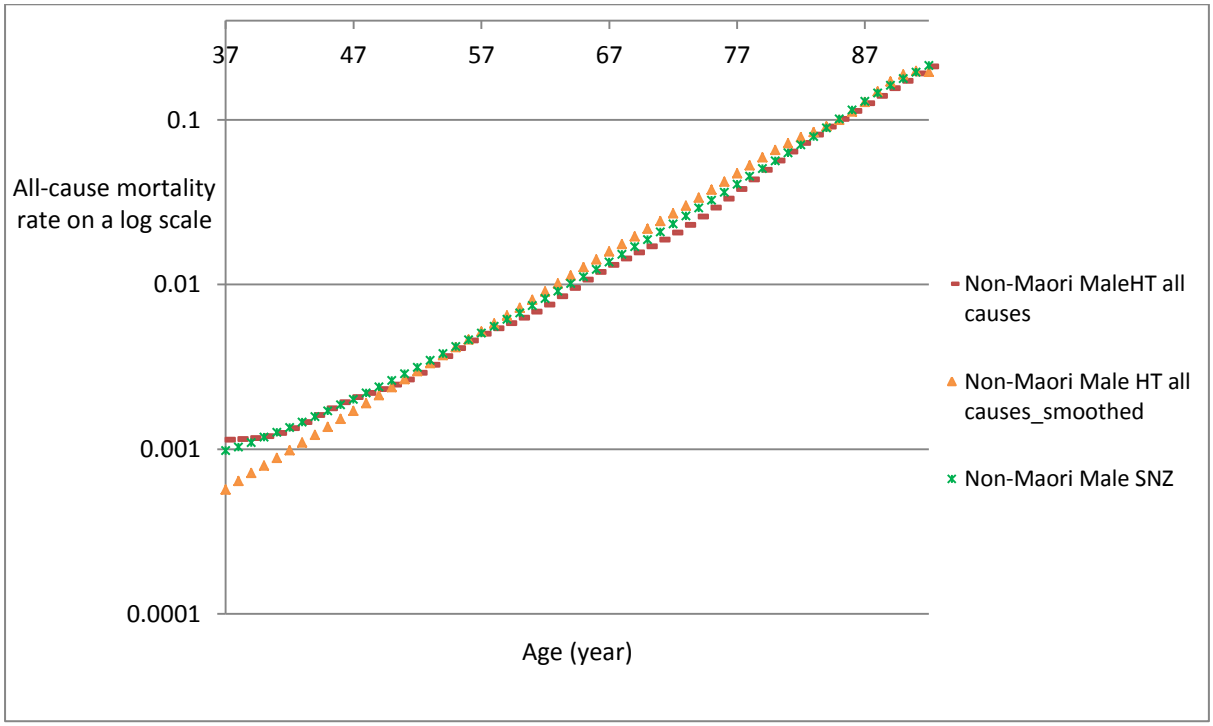
Both the above patterns are also seen for the all-cause mortality results for Māori females. However, there is some extra deviation from the SNZ data for the observed data in the ages covering the late 40s to mid-50s (and for older age groups for the smoothed data). For Maori males the all-cause mortality results for the smoothed data were slightly lower than the SNZ data at the younger ages and then slightly higher in the older ages.

Figure 2 Results showing a comparison between the age-specific all-cause annual mortality rate from the HealthTracker (HT) data (unsmoothed and smoothed) and the Statistics New Zealand (SNZ) estimates (for non-Maori males)



Source: Calibration\CVD Output 22Jan2014_CVD.xls\SNZdata!I15:AC46

Figure 3 Comparing all-cause annual mortality rates between HT and SNZ – log scale (non-Maori males)



Source: Calibration\ CVD Output 22Jan2014_CVD.xls\SNZdata!I15:AC46

Figure 4 Comparison between the age-specific all-cause annual mortality rates from the HealthTracker (HT) data (observed/unsmoothed and smoothed) and the Statistics New Zealand (SNZ) estimates (for Māori females, log scale)

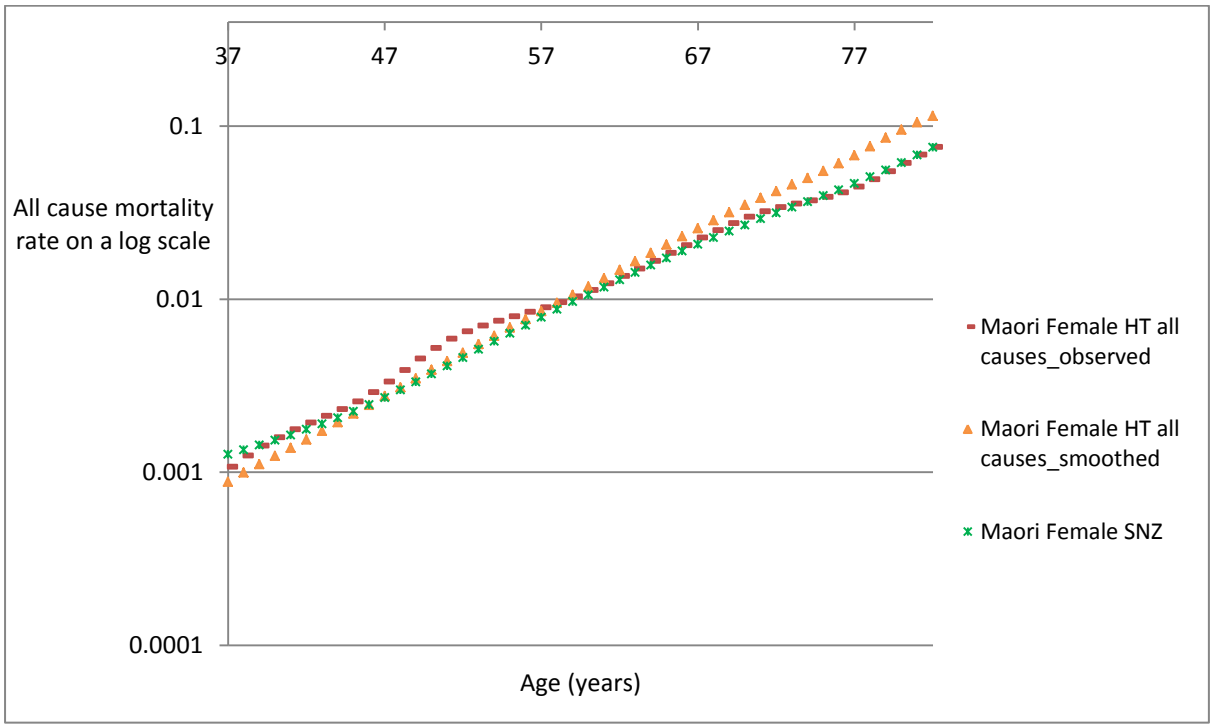
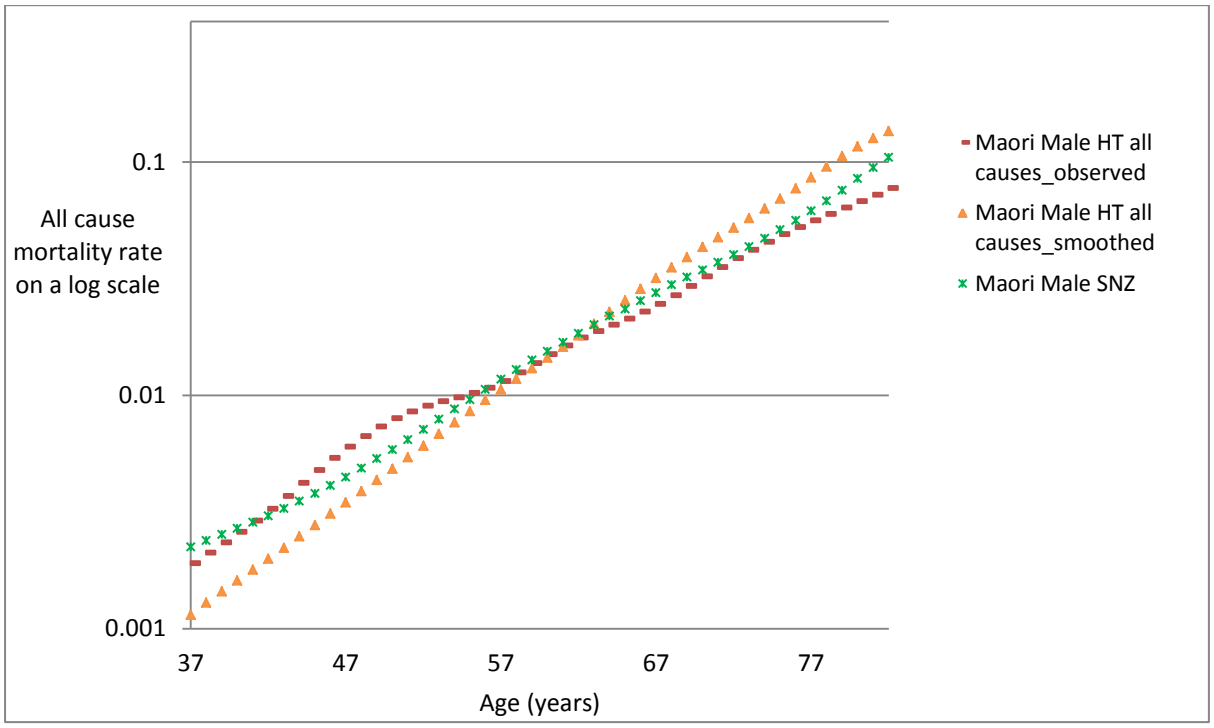


Figure 5 Comparison between the age-specific all-cause annual mortality rates from the HealthTracker (HT) data (observed/unsmoothed and smoothed) and the Statistics New Zealand (SNZ) estimates (for Māori males, log scale)

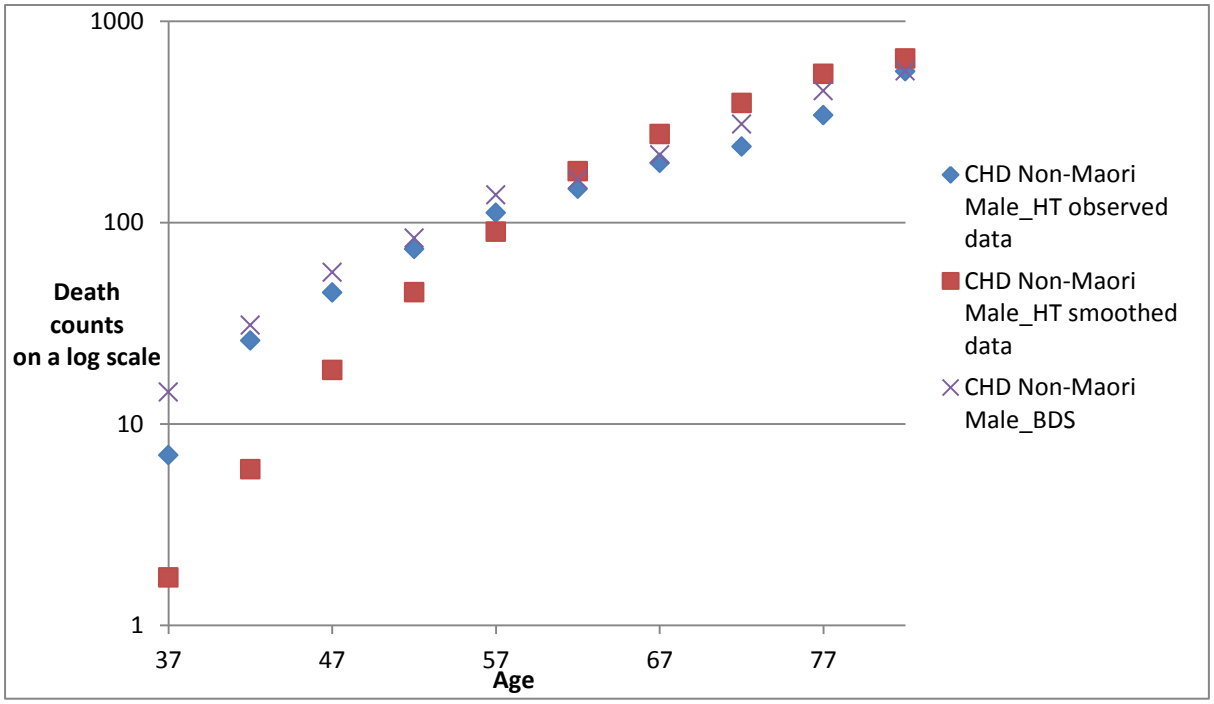


Source: Calibration\CVD Output 23Jan2014_CVD.xls\SNZdata!I15:AC46

The results below show that for CHD in non-Māori males, our estimates from HT are lower than for the NZBDS data in younger ages (< 60 years), but then tend to be slightly higher. These differences may be attributable to some definitional differences in CHD.

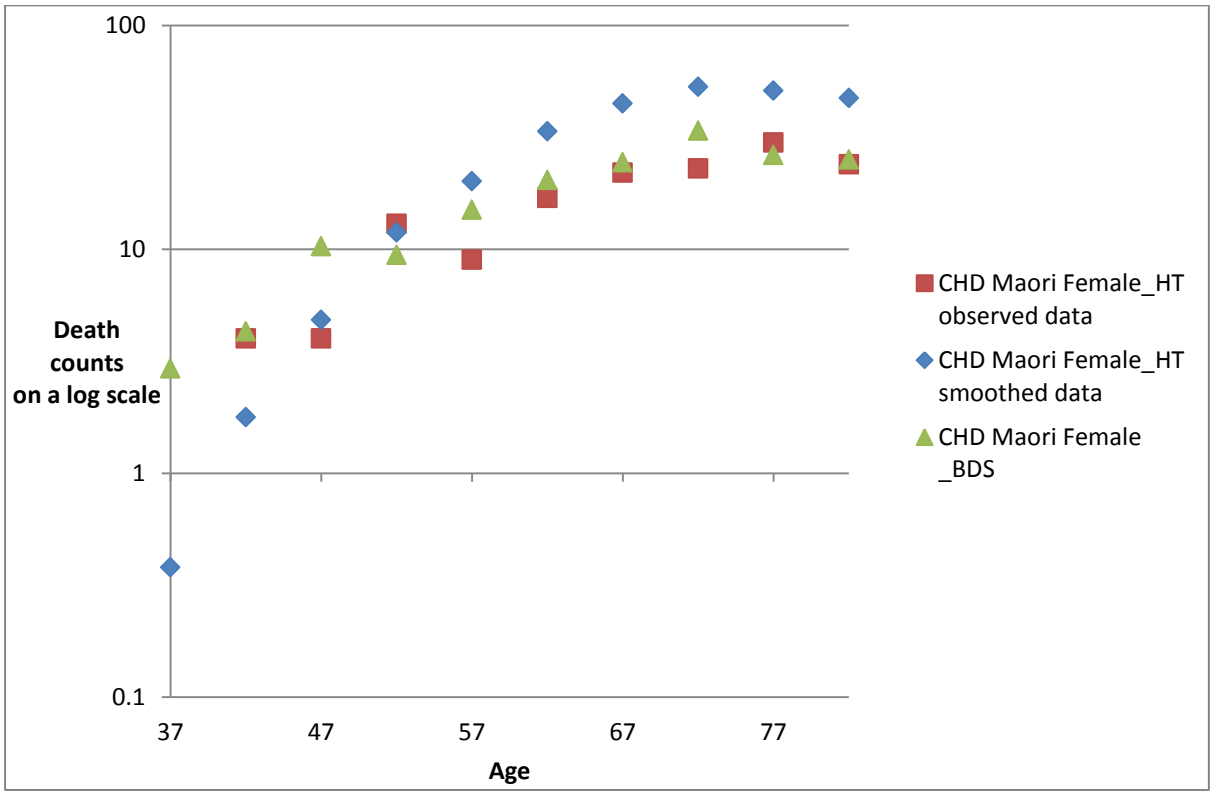
For CHD in Māori females, the same pattern was apparent, however with higher levels on older age groups for the smoothed data. Given the similarities with the observed data, it seems likely that the higher values at older ages are actually a consequence of the smoothing process.

