Centre of Methods and Policy Application in the Social Sciences

---

Primary Care in an Ageing Society: Developing the PCASO microsimulation model

Roy Lay-Yee
Janet Pearson
Peter Davis
Martin von Randow
Sanat Pradhan

Technical Report
December 2011

ISBN 978-0-473-20468-6 (online)

Centre of Methods and Policy Application in the Social Sciences
The University of Auckland

Postal address: Private Bag 92019
Auckland 1142
New Zealand

Location: 20 Wynyard St
Auckland 1010
New Zealand

Phone : +64 9 3737999

Fax: +64 9 3737986

Email: T.Crick@auckland.ac.nz (Administrator)

Website: www.Compass.auckland.ac.nz

Acknowledgements

The Primary Care in an Ageing Society (PCASO) project was funded by the Health Research Council of New Zealand. For their advice, we are grateful to Professor Ngaire Kerse, Associate Professor David O'Sullivan, Daniel Patrick, Professor Alastair Scott, and Andrew Sporle (the University of Auckland); and Annie Abello, Professor Laurie Brown, and Dr Sharyn Lymer (National Centre for Social and Economic Modelling [NATSEM], University of Canberra, Australia). For their peer review of this report, we wish to thank Michelle Gosse (Food Standards); Emmanuel Jo, Dr Jim Primrose, and Dr Martin Tobias (Ministry of Health); Dr Sharyn Lymer (NATSEM); and Vince Galvin (Statistics New Zealand).

Contents

Summary

The aim of the project was to establish a microsimulation model of the primary medical care system in New Zealand in its social context and to test the impact of demographic ageing, community support and practitioner repertoire.

Micro-level data were drawn from four sources: the New Zealand Health Survey (NZHS 1996/7 and 2002/3); a national survey of ambulatory care in New Zealand (NPMCS 2001/2); and the Australian National Health Survey (ANHS 1995). Data from the New Zealand surveys were statistically matched to create a representative synthetic base-file of over 13,000 individuals. Probabilities of health experiences and general practitioner (GP) use from the Australian health survey, and of GP activity from the New Zealand survey of ambulatory care, were derived. A microsimulation model was developed that applied these probabilities via a Monte Carlo process to create health histories for the individuals in the base-file. Final outcomes simulated were: the number of visits in a year, the distribution of health conditions, and GP activity levels. Policy scenarios were tested by changing characteristics of the synthetic population and by implementing counterfactuals on key attributes.

The model imputed a synthetic health history over a year to each individual. Verification showed that the model was able to reproduce expected results and was operating according to design specifications. The final outcomes produced by simulation were validated against data external to the model. Various scenarios, assuming moderate demographic ageing, were tested by a forward projection to 2021. These showed little change in model-predicted health care outcomes.

Using a microsimulation approach, we created from a number of different data sources a working model of primary medical care in New Zealand 2002 that has generated plausible results for key parameters. Furthermore, we were able to use the model to test a range of scenarios for demographic ageing. Model projections suggest limited change in system demand.

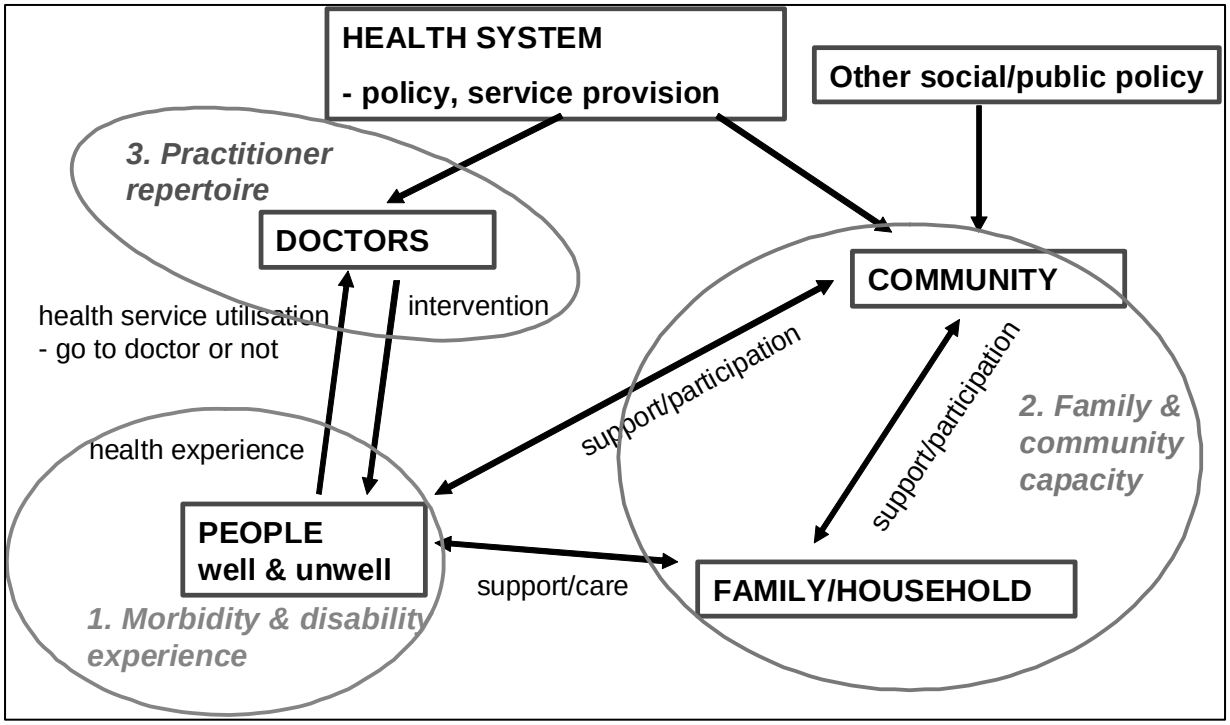
1. Introduction

The primary stimulus for this investigation is the appreciation that demographic ageing has major implications for the future of primary care *(Bryant et al 2004; Garces et al 2003; Lloyd-Sherlock 2000)*. Further the rationale is twofold – technical and policy-based. The technical rationale is that health services research in primary care has failed to deal with the sector as a complex, interconnected and evolving system understood within its broader social context *(Gabe 1991)*. The policy argument is that the primary care sector, as traditionally constituted in New Zealand and similar jurisdictions, is under challenge from a number of interconnected social trends, foremost of which is demographic ageing *(Moore et al 2003; Sox 2003)*. A rapidly ageing population has considerable implications for public health expenditure *(McGrail et al 2002; Ministry of Health 2004a; OECD 2006)*.

Microsimulation is based on the modelling of individual behaviour and allows for a more disaggregated approach to scenario building *(Gilbert et al 2005)*. Thus we have used microsimulation to mimic the heterogeneity of the population and the complexity of relationships in the primary health care setting *(Brown et al 2002; Complex Systems Modelling Group 2010; Fone et al 2003, Gupta & Harding 2007; Wolfson 1994)*. It can draw together diverse data from the real world to create an artificial one upon which virtual experiments can be carried out. Microsimulation operates at the level of individual units, in our case these are persons from a representative real-world sample. Each person has a unique identifier and a set of associated attributes as a starting point, for example, age, gender, ethnicity, and health state. A set of rules, here derived from statistical analysis, is applied in a stochastic manner to these persons to simulate changes in state or behaviour. Essentially, this process generates a set of synthetic health histories for our base sample of persons. The substantive output from such a model comprises estimates of the resulting outcomes including both aggregate and distributional effects. Furthermore, modifications of influential factors can be undertaken to test hypothetical 'what if' scenarios on a key down-stream outcome of policy interest such as health service use.

The overall aim of this report is to describe the construction of a microsimulation model designed to address the policy linkages of the three major components of demographic ageing – the pattern of morbidity and disability associated with the extension of the life span, the formal sector of care (as represented here by the role of the general practitioner (GP), and the informal sector of community and family support (shown schematically in Figure 1.1).

Figure 1.1 Model of primary health care



In order to address the interplay of these three distinct social sub-systems in an empirical and realistic manner, a diversity of data sources (outlined in Table 1.1) is required because a single data set with complete coverage of all system components is not available. Thus, the National Health Survey (NZHS) provides a representative sample of the population of New Zealand, together with details on household composition, an important feature of the informal sector. The Australian Health Survey (ANHS) gives details on the morbidity experience and health care utilisation patterns of a population survey with many attributes similar to those in New Zealand (see Table 1.2). Finally, the National Primary Medical Care Survey (NPMCS) gives information on the interaction between patients (GP users) and their GPs (patient visits), as well as patterns of practitioner behaviour (GPs).

Table 1.1 New Zealand and Australian data sources

Study	**National Health Surveys (NZHS)**	**National Health Survey (ANHS)**	**General Practice Survey (NPMCS)**	**General Practice Survey (NPMCS)**
Country	New Zealand	Australia	New Zealand	New Zealand
Year	1996/7 (children) 2002/3 (adults)	1995	2001/2	2001/2
Sample	Children & adults	Children & adults	Patient visits	Doctors (GP)
N	13,548	53,828	9,272	244
Model Component	Community	Morbidity; Community	Morbidity; Practitioner	Practitioner

Table 1.2 Demographic characteristics of New Zealand and Australian data sources (non-institutionalised population)

	Synthetic base file (NZHS 2002/3 adults plus NZHS 1996/7 children)*	Australian Health Survey (ANHS) 1995
Age group		
0-4	8.1	7.2
5-14	16.3	14.3
15-24	13.1	15.0
25-34	13.9	15.7
35-44	15.5	15.2
45-54	13.1	12.4
55-64	8.9	8.4
65-74	6.1	7.5
75+	5.1	4.5
Gender		
Female	51.2	50.2
Male	48.8	49.8
Household type		
Do not live with adult *	9.7	11.1
Live with adult partner	47.0	47.1
Live with adult but not partnered (excl. children)	19.0	41.8
Children	24.4	

* Adult is person aged 15 years or over.

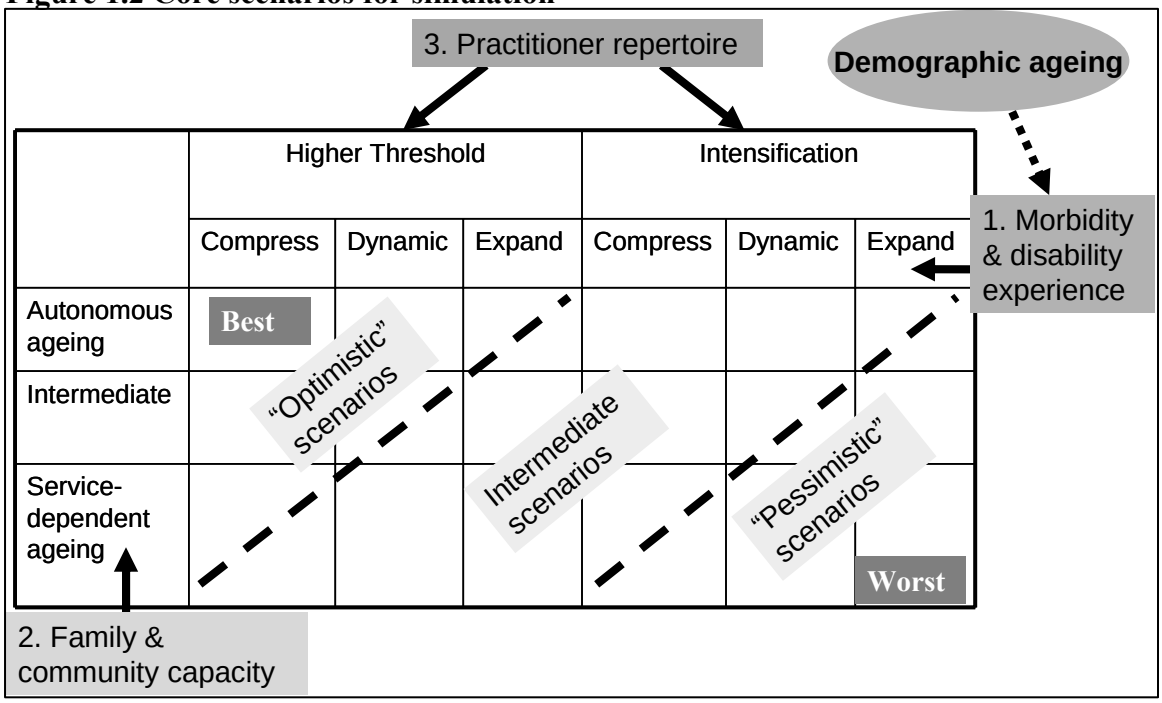
The logic is that people undergo health experiences (morbidity and disability sub-system, ANHS data) which they may take to their doctor, who will respond in various ways (formal sector, NPMCS data). This takes place in the context of a family or household, within a broad community, providing a level of social support which may promote individual health and care at home, thus mitigating the need to visit the doctor (informal sector, ANHS and NZNHS data) *(Ostberg & Lennartsson 2007; Prior & Hayes 2003; Van Houtven & Norton 2004)*. These components can in turn be related to scenarios (Figure 1.2):

(a) profile of morbidity and disability associated with demographic ageing *(OECD 2009; Swedish National Institute of Public Health 2006)*, as reflected in contrasting predictions of expansion and compression *(Graham et al 2004; Jagger et al 2006; Ministry of Health 2004b)*;

(b) "healthy ageing", as reflected in the potential of family and community capacity to assist in coping (autonomy, dependency, intermediate); *(Aboderin 2004)*

(c) the impact of changes in health service delivery, such as, technology and changes in practitioner repertoires (intensification, higher threshold of intervention) *(Davis et al 2000, 2002).*

Figure 1.2 Core scenarios for simulation



We applied the microsimulation framework to a static model of the primary care system as it was in the year 2002, and extrapolated to a year in the future, 2021, by reweighting the data *(Davis et al 2010; Pearson et al 2011)*. This report is a technical account of: the data used and data synthesis undertaken; the statistical models producing parameters for input; the microsimulation architecture; verification and validation checking procedures; and finally the outputs of the microsimulation model (see Figure 1.3). In particular, the different developmental stages of the microsimulation architecture will be presented.

For each individual in the synthesised base file, a health history over a year was created by firstly imputing health experiences and any visits to the doctor, and secondly imputing associated doctor activity (see Figure 1.4). By applying these rules, the model simulates outcomes based on probabilities and random allocation. This static model can be thought of as modelling a representative cross-section of the New Zealand population of 2002, as the inputs to the models were derived from data of approximately that period. The model was internally verified and externally validated to an acceptable level via an iterative process. It could then be used to project forward in time, and policy-sensitive factors could then be varied. Data manipulation and model implementation were programmed using SAS software *(SAS 9.1 and 9.2, SAS Institute Inc, Cary, NC, USA)*.

Figure 1.3 The model: data synthesis, simulation and scenario testing

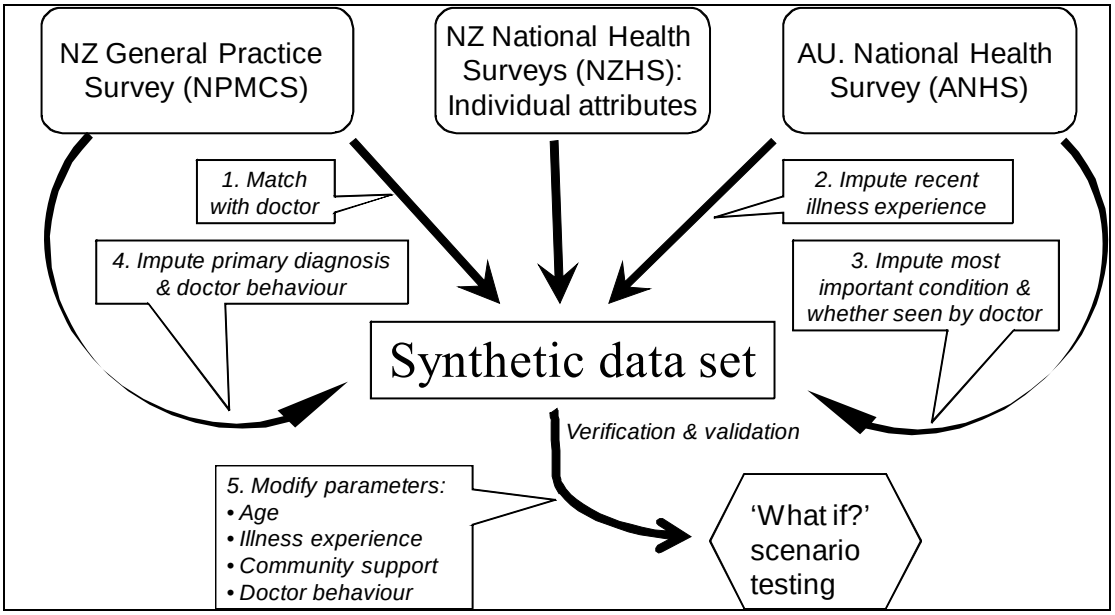


Figure 1.4 Creating a health history for individuals in the base file

Synthesised base file + imputed + imputed

NZ Health Surveys 1996/7 (children) & 2002/3 (adults) [n=13,548]	**NZ GP Survey 2001/2: Doctor & Practice (via patient visits) [n=244 GPs]**	**OZ Health Survey 1995 [n=53,828]**	**NZ GP Survey 2001/2 : Patient visits [n=9,272]**
Age		Age	Age
Gender		Gender	Gender
Ethnicity			Ethnicity
Deprivation			Deprivation
Number of visits in last 12 months			Number of visits in last 12 months
Living arrangements		Living arrangements	
Long-term conditions		Short-term & long-term condition categories	Primary diagnosis categories
		Go to doctor	
		1st listed reason for last visit in last 2 weeks	
		Number of visits in last 2 weeks	
	Doctor age, gender, ethnicity, etc		Doctor actions
	Practice type, location, number of doctors		

Note: Variables under the 'imputed' columns are un-shaded where they are link variables, and shaded where they have been added to the base file.

2. Data sources

The model used data from multiple sources: New Zealand Health Survey (NZHS, 1996/7 and 2002/3) *(Ministry of Health 1999, 2004)*, National Primary Medical Care Survey (NPMCS, 2001/2*) (Raymont et al 2004)*, and Australian National Health Survey (ANHS, 1995) *(Australian Bureau of Statistics 1996, 1997)* (see Table 1.1).

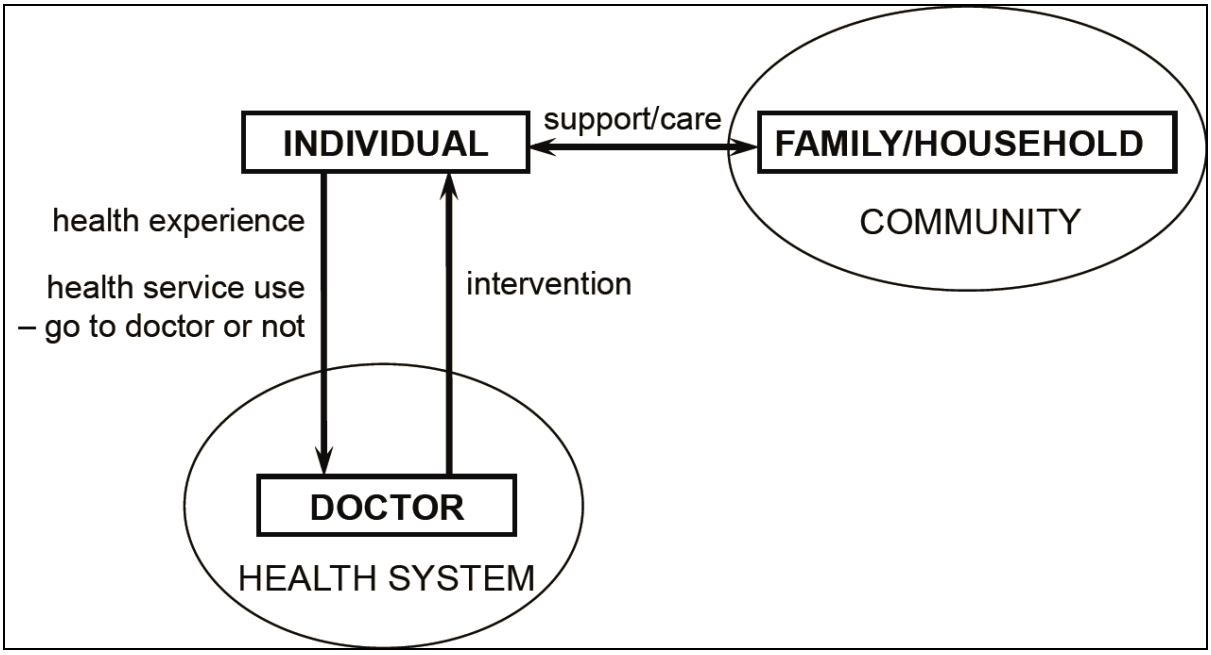
Micro-level data from the NZHSs (originally including the non-institutionalised only), weighted to be representative of the population using Census 2001, and NPMCS, representative of patients, that is, GP users, were statistically matched to create a representative synthetic base-file of 13,548 individuals each with an assigned general practitioner (GP). The ANHS was used to provide information on population levels of recent health conditions (that is, all conditions occurring in the last 2 weeks that were either seen or not seen by a GP), and GP use. NPMCS was used as the source of GP and practice information that was statistically matched with the base file, and as the data for predictive logistic regression models that derived probabilities of GP actions. These data sources were selected because of their availability, utility, quality, and compatible time periods and variable specifications. In the absence of information on recent illness in the NZHS 2002/3, it was decided to use the ANHS 1995 occurrence rates of conditions. The demographic makeup, with respect to age, gender and household type profiles, of the two countries was comparable (see Table 1.2).

3. Data Synthesis

3.1 Overview

Microsimulation permits a more disaggregated approach to model building and prediction. It also lends itself to combining diverse data sets to create realistic approximations of policy circumstances of strategic importance (*Abello et al 2008; Edwards & Clarke 2009; Martini & Trivellato 1997; Morrissey, et al 2008; Sutherland, et al 2002*). In the case of demographic ageing, model builders have been able to insert key social and demographic variables from existing data. However, there have been only limited attempts to incorporate practitioner behaviour, which is likely to be a key determinant of future cost and service outcomes (see Figure 3.1).

Figure 3.1: Sub-model of primary health care



Our model was based on publicly available data from the New Zealand Health Survey (NZHS), and data from the National Primary Medical Care Survey (NPMCS) that included practitioner information *(Raymont et al 2004)*. In particular we were able to allocate practitioners and their characteristics to members of a population sample *(von Randow et al, in press 2011)*. This is consistent with a general philosophy of combining data from different sources and adding value to a base-file by imputing values *(Abello et al 2008; Alegre et al 2000; Smith et al 2009)*.

We used the following procedure to carry out statistical matching:

1. Identify common variables on compatible scales

2. Divide data into 'cells' based on selected common variables
3. Choose between constrained and unconstrained matching
4. Apply a distance function within each cell with remaining common variables, and solve the transportation problem to assign matches within cells.

The method of statistical matching is well established in the literature *(Rodgers 1984; Rässler 2002)*. Its application here is more specific to the requirements of the associated microsimulation model *(Australian Bureau of Statistics 2004; Cohen 1991)*. Variables are not added directly to the existing data set; patients are each allocated an appropriate GP based on their being similar to that GP's actual patients.

3.2 Statistical matching

SAS software was used for data manipulation and to perform the statistical matching *(SAS Institute 2003)*. The data sources used for matching are summarised in Table 3.1.

Table 3.1: Data sets used for statistical matching

Data set	Details	N	Use
New Zealand Health Survey (NZHS) 2002/03	Repeated cross-sectional surveys on a representative sample of the New Zealand population (only 15 and over for 2002/03)	12,563 used (no missing)	Adult GP user records based on the presence of doctor visits in last year
New Zealand Health Survey (NZHS) 1996/97		1,019 used (children)	Child GP user records weighted up to 2001 Census proportions by age group, gender, ethnicity
National Primary Medical Care Survey (NPMCS) 2001/02	Cross-sectional survey on a representative sample of New Zealand general practitioners, and their patients	7,714 used covering 242 GPs (usual GP hours, no missing)	All ages patient visit records with associated doctors to be matched to NZHS GP users

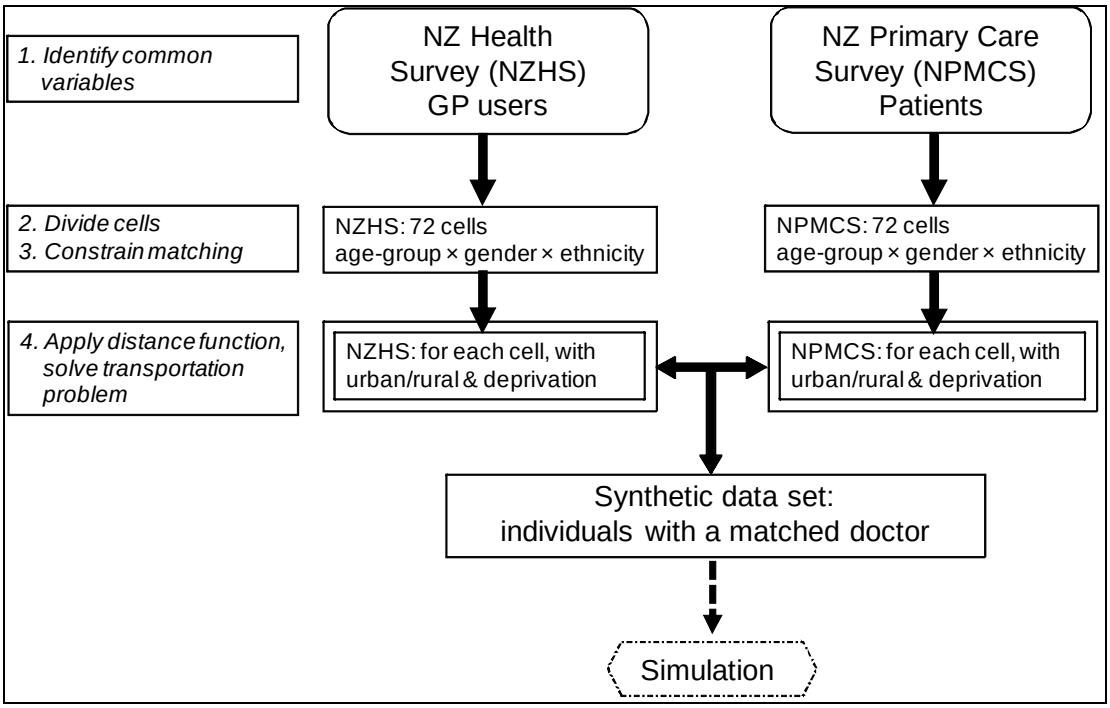
The sample weights of the 1,019 child (aged under 15 years) records in NZHS 1996/97 were adjusted so they could be appropriately appended to the records in NZHS 2002/03 (which surveyed 12,563 adults, but did not survey children). This involved generating weight multipliers from the 2001 New Zealand Census of Population and Dwellings. Proportions were matched by age group (0–4, 5–14), gender (Male, Female) and ethnicity (European, Māori, Pacific, And Other). The reweighted child records were appended to the NZHS 2002/03 data set, with as many NZHS 1996/97 variables as were compatible. Respondents reporting at least one GP visit in the last year were defined as 'GP users', and these were each assigned a GP via statistical matching.

There were 7,714 records in the NPMCS data set related to patient visits that took place during usual GP hours: Monday through Friday, 8am–6pm. NPMCS record weights were originally

calculated to be representative of all visits in a two-week period, while NZHS weights were representative of one visit in a twelve-month period. To adjust appropriately, NPMCS weights were multiplied by 26 (fortnights in a year) and divided by 6.58 (average GP visits per patient in the last year).

These data sets then became the subjects of the statistical matching. The process is summarised in Figure 3.2 and described in detail below, following the incremental steps identified earlier.

Figure 3.2: The statistical matching process



1. Identify common variables on compatible scales

Linking data sets without identifiers relies on there being variables in common. The usefulness of such variables depends on how many separate values can be compatibly defined *(Zaidi & Scott 2001)*. In the current case, the only common variables were: age, gender, ethnicity, urban/rural location ('urban' defined as population ≥30,000), and deprivation status of area of residence (NZDep deprivation quintile) *(Salmond et al 1998)*.

2. Divide data into 'cells' based on selected common variables

Using some of the common variables to divide up the data gives some base level of accuracy and ensures a certain standard for the subsequent matching *(Australian Bureau of Statistics 2004; Zaidi & Scott 2001)*. All cells created must be populated, and categories may need to be aggregated *(Abello et al 2008)*. The NZHS and NPMCS data sets were divided into 72 cells on age group, gender and ethnicity (9×2×4) as summarised in Table 3.2.

Table 3.2: Variables used to divide the data sets

Age group	Gender	Ethnicity
0–4, 5–14, 15–24,	Male	New Zealand or Other European
25–34, 35–44,	Female	NZ Māori
45–54, 55–64,		Pacific
65–74, 75 and over		Other

Thus each NZHS record would be matched with an NPMCS one from the same demographic sub-group. The statistical matching was constrained within these bounds.

3. Choose between constrained and unconstrained matching

Constrained matching preserves the marginal distributions of variables unique to each original data set *(Rodgers 1984)*. The alternative, unconstrained matching, puts no limitations on the number of times any record can be matched to others – distances between matched records may thus be more optimal, but for our case the distribution of GPs in the matched data set would have diverged from the original *(Australian Bureau of Statistics 2004)*.

As the intention was to link GPs in NPMCS via their patients to GP users in NZHS as completely as possible, there needed to be at least as many records in the NPMCS data set as in the NZHS according to sample weighted totals. Where this was not the case, NPMCS record weights were multiplied by a constant per cell to make the weighted totals equal those in NZHS *(Sutherland et al 2002)*. This allowed us to use constrained matching – all of the records in both data sets could be fully utilised.

4. Apply a distance function in each cell with remaining common variables, and solve the transportation problem to assign matches within cells

The variables in common that are not used to form cells are eligible for inclusion in the distance function. The selection can be tailored for specific aims, and variables can also be weighted differently based on which are deemed most important. Alegre, et al. *(2000)* concluded that the selection process is arbitrary, or at least that there is no method of optimisation. In our case, only deprivation and location were eligible for inclusion in the distance function, and we gave them equal weighting.

The distance measure is another consideration. We used Euclidean distance which was the default in the SAS software. The distance function defines per cell the statistical distance between each pair of records across the two data sets; the sampling weights indicate how many people each record represents in the whole population. For matching purposes, the data sets are defined as ‘donor’ and ‘recipient’; for records in the recipient data set, the weights show how

many records need to be matched to them (demand), and for the donor data set, how many they have on offer (supply).

These two pieces of information (distance and weight) are used to solve the transportation problem *(Australian Bureau of Statistics 2004)*. This is a linear programming optimisation exercise originating in the idea of minimising costs given supply and demand constraints in shipping ore from iron mines to factories. This is analogous to our situation as the record weights in our data sets correspond to replications of record combinations (matches), that is, how much mines can ship and factories can receive. The costs of shipping are represented by the distances from the distance function. The aim is to minimise the overall cost (distance) of shipping the records available in each cell of the donor data set to satisfy the demand of the recipient, that is, to minimise:

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\left(d_{ij} \times w_{ij}\right)$$

where d_{ij} is the distance between cases i and j in the recipient and donor data sets respectively, and w_{ij} is the weight to be allocated to records in the matched file based on those particular cases. The weights for our data sets were aligned so that demand and supply would be used up completely in the matching process. Weights were continuous not integer, and the aim was to assign a GP to each existing patient, not expand the data set. The concept of constrained matching was adapted to fit that purpose.

The aim was to link each GP user from NZHS with a single GP from NPMCS. An issue arising from this was that weighting resulted in a number of cases of one-to-many matches where GP users were assigned to multiple NPMCS records, and thus sometimes to multiple GPs. To solve this, we defined the distance between an NZHS GP user and an NPMCS GP as the average of the distances between that GP user and each of that GP's patients, and for each GP user matched one-to-many, a single GP was allocated based on the criterion of minimum distance.

3.3 Diagnostics

We compared the original NZHS records with their matched patient records from NPMCS to evaluate the performance of the distance function. Cross-tabulations were produced for each data cell, to assess the quality of a match *(Abello et al 2008)*. Overall, 96 percent of matches for deprivation, 98.7 percent for location, and 95.1 percent for both variables combined, occurred on the diagonal, that is, exact value matches. Similarly, correlations between NZHS records and their matches were 0.97 for both variables. Table 3.3 gives an example of this analysis focusing on the results of the matching process for the cell representing 15–24-year-old NZ Māori females. Combinations of deprivation and location, used in the distance function, are shown with

the resulting numbers of record pairs matched. Numbers on the marked diagonal represent exact matches which accounted for 93.4 percent of all matches in this cell.

Table 3.3 Distance function variables compared for matches – 15–24-year-old NZ Māori females

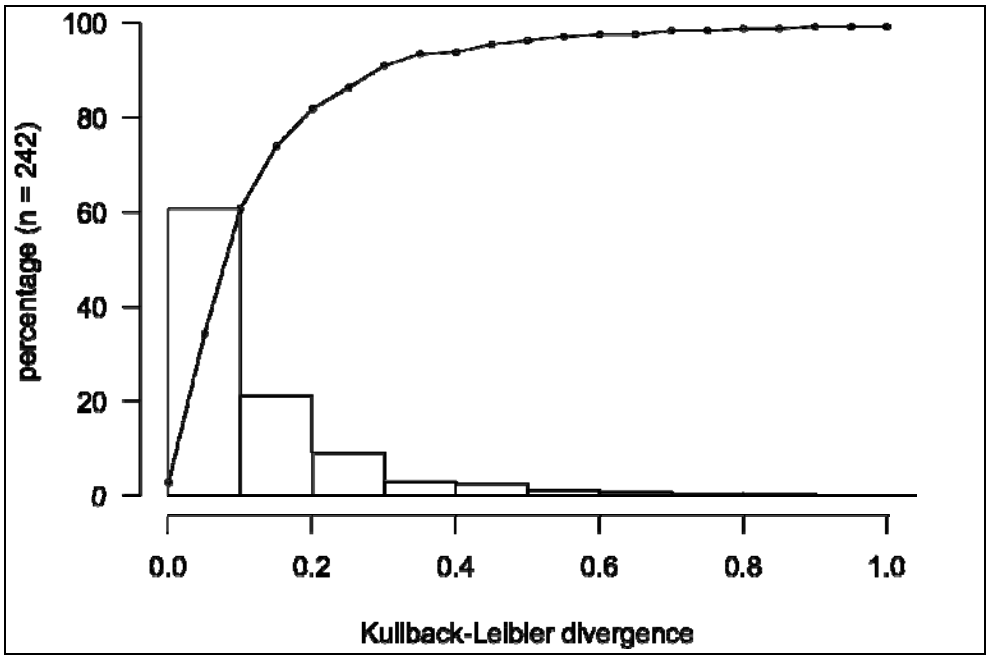
		NPMCS: NZDep, Urban/Rural									
		1, 1	1, 2	2, 1	2, 2	3, 1	3, 2	4, 1	4, 2	5, 1	5, 2
NZHS: NZDep, Urban/Rural	1, 1	2	0	1	0	0	0	0	0	0	0
	1, 2	0	2	0	0	0	0	0	0	0	0
	2, 1	0	0	7	0	0	0	0	0	0	0
	2, 2	0	0	0	1	0	2	0	0	0	2
	3, 1	0	0	1	0	12	0	0	0	0	0
	3, 2	0	0	0	0	0	8	0	0	0	0
	4, 1	0	0	0	0	0	0	34	0	0	8
	4, 2	0	0	0	0	0	0	0	14	0	5
	5, 1	0	0	0	0	0	0	0	0	98	0
	5, 2	0	0	0	0	0	0	0	0	0	100

We also compared the distribution of these characteristics among individuals who were allocated to GPs (NZHS) with that of the patients that had actually been to see the GPs (NPMCS). We used the Kullback-Leibler divergence measure (K-L) for the discrete case (*Afgani et al 2008*). This is defined as

$$D(p||q) = \sum_{x \in X} p(x) log \frac{p(x)}{q(x)}$$

where X is the set of possible outcomes for probability distributions p and q (those for NPMCS and NZHS respectively in our case). The equation evaluates to zero where the distributions are identical. Figure 3.3 shows the distribution and cumulative percentage of K-L divergences among GPs; over 90 percent of GPs had an acceptable K-L divergence < 0.3.