Figure 6 Comparison between the age-specific annual *CHD death counts* from the HealthTracker (HT) data (observed and smoothed) and the NZBDS estimates (for non-Maori males)



Source: Calibration\ Calibration_Inputs_NZ Salt 1st&2nd_7_27Nov2013_TB_17Dec.xls \Epi!L297:P311

Figure 7 Comparison between the age-specific annual *CHD death counts* from the HealthTracker (HT) data (observed and smoothed) and the NZBDS New Zealand estimates (for Maori females)



In terms of stroke deaths in non-Māori males, we note that the HT data produced consistently higher results than the NZBDS for the smoothed data (see below). The HT results were also more stable – especially the smoothed data.

For stroke deaths in Māori females (see below), the above two comments also applied – albeit with much more stable results than those from the NZBDS when considering the smoothed data.

For stroke deaths in Māori males (see below), the data were fairly unstable – but again the smoothed data involved higher values.

Figure 8 Comparison between the age-specific *total stroke death counts* from the HealthTracker (HT) data (observed and smoothed) and the NZBDS estimates (for non-Maori males)

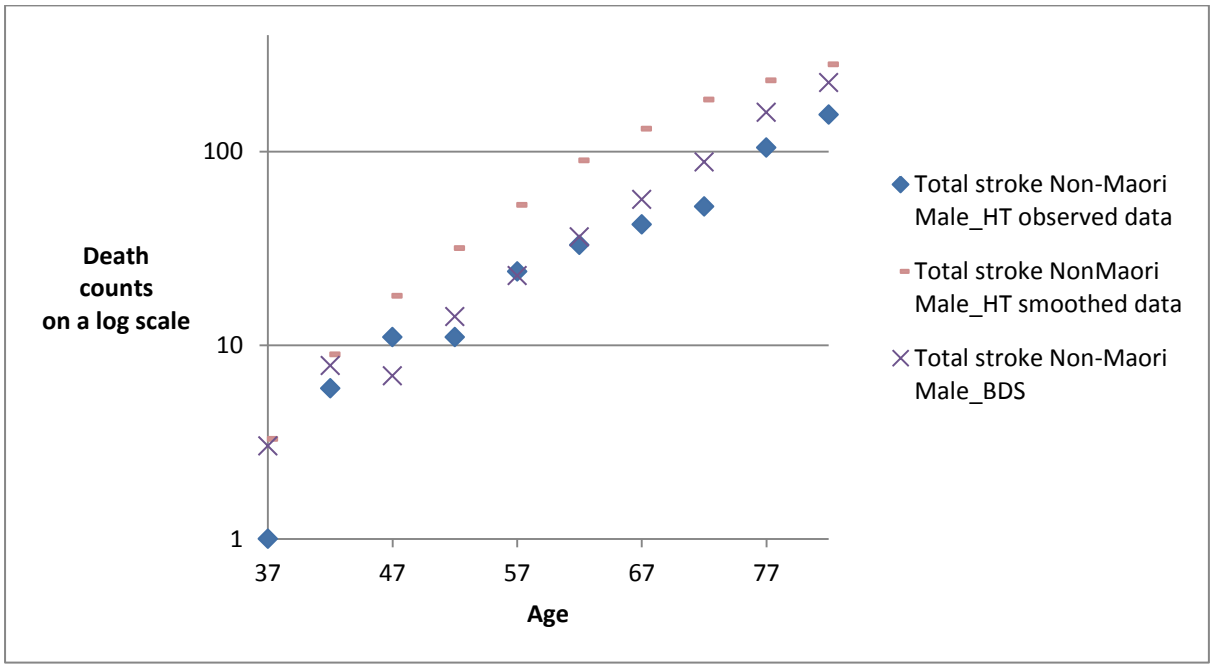


Figure 9 Comparison between the age-specific *total stroke death counts* from the HealthTracker (HT) data (observed and smoothed) and the NZBDS estimates (for non-Maori females)

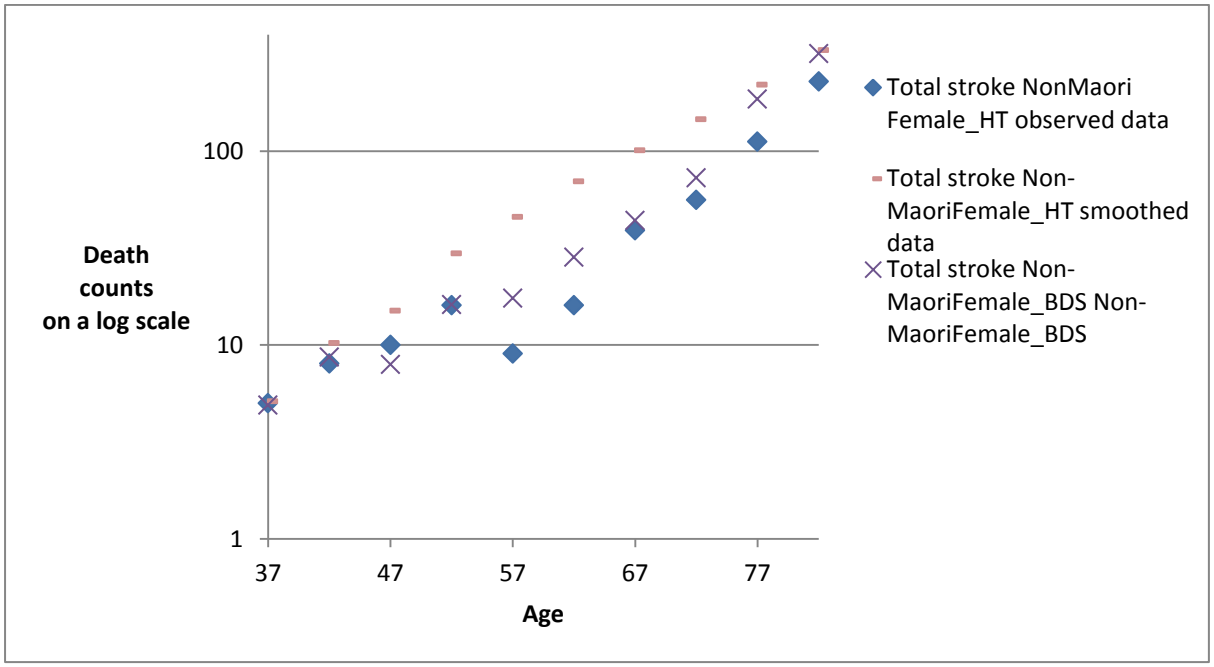


Figure 10 Comparison between the age-specific *total stroke death counts* from the HealthTracker (HT) data (observed and smoothed) and the NZBDS New Zealand estimates (for Maori males)

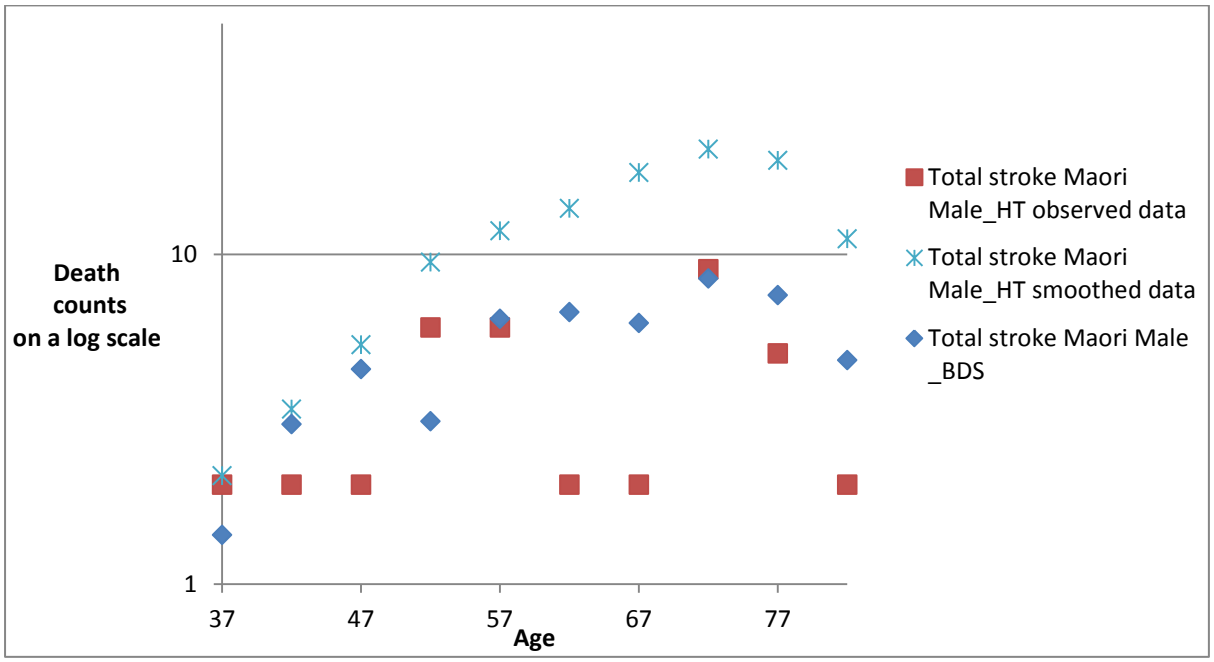
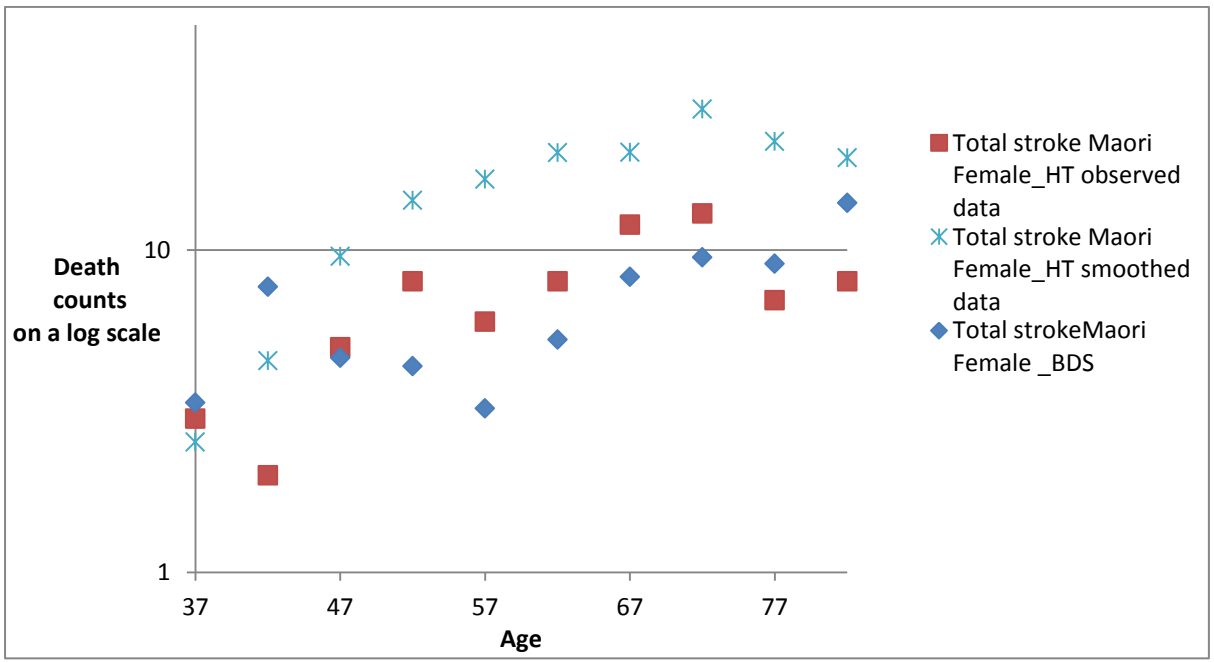


Figure 11 Comparison between the age-specific *total stroke death counts* from the HealthTracker (HT) data (observed and smoothed) and the NZBDS New Zealand estimates (for Maori females)



Objectives 2 & 3: Assessing Epidemiological Coherence of Input Parameters from HT (Using DisModII) and considering the impact of Smoothed Input Data

The epidemiological software tool DisModII was used to check the internal consistency of the key input parameters from Health Tracker. That is, the input 'set' of incidence, prevalence, case fatality and mortality (i.e. with alive population as the denominator) should be epidemiologically consistent.[1] There is a risk that they are not fully consistent, due to the look back time that we could use in HealthTracker not being long enough for 'equilibrium' to be reached in terms of incidence, prevalence and case fatality actually producing the true CVD mortality rates in the population. Put another way, we may be missing some prevalent cases due to inadequate look back.

DisModII has a range of functions we exploit or explore here: smoothing of inputs, smoothing of outputs, and weighting of inputs (e.g. we may trust more the robustness of the mortality input, and hence give that more weight). Note that with smoothing of inputs, we also undertook our own regression smoothing of HeathTracker inputs in SAS before passing them to DisModII (and before passing them to the Markov model in TreeAge).