Figure 3.3 Kullback-Leibler divergences across GPs for NZDep quintile



3.4 Sensitivity testing

To assess the performance of the statistical matching process, we considered how much of a difference to correlation results each incremental step made (see Table 3.4):

1. Random allocation of matches
2. Data sets divided into cells but random allocation within each cell
3. Data sets divided into cells and distance function applied within each cell.

We aimed to produce a similar distribution of patients among GPs to that originally observed in NPMCS. The 242 GPs were first sorted and thus ranked on the proportion of all patient visits that were theirs for the original NPMCS data. The same was then done for the matched data set (using Euclidean distance). The difference in rank between these two sets was calculated for each GP, and the average of the absolute values of these differences across all GPs was taken as a measure of similarity. Table 3.4 shows these measures, and again compares them with those for random and simple cell-based matching, and also for other distance measures. Higher correlation and lower rank difference values indicate better matching.

Table 3.4 Improvements made by statistical matching steps

Statistic	Random	Cells	Distances		
Correlations: NZHS vs NPMCS matches					
NZDep Quintile	0.004	0.19		0.98	
Urban/Rural	0.003	0.08		0.97	
Distribution of patients among GPs			Gower's	Manhattan	Euclidean
Average absolute rank difference	53.46	50.17	47.63	31.21	31.66

Clearly the addition of the distance function was the most influential component in the matching process. Euclidean distance turned out to be a better choice than Gower's for our data; Manhattan distance was better again, but negligibly so.

4. Statistical Analysis

4.1 Overview

To inform our simulation, we aimed to develop and integrate predictive models of clinical activity based on data from the 2001 National Primary Medical Care Survey (NPMCS), a nationally representative sample of general practitioners (GPs) and their patient visits in New Zealand. We sought to produce models that were explanatory based on underlying theory and that were empirically shown to be predictive.

For this purpose, we employed multilevel models (*Bryk and Raudenbush 1986; Goldstein 1987, Snijders & Bosker 1999)*. Multilevel models offer a number of advantages over the traditional models. Firstly, they provide a convenient framework for analysing multilevel data. Such a framework supports a methodical analysis of how individual covariates and interactions among covariates measured at various levels of a hierarchical structure affect the outcome variable. Secondly, multilevel models account for the biases in parameter estimates resulting from clustering. Ignoring the multilevel composition and the clustering within the levels can result in biases in parameter estimates as well as their standard errors. When observations are clustered in higher-level units, they are no longer independent and the assumption of independence for regression models is violated. By taking account of the clustering and providing appropriate standard errors, multilevel models yield more accurate confidence intervals and significance tests.

4.2 Modelling process

The key parameters of clinical activity, that is, the probabilities of an investigation, prescription, non-drug treatment, follow up, and referral would be estimated using multi-level logistic regression models which take account of the multi-stage sampling scheme, and the variability between GP's as well as between patients who are nested within the practice *(Davis et al 2002)*. Patient visits (level 1) were considered to be nested within practices (level 2). Practitioner identity was introduced as a random effect in our base intercept-only model and variance components monitored as covariates were gradually added to subsequent models. The outcome variables were modelled as binary, that is, taking a particular action (1) or not taking that action (0). The potential predictor variables consisted of both patient and doctor variables as well as the practice characteristics. In order to improve predictivity, the outcome measures for the GP actions prescription and non-drug treatment are each split into two depending on the number of diagnoses variable and a separate model was fit for each of these, that is, a 'single' model for prescription and non-drug treatment if the number of diagnoses equals 1 and a 'multiple' model if the number of diagnoses is more than 1. We used the SAS Glimmix procedure *(SAS 9.2, SAS*

Institute Inc, Cary, North Carolina, USA; Schabenburger 2005) to implement the multilevel models.

Data description

The National Primary Medical Care Survey (NPMCS), carried out over 2001/02, was a national survey of ambulatory care in New Zealand *(Raymont et al 2004)*. It involved a nationally representative, multistage, probability sample of general practitioners (GPs) and their patient visits. The variables required for the statistical modelling process were obtained using the dataset derived from a combination of the patient, visit, practitioner, and practice questionnaires. The units of analysis are the patient visits to the GP. The main outcome measures are GP actions made in response to a patient visit, that is, investigation, prescription, non-drug treatment, follow up, and referral. The potential predictor variables consist of patient and GP (or doctor) variables as well as the practice characteristics. Table 4.1 lists the full set of variables used for the modelling process.

Table 4.1 Description of variables required for the modelling process

Variables	Description
Outcome	
Clinical Activity	
Follow-up	GP requesting patient for follow-up visit
Prescription	GP prescribing a drug to patient
Referral	GP referring patient to a specialist
Investigation	GP orders an investigation or test
Non-drug treatment	GP recommending a non-drug treatment
Predictor	
Patient	
Patient Age (in years)	Less than 25 / 25 - 44 / 45 - 64 / 65+
Patient Gender	Female / Male
Patient Ethnicity	Asian / European / Maori / Other / Pacific
NZ Deprivation Index	1,2 (lowest) / 3,4 / 5,6 / 7,8 / 9,10 (highest)
Number of visits in last 12 months	Less than 3 / 3-5 / 6-11/ 12+
Primary diagnosis (as recorded by GP)	Refers to the 17 medical conditions (table 1)
Doctor	
Doctor Age (in years)	Less than 35 / 35 - 44 / 45 - 54 / 55 - 64 / 65+
Doctor Gender	Female / Male
Doctor Ethnicity	Asian / European / Maori / Other / Pacific
Workload (in hours)	Less than 8 / 8 - 14 / 15 - 21 / 22 – 28
Practice	
Practice Type	HCA (community-governed) / PRI (private)
Practice Location	Urban / Rural
Number of doctors in the practice	1 / 2 - 3 / 4+

Data preparation

The following steps were taken to prepare data and select predictor variables for the modelling process:

- The frequency tables of each doctor action were examined in order to ensure it did not contain excessive missing data and that the proportion of 'yes' or 'no' for that particular doctor action was not disproportionate.
- Potential predictor variables were identified using subject-matter knowledge and possible predictive ability while considering their derivability and consistency across other data sources required for the simulation process.
- For the modelling process, any continuous variable out of the selected group was categorised in order to avoid fitting higher-order (polynomial) terms for variables which may be non-linear in the logit. This maintains theoretical interpretability though there is loss of information.
- The association between each potential categorical predictor variable and the doctor actions was evaluated with bivariate analysis using chi-square tests.
- Prior to the model development process, the choice of scale or categories of each potential predictor variable was informed by existing literature (NPMCS 2001).
- In order to ensure comparability across other data sources in the simulation, predictor variables and their groupings were kept consistent throughout the process.

4.3 Variable selection

We carried out an exhaustive search from all potential candidate variables by minimising the standard statistical criterion to pick the best models for each doctor action. The 'exhaustive search' variable selection procedure, as its name implies, searches all possible subsets and selects the one with the best evaluation criterion. It is the only technique that is guaranteed to find the best subset using a given criterion. The only drawback is that if the number of covariates is large, this method can be computationally intensive and time consuming. Since the objective of this project is to find the best predictive models using all our candidate predictors, we have chosen the exhaustive search procedure as our variable selection method. The proposed approach for variable selection is detailed below:

- Start with the full model using all our candidate predictors
- Pick the best sub-models by minimising the Akaike Information Criterion (AIC) and/or Bayesian Information Criterion (BIC) using the integral approximation method of likelihood estimation in SAS version 9.2.
- Evaluate performance of the best fitting sub-models by internally validating the results of the outcome distribution

Table 4.2 shows the variable codes: patient variables are coded from X1 to X6, doctor variables coded from Y1 to Y4, and practice variables coded from Z1 to Z3.

Table 4.2 Variable codes for the predictor variables

Code	Variable
	Patient
X1	Patient Age
X2	Patient Gender
X3	Patient Ethnicity
X4	NZ Deprivation Index
X5	Number of visits in last 12 months
X6	Primary diagnosis
	Doctor
Y1	Doctor Age
Y2	Doctor Gender
Y3	Doctor Ethnicity
Y4	Workload
	Practice
Z1	Practice Type
Z2	Practice Location
Z3	Number of doctors in practice

We chose the 'best' sub-models by minimising the BIC. The BIC applies a greater penalty for models with more parameters and tends to favour more parsimonious models than the AIC. The variable codes indicate which variables are included in the selected sub-models.

The variables included in the final predictive models are (also see Table 4.3):

1. X1 to X6, Y1, and Y2 for Referral
2. X1 to X6, and Y1 for Prescription (Single Diagnosis)
3. X1 to X6, Y1, and Y2 for Prescription (Multiple Diagnoses)
4. X1 to X6, Y1, Y2, and Z3 for Non-drug treatment (Single Diagnosis)
5. X1 to X6, Y1, and Y2 for Non-drug treatment (Multiple Diagnoses)
6. X1 to X6, Y1, and Z2 for Follow-up
7. X1 to X6, Y1, and Y2 for Investigation

4.4 Internal validation

In the final phase of the modelling process, we partitioned the data into subsets such that the statistical models were initially generated using the training set and then validating/confirming these models using the testing set. A *randomised hold-out validation* was carried out where observations were chosen randomly from the initial sample to form the validation data, and the remaining observations were retained as the training data. The first step in this process was to

'clone' the survey data to the actual population using weights associated with each observation. Here, 9,272 observations used in the analysis were now cloned to 264,272 observations representing the actual population to be used in the validation process.

For this validation method, a randomly selected subset of the data (40%) from the cloned population was kept out of the modelling process.

Similar to the earlier modelling process using the entire dataset, we used this training data set to generate multi-level logistic regression models for each doctor action. The model selection process included minimising the AIC and/or BIC in order to get the best predictive model. The final model was similar to that where the entire dataset was used. For validation purposes, the models derived from the training set were applied to the testing set. The outcome distribution derived from these compared closely to the actual distribution of clinical activity level in NPMCS 2001.

4.5 Predictive equations

The models selected in the variable selection and validation process were integrated into the microsimulation architecture by calculating the predicted probabilities of clinical activity, that is, investigation, prescription (for single and multiple diagnoses), non-drug treatment (for single and multiple diagnoses), follow-up, and referral using the logistic regression coefficients for each predictor variable over its range.

The general fitted relationship for the logistic model of a clinical activity c is

$$logit(\pi_c) = \beta_0 + \sum_{j=1}^{J} \beta_{1j} \times x_{1j} + \sum_{j=1}^{J} \beta_{2j} \times x_{2j} + \cdots + \sum_{j=1}^{J} \beta_{kj} \times x_{kj}$$

where π_c is the probability of the GP taking a particular action, β_0 is the intercept for the action c, $x_{1j}, \ldots, x_{kj}$ are indicator variables which only take on the values 0 or 1, $\beta_{1j}, \beta_{2j}, \ldots, \beta_{kj}$ are the regression coefficients corresponding to each level of each predictor variable, and J represents the total number of levels for the k^{th} predictor variable.

The predicted probability of each clinical activity is obtained by back-transforming

$$\pi_c = \frac{e^{logit(\pi_c)}}{1 + e^{logit(\pi_c)}}$$

which can also be expressed as

$$\pi_c = \frac{1}{1 + e^{-logit(\pi_c)}}$$

The probability of each clinical activity can be calculated by substituting values of the regression coefficients (β_{kj}) for the k^{th} predictor variable with J levels (Table 4.3).

Table 4.3 Regression coefficients for the predictive equations

Predictor Variables	Regression Coefficients (β_{ij})						
	Follow-up	Prescription (Single)	Prescription (Multiple)	Referral	Investigation	Non-Drug (Single)	Non-Drug (Multiple)
Intercept	2.59	0.41	1.61	-1.09	-0.08	1.83	0.12
Patient Age							
0 – 24 years	-1.04	-0.09	-0.09	-0.37	-0.71	-0.17	-0.15
25 – 44 years	-0.63	0.04	-0.16	0.37	0.22	-0.01	0.68
45 – 64 years	-0.36	-0.04	0.23	-0.01	0.27	-0.02	0.32
65+ years	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Patient Gender							
Male	0.02	0.05	-0.18	0.04	-0.18	0.08	-0.01
Female	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Patient Ethnicity							
NZ European	-0.46	-0.54	-0.31	0.27	0.69	0.08	0.32
Maori	-0.01	-0.10	-0.21	0.31	0.41	-0.09	1.12
Asian	0.29	-0.04	-0.31	0.32	0.39	0.17	0.86
Other	0.28	-0.20	-0.45	0.03	0.40	0.32	0.64
Pacific Islander	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Primary Diagnosis							
Infectious	-1.19	-0.10	0.52	-2.54	-0.81	-0.97	-1.38
Digestive	-0.22	0.07	0.69	-0.68	-0.62	-1.01	-0.29
Pregnancy	-0.57	-0.89	-1.28	-0.23	-0.52	-0.75	5.06
Skin	-0.59	0.68	0.43	-1.52	-1.37	-1.39	-0.65
Musculoskeletal	-0.07	-0.27	0.77	0.17	-0.87	-0.65	-0.15
Congenital	1.30	-0.19	5.35	0.69	-2.17	-0.82	5.48
Symptoms/Signs	-0.19	-0.85	-0.54	-0.54	-0.52	-0.31	-0.15
Injury	-0.21	-0.89	0.02	-0.29	-1.70	-0.20	-0.18
Unspecified	-0.50	-1.51	-0.50	-1.11	-1.02	-0.82	-0.56
Neoplasm	0.60	-2.36	-0.23	-0.81	-1.10	0.50	0.55
Endocrine	0.07	0.41	1.44	-0.64	-0.20	-0.99	-0.72
Diseases of blood	1.10	-2.52	0.28	-1.37	1.34	-1.03	0.09
Mental disorders	0.78	0.40	0.67	-0.54	-1.37	-1.10	-0.81
Nervous system	-0.33	0.46	0.39	-1.02	-1.98	-1.78	-0.82
Cardiovascular	0.32	0.63	1.20	-0.94	-0.87	-1.39	-1.26
Respiratory	-1.02	1.22	1.23	-2.10	-1.55	-2.02	-1.24
Genitourinary	0.00	0.00	0.00	0.00	0.00	0.00	0.00
No. of Visits							
2 or less	-1.14	0.06	-0.62	0.13	0.34	0.15	-0.13
3 – 5	-0.83	0.44	-0.04	-0.16	0.08	-0.07	-0.22
5 – 11	-0.62	0.13	0.01	-0.01	-0.12	0.07	-0.28
12 or more	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Predictor Variables	Regression Coefficients (β_{ij})						
	Follow-up	Prescription (Single)	Prescription (Multiple)	Referral	Investigation	Non-Drug (Single)	Non-Drug (Multiple)
Deprivation Index							
1	0.17	0.06	0.38	0.07	0.04	-0.27	0.11
2	0.07	0.14	0.02	-0.13	-0.16	-0.15	-0.17
3	0.21	0.08	0.18	-0.13	-0.08	-0.03	0.13
4	0.07	0.01	0.03	0.16	-0.02	0.29	0.19
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Doctor Age							
<24 years	-0.66	-0.37	-0.09	0.69	0.01	1.40	3.14
35 – 44 years	-0.20	0.14	-0.04	0.31	-0.03	0.98	2.22
45 – 54 years	-0.26	-0.02	0.14	0.33	0.00	0.71	1.86
55 – 64 years	-0.21	0.01	0.46	-0.41	0.18	0.29	1.22
65+ years	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Doctor Gender							
Male	NP	NP	0.35	-0.62	-0.39	-0.68	-1.13
Female	NP	NP	0.00	0.00	0.00	0.00	0.00
No. of Doctors							
1	NP	NP	NP	NP	NP	-0.79	NP
2 - 3	NP	NP	NP	NP	NP	-0.97	NP
4+	NP	NP	NP	NP	NP	0.00	NP
Practice Location							
Rural	-0.67	NP	NP	NP	NP	NP	NP
Urban	0.00	NP	NP	NP	NP	NP	NP

*NP – Variables are not predictive for the given model

We also computed 68% confidence intervals (+- 1 standard error) around the predicted probabilities from the multi-level logistic models. Similar to the integration of the predicted probabilities, the upper and lower predicted probabilities of each clinical activity are calculated by adding and subtracting the standard error of regression coefficients for each explanatory variable over its range. The upper and lower predicted probabilities of each clinical activity is obtained by back-transforming

$$\pi_u = \frac{e^{logit(\pi_u)}}{1 + e^{logit(\pi_u)}}$$

$$\pi_l = \frac{e^{logit(\pi_l)}}{1 + e^{logit(\pi_l)}}$$

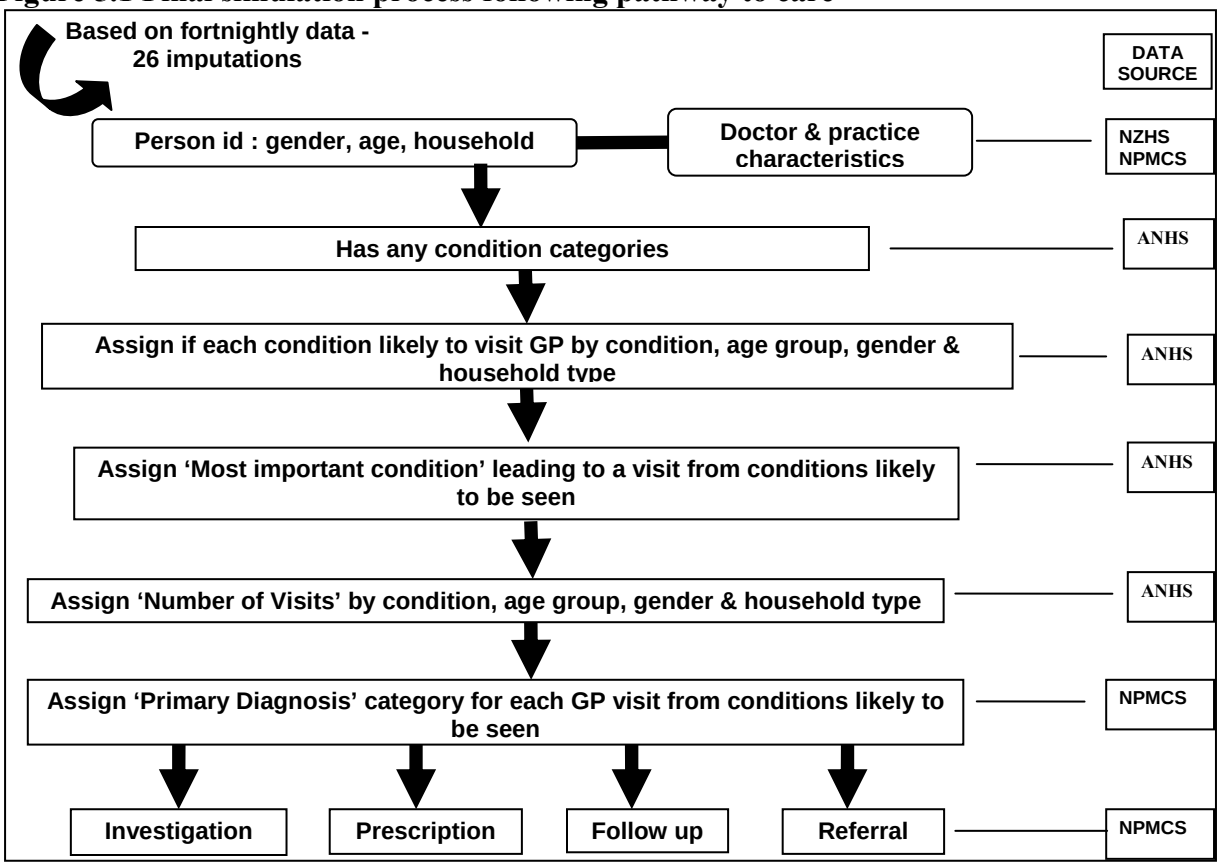
where π_u and π_l are the upper and lower probabilities of the GP taking a particular action.

5. Microsimulation: Construction

5.1 Overview

The construction of the simulation model followed a "pathway to care" sequence (see Figure 5.1). Effectively, a health history was created for each person with as realistic a linkage as possible from one health event to the next. However, this was not always practicable depending on the availability and nature of the data. Thus a representative sample of the non-institutional population, with assigned doctors, has a characteristic symptom or illness experience or other reason, which may be taken to the family doctor or GP, among which will be a leading reason for visit, of which there may be more than one in any given fortnight, to which the GP responds with a diagnosis and various actions. The steps involving the use of ANHS 1995 data draw on earlier work by the National Centre for Social and Economic Modelling (NatSEM) *(Abello et al 2008; Lymer et al 2006; Lymer et al 2008).*

Figure 5.1 Final simulation process following pathway to care



5.2 Synthetic base file

It was decided to use the NZHS 2002/3 as the core of our base file (instead of the New Zealand Census 2001) as this set contained most of the required variables and we had ready access to micro-level data. As this survey did not include children under the age of 15, information about children had to be brought in from an earlier survey (NZHS 1996/7). Additionally, the survey only interviewed adults in permanent private dwellings, and so the original survey weights had to be adjusted (via comparison to the Census 2001) in order to make the data representative of the entire New Zealand population. The ANHS 1995 was used to provide information on population levels of recent health conditions and GP utilisation. The NPMCS 2001/2 was used as the source of GP and practice information that was statistically matched with the NZHSs, and as the data for statistical models that derived probabilities of GP actions.

5.2.1 Assigning a doctor profile to each patient

As described above, a likely GP profile for each person was matched in beforehand. It was decided to do this rather than assigning a GP via probabilities during the simulation as it was thought that this would provide a better distribution of the doctors among the population of potential patients; we did not want to reduce the value of the information we already had by using probabilities, especially as most people are stable over time in their choice of type of doctor. It would also have been impracticable and cumbersome to continually match new doctors during the simulation process. It was also decided not to include a doctor choice component in the simulation as it was not the most important part of the project and it was difficult to get an idea of the number of patients a doctor could see as we only had information on the number of patients per half day, and there were issues with this variable, for example, it had been imputed for many doctors. Also we wanted to prepare the base-file (including GP and practice characteristics) using statistical matching and then validate the results of matching, all prior to simulation. It would have been challenging to validate if we allocated GPs 'on demand' within the simulation.

5.2.2 Using Australian data

In the absence of population-level information in the NZHS 2001/2, it was decided to use the ANHS 1995 rates of population conditions as input into the simulation. Table 5.1 shows that the demographic makeup of the two countries was similar so it was assumed that health profiles would also be comparable.

Table 5.1 Comparison of New Zealand and Australian health surveys

	New Zealand Health Survey 2002/3 distribution (%)*	Base file distribution (including children) (%)*	Australian Health Survey 1995 (%)*
Age group			
0-4	NA	8.1	7.2
5-14	NA	16.3	14.3
15-24	17.3	13.1	15.0
25-34	18.3	13.9	15.7
35-44	20.6	15.5	15.2
45-54	17.3	13.1	12.4
55-64	11.8	8.9	8.4
65-74	8.0	6.1	7.5
75+	6.8	5.1	4.5
Gender			
Female	51.9	51.2	50.2
Male	48.1	48.8	49.8
Household type			
live W ITHOUT someone >=15yrs age	12.8	9.7	11.1
live WITH someone >=15yrs age & partnered (husband/wife or de facto, boyfriend/ girlfriend)	62.1	47.0	47.1
live WITH someone >=15yrs age & NOT partnered	25.1	19.0	41.8
children <15yrs age		24.4	

* Non-institutionalised populations.

5.2.3 Inability to use income, occupation/employment and GP use variables

There were no Australian data available by income or occupation/employment status that were compatible with or could be matched to the NZHS base file.

Further, we decided it would be best not to use the gross income variable in the NZHS as it had various problems. Firstly it was a categorical variable with income bands of different width, making it difficult to model. Secondly there were unusual categories present, that is, 0.4% of people had a 'loss', 5% of people had 'zero' income, and 35% of people supposedly had a gross income of '$15,000 and under'; this does not make sense given that the question asked for gross before tax income which by definition should be at least zero. There was also a significant amount of missing data with 3.3% of respondents not specifying or refusing to answer, and 5.9% saying they did not know.

Information on occupation also could not be used due to too much missing data – 35.13% of people did not answer the question due to their responding that they had not worked (or been absent from work) in the last 7 days. Neither could we use information on employment status for the same reason (35.13% did not work in the last week). In addition since we were unsure about the validity of using the '7 day' period questions ('work/not work in the last 7 days' or 'in the last 7 days did you have one job or more than one job') as an indication of employment status in general over the last year. Neither could we use 'how many hours do you usually work each week in paid employment' due to it not being a required question for the 33.15% of people who said they had not worked in the last week.

We would have liked to use variables describing reasons around 'why/why not able to visit the doctor', but we were unsure of the appropriateness of doing this, given the time frame of the question ('in the last twelve months, has there been any time when you needed to see a GP or family doctor … but didn't..?' and 'the last time this happened what was the reason?') did not match up to the time frame being used in the simulation, that is, modelling each 2 week period.

5.2.4 Variables in the base file used for simulation

From the synthetic base file, the following variables were used in the simulation: age group, household type, number of GP visits in the last 12 months, gender, ethnicity, and deprivation of the potential patient; and gender, age group, ethnicity and workload of the doctor; and practice location and number of doctors in the practice. Survey weight variables both for baseline (2002/3) and for a 2021 projection were also used.

5.2.5 Definition of variables in the base file

Age group: 0-24, 25-44, 45-64, and 65+.
Gender: male, female.
Ethnicity: Maori, Pacific, Asian, Other, and European.
Household type:
1 = do not live with adult (person aged 15 or over)
2 = live with adult partner (husband/wife or de facto, boyfriend or girlfriend)
3 = live with adult but not partnered.
Recent illness is defined as a condition occurring in the last 2 weeks including both short-term and/or flare-ups of long-term ones. Conditions were classified according to 17 broad categories:

1. Infectious and parasitic diseases
2. Neoplasms
3. Endocrine/nutritional/metabolic/immunity disorders
4. Diseases of blood and blood forming organs
5. Mental disorders

6. Nervous system/sense organ diseases
7. Cardiovascular/Circulatory diseases
8. Respiratory system diseases
9. Digestive system diseases
10. Genitourinary system diseases
11. Complications of pregnancy/childbirth/puerperium
12. Skin and subcutaneous tissue diseases
13. Musculoskeletal and connective tissue diseases
14. Congenital anomalies
15. Symptoms, signs and ill-defined conditions, and disability not elsewhere classified
16. Injury and poisoning
17. Not an illness, non-symptomatic, and 'not stated'.

Practitioner: age, gender, and ethnicity.
Practice: type, location, and size.

5.2.6 Definition of imputed variables

Most important condition (MIC): the condition category, chosen from the conditions present in a fortnight (that were allocated as likely to be seen by a GP) for a person, deemed to be the most important in predicting how many doctor visits that person will have for the fortnight.
Primary diagnosis (PD): the condition category deemed to be the main reason, out of all the conditions a person has in a fortnight (that were allocated as likely to be seen by a GP), for any given visit.
GP clinical activity: outcome related to a visit (yes/no): investigation, prescription, non-drug treatment, follow up, and referral.

5.3 Random assignment of characteristics (see Appendix 9.2 for further information)

Via a Monte Carlo process, r**andom numbers** were used throughout the simulation to convert probabilities, whether from tables (Australian data), or from statistical models (NZ data) into characteristics for an individual. A random number from a uniform distribution between 0 and 1 was first assigned. If that random number was less than or equal to the probability then the characteristic was deemed to be present.

A **cumulative distribution function** can be created in order to assign a characteristic where there are probabilities related to multiple categories of outcome. Each of these probabilities can be standardised so that it is converted to a scale from 0 to 1. This can be done by dividing each probability by the sum of all the probabilities to give new probabilities. Consequently a random

uniform number on the same interval (0, 1) can be compared to the new probabilities and so used to assign the characteristic.

The simulated estimates presented in this paper are the average results of 100 runs with a different random seed specified for each run.

5.4 Trialling different simulation processes

Many different processes (or groups of steps that make up a simulation run) were tried iteratively and then compared to benchmarks to assess whether the process created a good representation of reality.

We will describe the methods that were used for the three main options trialled to give an indication of the development of the simulation (see Table 5.2; see section 5.6.6 below for more detail). The final process decided upon was one that, as will be shown, was plausible when validated against external data.

Table 5.2 Simulation process options and their main points of difference

Points of difference	Option		
	Preliminary		Final
	A	B	C
Link between condition and visiting the doctor	yes	no	yes
Distinguish between 'GP' and 'GP/Specialist' visits	yes	yes	no
Good estimate of average GP visits	no	yes	yes

Option A. The first process tried linking whether a person had a visit or not directly to their condition category, but modelled the likelihood of seeing a GP or of seeing a 'GP and Specialist' separately. This resulted in an overestimation of the average number of visits and was abandoned.

Option B. The second process assigned whether a person saw the GP, 'Both the GP and the Specialist' or neither, depending on their demographic profile, and was not linked to their conditions at all. This method produced plausible results, but was deemed unsatisfactory due to the lack of linkage between visiting and the conditions a person had.

Option C. The third and **final process** returned to linking whether a person visited their GP or not directly to the conditions they had, but did so by assuming that even though 'GP and Specialist' type individuals had a higher number of visits, that because they represented such a small percentage of people, treating them as 'GP' type individuals would not be detrimental to the outputs. It looked at each fortnight in the year in turn, and simulated if each person had any of the 17 broad condition categories in each fortnight. Each allocated condition category was then assigned whether or not it was likely to be seen by a GP. If any were deemed likely, then the person was allocated as having at least one visit in that fortnight. If assigned at least one visit, then the actual number of visits for that fortnight was assigned based on an assigned 'most important condition' out of those present and likely to see a doctor. Based on a cumulative distribution function of the probability of appearing as a primary diagnosis (based on NPMCS information), a primary diagnosis was assigned for each visit (each condition category present could be used more than once in the fortnight). Based on this primary diagnosis, GP activities could be assigned for each visit. This resulted in valid results as well as a satisfactory model.

5.5 Data manipulation

This section details further definitions of data items, and then describes specific data tables and manipulation required for inputs to the final simulation process.

5.5.1 Some further definitions: illness conditions and derived household type

In reference to the ANHS data:

- A **recent (either ST or STLT) condition** was any condition reported in the ANHS questions asking about contact with the health system in the last 2 weeks (for example, outpatients, day clinic, hospital, GP), or any other condition in the last 2 weeks that had caused reduced activity, or any conditions reported that had to take medications for (less than 6 months). Whether a condition was labelled as STLT or ST depended on whether the person had an LT condition in the same ANHS specific category – if they did then the recent condition was labelled STLT, otherwise just ST.
- A **long-term (LT) condition**, refers to any condition reported or any conditions reported that had to take medications for (lasting 6 months or more), that was also not able to be recorded as STLT (which only includes recent flare-ups of long-term conditions).

H**ousehold type** was originally defined as:
1=live *without* someone >=15yrs age;
2=live *with* someone >=15yrs age and partnered (husband/wife or de facto, boyfriend or girlfriend);
3=live *with* someone >=15yrs age and *not* partnered, where 'someone >=15yrs age' is the definition of an adult.
We subsequently derived a second version of household type:
1=child,
2=partnered adult,
3=unpartnered adult.

5.5.2 Deciding on sub-groups (predictors) to use for ANHS inputs

In using the ANHS data, we decided to break down the outcomes by age-group, gender, and household type because they were thought to be theoretically influential. Further, age group and household type were crucial to our core scenarios on demographic ageing and social support respectively (see Appendix 9.4). However the patterns were not clear for household type.

5.5.3 Checking of zero cells/profiles in tables

Some input tables were found to have cells with zero numbers, and so ways of collapsing the table into sensible categories were investigated.

Tables were initially produced with age broken into the following categories: 0-4, 5-14, 15-24, 25-34, 35-44, 45-54, 55-64, and 65-74. However, this proved to be too fine grained, as tables with zeros within their profile were produced. Therefore the age groups were further collapsed into the four categories as they currently stand, that is, 0-24, 25-44, 45-64, and 65+. This was collapsed as much as possible, given the emphasis of the model on older people and the need to retain this information.

For the condition prevalence rate tables, as we were unable to reduce the 17 categories to any meaningful new configuration, it was decided that as long as the profiles of people (that is, age group, gender and household type) had numbers present, that we would allow zero cells for condition categories known to be rare or infrequent (that is, those that made up less than 5% of the conditions presented in NPMCS). Thus, 16 of the 17 categories were able to be split up by age group, gender, and household type as required ('complications of pregnancy and/childbirth/puerperium' category was only split by household type as it is not relevant to males or older age groups). The zero cells created are not by any means ideal, but given that we wanted to include age group, gender and household type as sub-groups, and that we were limited to just the use of tabulated data we went ahead with this method. Using a table broken down as

far as possible so as to have no zero cells appearing, and then weighting up or down depending on the remaining sub-groups, was also considered, but it was thought more accurate to use the actual data directly for these three important sub-groups which were to be used in the core scenario testing.

Similarly for the number of visits tables – not all 'number of visits' cells – from 1 up to 10 – had people represented in them for all conditions. When deciding on what breakdowns were possible, all conditions had to have at least one cell with people in it, and this was deemed sufficient. Once it was decided to discard the use of 'type of doctor visit', and to allow different breakdowns for different conditions, as numbers allowed, the number of visits table was able to be broken down by age group, gender and household type depending on condition category.

Also, for the probability 'saw doctor' for each condition category – in deciding what breakdowns were possible, we allowed zero numerators, but not zero denominators.

5.5.4 Ensuring data groupings matched up across statistical models and base file

Much time had to be devoted to ensuring that the categories within variables in the models being built using NPMCS data were the same as available in the base file being produced. This consisted of considering what categories were available in the NZHS for the variables of interest, and what was available for children in the 1996 /7 NZHS. These two datasets that had to have their variables (and therefore the categories that they consisted of) matched to produce the base file.

5.5.5 Imputation of missing data in base file

For the purposes of the simulation it was required that the base file be complete with no missing values on the variables used for prediction. The NZHS and NMPCS variables were each examined; there was little missing data and only deprivation (NZDep) and 'number of visits to doctor in last year' had to have some of their values imputed. We used a mean-based approach where the mean of the variable of interest for individuals in a particular demographic profile was imputed to similar individuals who had missing data. Imputations were made after the data matching process had been completed so as to reduce computational burden and avoid unnecessary complication.

5.5.6 Reallocation of 'non-symptomatic' and 'ill-defined' condition categories

Primary diagnoses for each visit in each fortnight for each person needed to be assigned. As primary diagnosis would be used in turn to assign the likely actions for each visit, we had to address discrepancies in the categorisation (into 17 groups) of conditions between that reported

by the population in the Australian data we were using, and that which the GP would actually apply when diagnosing the conditions, and investigate any impact on the simulation outcomes. Initial validation of the percentage distribution of condition categories seen over the year, showed that, in comparison to what we expected to be the case (given by NPMCS), the 'Symptoms signs and ill-defined conditions and Disability NEC' category (what we call the 'Sx' [Symptoms] category) was being over estimated. Both the 'Sx' category and the 'Not an illness, non-symptomatic' category (what we call the 'NI' [Not illness] category) contained items that may be classifiable elsewhere by a GP as belonging in one of the other more well defined diagnosis categories. On this basis it was decided to try and reallocate the 'Sx' and 'NI' categories to see what improvement, if any, this could make. Appendix 9.3 contains a discussion outlining the investigations carried out on methods of reallocating 'Sx' and 'NI' categories.