Below are plots of inputs (stepped lines as by five-year age group) and outputs (smooth curves) from DisModII, for non-Maori males, starting at age group 37 years (4 inputs: incident rate, prevalent rate, integrated case fatality rate (CFR), and the type of CVD specific mortality rate (CHD, ischaemic stroke etc); with weighting on the mortality results (at the 100% setting in DisModII). The first figures for each demographic group are those using the observed (unsmoothed) data from HT. Subsequent results are using smoothed inputs (i.e. incident rate, prevalent rate, case fatality rate pre and post 28 days, disease specific mortality rate, and background mortality rate were smoothed using Poisson regressions with a log-linear functional form in SAS 9.3. All the rates were smoothed by age, sex and ethnicity with main effects only and weighted by person time or death counts where appropriate; see the Main Technical Appendix for details). For all the DisModII work shown below we set remission rates to zero (ie, assuming that the atherosclerotic processes involved in CVD are chronic conditions that do not remit).

Regarding the CFR, for the purposes of examining the results in DisModII we needed to produce an integrated single CFR from combining the values we obtained from HT (representing separate pre-28 day CFR and the post 28-day CFR). This was done by dividing all CVD deaths in 2010 from HT for prevalent cases from HT in 2010. CVD deaths were identified using death certificates. This included people who died in the community, and excluded people who got CVD but died from other causes.

In interpreting the figures below, the thinner and stepped lines are inputs "(i)" to DISMOD, while the thicker and smoothed lines are DISMOD outputs "(o)".

Results from use of DisModII and interpretation

Coronary heart disease (CHD) in non-Maori males

The two figures directly below show DisModII inputs and outputs for CHD in non-Maori males. There are not major differences between the use of observed and smoothed input data. In both cases the prevalence outputs are increased (as we expected given the limited HT look back period). However, the shape of the smoothed output prevalence data is more plausible (than the observed data inputs) in that it does not have a major downturn on the older age groups (ages 85+ years). For the smoothed data, the increase in the CFR with age is less steep (than for the observed data), which might be more plausible.

Figure 12 CHD DisModII inputs and outputs for non-Maori males using Health Tracker observed data (for this and subsequent graphs, all values should be read off from the Y1-axis; except for the CFR and the prevalence on the Y2-axis)

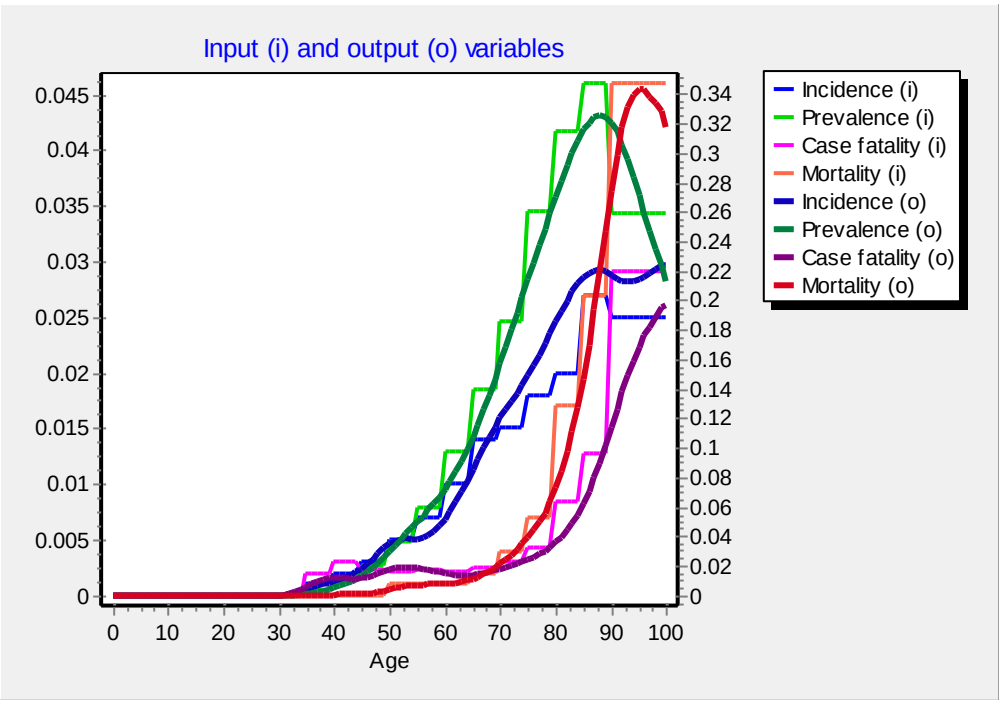
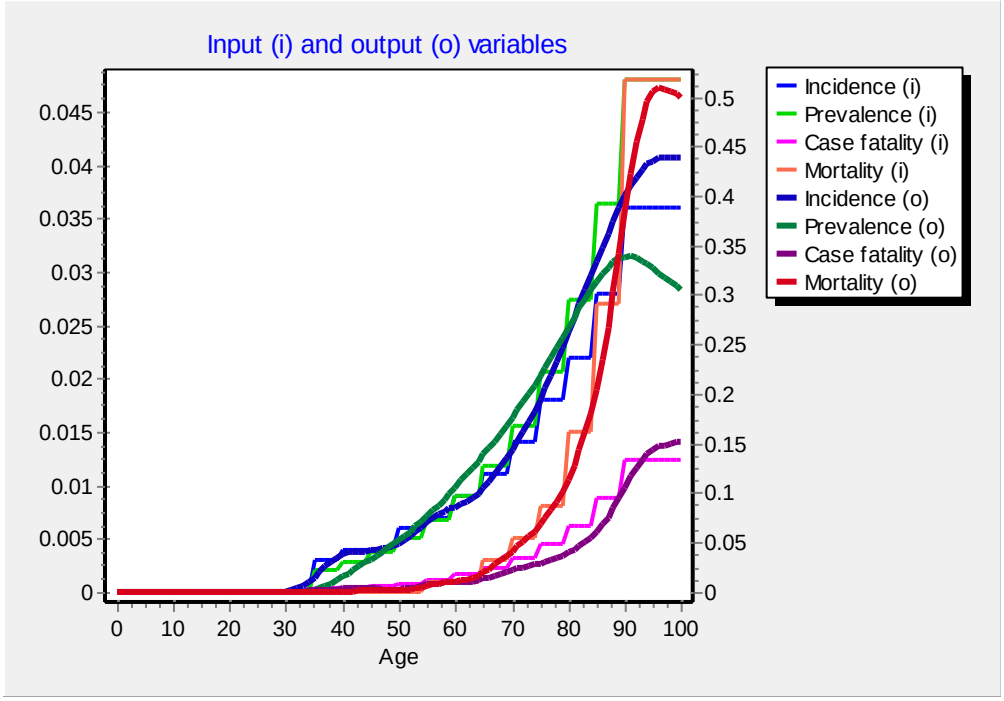


Figure 13 CHD DisModII outputs for non-Maori males, as per the previous figure except for the use of smoothed input data



Ischaemic stroke

The two figures directly below show DisModII inputs and outputs for ischaemic stroke (IS) in non-Maori males. There are not major differences between the use of observed and smoothed input data. In both cases the prevalence outputs are increased (as expected), but more so for the smoothed data. The shape of the outputs generally seem more plausible for the smoothed input data, eg, see the unusual shape for the incidence with the observed input data (which declines in the 80s and then goes up again in the 90s).

Figure 14 Ischaemic stroke related DisModII inputs and outputs for non-Maori males using Health Tracker observed data

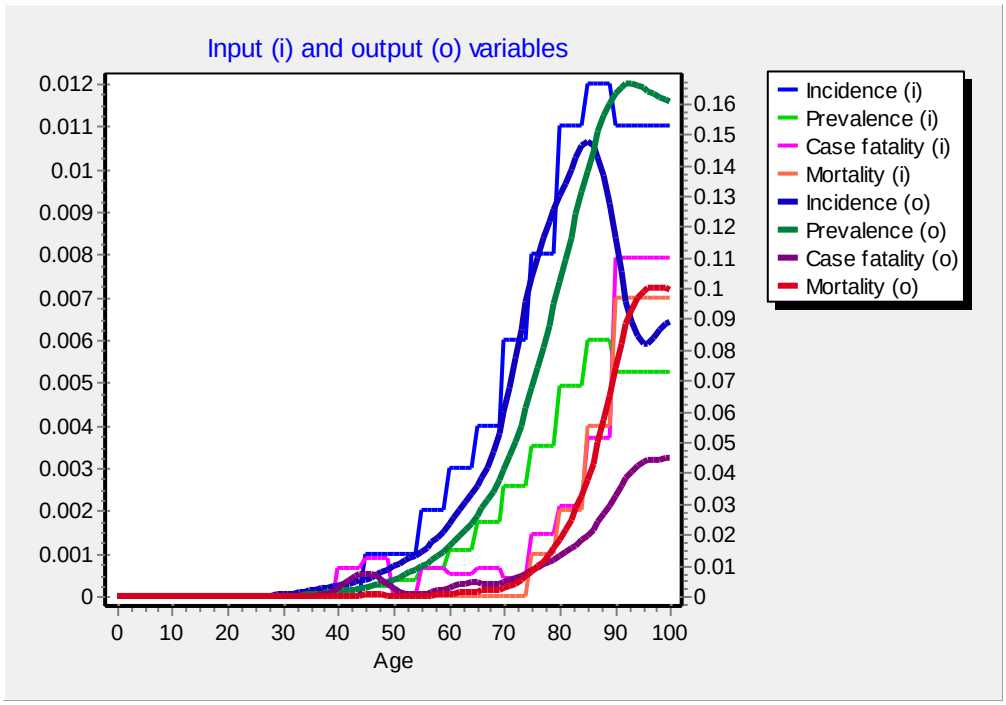
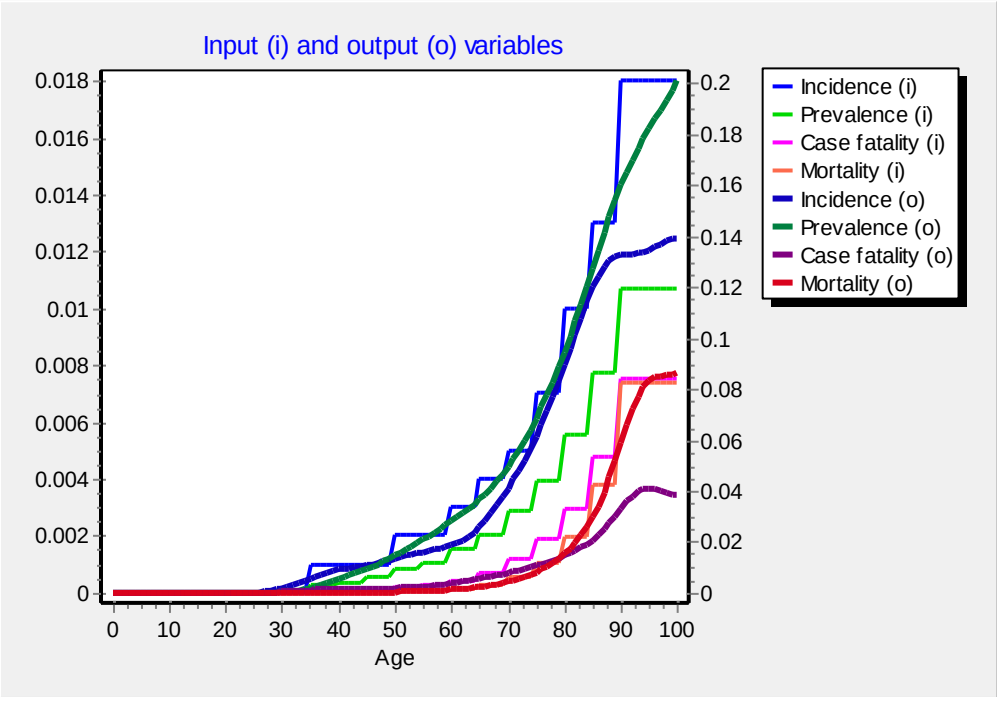


Figure 15 Ischaemic stroke DisModII inputs and outputs for non-Maori males, as per the previous figure but using smoothed input data from HT



Haemorrhagic stroke

The two figures directly below show DisModII inputs and outputs for haemorrhagic stroke (HS) in non-Maori males. In both cases the prevalence outputs are increased (as expected), but the results from the smoothed data do not drop down as much in the 85+ ages (i.e., the smoothed data appears to be more epidemiologically plausible). Other results from using the smoothed input data also look more plausible (than the observed data), such as the lower maximum value for the CFR.