5.5.7 Form of input information used in final simulation process (Option C – also see section 5.4 above)

5.5.7.1 Condition rates

Medical conditions were modelled in the simulation via 17 broad categories which were harmonised between ANHS and NPMCS data.

Rates of one or more population-level **recent conditions** for each of the 17 categories (per 1000 population, during fortnights, spread across the year of collection) were derived from table data from the ANHS. We had separate rates of recent conditions depending on whether the condition was labelled as 'ST only' or whether it was 'STLT'. Rates of one or more population-level **long-term conditions** for each of the 17 categories were also derived from the ANHS. The 'LT' category includes long-term *only* conditions (that is, those *without* a recent flare-up/maintenance of the long-term condition), but excludes 'STLT' conditions. Rates of 'ST only', 'STLT', and 'LT' conditions were also broken down by age group, gender, and household type where numbers allowed.

We also wanted to incorporate a measure of deprivation. The Index of Relative Social Deprivation, part of the Australian SEIFA index, is comparable to the New Zealand NZDep index as area measures (the 'collection unit' is equivalent to the 'census area unit'). However, small numbers in the data table prevented breaking down further by SEIFA categories (and aggregating categories would have negated the benefit of using SEIFA). We therefore investigated whether the condition rates (by sub-groups) could be weighted up or down depending on the SEIFA category. Potential weights were calculated by dividing the condition rate for each SEIFA category in turn by the condition rate overall. There was no consistent pattern in the resulting weights across the SEIFA categories, for either 'ST' or 'STLT' conditions, and so it was decided not to adjust the condition rates by these weights.

Seasonality factors for the 17 condition categories were calculated from 2001/2 NPMCS data. The weighted proportion of each category present in each month of the year in NPMCS (including all four possible diagnoses per visit in the tally) was standardised by the average proportion over the twelve months for that category (Table 5.3). In basing the seasonality factors on NPMCS data, the only information available, we were assuming that the seasonality factors for conditions seen by the GP were the same as for population-level conditions as a whole.

Table 5.3 Proportion of all conditions in NPMCS in each month, divided by average proportion for category for year

	Month											
Condition category	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Infectious and parasitic diseases	1.03	1.15	0.91	0.42	0.93	0.89	1.12	1.33	0.80	1.13	1.11	1.18
Neoplasms	1.10	1.14	0.83	1.78	0.64	0.86	0.64	1.08	0.84	1.15	0.87	1.06
Endocrine/nutritional/metabolic/immunity disorders	1.16	0.92	0.98	0.80	1.00	1.25	1.29	1.18	0.76	0.69	1.16	0.82
Diseases of blood and blood forming organs	1.63	1.06	1.31	0.00	1.11	0.53	1.17	0.35	1.99	1.09	0.66	1.11
Mental disorders	1.26	0.97	0.82	0.78	1.15	1.29	0.89	0.96	0.78	0.94	0.77	1.40
Nervous system/sense organ diseases	0.72	1.08	0.98	0.95	1.02	0.91	1.01	1.18	1.09	1.00	1.09	0.97
Cardiovascular/Circulatory diseases	1.12	1.06	0.99	1.22	0.93	0.93	1.10	0.93	0.72	0.89	1.22	0.89
Respiratory system diseases	0.63	0.78	0.91	0.94	1.09	1.26	1.22	1.36	1.40	0.85	0.87	0.71
Digestive system diseases	0.75	0.79	0.96	0.88	1.18	0.98	1.01	0.93	1.29	0.98	1.08	1.18
Genitourinary system diseases	1.11	1.29	0.87	1.14	0.85	1.04	1.10	0.88	0.86	0.90	1.02	0.92
Complications of pregnancy/childbirth/puerperium	0.00	1.16	1.78	0.17	1.31	1.49	2.14	0.04	1.69	0.05	0.48	1.70
Skin and subcutaneous tissue diseases	1.29	1.31	1.25	0.79	0.95	0.68	1.02	0.85	0.94	1.02	0.88	1.01
Musculoskeletal and connective tissue diseases	1.03	1.05	0.99	0.97	0.90	0.75	0.79	0.97	0.94	1.37	1.27	0.96
Congenital anomalies	1.11	1.58	0.00	2.36	0.29	1.86	0.56	0.74	0.96	0.42	0.51	1.61
Symptoms, signs and ill-defined conditions & Disability NEC	0.76	0.82	1.19	0.84	1.18	0.94	0.65	0.87	1.17	1.05	1.10	1.41
Injury and poisoning	1.07	0.87	1.08	0.90	0.95	1.11	1.16	0.91	1.04	1.01	0.96	0.94
Not an Illness, non-symptomatic	1.15	1.00	1.07	1.24	1.01	0.94	0.79	0.78	0.90	1.11	0.92	1.09

5.5.7.2 Summary of input items

1. Probability of a condition category being likely to be seen by a GP

We estimated the probability seeing a doctor/ specialist in that same fortnight for each condition category of a person, who had that condition category in a fortnight. The estimation was done by using the ANHS to calculate the proportion of people with each condition category of interest, who said that their last visit to a doctor/ specialist in the last 2 weeks was for that same condition category (among perhaps other things). Each person in the ANHS was asked how many times in the last 2 weeks they had consulted a doctor or specialist, and for the last such visit, what

conditions it was for. The visits reported could be classified in the survey as being: 'visit(s) to GP only', 'a combination of GP and Specialist visits' or 'visit(s) to Specialists only'.

We linked the chance of having at least one visit with the particular conditions that a person was assigned. The chance of people with each condition being seen by 'definitely a GP' in the last fortnight, and the chance for it being seen by a 'GP or Specialist – unsure' were combined into one probability. Separate probabilities were again given by age group, gender and household type where numbers allowed. There was no splitting up of the chance of visiting by ST or STLT conditions.

2. Probability of being the 'Most important condition leading to a visit' (MIC)

The 'most important condition leading to a visit' was initially based on a cumulative distribution function of the likelihoods (from ANHS data) of having a recent condition (that is, ST rate + STLT rate).

However, the distribution of the first listed reason (that is, one of the condition categories) for a visit in the ANHS was deemed a more appropriate method of assigning the MIC rather than using likelihoods as above. This was to enable a better link between the assigned MIC and the number of visits distribution as these were linked in the structure of the ANHS. The 'first listed reason' was identified if present otherwise 'check-up/investigation' if ticked was used. The distribution was calculated just for those who had a doctor's visit in the survey.

The distributions for 'definitely GP' and 'either GP or Specialist – unsure' types of visits were combined. This was to align with the allocation of the number of visits (partly predicted by the MIC) which ultimately was not based on visit type.

3. Probability of possible number of doctor visits

The distribution of the number of 'definitely GP' and 'either GP or Specialist – unsure' visits (that is, combined) was derived for those who visited the doctor (ranging from 1 to 10 visits), depending on the MIC. This information also came from the ANHS, based on the distribution of the different number of visits reported for people who had each of the 17 broad categories as their MIC (assumed to be the first listed condition). The distribution could not be split up by all sub-groups across the board for all conditions as this produced profiles or cells with no visits. Thus the distribution was split up only by age group, gender and household type as numbers permitted (so that each profile was represented by at least one person).

4. Probability of a primary diagnosis

The probability that each condition category (out of all possible ones in the fortnight) was a primary diagnosis (assumed to be the first listed diagnosis in NPMCS) was calculated from models of 2001/2 NPMCS GP visits.

5. Probability of a doctor action

The probability for each of five possible doctor actions, based on the assigned primary diagnosis for the visit, was calculated from models of 2001/2 NPMCS GP visits.

5.6 Model implementation

This section details the methods used in implementing the simulation, their development and culmination in the final process adopted (Option C) that generated plausible results.

5.6.1 Modelling each fortnight

Every fortnight in the year 2002 was noted in the data set by recording its end date (that is, the first fortnight's end date would be 14th January 2002). From this, the month was derived to indicate the relevant seasonality adjustment. Each fortnight in the year (2002-2003) was then simulated in turn.

5.6.2 Recent condition allocation

Recent conditions in general were modelled using the 'ST' and 'STLT' rates by age group, gender, and household type. We continued to simulate these two types of recent conditions separately to retain flexibility so that, if we wanted to later include long-term conditions in the model, we would easily be able to link these to the recent conditions. The seasonality factor for the month related to the end of the fortnight was applied to the crude rates of recent condition categories. According to these adjusted rates, each person was assigned whether they had any of the 17 recent condition categories for each fortnight throughout the year for both 'ST' and 'STLT' conditions. Recent conditions were imputed by using each calculated rate as if it was a probability and comparing it to a random number from a uniform distribution on the interval (0, 1). If the random number for a person was less than or equal to the calculated rate then that person was assigned that they had that particular 'ST' or 'STLT' recent condition category in that fortnight (otherwise they were assigned that they had not). Having been assigned, 'ST' and 'STLT' categories were combined into one array (or list) of recent conditions for each fortnight.

5.6.3 The simulation process

We trialled three options (A, B, C) before deciding on option C as the final simulation process. Each component of the process is described within the various options (also see section 5.4 above).

5.6.3.1 Option A (preliminary)

1. Modelling seeing the doctor

Condition categories were assigned to be present or absent for each individual in each fortnight, and if so they were then assigned whether or not they were likely to be seen by 'definitely only a GP' or by 'both a GP and a specialist' (termed 'doctor type') or likely not to be seen by any doctor at all. This was an attempt to link the particular condition category to the chance of having at least one visit (*Lymer, Brown, Harding, 2008).*

First we created a cumulative distribution function (CDF) of the mutually exclusive probabilities of being seen by 'definitely only a GP' and by 'both a GP and a specialist' for the condition category in question for each fortnight. Each probability was based on age group, gender, and household type, and did not distinguish between whether the condition was an 'ST' or an 'STLT'. The Australian data showed that the 'doctor type' probabilities for 'ST' conditions (Both=0.147, GP=0.011) were not much different to that for 'STLT' conditions (Both=0.105, GP=0.018). This CDF was then compared to a random draw from a uniform distribution on the interval (0, 1). The condition was assigned to be likely to be seen by 'definitely a GP only' if the random number was less than or equal to the probability. And to be likely to be seen by 'both a GP and a specialist' if the random number lay between the probability of being seen by 'definitely only a GP' and the sum of the two probabilities together. Otherwise, if the random number was greater than this summation, the condition was assigned to be not likely to be seen by a doctor. If any conditions were likely to be seen by either 'doctor type' then that person was allocated as having at least one GP visit in that fortnight. If any conditions were allocated as 'saw GP' or 'saw GP or Specialist', one of this bundle was chosen as 'the most important condition leading to a visit' (MIC), and the numbers of visits were assigned based on 'doctor type' and this MIC.

This simulation process tended to overestimate the average number of doctor visits in a year. It was suspected that modelling the link between having a specific condition category and 'doctor type' of visit was not reproducing the correct mix of visits (that is, number of 'GP' vs. 'GP and specialist' vs. 'no visit'). In actuality, each person should have the same 'doctor type' for each fortnight for all their conditions; it should not be allowed to be different depending on the condition. Also the chance of having a visit (that is, not having doctor type = 'none') was

artificially increased as each and every condition was given a chance rather than being considered as a one-off event applied to the person. The average 'doctor type' distribution for a fortnight was: Both=2.29%, GP=21%, None=76.7%, compared to the actual distribution in the ANHS1995: Both=1.8%, GP=19.2%, None=79%. As can be seen, the simulated results not only had less people on average having no visit (type=None) than the Australian data, but it also had a larger percentage of people on average having 'Both' visits. This may be significant as the mean number of visits for type=Both was 2.56 (which would translate to an estimate of 0.824*2.56= 2.11 GP visits) whereas the mean number of visits for type=GP was only 1.23.

2. Modelling the 'most important condition leading to a visit' (MIC)

For those people who were allocated at least one doctor visit, a 'most important condition category leading to a visit' (MIC) had to be designated for each array of conditions in each fortnight for each person. This MIC was the basis for allocating the number of visits in each fortnight. If there was only one condition present, then obviously that condition was designated as the MIC. In the case of more than one condition, the MIC was allocated based on a cumulative distribution function of the probability of each condition in the array being the MIC (*Lymer, Brown & Harding 2008).*

The MIC was chosen from the bundle of conditions that had previously been assigned as likely to have been seen by the doctor. The MIC was allocated according to a cumulative distribution function of the likelihoods of a person having each of the conditions in the array. Each likelihood was calculated as being equal to the 'ST' + 'STLT' rates depending on age and gender. The rationale was that the rate of a condition experienced by people in the population may reflect how important each condition was in terms of it leading to a doctor's visit.

3. Modelling the number of GP visits

The allocation of the number of visits in each fortnight was based on ANHS tables of distributions of the number of 'GP only' and 'GP and specialist' visits respectively in the last 2 weeks by the first listed reason for the last visit in the last 2 weeks.

If the assigned 'most important condition' (MIC) had a 'doctor type' of 'definitely GP' then the table of 'GP only' visits was used; otherwise if the 'doctor type' was 'GP or specialist – not sure' then the table of 'GP and specialist' visits was used. If 'ST' and 'STLT' components of the designated MIC had both been assigned then the STLT component was checked to see if it had been assigned as likely to be seen by the doctor. If so, the doctor type related to the STLT was used.

If the doctor type for the person was 'GP only' then the number of GP visits was that assigned directly from the distribution in the table. However, if the doctor type was 'both GP and Specialist' then the number of GP visits was based on the tabled value multiplied by the ratio of 'GP' to 'GP and Specialist' visits estimated by Australian Medicare data to be 0.824 *(Lymer et al, undated)* and rounded to the nearest visit.

An alternative estimate of the proportion of 'GP' to 'GP plus Specialist' visits was also considered. Using NPMCS data, this was the ratio of the number of visits with a referral to the total number of visits, assuming that to be able to see a specialist a person would first have to be referred by a GP. This proved unhelpful as it gave rise to an unrealistic number of GP visits.

4. Modelling primary diagnosis (same for Options A, B and C)

Discrepancies between categories assigned by the simulation model and those likely to be diagnosed by the GP were first resolved by reallocating the miscellaneous 'Symptoms' category. A primary diagnosis could then be assigned to each visit in each fortnight using the probability that each of the 17 condition categories was the primary diagnosis in the NPMCS dataset (assumed to be the first listed condition category on the NPMCS survey form). In earlier versions, these probabilities were exactly the same for all condition categories and record profiles (as defined by predictors in the model for primary diagnosis). For the final version, these probabilities were allowed to vary within proscribed limits in an attempt to be more realistic and increase variability in outcome. For each condition category (by individual and fortnight), we assigned a random number (RN) on the interval U (0, 1). The probability for that individual in that fortnight of that condition being the primary diagnosis was then allocated as the ((upper limit - lower limit)*RN + (lower limit)). This gave a random probability within the designated upper and lower limits.

A cumulative distribution function (CDF) of the probabilities of being a primary diagnosis was then formed. This was based on the particular array of conditions that a person had been allocated in a fortnight; the final version made a further restriction to just those conditions likely to have been seen by the GP. Again, this CDF was compared to a random number on the interval U(0, 1) in a stepwise fashion (as for assigning the most important condition). The probability was 0 for a condition category if the person did not have that condition category as being likely to have been seen by the doctor in that fortnight. The same CDF was used repeatedly for each person and fortnight until all visits for the fortnight had been assigned a primary diagnosis. The procedure initially assigned 10 primary diagnoses, one for each of a possible maximum of 10 visits, and then deleted unnecessary primary diagnoses over and above the actual number of visits assigned for that person in that fortnight.

5. Modelling doctor actions (same for Options A, B and C)

Finally given the primary diagnosis, the probabilities of doctor actions - an investigation, a drug treatment, a non-drug treatment, a follow-up, or a referral - were calculated for each visit. In earlier versions, there was just one possible probability for each primary diagnosis, and each record profile (depending on predictors in the model for the doctor action). The final version allowed each of these probabilities to vary within proscribed limits for each primary diagnosis and doctor action. The random probability for a person-visit of receiving a particular doctor action for a particular primary diagnosis was assigned as being the ((upper limit - lower limit)*RN + (lower limit)) where RN was a random number on the interval U(0, 1).

Actual doctor actions were estimated by comparing each of the assigned probabilities with a random uniform number on the interval (0, 1). Each visit was then recorded as the doctor having given the patient an investigation (or not), a drug treatment (or not), a non-drug treatment (or not), a follow-up (or not), and a referral (or not),

5.6.3.2 Option B (preliminary)

1. Modelling seeing the doctor

Visiting by 'doctor type' was assigned to people with at least one recent condition regardless of the particular condition category. It was assumed that the common factor - both condition and doctor type were assigned according to the person's age group, gender and household type – would preserve the link between the condition category and whether they visited the doctor. In other words, the person was assigned whether they had at least one doctor visit (rather than assigning a doctor type for each condition) and then all their conditions were searched to see if any were likely to be seen by the doctor.

2. Modelling the 'most important condition leading to a visit' (MIC)

To better approximate reality, we based the MIC likelihoods on the distribution of the first listed reason (that is, one of the condition categories) for a doctor's visit as recorded in the ANHS data. We used separate distributions by 'doctor type', that is, for 'GP only' and 'GP and Specialist' users respectively, for the last fortnight. The MIC was taken from the entire bundle of conditions as there was no assigning of which conditions were likely to be seen by the doctor; whether someone was seen by a doctor was not linked to what conditions they had.

3. Modelling the number of GP visits

A ‘doctor type’ was assigned for each person with any condition(s) rather than for each condition and then taking the type for the fortnight to be the one associated with the ‘most important condition leading to a visit’ (MIC). If the doctor type was ‘None’ for the person then no visits were assigned; if they had been assigned ‘GP only’ then they were assigned a certain number of ‘definitely GP’ visits; and if they were assigned ‘GP and Specialist’ then they were assigned a certain number of ‘GP and specialist’ visits. The number of visits was based on the distribution of the relevant visits by doctor type for the ‘first listed reason’ (ANHS data) that was the same category as the assigned MIC. This was done by first deriving probabilities of having 1 visit or 2 or 3 visits up to a possible 10 visits. A cumulative distribution function of these probabilities was then created and compared in a stepwise fashion to a random number from a uniform distribution on the interval (0, 1).

4. Modelling primary diagnosis

This was modelled in the same way as for Options A and C.

5. Modelling doctor actions

This was modelled in the same way as for Option A and C.

5.6.3.3 Option C (Final process)

1. Modelling seeing the doctor

‘Doctor type’ had been initially taken into account because of concerns that differences in visit rates between ‘GP and Specialist’ and ‘GP only’ types would impact on the ability of the simulation to hit the benchmark number of visits. However, as only 1.8% of people in the ANHS survey had a ‘doctor type’ of ‘GP and specialist’, we could disregard ‘doctor type’ and instead use combined input data.

We returned to linking the chance of having at least one GP visit (regardless of ‘doctor type’) with the particular condition(s) that a person was assigned. If any conditions were likely to be seen by a GP then that person was allocated as having at least one GP visit in that fortnight. One of those ‘likely to be seen by a doctor’ conditions was then chosen as ‘the most important condition leading to a visit’ (MIC), providing the basis on which the number of visits was assigned.

2. Modelling the 'most important condition leading to a visit' (MIC)

As in Option B, the distribution of the first listed reason for a doctor's visit was used to assign the MIC. However 'GP only' and 'GP and Specialist' users were combined into one distribution. Also the MIC was taken just from those conditions previously assigned as likely to be seen by the doctor.

3. Modelling the number of GP visits

The number of visits was assigned in a similar way to Option B except that doctor type was not taken into account and the combined distribution of number of visits for 'GP only' and 'GP and Specialist' (from ANHS) was used. The resulting visits were all assumed to be GP visits.

4. Modelling primary diagnosis

This was modelled in the same way as for Options A and B.

5. Modelling doctor actions

This was modelled in the same way as for Options A and B.

5.6.3.4 Use of 'number of diagnoses' in modelling doctor actions

Two doctor actions, prescription and non-drug treatment, were being underestimated presumably due to their being more tightly linked to the specific diagnosis category rather than just to the primary diagnosis. To increase accuracy, it was decided to use two separate models depending on the 'number of diagnoses', either single or multiple. Thus for visits with a single diagnosis, there would be a direct link between the only diagnosis and the chance of a prescription/non-drug action. In order to make use of these two models, an indicator variable for whether a visit was for a single diagnosis or multiple diagnoses was created.

The raw simulated distribution of the number of distinct condition categories (see above) did not match the expected distribution (that is, the distribution in NPMCS). However this match improved if we assumed that on average one imputed condition category in the fortnight was not associated with the GP visit (that is, it was taken to another type of health professional or not at all). Accordingly, we decided it was more appropriate to use a modified 'number of diagnoses' indicator variable, that is, if the simulated number of distinct condition categories was >=3 (rather than the more obvious >=2) then the indicator variable was assigned as 'multiple' and otherwise 'single'.

The use of dual models (one for single-diagnosis and one for multiple-diagnoses visits) for prescription and non-drug actions respectively resulted in improved estimates.

We also tried and abandoned using the 'single/multiple' diagnosis indicator (where significant) as a predictor in all doctor action models. It did not result in much shift in predictivity of the prescription model, and more importantly it resulted in an extremely large error for the referral outcome (compared to NPMCS data).

5.6.4 Multiple runs to estimate mean and variance of outcome

In the final process (Option C), 100 separate runs of the simulation were undertaken using a different seed for random number generation in each run. This allowed 95% confidence intervals to be estimated around a mean outcome. This accounted for stochastic variation due to the simulation process but not other sources of variation related to the data (for example, sampling error) or our predictive models (for example, statistical precision). The confidence intervals were all very narrow (possibly but unlikely due to auto-correlated random number streams).

5.7 Verification

Verification, or quality assurance, involved internal checking of the SAS code throughout the simulation program to ensure that variable creation and key components and steps were working properly, and results were being produced as expected.

The process for verification included:

1. Checking syntax for coding errors throughout the program.

2. Checking examples of each type of variable being created during the simulation process to ensure they were being produced as expected.

3. Checking the output of key complex steps in the simulation process to ensure these components were working as expected for that step. In particular, some of the more complicated imputations were investigated and found to be working as required, for example, conditions, 'most important' illness category, number of visits, primary diagnosis for each visit (equal numbers), actions depending on primary diagnosis were tested.

4. Comparing the distributions resulting from the simulation process at each step with the distributions in the input data from the ANHS. This provided a means of identifying where any errors in the program might be, and also gave reassurance that the random number generator was working as expected.

5. Simulated results for key variables, such as 'recent condition' occurrence rates and 'primary diagnosis' distribution, were found to be similar to their ANHS 1995 and NPMCS 2001/2 survey benchmarks respectively as would be expected.

The following list provides an indication of the output considered to verify the microsimulation model (see Appendix 9.5):

1. ST and STLT rates of conditions
2. Average percentage with >=1 condition in a fortnight
    a. overall
    b. by age group
    c. by household type
    d. by gender
3. Most important condition distribution
4. Number of visits distribution
5. Average percentage with >=1 visit in fortnight
    a. overall
    b. by age group
    c. by household type
    d. by gender
6. Average percentage with no visit in fortnight (for those with a condition)
    a. overall
    b. by age group
    c. by household type
    d. by gender
7. Average percentage with no visit in fortnight (for all people)
    a. overall
    b. by age group
    c. by household type
    d. by gender

5.8 Validation

In order to externally validate the simulation model's ability to produce output close to what would happen in the real world, results at each stage were compared to benchmark data from the NPMCS 2001/2. Comparison was made at an aggregate level (see Tables 5.11-5.18), and where appropriate and possible by age group, gender, ethnicity and household type (see Appendix 9.6).

The following list provides an indication of the outputs considered for validation of the microsimulation model:

1. Prevalence of illness (for GP-users, and for population)
   a. by broad condition category (e.g. percentage of all conditions that are 'respiratory')
   b. by age-group
   c. by gender
   d. by household type
2. Average number of visits to GP per annum (for GP-users, and for population) and percentage with >= 1 visit per year
   a. overall
   b. by age-group
   c. by gender
   d. by household type
3. Primary diagnosis distribution
   a. overall
   b. by age-group
   c. by gender
4. Distribution of doctor actions
   a. percentage of visits where a test given
   b. percentage of visits where a prescription given
   c. percentage of visits where a non-drug treatment given
   d. percentage of visits where a follow-up given
   e. percentage of visits where a referral given

   Each action was considered by the following categories
   a. overall
   b. by primary diagnosis of visit
   c. by age-group
   d. by gender
   e. by ethnicity

5.8.1 Validation - aggregate

5.8.1.1 Validation – aggregate: Visits

When comparing the average number of GP visits in a year to benchmarks, the simulation model was overestimating for both GP users (patients) and for the population respectively (Tables 5.4 and 5.5).

Table 5.4: Mean (95% confidence limits) number of visits per year for persons visiting the doctor (over 100 runs)

GP Survey (NPMCS)	Simulation	Absolute error
6.6	7.5 (7.4, 7.5)	0.9

Table 5.5: Mean (95% confidence limits) number of visits per year for whole population (over 100 runs)

NZ Target	Simulation	Absolute error
5.5	7.4 (7.4, 7.4)	1.9

After verifying that the simulation process was working as expected, we identified two possible reasons for this over-estimation of the number of visits.

1. There were differences in rates of visiting a doctor between the Australian and New Zealand contexts (see Table 5.6 below).

2. Modelling the outcome for a period of one year based on 2-weekly input data created too many people in the population with at least one visit in a year (see Table 5.6). In order to model a year's worth of conditions and GP activity using the 2-week snapshot data available, we had to assume (necessarily but incorrectly) that each of the 26 fortnights was independent of the others. Although the ANHS was administered throughout the whole year, and so the 2-week information should have been representative of a whole year, we would have expected that people who visited the doctor regularly would have been underestimated by these data, and people that visited seldom would have been overestimated. It appears that these effects did not cancel each other out, and that seldom users must constitute the majority of the population.

In both instances, we needed to adjust the average number of GP visits per year.

Table 5.6 Percentage of people with >=1 GP visit per year

	NZHS 2002-3 – over year	Simulation 1st run - over year	Base file*
Percentage with >=1 visit	80.8	99.3	80.7

* Adults and children

5.8.1.2 Alignment of the average number of GP visits per person per year

1. Estimating average number of GP visits per year for Australian GP users

We estimated the expected number of visits per year for GP users based on ANHS (1995) data. The number of visits in the last 2 weeks for a GP and/or Specialist had to be multiplied by 26 to obtain the result for a year, and then adjusted to relate only to the GP. The resulting estimate was 7.2 visits per year for Australian GP users. This estimate and other targets are shown in Table 5.7 below.

Table 5.7 Number of visits to GP per annum: comparing Australia & New Zealand

	Population	GP users
Australia:		
Best target estimate	6.0	7.2 (ANHS 1995)
Simulated estimate (raw)	7.4 *	7.5 *
New Zealand:		
Best target estimate	5.5	6.6 (NPMCS 2001)
Simulated estimate (aligned)	5.3	6.7

* Mean of 100 runs

2. Adjusting for use of Australian (AUS) data to model New Zealand (NZ) GP users

Our best estimate of the average number of visits per year for GP users in AUS was about 7.2 (using ANHS 1995). Our PCASO simulated estimate of 7.5 compared well to that target (7.2). However, our best estimate of the average number of visits for GP users in NZ was 6.6 (based on NPMCS). This indicates that Australians in 1995 had a higher rate of visiting their GP than New Zealanders in 2002/3. Taking these two targets (6.6 for NZ, 7.2 for AUS) enabled us to calculate an NZ-AUS adjustment factor of 6.6/7.2 = 0.9 (to 1 decimal place). Therefore, the adjusted simulated average number of visits per year for NZ GP users = 7.4544 * 0.9 = 6.7 (to 1 decimal place), which was only about 0.1 visit off the NZ target of 6.6 (see Table 5.7).

3. Adjusting for use of 2-weekly data

Over and above the effect of using Australian rather than NZ data, having to assume that each fortnight in the year was independent of all other fortnights resulted in a simulated overestimate of the average number of GP visits in the population. We were not able to model the dependence between fortnights, for example, someone visiting in a particular fortnight would decrease their chance of visiting in the next fortnight). Our best estimate for the average number of visits per year for the Australian population was 6.0 compared to our simulated estimate of 7.4 (Table 5.7). We had assumed incorrectly that the overestimation of seldom users in the Australian 2-week

data would be negated by the underestimation of regular users. The unbalanced effect of seldom users (79% of people had no visit in the last fortnight) seemed to be causing the discrepancy in the average number of visits.

Our simulated percentage of people with at least one visit in a year was 99.3% versus the true value of 80.7% given in the base file (NZHS adults and children) (see Table 5.6). These figures enabled us to calculate an adjustment factor of 80.7/99.3 = 0.8 (to 1 decimal place) to account for the use of 2-week data. Applying this 2-week data adjustment factor to the simulated average number of visits per year gave an estimate of 7.4109 * 0.8 = 5.92872 or 6.0 (to 1 decimal place.) which was similar to the Australian target.

4. Alignment of the NZ average number of GP visits per year: adjusting for use of (1) Australian, and (2) 2-weekly data

Adjusting for both the use of Australian and 2-weekly data gave a final aligned simulated estimate of 7.4109 * 0.9 *0.8 = 5.3 (to 1 decimal place) which was close to the NZ target of 5.5 visits per year for the population (Table 5.8).

Table 5.8 Mean number of GP visits per year (over 100 runs)

Simulation 1st run - mean visits per year (95% CL)		Aligned simulation - mean visits per year (95% CL)		NZ target - mean visits per year	
Population	**GP users**	**Population**	**GP users**	**Population**	**GP users**
7.4 (7.4, 7.4)	7.5 (7.5, 7.5)	5.3 (5.3, 5.3)	6.7 (6.7, 6.7)	5.5	6.6

In summary, when comparing the NZ average number of GP visits in a year, the model overestimated for GP users (expected 6.6, simulated 7.5) and more so for the NZ population (that is, GP users and non-users) (expected 5.5, simulated 7.4). The reasons for this overestimation of visits were: (1) higher rates of visiting in AUS than in NZ; and (2) based on 2-weekly data multiplied up 26 times assuming independence between fortnights, resulting in too many people having at least one visit in the year. Therefore we aligned the average number of visits in the year such that the simulation produced estimates of 6.7 for GP users and 5.3 for the population.

5.8.1.3 Validation – aggregate: Conditions

It was reasoned that the most comparable set of conditions from the simulation to that of NPMCS was for those people in each fortnight that had at least one GP visit.
The aggregate distribution of condition categories for the simulated results against the New Zealand benchmark (NPMCS diagnoses) are displayed side by side and can be seen to be similar

in the rank order by frequency and in the percentages for particular categories with an average error of only 1.4 (Table 5.9). For example, the most frequent category (respiratory system disease) contributed 16.0% versus 14.8% respectively.

Table 5.9 Morbidity experience of GP users - all condition categories

Condition category	GP Survey (NPMCS)	Simulation –mean (95% CL)*	Absolute error
Infectious and parasitic diseases	4.3	2.6 (2.6, 2.6)	1.7
Neoplasms	2.4	1.2 (1.2, 1.2)	1.2
Endocrine/nutritional/metabolic/immunity disorders	4.1	5.4 (5.4, 5.4)	1.3
Diseases of blood & blood forming organs	0.5	0.7 (0.6, 0.7)	0.2
Mental disorders	5.0	3.0 (3.0, 3.0)	2.0
Nervous system/sense organ diseases	8.2	6.1 (6.1, 6.1)	2.1
Cardiovascular/Circulatory diseases	9.3	9.7 (9.7, 9.7)	0.4
Respiratory system diseases	14.8	16.0 (16.0,16.1)	1.2
Digestive system diseases	4.4	6.8 (6.8, 6.8)	2.4
Genitourinary system diseases	4.6	3.3 (3.3, 3.3)	1.3
Complications of pregnancy/childbirth/puerperium	0.3	0.1 (0.1, 0.1)	0.2
Skin and subcutaneous tissue diseases	6.7	5.8 (5.7, 5.8)	0.9
Musculoskeletal & connective tissue diseases	5.7	9.4 (9.4, 9.4)	3.7
Congenital anomalies	0.2	0.1 (0.1, 0.1)	0.1
Symptoms, signs ill-defined conditions & disab nec	3.5	3.5 (3.5, 3.5)	0.02
Injury & poisoning	7.1	5.0 (4.9, 5.0)	2.1
Not Illness/unspecified	19.0	21.4 (21.4, 21.4)	2.3
Total	**100%**	**100%**	
		Average error	1.4

* 100 simulation runs

The subset of conditions that had been deemed likely to see the GP rather than all conditions is shown in Table 5.10. This does not tie up as well to the NPMCS distribution with an average error of 2.02 (increased from 1.37). This reinforces our decision to consider the entire set of conditions when speaking about conditions presented to the GP.

Table 5.10 Morbidity experience of GP users – conditions deemed likely to be seen by a GP

Condition category	GP Survey (NPMCS)	Simulation 1st run	Absolute error
Infectious and parasitic diseases	4.3	2.8	1.5
Neoplasms	2.4	1.4	1.0
Endocrine/nutritional/metabolic/immunity disorders	4.1	3.2	0.9
Diseases of blood & blood forming organs	0.5	0.3	0.2
Mental disorders	5.0	2.3	2.7
Nervous system/sense organ diseases	8.2	6.2	2.0
Cardiovascular/Circulatory diseases	9.3	7.0	2.3
Respiratory system diseases	14.8	20.8	6.0
Digestive system diseases	4.4	4.3	0.1
Genitourinary system diseases	4.6	3.0	1.6
Complications of pregnancy/childbirth/puerperium	0.3	0.1	0.2
Skin and subcutaneous tissue diseases	6.7	5.4	1.3
Musculoskeletal & connective tissue diseases	5.7	8.5	2.8
Congenital anomalies	0.2	0.1	0.1
Symptoms, signs ill-defined conditions & disab nec	3.5	2.2	1.3
Injury & poisoning	7.1	5.2	1.9
Not Illness/unspecified	19.0	27.4	8.3
Total	**100%**	**100%**	
		Average error	2.0

5.8.1.4 Validation – aggregate: Doctor actions

The simulated levels of different types of GP activity for all visits are outlined below and can be seen to be similar to those for the NPMCS benchmark with an average absolute error of 2.2 (Table 5.11).

Table 5.11 Percentage of visits per year with each type of doctor activity

Doctor activity	GP Survey (NPMCS)	Simulation Percent (95% CL) *	Absolute error
Investigation	24.9	27.8 (27.7, 27.8)	2.9
Prescription	66.2	64.5 (64.4, 64.5)	1.8
Non-drug treatment	62.1	62.6 (62.6, 62.7)	0.6
Follow-up	57.3	60.3 (60.3, 60.4)	3.1
Referral	15.9	18.3 (18.3, 18.4)	2.5
	Average error		2.2

* 100 simulation runs

5.8.1.5 Validation – aggregate: Summary

The main findings from comparing simulated estimates to benchmark data (NPMCS GP survey) from the same year (2002) are summarised in Table 5.12. The simulations are the average results of 100 runs with a different random seed specified for each run. The simulated estimates are very similar to the corresponding benchmarks with an average absolute error of 1.4 for condition categories, 2.1 for GP activity, and 0.1 for GP visits per annum.

Table 5.12 Visit rates and morbidity experience of GP users, and GP activity per year for synthesised data compared with NPMCS data

	Synthesised data 2002	NPMCS 2001/2	Absolute error
Mean number of visits per year	6.7	6.6	0.1
Morbidity: Top 10 condition categories	**Percent of all conditions**		
Respiratory system diseases	16.0	14.8	1.2
Cardiovascular/circulatory diseases	9.7	9.3	0.4
Musculoskeletal and connective tissue diseases	9.4	5.7	3.7
Digestive system diseases	6.8	4.4	2.4
Nervous system/sense organ diseases	6.1	8.2	2.1
Skin and subcutaneous tissue diseases	5.8	6.7	0.9
Endocrine/nutritional/metabolic/immunity disorders	5.4	4.1	1.3
Injury and poisoning	5.0	7.1	2.1
Genitourinary system diseases	3.3	4.6	1.3
Mental disorders	3.0	5.0	2.0
…	…	…	…
Total	**100%**	**100%**	
		Average error	1.4
GP activity	**Percent of visits**		
Investigation	27.8	24.9	2.9
Prescription	64.5	66.2	1.7
Non-drug treatment	62.6	62.1	0.5
Follow-up	60.3	57.3	3.0
Referral	18.3	15.9	2.4
		Average error	2.1

5.9 Possible enhancements

There are six possible enhancements that will be considered further in this next section.