Figure 16 Haemorrhagic stroke DisModII outputs for non-Maori males, using observed data from HT

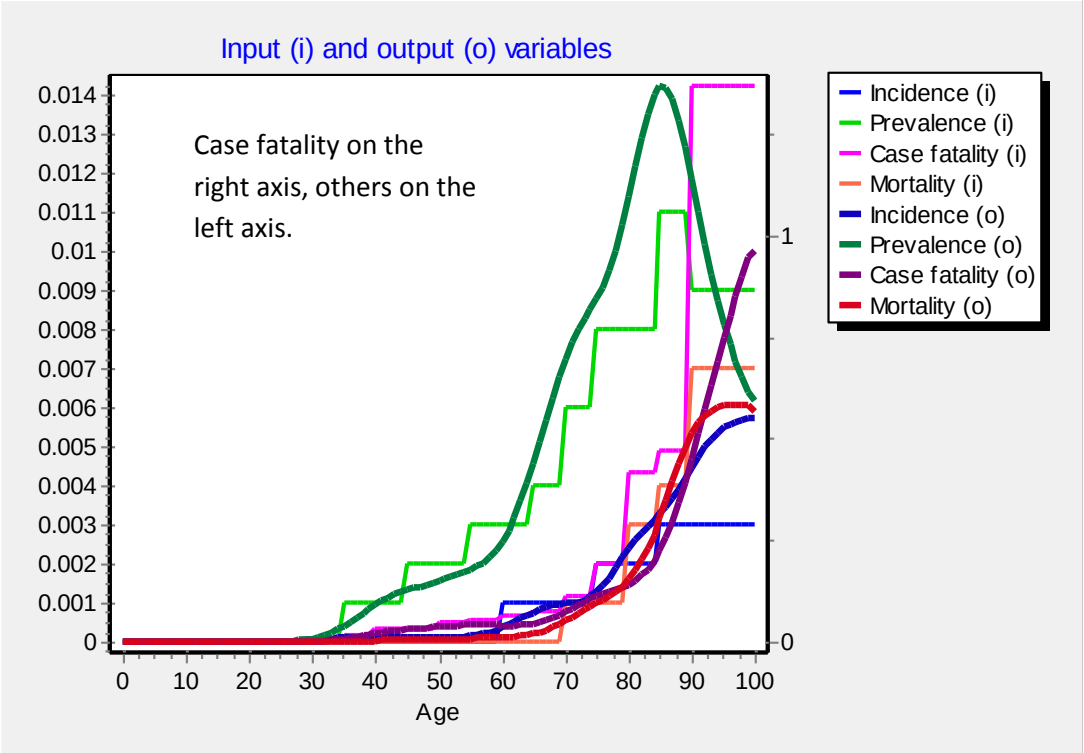
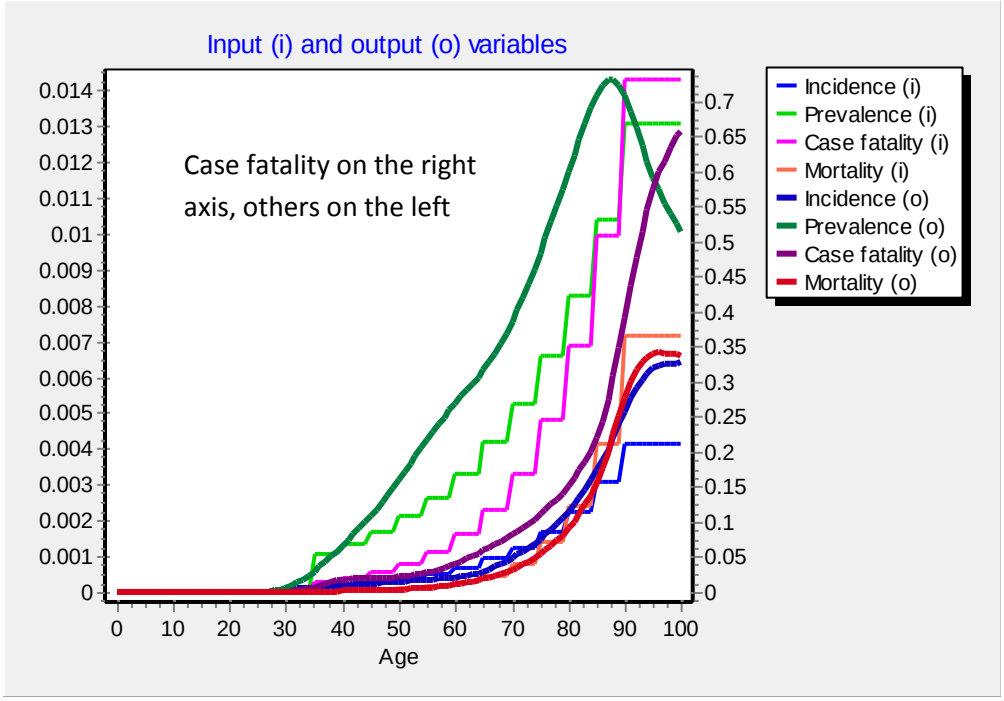


Figure 17 Haemorrhagic stroke DisModII outputs for non-Maori males, using Health Tracker smoothed data



Coronary heart disease (CHD) in Maori females

The two figures directly below show DisModII inputs and outputs for CHD in Maori females. There are not major differences between the use of observed and smoothed input data. Nevertheless, for mortality the match between input and output data is closer for the smoothed input data (and there is less of a steep decline for older ages). Similarly, for the smoothed data, various parameters show more epidemiologically plausible patterns (e.g., incidence at older ages and prevalence at older ages). For the smoothed data, the prevalence outputs are increased compared to little change for the observed data (with the former being more of what was expected by us given the limited HT look back period).

Figure 18 CHD DisModII outputs for Maori females, using observed data from HT

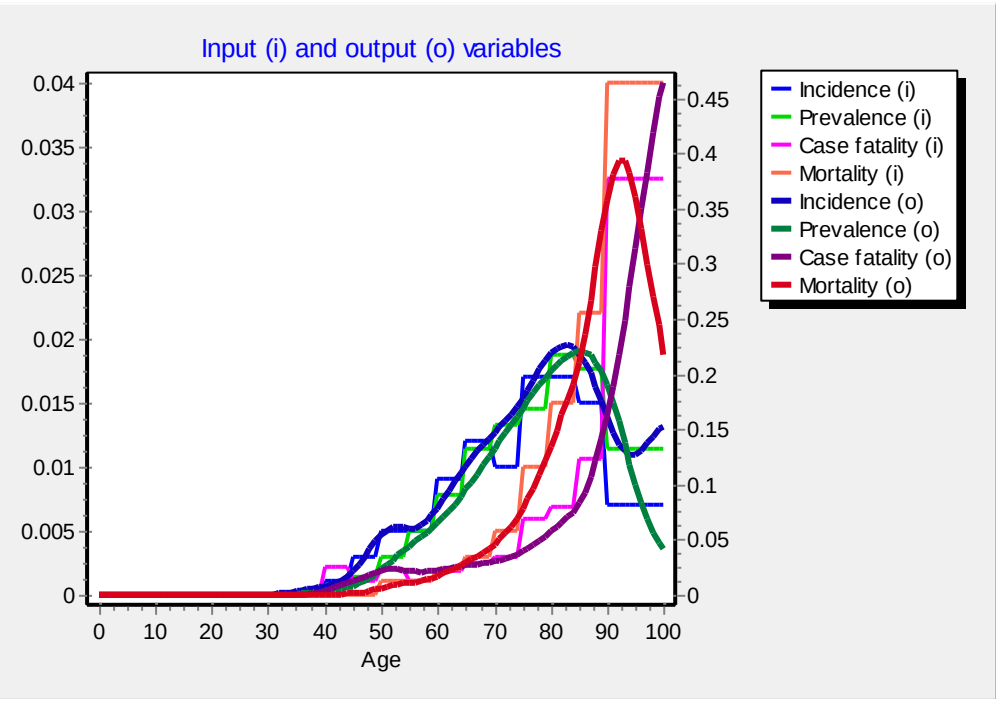
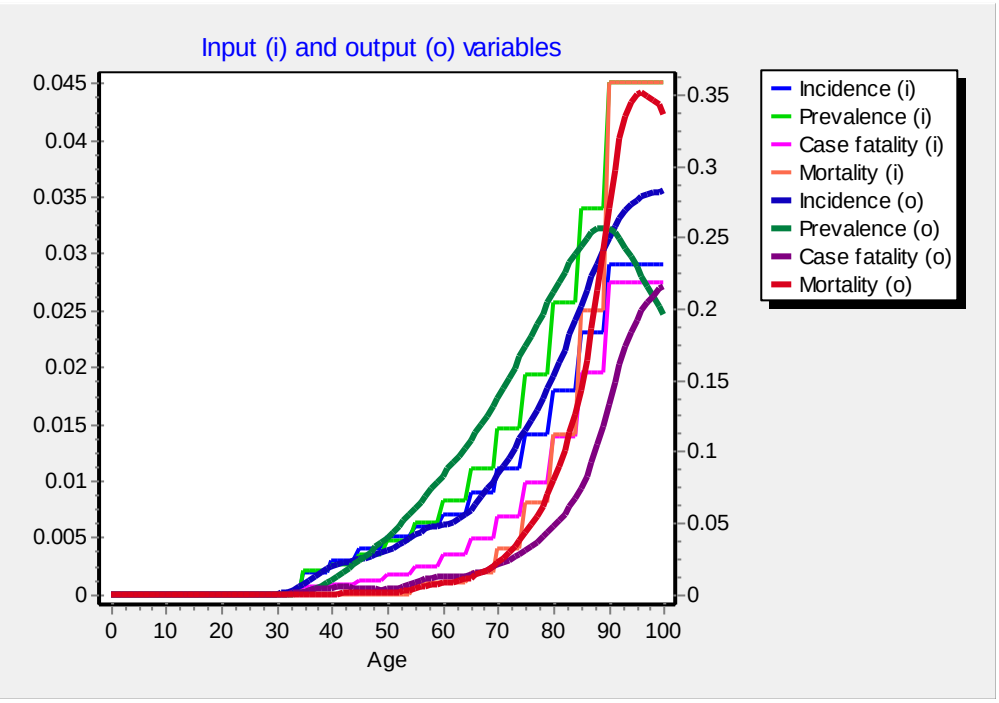


Figure 19 CHD DisModII outputs for Maori females, using smoothed data from HT



Ischaemic stroke in Maori females

The two figures directly below show DisModII inputs and outputs for ischaemic stroke (IS) in Maori females. The outputs are generally more stable and epidemiologically plausible for the smoothed data inputs (compared to the observed data inputs). See for example the unstable incidence output in the observed data and to a lesser extent the CFR data. The mortality data output is also more plausible for the smoothed input data (i.e., not plateauing at older ages), and there is a closer match between the input and output data (for the smoothed data).

Figure 20 Ischaemic stroke DisModII outputs for Maori females, using observed data from HT

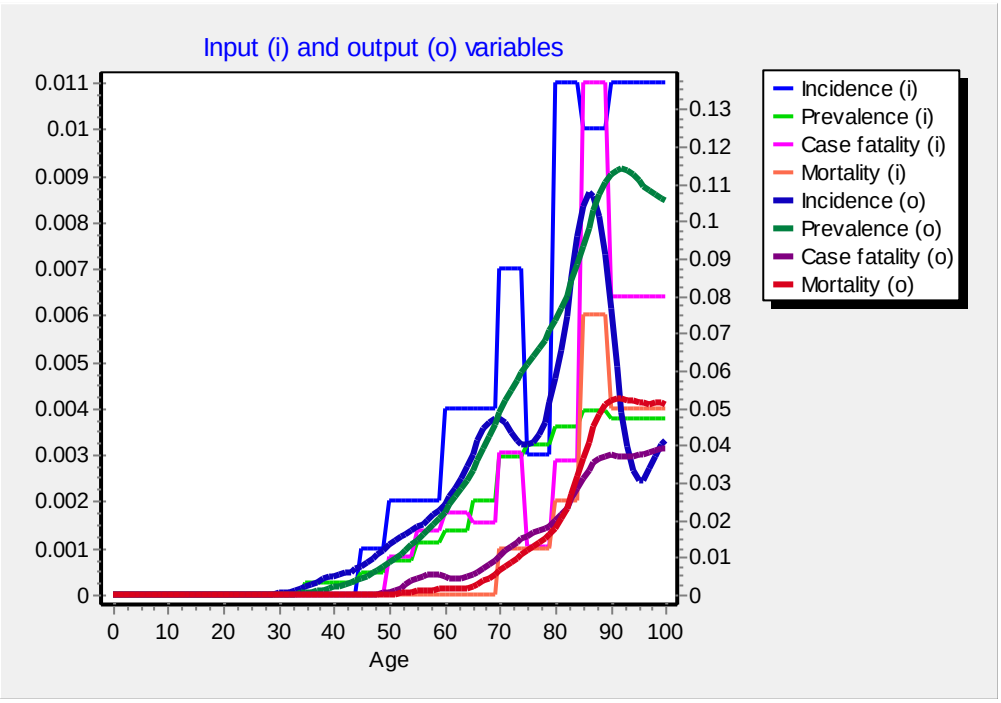
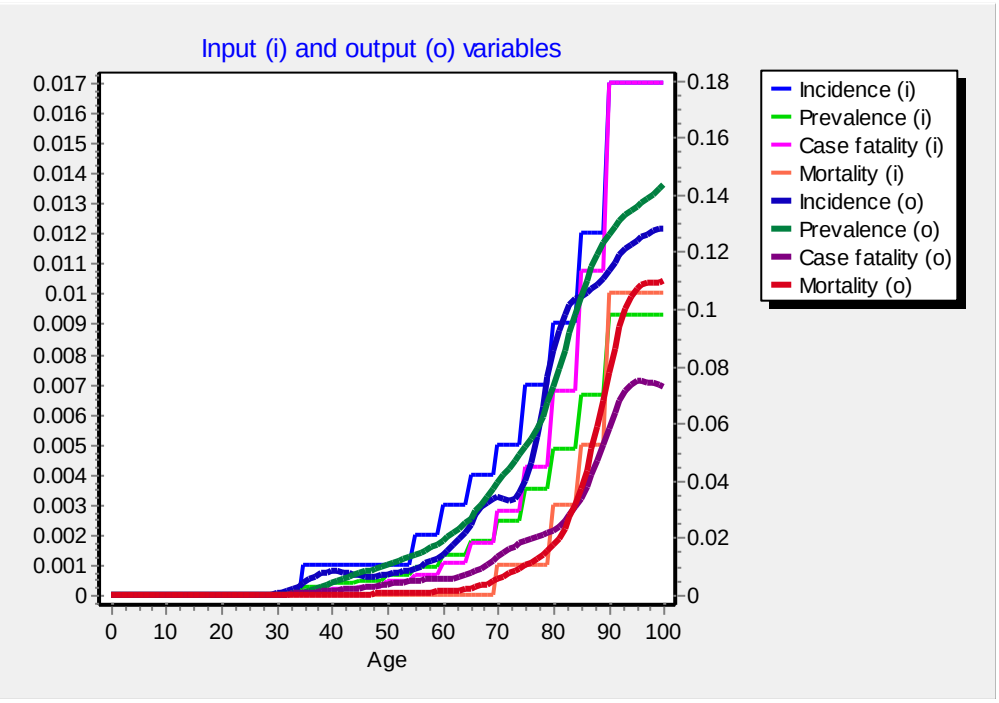


Figure 21 Ischaemic stroke DisModII outputs for Maori females, using smoothed data from HT



Haemorrhagic stroke in Maori females

The two figures directly below show DisModII inputs and outputs for haemorrhagic stroke (HS) in Maori females. The outputs appear to be far more stable and epidemiologically plausible for the smoothed data inputs (compared to the observed data inputs). Indeed, some of the observed data inputs result in non-credible results eg, the decline in mortality after age 85 years and the marked decline in CFR at older age groups.

Figure 22 Haemorrhagic stroke DisModII outputs for Maori females, using observed data from HT

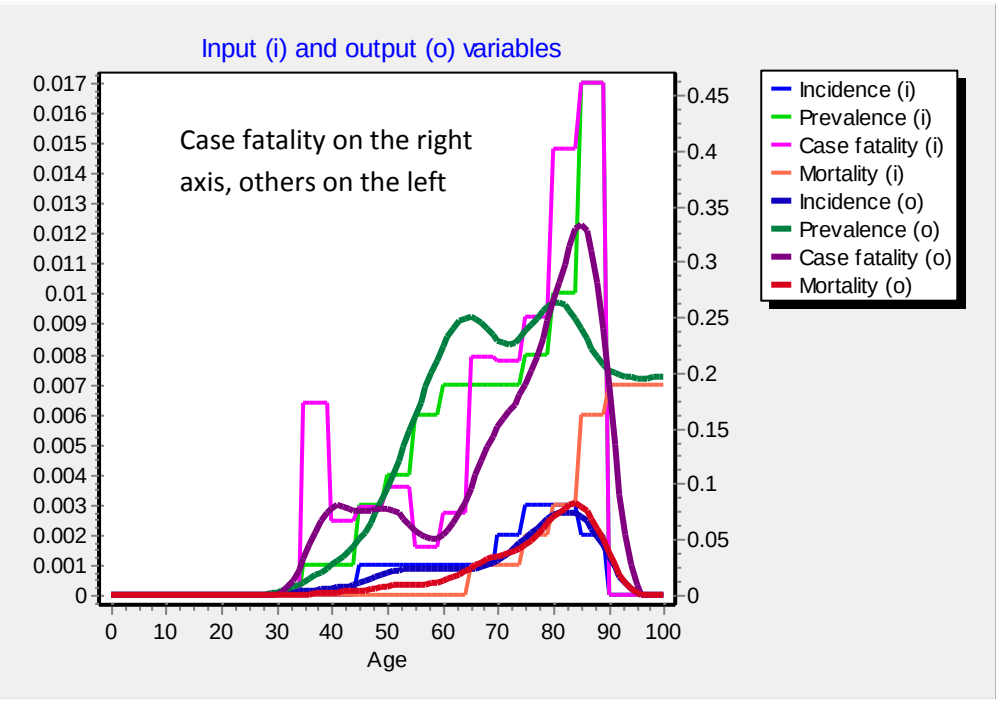
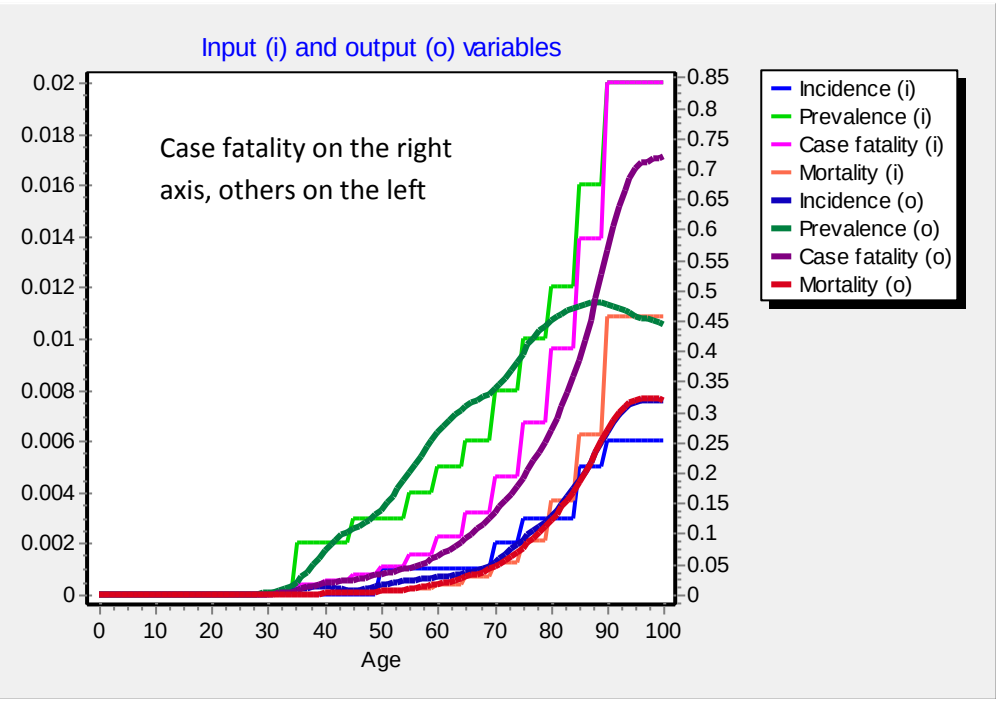


Figure 23 Haemorrhagic stroke DisModII outputs for Maori females using smoothed data from HT



Objective 4: To Validate the CVD Model Outputs Against Observed Mortality Rates

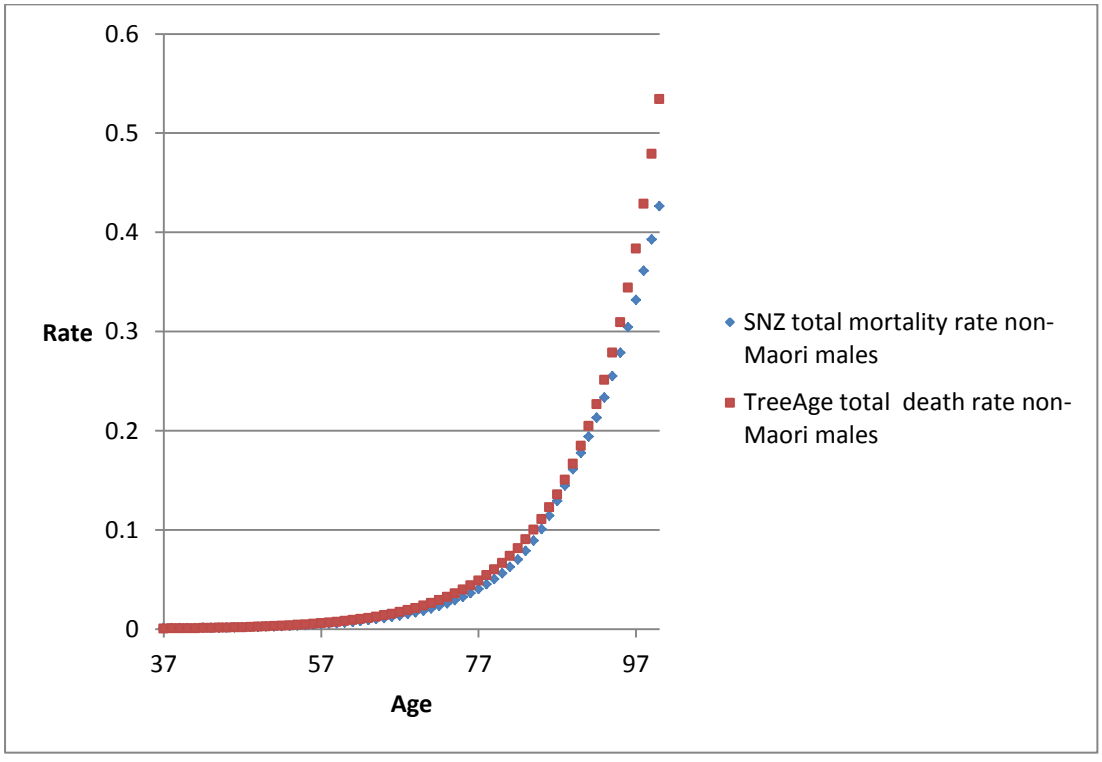
In this Section we detail the CVD Model outputs from the TreeAge based model, albeit with no consideration of trends and for the baseline (no interventions). We compare these outputs with Statistics New Zealand data for all-cause mortality (for the year 2011).

All-cause mortality

In Figure 24 below is the output over annual cycles of the TreeAge model of all-cause mortality rates, meaning that 'age 57' equates to the 20th annual cycle for this 37 year old cohort. The denominator is simply the remaining people in the cohort within the Markov model. The numerator is the number of deaths in the last cycles.

For a calibration scenario of no annual percentage change in future mortality rates (but selecting the older age-group mortality rates as the 37 year old cohort ages), and no change in CHD and stroke incidence rates, we should expect that the TreeAge model output for all-cause mortality rates to closely agree with the SNZ results. Indeed, there seems to be reasonably good agreement (albeit with some divergence at very old ages ie, at 90+ years). Similarly, there is reasonably good agreement when the different levels of future reductions are factored in.

Figure 24 All-cause mortality rate comparison between TreeAge output and SNZ, for non-Maori males starting at age 37, without time trends in changing mortality rates (note: mortality reduction trend that we used in the CVD model in TreeAge is 1.75% & 2.25% per annum for non-Maori and Maori, respectively; reduction trend in SNZ is 2% per annum)



Source: Calibration\ Objective 4-CHD&stroke death rate_18Feb2014\nM37!

Figure 25 All-cause mortality rate comparison between TreeAge output and SNZ, for non-Maori males starting at age 37, without time trends in changing mortality rates (note: mortality reduction trend that we used in the CVD model in TreeAge is 1.75% & 2.25% per annum for non-Maori and Maori, respectively; reduction trend in SNZ is 2% per annum), on a log scale

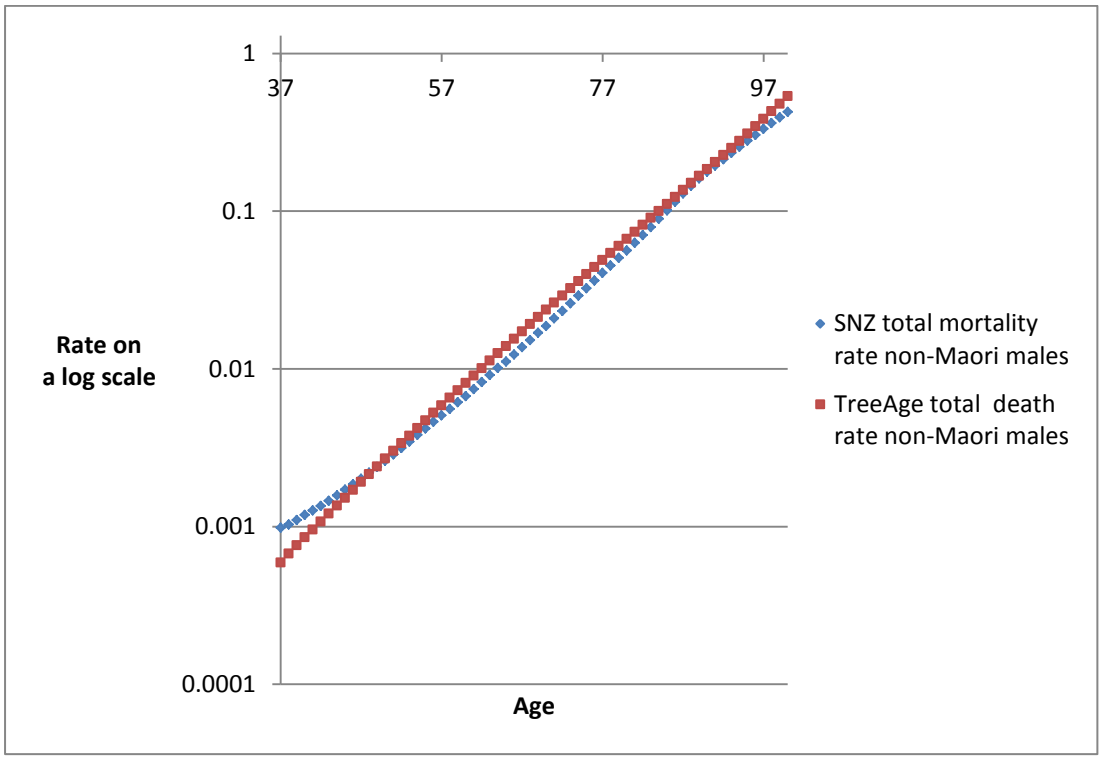
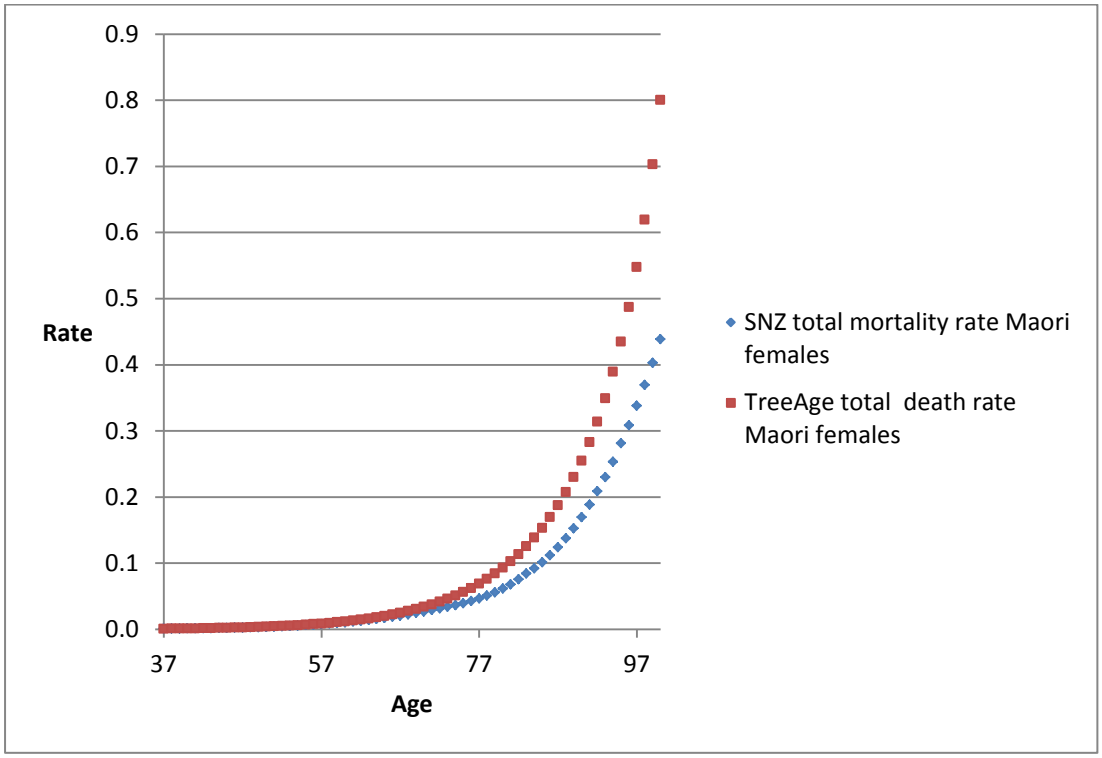
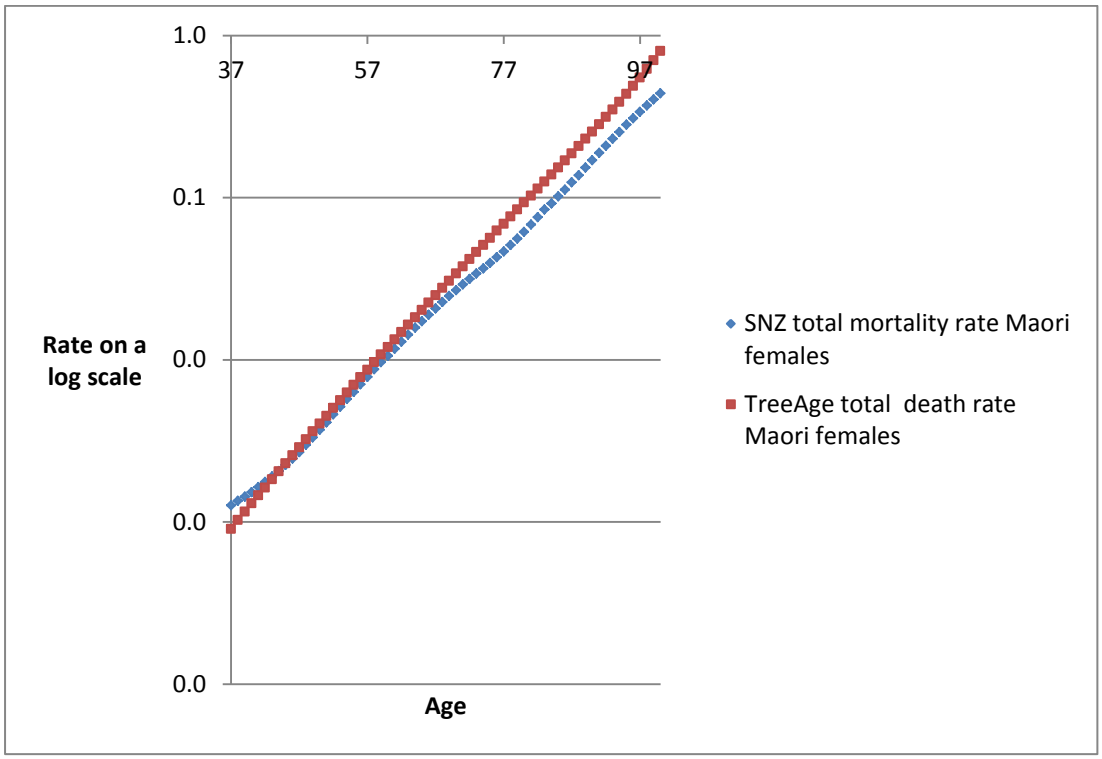