5.9.1 Use of 'Long-Term condition' information

This would apply if it is found that there are different rates of short-term ('ST') conditions among people with and without a long-term condition (defined as at least one 'STLT' or 'LT') (ANHS 1995). There are issues around the construction and application of such a 'long-term condition' variable. The three key issues are:

1. The 'ST' rate could be given separately for those people who had and did not have a long-term condition in the same broad category (1-17) rather than at the specific level. For example, a person who has chronic asthma (an 'LT' or 'STLT' respiratory condition) may be more likely to contract a chest infection (an 'ST' respiratory condition). By-variables would be used if there were sufficient numbers in the resulting profile groups, categories might need to be aggregated, or perhaps multi-tiered tables produced where different by-variables are used for different conditions.

2. The chance of having a recent 'STLT' in a fortnight could only be given to those already assigned as having a long-term condition.

3. The type of doctor visit ('GP', 'GP and Specialist' or 'None') and the chance of each condition being seen by the doctor could be assigned according to whether a person had a long-term condition. This could account for the possibly increased chance of visiting a doctor in the presence of a long-term condition.

4. Which long-term conditions would be included? Is it possible to select certain conditions only? What is the connection of the 'LT' status to the 'STLT' and 'ST' status of conditions? These are questions that need to be addresses for implementation.

We have not implemented a 'long-term condition' variable in the simulation model given it would need to be based on a combination of assigned 'LT' status and recent 2-week 'STLT' status. How would 'long-term condition' status in the simulation be calculated? According to 'STLT' status in the current fortnight being simulated or whether a person ever gets simulated as having an 'STLT' throughout the entire year? We only had available data to try the former but then people would be less likely to have a 'long-term condition' in early fortnights (for example, in January) than later fortnights in the year. As this did not make sense, we decided not to use a 'long-term condition' variable to assign recent conditions or doctor visiting.

Future investigation could determine what percentage of the 'long-term condition' variable was due to recent 'STLT'; if the percentage was small it could be argued that we could just use the less time-dependent 'LT' variable directly. The use of 'LT' as a by-variable for tables might result in small cell numbers. If so and if a pattern did exist by 'LT' status then rates could be weighted up or down accordingly.

5.9.2 Cloning

For accuracy of imputation, records can be cloned, that is, replicates of the same record in the dataset can be created. In their MediSim model, using data with survey weights attached to each record, Abello et al (*2008*) cloned to have a maximum weight of 200 per record when imputing short-term conditions. The maximum weight depended on the rarity of whatever was being imputed. In this case, given that the sum of weights within each sub-group was 10,000, and the prevalence of a short-term condition was 0.02, the maximum weight would be 200 (10,000 x 0.02 = 200). Thus a record with a survey weight of 672.5, given a maximum target weight of 200, would be replicated to 4 records having weights of 200, 200, 200 and 72.5 respectively.

If we wanted to clone, we would first need to decide on an appropriate maximum weight in the dataset, and then clone any records with a weight greater than the value. For any given outcome (for example, short-term conditions), we would need to calculate the average prevalence/probability across the cells created by cross-tabulating the important factors in predicting that outcome (for example, age group, gender, household type). Then, assuming the sum of the weights equalled the population size, we could then take the average prevalence multiplied by the population size as the maximum weight allowed for any one record. This would ensure, on average, that any one record would not prevent variation we would expect in the outcome, that is, if a record represented more than the number of people in the cell in the population (or weighted dataset) that we would expect to have the outcome (according to prevalence) then we would be at risk of incorrectly estimating the outcome.

Due to the complexity of having to derive a maximum weight for every outcome to be imputed, it was decided to trial cloning on an earlier version of the model using an arbitrary maximum weight of 100. Simulated results showed much larger errors in the outcomes when compared to external benchmarks. Cloning also exacerbated the problem posed by some of our input probabilities/rates being based on tabulated data with small cell numbers. Therefore we abandoned cloning at that point. However, cloning on the final model could be further trialled where the maximum weight was calculated based on prevalence rather than being set arbitrarily.

5.9.3 Random number seed usage when performing multiple runs

The seed value specifies where to start drawing randomly from a pre-determined sequence of numbers. To ensure that narrow confidence intervals for our simulated estimates are not due to autocorrelations, we need to consider using a list of seed values that will produce non-overlapping streams of random numbers.

5.9.4 Sensitivity analysis of variables

Preliminary testing showed little change in results. Further investigation of the effects of changing inputs, that is, data and/or parameters, to the simulation model would be desirable. This would test the model's inherent robustness on the one hand and its responsiveness to meaningful change on the other hand.

5.9.5 Testing for difference in outcome for scenarios compared to status quo

Statistical testing of significance for a simulation model is a complex exercise and beyond the scope of this project. Various sources of variation must be taken into account, for example, sampling error, prediction error, and stochastic variation (confidence intervals calculated from multiple runs only attempts to address the last). Bootstrap methods could be considered to address this issue.

5.9.6 Allowing variation around probabilities for doctor actions (derived from multilevel model)

The general form of a multilevel model is:

$y_{tj} = X_t\beta + \gamma_j + R_{tj}$, where $\gamma_j \sim N(0, \tau^2)$ and the residual $R_{tj} \sim N(0,\sigma^2)$.

γ_j is the random group component.

$X_t\beta$ here contains a 'γ_0' term, the average intercept across doctors (β would be γ_0 here, and X_t would be 1) .

For a binary outcome (say y^*_{tj}), distributed ~ as $bin(1, \pi_{tj})$, using a continuous latent variable framework, we can think of there being a y_{tj} such that $y^*_{tj} \begin{cases} 1 \text{ if } y_{tj} > 0 \\ 0 \text{ otherwise} \end{cases}$.

The random effects model for the log odds of the probability, π_{tj}, of y^*_{tj} being equal to 1 is as follows:

$[logit(\pi)_{tj}] = ln\left(\frac{\pi}{1-\pi}\right) = X_t\beta + \gamma_j$ where $\gamma_j \sim N(0, \tau^2)$.

As the logit transform's inverse is the logistic function, we can also write this as:

$$\pi_{ij} = logistic\left(\gamma_0 + X_i\beta + \gamma_j\right) = \frac{\exp\left(X_i\beta + \gamma_j\right)}{1 + \exp\left(X_i\beta + \gamma_j\right)}.$$

Note the lack of a level-1 error term in this model for the log odds; for a binomially distributed success/failure outcome, the variance, $\pi_{ij}(1 - \pi_{ij})$, is already inherent/fixed in the model, given the probability π_{ij} being modelled. However, there is still the level-1 error term R_{ij} in the model for the underlying process y_{ij}, and in order for the model of the log odds to be a logistic regression model, R_{ij} is fixed as a standard logistic variable with mean 0, and scale 1, that is, the variance is $\pi^2/3$ *(Snijders & Bosker, 1999, p 223)*.

We suspect that the effect of the variance due to the R_{ij} term is mediated in the actual use of the probability π_{ij}, that is, when we compare it to a random uniform number to create a yes/no outcome y_{ij}. Only the γ_j term needs to be simulated directly (that is, obtaining a random number from an N(0, τ^2) distribution to be the γ_j) if we want to use all the information available from the model.

The γ_j is a random doctor component that perturbs the intercept up or down for any individual random doctor according to the variation observed in the data among doctors (having taken account of predictors). This random doctor component was not used in the simulation – the expected value across all doctors, that is, $X_i\beta$, was used instead.

Variation around probabilities for doctor actions was implemented in the simulation process as follows (also see Section 4.5):

(a) (i) We used 68% confidence intervals (that is, ± 1 * standard error from a normal distribution) based on the variation of the betas (variation due to the model being derived from sample data) including that due to the beta for the average intercept across doctors, γ_0. We did not take into account the random doctor variation component or the lowest level residual variance (which is inherent in the probability).

$$Upper\ probability\ limit = logitistic\ (y_{ij}) = logistic\ \{X_i\ (\beta + (1 * se_\beta))\}$$

$$Lower\ probability\ limit = logistic\ (y_{ij}) = logistic\{X_i\ (\beta - (1 * se_\beta))\}$$

(ii) We first assigned the probability ('PROB') based on a random draw from a uniform distribution on the interval (lower probability limit, upper probability limit). We then compared this probability to another random draw from a uniform distribution ('RU') on the interval (0, 1). If RU<=PROB then the person was assigned as having the outcome and otherwise not.

(b) For the core scenarios involving doctor low/high intervention levels, the probability of each doctor action was set at the lower/upper 68% confidence limits respectively. This was deemed to be an appropriate way to set arbitrary yet plausible levels of low and high intervention rates. A future enhancement addressing variation around probabilities of doctor actions would be to incorporate the random doctor component in the simulation process.

6. Microsimulation: Application

6.1 Scenario testing

This was carried out through simulating a potential outcome by manipulating a variable of interest while holding other variables constant and observing change to the outcome. Three key scenarios were tested as part of this project.

1. Impact of demographic ageing by forward projection: The model was extrapolated to 2021 by re-weighting the 2002 population via Statistics New Zealand's mid-range projection for 2021 by age, gender and ethnicity, assuming medium birth, mortality and migration rates. This allowed consideration of the question: What if the 2002 population looked demographically like the 2021 population with everything else remaining the same? This represents a purely demographic effect, that is, parameters remain the same, but the numbers at different age levels change. Thus what impact does this have on morbidity experience, the number of GP visits and GP activity?

2. Family & Community Support: There were limited data available for our measure of support so the proxy of 'being partnered or not' - a re-categorisation of household type - has been used. The availability of support is hypothesised to affect the level of an individual's health service use. What if, in 2021, we changed the proportion of individuals who were partnered? What impact does this have on the number of GP visits?

3. Practitioner repertoire: It is hypothesised that a doctor's age - a proxy for training, experience, and practice style - affects their activity level. What if, in 2021, we changed the proportion of older doctors (aged 45 years and over). What impact does this have on GP activity levels?

6.1.1 Scenario testing: Results

6.1.1.1. Projection to 2021: Demographic ageing

We re-weighted the 2002 population via a Statistics New Zealand projection to 2021, by age, gender and ethnicity, assuming medium birth, mortality and migration rates. The simulation model was re-run on the re-weighted data with everything else remaining the same. There were only slight changes, mostly in the expected directions, in all 3 outcomes, that is, in the number of visits (from 6.7 to 6.9 for GP users (patients) and from 5.3 to 5.5 for the population overall), in the distribution of condition categories and in GP activity levels (Tables 6.1 – 6.3).

Table 6.1 Age projection: Visit rates and morbidity experience of GP users, and GP activity per year as for 2002 (synthesised data) and as projected to 2021

	Synthesised data 2002 *	Projection 2021 *	Absolute change
Mean number of visits per year	6.7	6.9	0.2
Morbidity: Top 10 condition categories	**Percent of all conditions**		
Respiratory system diseases	16.0	15.3	0.7
Cardiovascular/circulatory diseases	9.7	10.7	1.0
Musculoskeletal and connective tissue diseases	9.4	9.7	0.3
Digestive system diseases	6.8	6.8	0.1
Nervous system/sense organ diseases	6.1	6.0	0.1
Skin and subcutaneous tissue diseases	5.8	5.6	0.2
Endocrine/nutritional/metabolic/immunity disorders	5.4	5.8	0.4
Injury and poisoning	5.0	4.7	0.2
Genitourinary system diseases	3.3	3.3	0.0
Mental disorders	3.0	2.9	0.0
…	…	…	…
Total	**100%**	**100%**	
		Average change	0.2
GP activity	**Percent of visits**		
Investigation	27.8	28.3	0.5
Prescription	64.5	64.4	0.1
Non-drug treatment	62.6	62.5	0.1
Follow-up	60.3	61.3	1.0
Referral	18.3	18.2	0.2
		Average change	0.4

* The simulations for the synthesised data for 2002 and the projection to 2021 are the average results of 100 runs with a different random seed specified for each run.

Table 6.2 Average number of visits per year for GP users

	Sim 2002	Sim 2021
Number of visits per year mean (95% CL)	6.7 (6.7, 6.7)	6.9 (6.9, 6.9)

Table 6.3 Morbidity experience of GP users & GP activity per year: 2002 vs. 2021

	Simulation 2002	Simulation 2021	
Morbidity: condition category	**Percent of all conditions (95% CL) ***		**Absolute change**
Respiratory system diseases	16.0 (16.0 ,16.1)	15.3 (15.3, 15.3)	0.7
Cardiovascular/circulatory diseases	9.7 (9.7, 9.7)	10.7 (10.7, 10.7)	1.0
Musculoskeletal & connective tissue diseases	9.4 (9.4,9.4)	9.7 (9.6, 9.7)	0.3
Digestive system diseases	6.8 (6.8, 6.8)	6.8 (6.8, 6.9)	0.1
Nervous system/sense organ diseases	6.1 (6.1, 6.1)	6.0 (6.0, 6.0)	0.1
Skin & subcutaneous tissue diseases	5.8 (5.7, 5.8)	5.6 (5.5, 5.6)	0.2
Endocrine/nutritional/metabolic/immunity disorders	5.4 (5.4, 5.4)	5.8 (5.8, 5.8)	0.4
Injury & poisoning	5.0 (4.9, 5.0)	4.7 (4.7, 4.8)	0.2
Genitourinary system diseases	3.3 (3.3, 3.3)	3.3 (3.3, 3.3)	0.05
Mental disorders	3.0 (3.0, 3.0)	2.9 (2.9, 2.9)	0.04
Infectious & parasitic diseases	2.6 (2.6, 2.6)	2.5 (2.5, 2.5)	0.1
Neoplasms	1.2 (1.2, 1.2)	1.3 (1.3, 1.3)	0.1
Diseases of blood & blood forming organs	0.7 (0.6, 0.7)	0.7 (0.7, 0.7)	0.01
Complications of pregnancy/childbirth/puerperium	0.1 (0.1, 0.1)	0.1 (0.1, 0.1)	0.01
Congenital anomalies	0.1 (0.1, 0.1)	0.1 (0.1, 0.1)	0.0
Symptoms, sign, ill-defined conditions, disability nec	3.5 (3.5, 3.5)	3.3 (3.3, 3.3)	0.1
Not Illness, non-symptomatic	21.4 (21.4, 21.4)	21.2 (21.2, 21.3)	0.2
Total	**100%**	**100%**	
		Average change	0.2
GP activity	**Percent of visits**		
Investigation	27.8 (27.7, 27.8)	28.3 (28.3, 28.3)	0.5
Prescription	64.5 (64.4, 64.5)	64.4 (64.4, 64.5)	0.03
Non-drug treatment	62.6 (62.6, 62.7)	62.5 (62.4, 62.5)	0.2
Follow-up	60.3 (60.3, 60.4)	61.3 (61.3, 61.4)	1.0
Referral	18.3 (18.3, 18.4)	18.2 (18.1, 18.2)	0.2
		Average change	0.4

* 100 simulation runs

Although our results showed little relative change proportionally, the contribution by the 65 years and over age-group increased as expected (from 20% in 2002 to 24% in 2021) which translates into not inconsiderable change in absolute numbers of visits (an extra six million visits for the New Zealand population) and thus consequent doctor activities (Table 6.4).

Table 6.4 Contribution to GP visits by age group

					Est. total mean visits	Aligned total mean visits	Population	Est. no. of visits †
2001								
Age group	**0-24**	**25-44**	**45-64**	**65+**				
Distribution (%)	36.0	29.7	22.2	12.1				
Simulated mean visits *	6.0	6.7	8.2	12.2				
Contribution to total mean visits	2.2	2.0	1.8	1.5	7.4	6.7	3737553	20145934
2021								
Age group	**0-24**	**25-44**	**45-64**	**65+**				
Distribution (%)	34.0	26.4	24.5	15.2				
Simulated mean visits *	6.0	6.6	8.2	12.1				
Contribution to total mean visits	2.0	1.8	2.0	1.8	7.6	6.9	4718029	26151857

* 1st run

† Mean x population x % likely to visit (assuming 80.6% of people have at least one visit in a year - based on NZHS 2002/3)

6.1.1.2. Scenario testing: counterfactuals

Family Support

After projecting to 2021, as an index of the influence of the informal sector, the proportion of people who were partnered and unpartnered was changed; again there was a slight decrease in the average number of visits per year for persons visiting the doctor as partnership levels were increased. There was a narrow range of difference, a decrease of 0.05 visits, in the results between the extreme counterfactuals of all adults being partnered versus all adults being unpartnered. This suggests this variable has limited impact on this model outcome (see Table 6.5).

GP Repertoire

After projecting to 2021, we also changed the proportion of older GPs (aged 45 years and over). Counterfactual results were derived by weighting up the desired type of records for the scenario to be 100% of the data (for example, for the '0% Older GPs' column, we weighted up records

which had a GP who was under 45 to be 100% of the data). This was done by giving them a weight equal to their original survey weight multiplied by '1/the original proportion of records with the counterfactual profile'. Records that did not fit the counterfactual profile (for example, in the above example, those with a GP aged 45+) were given a weight equal to zero.

The benchmark NPMCS data show that we would expect increasing the percentage of older GPs should decrease the percentage of visits with a non-drug treatment and referral respectively, and increase the percentage of visits with an investigation, follow up and prescription respectively. The simulated extreme counterfactuals show results in the expected direction, except perhaps for investigation (see Table 6.5).