Figure 26 All-cause mortality rate comparison between TreeAge output and SNZ, for Maori females starting at age 37, without time trends in changing mortality rates (note: mortality reduction trend that we used in the CVD model in TreeAge is 1.75% & 2.25% per annum for non-Maori and Maori, respectively; reduction trend in SNZ is 2% per annum)



Source: Calibration\ Objective 4-CHD&stroke death rate_18Feb2014\Mf37!

Figure 27 All-cause mortality rate comparison between TreeAge output and SNZ, for Maori females starting at age 37, without time trends in changing mortality rates (note: mortality reduction trend that we used in the CVD model in TreeAge is 1.75% & 2.25% per annum for non-Maori and Maori, respectively; reduction trend in SNZ is 2% per annum), on a log scale



CVD mortality

In the graphs below (Figure 28 onwards) is the output over annual cycles of the TreeAge model of CVD specific mortality rates, meaning that 'age 57' equates to the 20th annual cycle for this 37 year old cohort. The denominator is simply the remaining people in the cohort within the Markov model. The numerator is the number of all CVD deaths in the last cycles.

Note that the CVD deaths from the BDS were reallocated for age 85+ as follows: a quarter of CVD deaths from aged 85-90, the rest was for 90+. And for the denominator population it was assumed that a third of the deaths were from aged 85-90, with the rest being in the 90+ group.

For a calibration scenario of no annual percentage change in future mortality rates (but selecting the older age-group mortality rates as the 37 year old cohort ages), and no change in CHD and stroke incidence rates, we should expect that the TreeAge model output for CVD mortality rates to closely agree with the BDS results. There appears to be a fairly close match in CVD mortality rates between the BDS results and the CVD Model output in TreeAge (albeit with minor differences in the late 60s and in the 70s).

Figure 28 CVD mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for non-Maori males

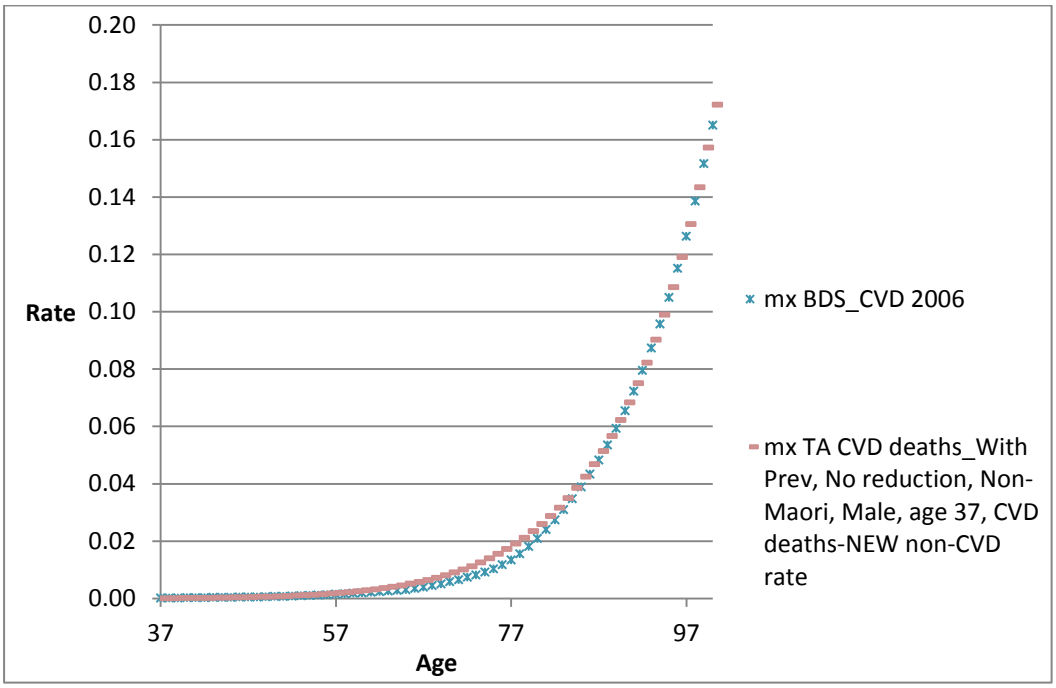


Figure 29 CVD mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for non-Maori males, on a log scale

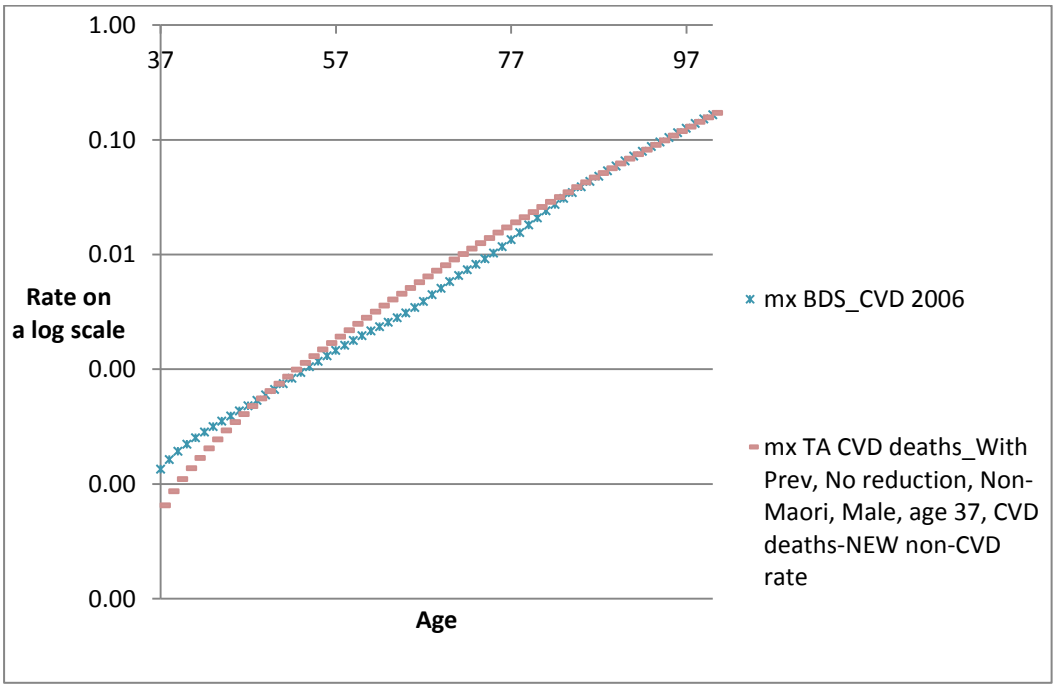
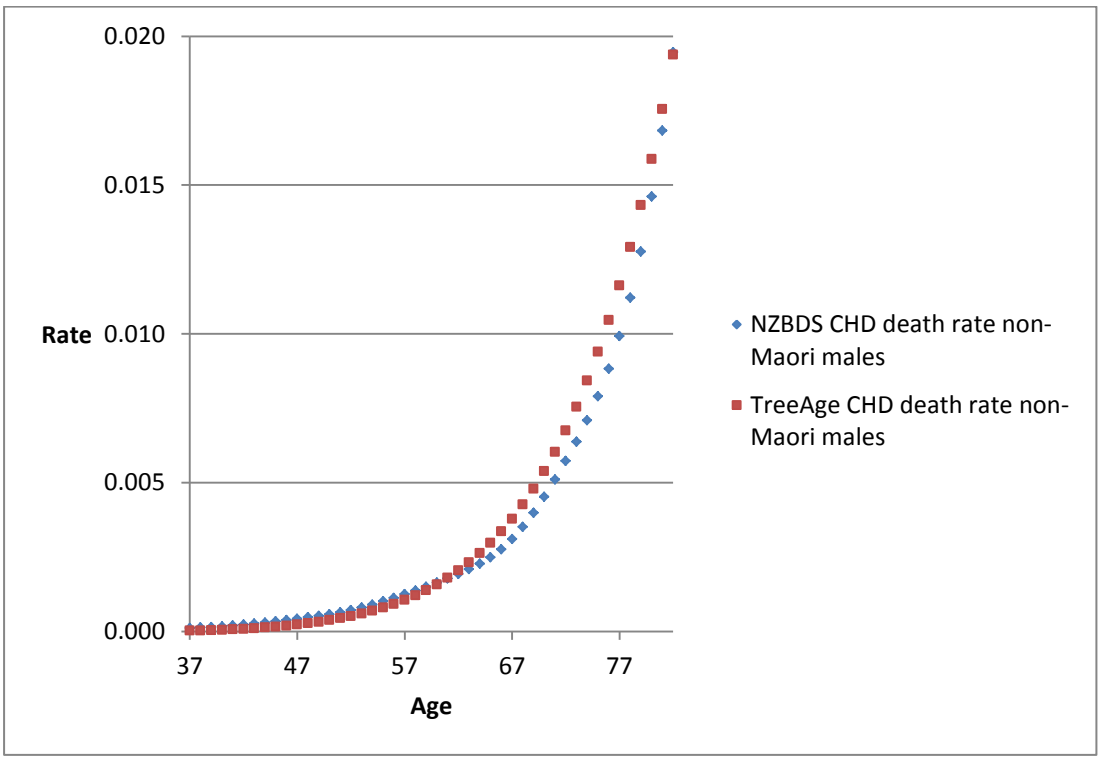


Figure 30 CHD mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for non-Maori males



Source: Calibration\ Objective 4-CHD&stroke death rate_18Feb2014\nM37!

The figures below suggest a fairly close level of matching between the NZBDS data and TreeAge outputs for CHD death rates, but less so for total stroke death rates. Also of note is that there is some gap between stroke death rates starting at age 37 and 67 (Figure 35), implying that the CVD Model in TreeAge is slightly overestimating the total stroke prevalence in the model.

Figure 31 CHD mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for non-Maori males, on a log scale

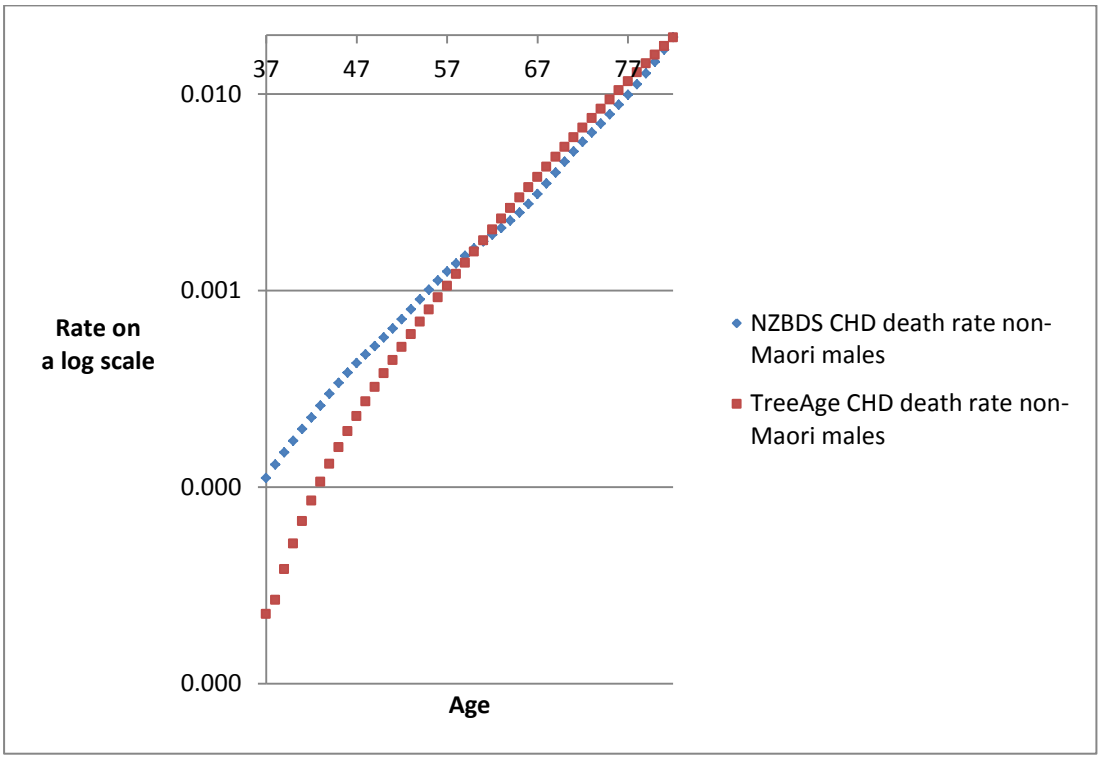


Figure 32 CHD mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 and 67 years for non-Maori males, on a log scale

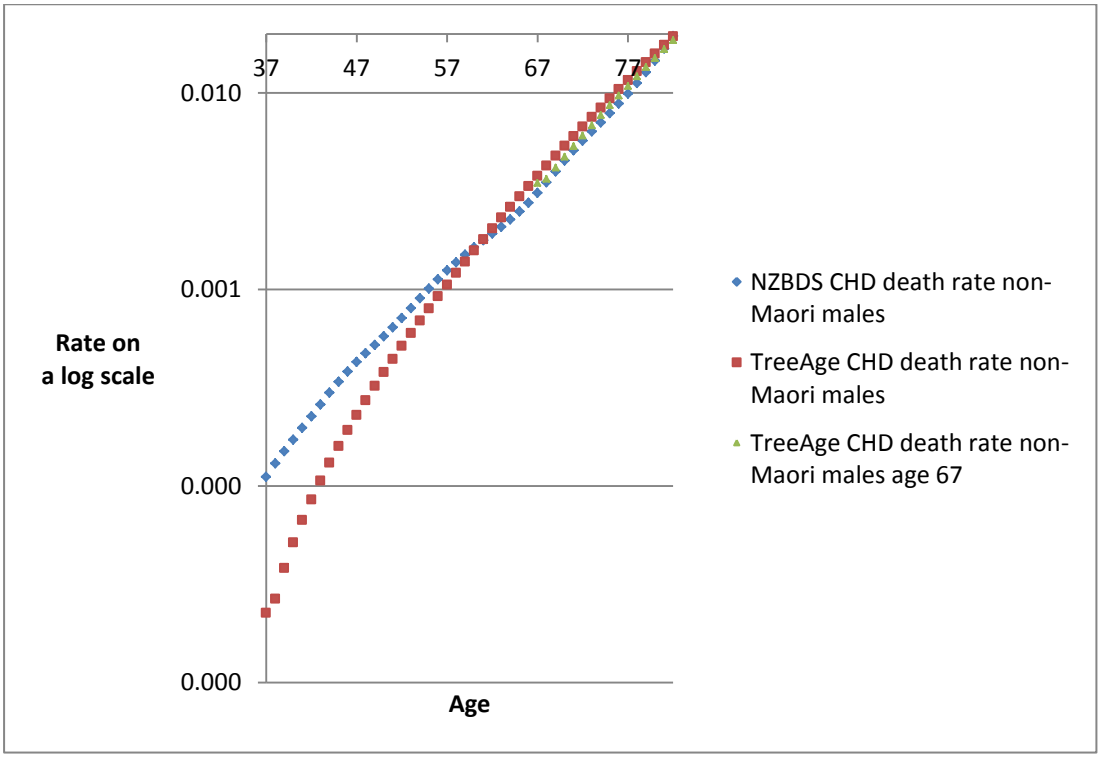


Figure 33 Total stroke mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for non-Maori males

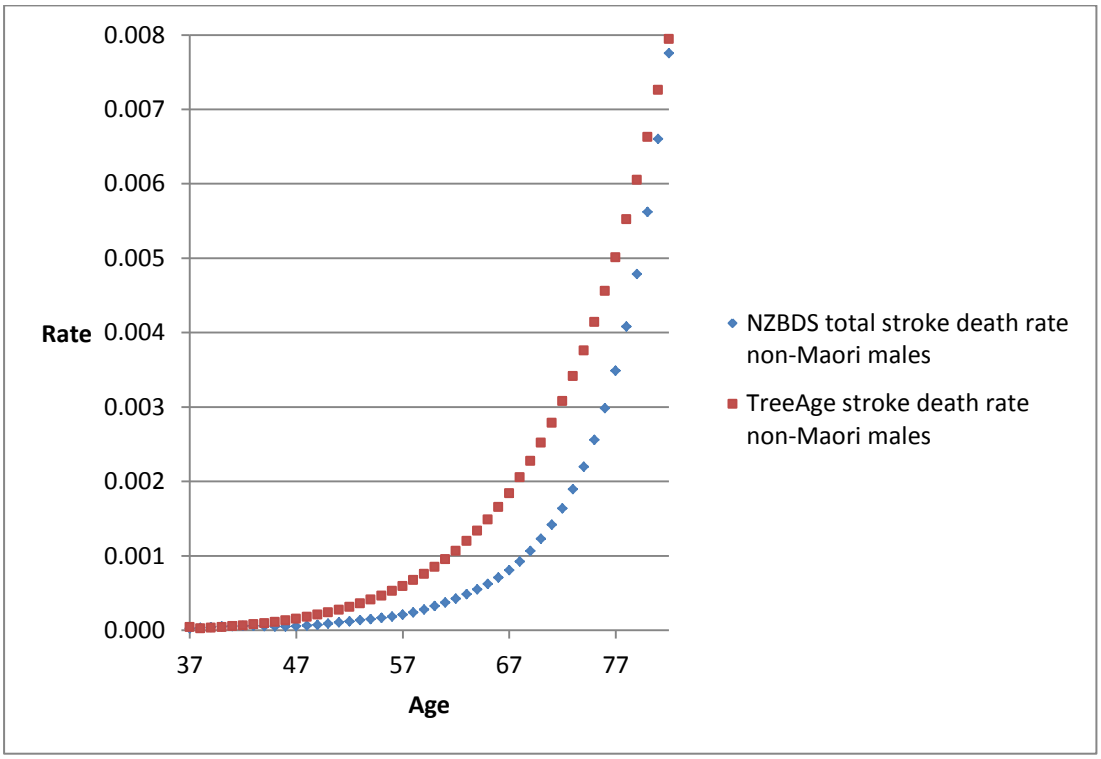


Figure 34 Total stroke mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for non-Maori males, on a log scale

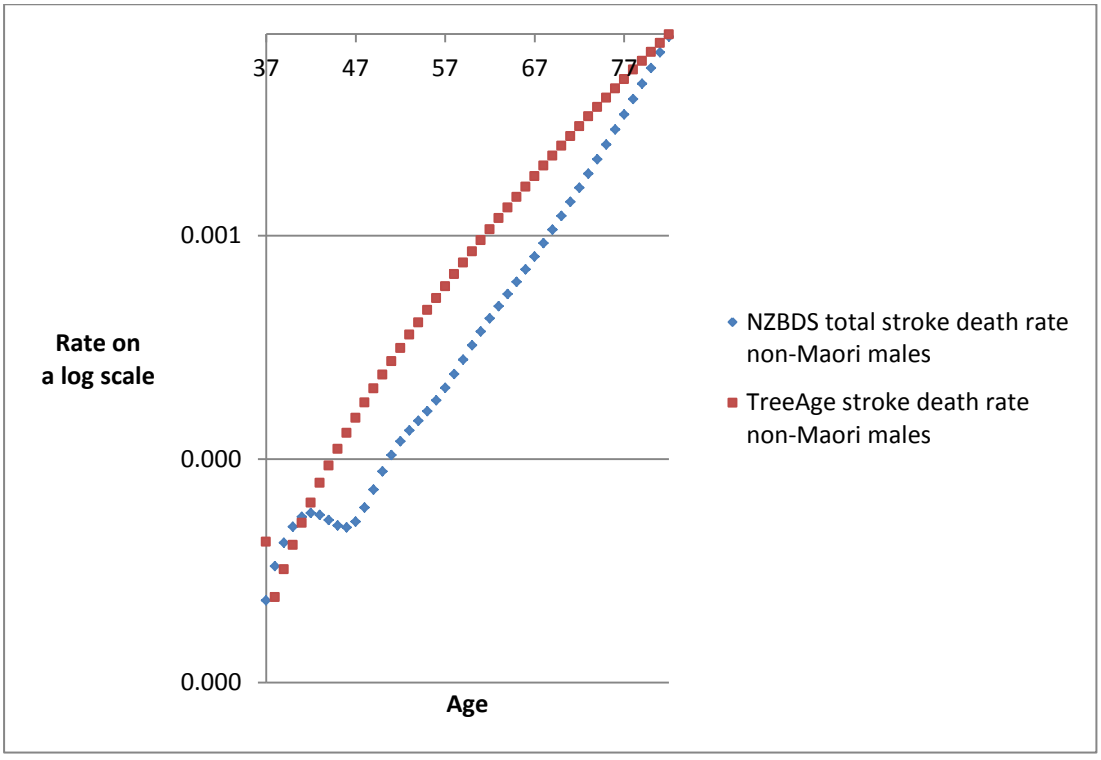


Figure 35 Total stroke mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 and 67 years for non-Maori males, on a log scale

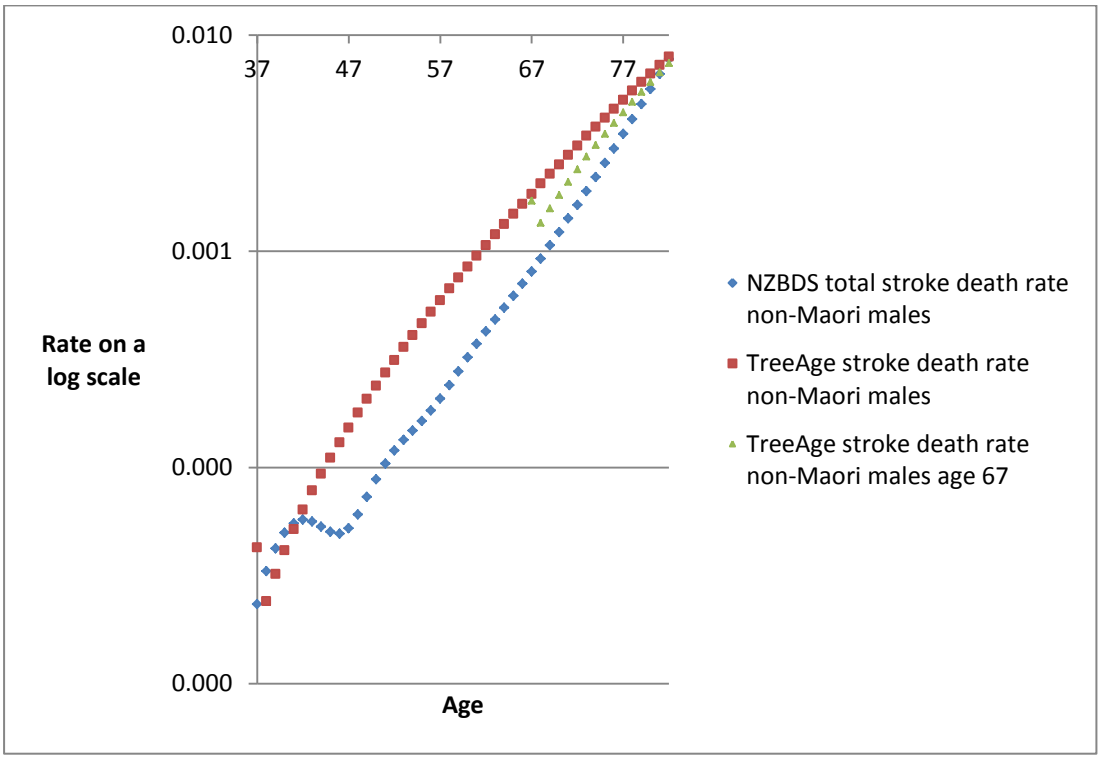


Figure 36 CHD mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for Maori females

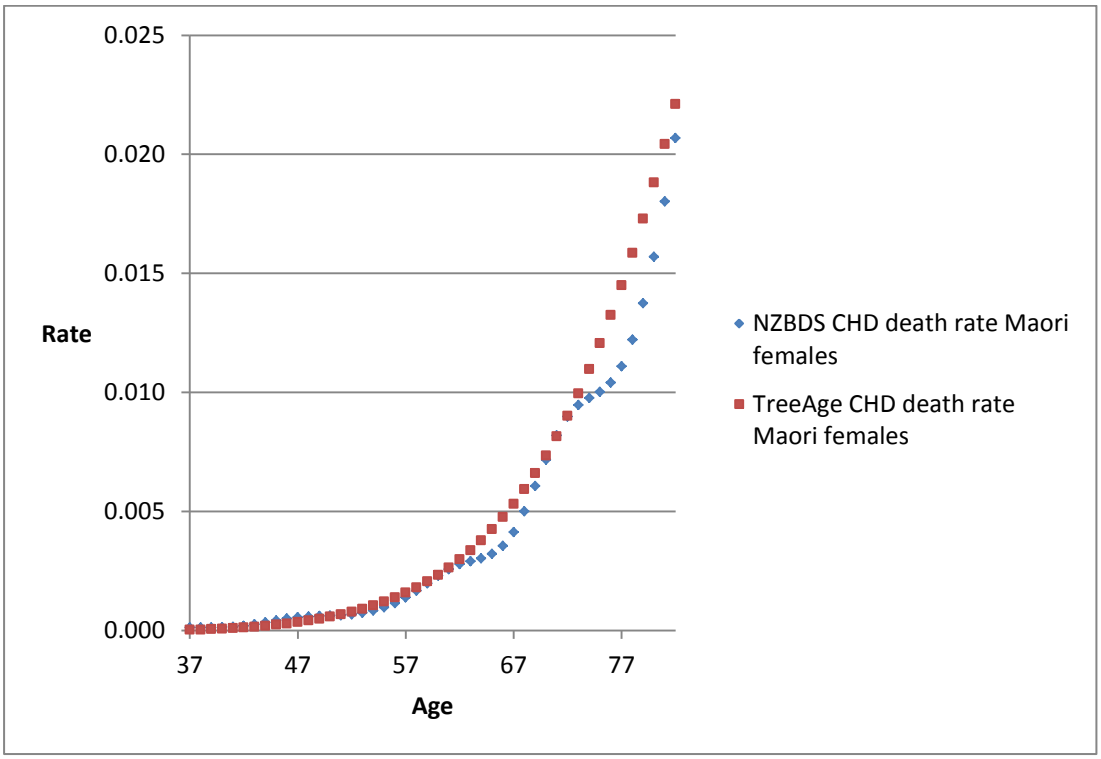


Figure 37 CHD mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for Maori females, on a log scale

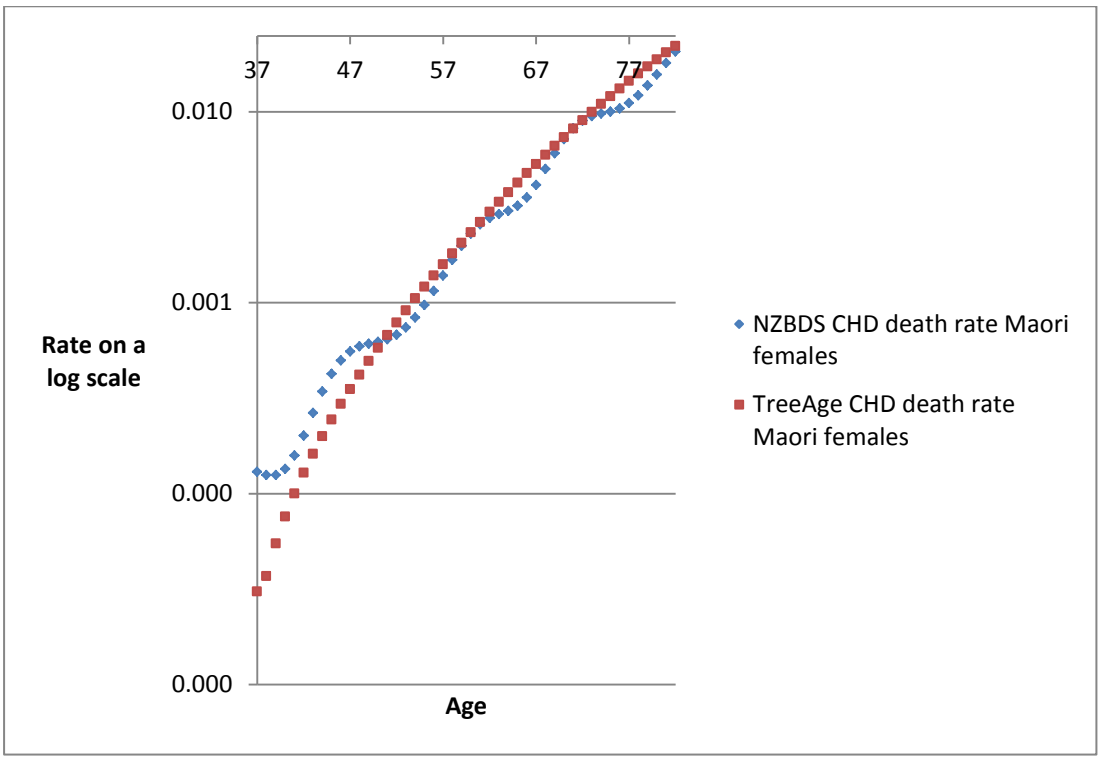


Figure 38 CHD mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 and 67 years for Maori females, on a log scale

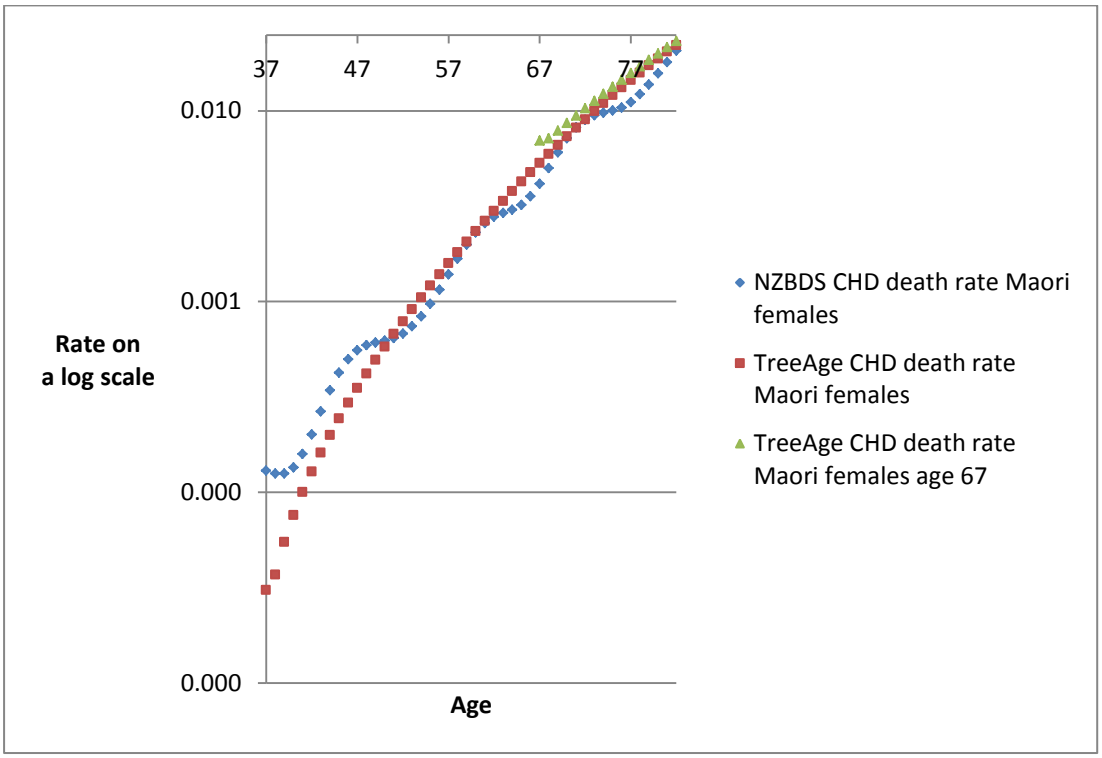


Figure 39 Total stroke mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for Maori females

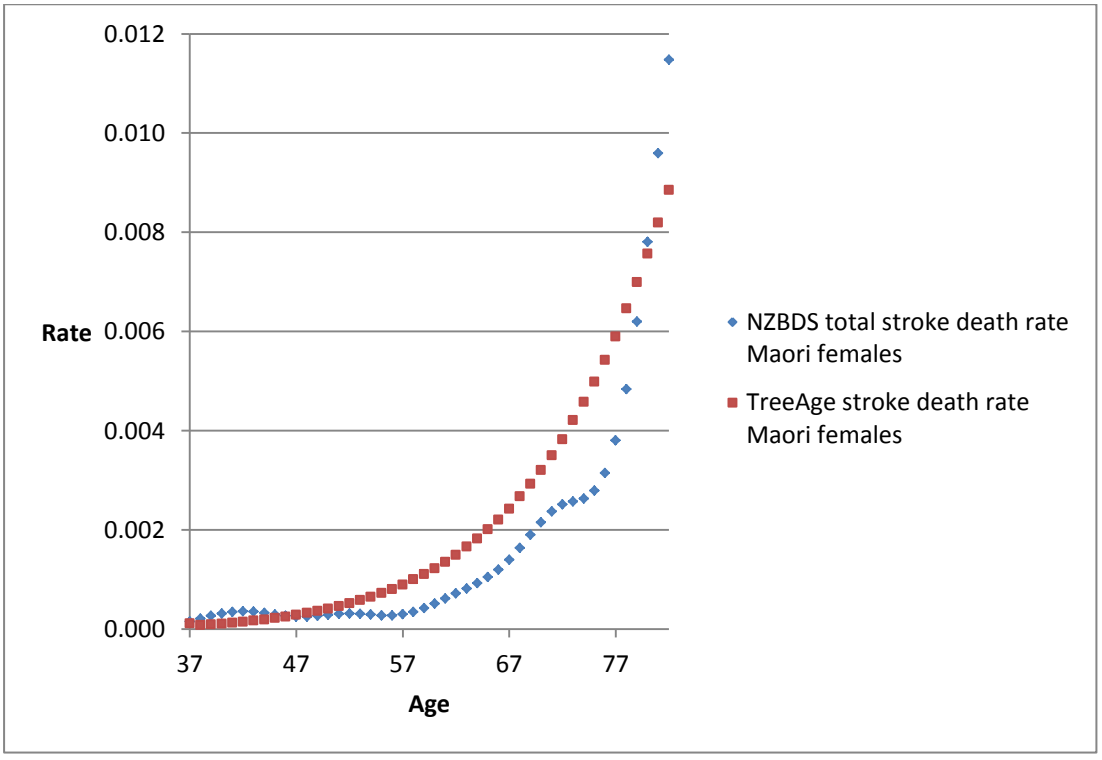


Figure 40 Total stroke mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 years for Maori females, on a log scale

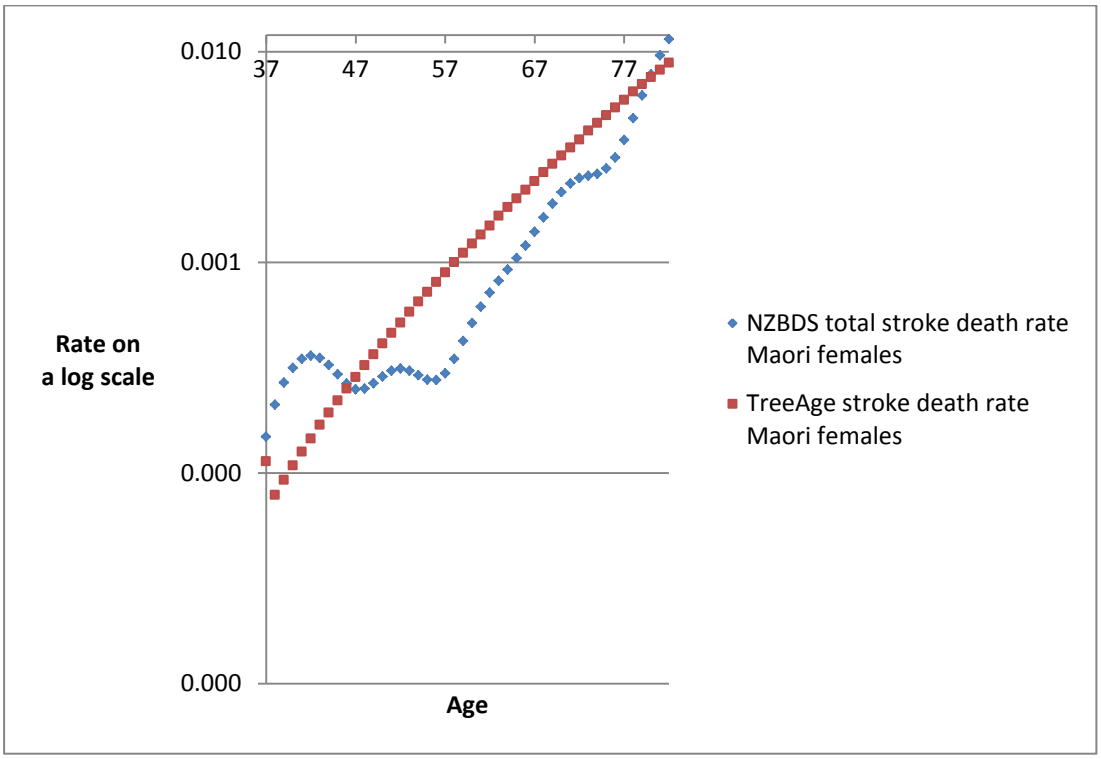
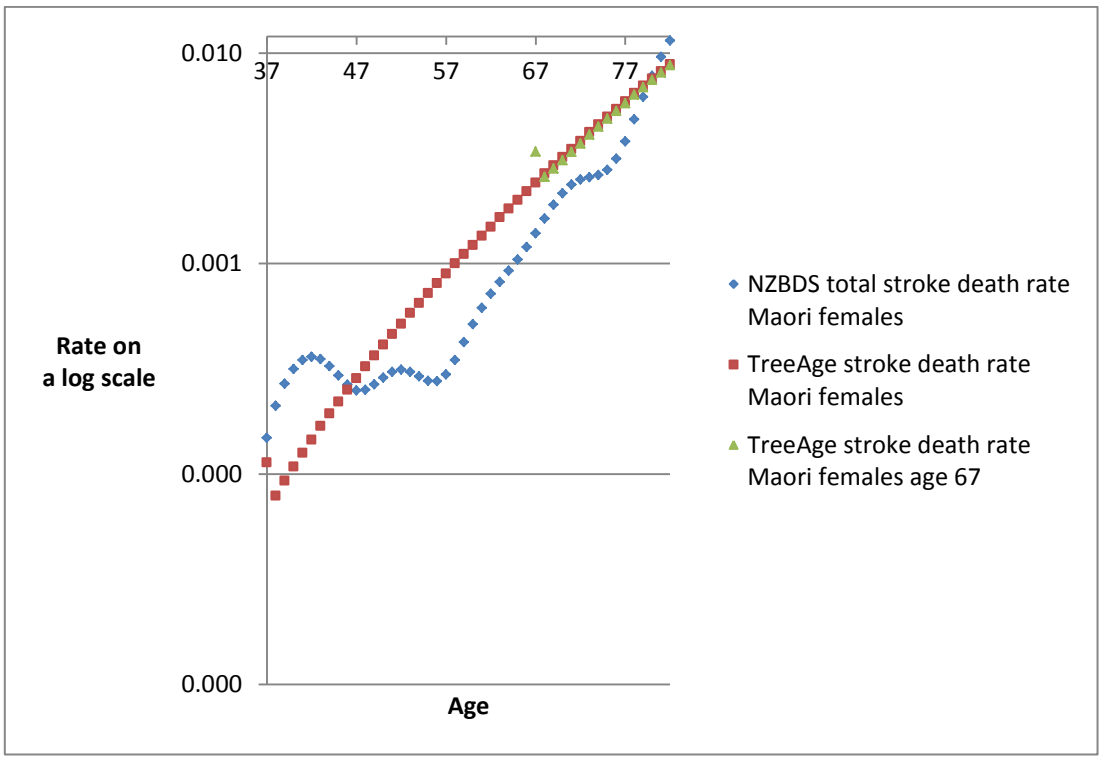


Figure 41 Total stroke mortality rate comparison between TreeAge output and NZBDS data, starting at age 37 and 67 years for Maori females, on a log scale



References

1. Barendregt J, Oortmarssen GJ, Vos T, Murray CJL. A generic model for the assessment of disease epidemiology: The computational basis of dismod ii. *Population health metrics* 2003;1:4.
2. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: A report of the ispor-smdm modeling good research practices task force-7. *Value in Health* 2012;15:843-850.