Table 6.5 Impact of changing levels of community support (via percent adults partnered), and GP repertoire (via percent older GPs): 2021

			Sim 2021	Sim 2021: Extreme counterfactuals	
	Sim 2002		Adults partnered = 61% of all adults	What if 0% adults partnered?	What if 100% adults partnered?
Average no. of visits per year for GP users – mean (95% CL)*	6.7 (6.7, 6.7)		6.9 (6.9, 6.9)	6.9 (6.9, 7.0)	6.9 (6.9, 6.9)
	NPMCS 2001/2				
	Younger GPs	Older GPs	Older GPs = 48% of all records	What if 0% older GPs?	What if 100% older GPs?
GP activity – mean (95% CL)*	Percentage of visits with each GP activity (%)				
Investigation	23.8	26.1	28.3 (28.2, 28.3)	28.5 (28.4, 28.5)	28.0 (28.0, 28.1)
Prescription	64.5	68.2	64.4 (64.4, 64.5)	63.9 (63.8, 64.0)	65.0 (64.9, 65.1)
Non-drug treatment	63.7	60.2	62.5 (62.4, 62.5)	66.4 (66.4, 66.5)	58.3 (58.2, 58.4)
Follow-up	55.5	59.2	61.3 (61.3, 61.4)	60.7 (60.7, 60.8)	62.0 (61.9, 62.1)
Referral	17.0	14.6	18.2 (18.1, 18.2)	20.1 (20.0, 20.1)	16.2 (16.2, 16.3)

* 100 simulation runs

Older people: worst case scenarios

We tested the impacts of worst case scenarios for the 65 years and over age group on average number of visits per year, and referral and prescribing rates (Tables 6.6 and 6.7). The 'worst morbidity experience' was defined as above the median number of visits per annum, 'living alone' indicated the worst social support, and a 'most interventionist' GP was defined as being at

or above the median percentage of the doctor activity in question. Further, the intervention scenarios were produced by both weighting up records which had a GP ranked as being a 'most interventionist' GP, and by re-running the simulation with these high ranking GPs being arbitrarily assigned the 68% upper confidence limit (+1 standard error) for the probability of that activity (usually the probability was allowed to vary for each visit).

Tables 6.6 and 6.7 show the results of posing three extreme scenarios and their impact on the average number of GP visits per year, and referral and prescribing rates. Under the first scenario of worst morbidity experience: the number of visits per year increased by a quarter (from 12.1 to 15.2), the referral rate remained virtually unchanged (17.8% to 17.9%), and the prescribing rate increased slightly (67.0% to 67.5%). Under the second scenario where social support was hypothesised to be completely absent: there was virtually no change in the number of visits (12.0 to 12.1) and in the prescribing rate (67.0 to 66.8), while the referral rate increased slightly (17.8 to 18.3). Under the third scenario where it was assumed that all GPs behaved like the most interventionist among them: the referral rate almost doubled (17.8 to 32.7) while the prescribing rate increased by nearly a third (67.0 to 87.4).

Table 6.6 Counterfactual analysis of morbidity, social support and GP referral among people aged 65+

	Sim 2002	Sim 2021	Sim 2021: Counter-factual
People aged 65+ with worst morbidity experience (above median no. of visits)			
	51.1%	**50.7%**	**100%**
Mean no. of visits per year for GP users aged 65+	12.1	12.0	15.2
Percent of visits referred on	18.0	17.8	17.9
People aged 65+ living alone			
	30.6%	**30.2%**	**100%**
Mean no. of visits per year for GP users aged 65+	12.1	12.0	12.1
Percent of visits referred on	18.0	17.8	18.3
People aged 65+ seen by the most interventionist GPs (above median for referral)			
	48.3%	**47.9%**	**100%**
	probability allowed to vary		probability set at upper 68% CL
Percent of visits referred on	18.0	17.8	32.7

Table 6.7 Counterfactual analysis of morbidity, social support and GP prescription among people aged 65+

	Sim 2002	Sim 2021	Sim 2021: Counter-factual
People aged 65+ with worst morbidity experience (above median no. of visits)			
	51.1%	**50.7%**	**100%**
Mean no. of visits per year for GP users aged 65+	12.1	12.0	15.2
Percent of visits prescribed	67.1	67.0	67.5
	People aged 65+ living alone		
	30.6%	**30.2%**	**100%**
Mean no. of visits per year for GP users aged 65+	12.0	12.0	12.1
Percent of visits prescribed	67.1	67.0	66.8
People aged 65+ seen by the most interventionist GPs (above median for prescribing)			
	54.7%	**55.3%**	**100%**
	probability allowed to vary		probability set at upper 68% CL
Percent of visits prescribed	67.1	67.0	87.4

6.2 Scenario mapping

We developed a scenario map to address the impacts of the core scenarios above by combining the influences of demographic ageing, support, and practitioner repertoire (Figure 6.1). The map covered changes in single factors and factors in combination, and implemented counterfactuals that ranged from optimistic to pessimistic.

Figure 6.1 Scenario mapping

Social support [2]	**Practitioner repertoire** [3]			
	Higher threshold		*Intensification*	
	Morbidity experience [1]		**Morbidity experience**	
	Compress	*Expand*	*Compress*	*Expand*
Autonomous ageing	+ + +	- + +	+ + -	- + -
Service-dependent ageing	+ - +	- - +	+ - -	- - -

1. Below (+) vs. above(-) median number of visits
2. 0% (+) vs. 100% (-) living alone
3. Probability of activity set at:-
    a. lower 68% CL & applied to GP-below-median-rate-visits(+)
    b. upper 68% CL & applied to GP-above-median-rate-visits(-)

6.2.1 Scenario mapping: Results

The following scenario mappings combine the effects of three counterfactual domains: morbidity experience, social support, and practitioner repertoire (Table 6.8-6.10).

Under the first scenario mapping of extremes of social support and morbidity experience: social support had no effect but the average number of GP visits per year doubled for higher morbidity (Table 6.8).

Table 6.8 Mean number of visits per year for GP users aged 65+ in 2021

Social support [2]	**Morbidity experience** [1]	
	Compress (+)	*Expand (-)*
Autonomous aging (+)	8.8	15.3
Service-dependent aging (-)	8.7	15.2

1. 'Compress (+)' signifies that all GP users have below the median number of visits; 'Expand (-)' signifies that all GP users have above the median number of visits.
2. 'Autonomous aging (+)' signifies that no GP users are living alone; 'Service-dependent aging (-)' signifies that all GP users are living alone.

The second and third scenario mappings add the effect of practitioner repertoire (Tables 6.9 and 6.10). Under the second scenario mapping: the prescribing rate for more interventionist GPs were nearly double that of less interventionist GPs - as was the number of visits associated with a prescription - while there was virtually no difference according to lower or higher morbidity (Table 6.9).

Table 6.9 Percentage of visits (average number of visits p.a.) prescribed for GP users aged 65+ in 2021

Social support [2]	Practitioner repertoire [3]			
	Higher threshold (+)		*Intensification (-)*	
	Morbidity experience [1]			
	Compress (+)	*Expand (-)*	*Compress (+)*	*Expand (-)*
Autonomous aging (+)	46.2% (= 4.1 visits p.a.)	47.0% (= 7.2 visits p.a.)	87.0 (7.7)	87.9 (13.4)
Service-dependent aging (-)	46.9 (4.1)	44.4 (6.7)	86.0 (7.5)	87.7 (13.3)

1. 'Compress (+)' signifies that all GP users have below the median number of visits; 'Expand (-)' signifies that all GP users have above the median number of visits.
2. 'Autonomous aging (+)' signifies that no GP users are living alone; 'Service-dependent aging (-)' signifies that all GP users are living alone.
3. 'Higher threshold (+)' signifies probability of practitioner activity set at level below the median rate; 'Intensification (-)' signifies probability of practitioner activity set at level above the median rate.

Finally under the third scenario mapping: the referral rate for more interventionist GPs was about six times that of less interventionist GPs - as was the number of visits associated with a referral (Table 6.10).

Table 6.10 Percentage of visits (average number of visits p.a.) referred for GP users aged 65+ in 2021

Social support [2]	Practitioner repertoire [3]			
	Higher threshold (+)		*Intensification (-)*	
	Morbidity experience [1]			
	Compress (+)	*Expand (-)*	*Compress (+)*	*Expand (-)*
Autonomous ageing (+)	5.5% (= 0.5 visits p.a.)	4.9% (= 0.7 visits p.a.)	32.6 (2.9)	32.4 (5.0)
Service-dependent ageing (-)	5.1 (0.4)	4.6 (0.7)	32.5 (2.8)	33.5 (5.1)

1. 'Compress (+)' signifies that all GP users have below the median number of visits; 'Expand (-)' signifies that all GP users have above the median number of visits.
2. 'Autonomous aging (+)' signifies that no GP users are living alone; 'Service-dependent aging (-)' signifies that all GP users are living alone.
3. 'Higher threshold (+)' signifies probability of practitioner activity set at level below the median rate; 'Intensification (-)' signifies probability of practitioner activity set at level above the median rate.

6.3 Software development

The original simulation programme was implemented in **SAS** *((http://www.sas.com)* though it has been progressively converted to **JAVA** *(http://java.com/en)* and ***R*** *(http://www.R-project.org)* for ease and speed of development and usage. The programme now operates within an open-source simulation package **ASCAPE** *(http://ascape.sourceforge.net)* whose interface we have adapted for our purposes. We call our software tool JAMSIM (**JA**VA for **M**icro **Sim**ulation) which is available at *http://code.google.com/p/jamsim* (*Mannion et al, in press, 2011*).

7. Conclusion

7.1 Data synthesis

The method of statistical matching used is well established in the literature, but as far as we can ascertain, this is the first attempt to allocate GPs from one existing data source to individuals in another without the benefit of unique identifiers. The use of existing data can be valuable for further analysis or modelling that otherwise would not have been possible, and alleviates the need for further expensive data collection. Data synthesis of this kind can bring together diverse data sets and thus extract more value from the constituent sets. However, there are necessarily practical assumptions and decisions made in the process of statistical matching. These tend to be driven by what is available in the data at hand while also trying to maintain rigour.

There were a limited number of common variables between our two data sets to employ for the matching process although this is not unusual in models of this kind. We were also implicitly assuming conditional independence (*Australian Bureau of Statistics 2004)*; that is, that all variation in GP choice can be explained by the matching variables. In justification, the variables we used were the key patient variables as well as being the only common ones. Results demonstrated a plausible relationship between the original and matched variables.

In cell division, the precision of matching depends on the number of common variables used, and if there are too many it is much harder to find exact or even close matches *(Zaidi & Scott 2001)*. With few common variables available in our case, it was relatively easy for two records to be judged as 'identical' in the matching process, whereas more information would have given rise to greater discrepancy. Within cells, nearest-neighbour matching via a distance function was beneficial and only moderately sensitive to the use of different distance metrics.

The data set created through statistical matching is of course synthetic, but diagnostic tests and sensitivity analyses showed the method performed well in our application. Furthermore, the addition of these matched data substantially improved the functioning of the microsimulation model.

This method of statistical matching produced accurate and robust results. We were able to create a synthetic data set that combined information from two disparate sources while preserving expected relationships among variables. This new data set could then be used for the novel purposes, which were hitherto not possible, of performing simulation and undertaking virtual experiments. The innovative aspects of our study were in using the process to match GPs to individuals to enhance the data set underlying a microsimulation model of the primary care process.

7.2 Statistical modelling

From NPMCS, general practitioners ordered an investigation in 25% of patient visits, wrote a prescription in 66%, gave non-drug treatment in 62%, requested follow up in 57%, and referred on in 16% of visits. The final models predicted similar probabilities of clinical activity to the actual with an average absolute error of 4.8 percentage points.

Clinical activity was mainly explained by the primary condition of the patient as diagnosed by the GP. This result was not unexpected as we would assume that the action a doctor takes would primarily depend on the diagnosis made. The patient characteristics were the next set of variables important for predicting the probability of a doctor action. Doctor and practice information contributed the least to the predictivity of the models.

Prior to the modelling process, it was decided to take an approach that balanced predictivity of the models with theoretical interpretability in the medical/sociological context. This approach meant that pure prediction tools were not used and interaction effects and other higher polynomial terms were not incorporated in the statistical models. There was a tension between producing the best predictive model while ensuring that the independent variables used had explanatory value and were theoretically interpretable.

The visit was taken as the unit of analysis so the doctor actions were assumed to be a characteristic of the visit rather than individual diagnoses that could have occurred within the visit. Thus, the link between say getting a prescription and the actual underlying therapeutic need may not always be as direct as would be desirable. The modelling process also did not account for multiple doctor actions per patient.

The variables that could be considered as predictors had to be restricted to and be in the same form as those that were available in the data base that was being used for simulation. The models did not account for severity of illness nor co-morbidities which inevitably reduced their predictive ability.

We were able to develop multilevel logistic regression models using the enhanced features of SAS 9.2 that gave good predictivity when assessed against empirical benchmark data. We selected the best model by minimising an information criterion and compared the output distribution to the benchmark. Note that models fitted using the SAS GLIMMIX procedure (multilevel model for a dichotomous outcome) did not take account of survey stratification.

Our models of clinical activity in NZ general practice are comparable to other models in the area of health care services utilisation *(Cutts et al 2005; Friedman et al 2007; Polen et al 2001; Pope 1988)*.

We were able to build statistical models of GP activity level using appropriate techniques that identified the best explanatory variables and gave good predictivity. In turn, the derived parameters were able to be applied to our simulation model of general practice though it is a limitation that the random doctor effect was not used.

7.3 Microsimulation

The microsimulation model followed a "pathway to care" process imputing in turn health conditions, the occurrence of a doctor visit, the most important condition, the number of doctor visits, the primary diagnosis, and associated GP activities.

7.3.1 Imputing health conditions

We were predominantly interested in modelling the occurrence of recent conditions as they show the immediate need for engagement with primary care services. Broad categories of health conditions were chosen based on earlier work using the ANHS 1995, and modified to ensure compatibility with the NPMCS 2001/2.

Occurrence rates of recent conditions in the population (during fortnights, spread across the year of data collection) were derived from tabulated data using the ANHS 1995. These rates were broken down by the 17 condition categories, age group, gender, and household type. Up-rating of these 1995 rates to 2002 level was considered but ultimately not deemed necessary after successful validation of the model (without up-rating). However, approximate adjustments for seasonality were made for the 17 condition categories by using standardised proportions of each condition category present in each month of the year that pertained in the NPMCS 2001/2 visits data.

Each fortnight in the year (2002) was then simulated in turn. Each person was assigned, using Monte Carlo simulation, whether they had each one of the 17 recent condition categories for each fortnight throughout the year culminating in a list of such condition categories. Each condition category was simulated independently, that is, there was no explicit modelling of the co-occurrence of clinically related condition categories.

7.3.2 Imputing if a doctor visit occurred

We assigned, using Monte Carlo simulation, for each condition that a person had in the fortnight whether that person would be likely to choose to see a doctor. Thus each person was allocated a series of probabilities of visiting the doctor based on imputed conditions. If any of these conditions were allocated as likely, then the person was deemed to have seen their GP for that fortnight. The probabilities were based on condition category, the person's age group, gender and household type. They were derived from the proportion of people with each recent condition category in the ANHS 1995 who had that condition listed as one of the matters they saw their doctor for at their last visit in last 2 weeks.

7.3.3 Imputing the 'Most important condition leading to a visit'

For those people who were allocated at least one doctor visit, the 'most important condition category leading to a visit' had to be designated for each array of conditions in each fortnight. This was for the assignment of the number of visits in each fortnight. If there was only one condition present, then that condition was designated as the 'most important'. In the case of more than one condition category being present, the 'most important' was assigned based on a cumulative distribution function (CDF) of the probability of each condition likely to have been seen by a doctor in the array. The probabilities of the conditions were based on the distribution of the first listed reason for a visit using data from the ANHS 1995. This was to enable a better link between the assigned 'most important condition' and the number of visits distribution both based on the same survey data. A CDF was then derived and used to assign, via Monte Carlo simulation, which condition category for a person was 'the most important condition leading to a visit' taken from the array of conditions likely to have been seen by a doctor.

7.3.4 Imputing number of doctor visits

Data from the ANHS 1995 on the distribution of number of visits (ranging from 1 to 10 visits) by doctor type and depending on the 'most important condition category leading to a visit in the fortnight' was used to assign the number of doctor visits. This distribution was split up by age group, gender and household type where numbers permitted. From this distribution, probabilities were derived of having 1 or 2 or 3 visits up to a possible 10 visits. A cumulative distribution function of these probabilities was then created and used, via Monte Carlo simulation, to assign the number of visits.

7.3.5 Imputing primary diagnoses

Primary diagnoses for each visit in each fortnight for each patient then needed to be assigned. The probability of being a primary diagnosis (assumed to be the first listed diagnosis) was

derived for each condition category (out of all possible ones in the fortnight deemed as likely to have been seen by the doctor). These probabilities were produced via statistical modelling of GP visits using NPMCS 2001/2 data. Given the particular array of conditions likely to see a doctor that a person had been allocated in a fortnight, a cumulative distribution function was firstly made of these probabilities and then used to assign a primary diagnosis. As the primary diagnosis would be used to assign the likely GP actions for each visit, harmonisation was necessary between the categorisation of conditions in the Australian illness data and New Zealand GP visit data. Initial simulated results of the distribution of condition categories seen over the year showed that, in comparison to the NPMCS benchmark, the 'Symptoms ...' category was being over-estimated. This 'Symptoms …' category contains items that could be possibly diagnosed by the GP as being in one of the other more well-defined categories and so it was decided to re-distribute this category accordingly.

7.3.6 Imputing associated GP activity

Given the primary diagnosis, the probabilities of a GP action (investigation, prescription, non-drug treatment, follow-up, or referral) were calculated for each visit via predictions from a series of multilevel logistic regression models using NPMCS 2001/2 data that included both patient (demographics and health status) and doctor characteristics. Each of these probabilities was then used to randomly assign whether each visit was recorded as having been associated with each GP action.

7.3.7 Assumptions of the simulation model

1. We assumed that Australian levels of two-weekly condition occurrence were similar to those in New Zealand.

2. We are assumed that each fortnight in the year was independent of every other fortnight.

3. Any future use of the LT rate would assume that if a person had a long-term condition, defined as lasting six months or more, it would apply for the whole year.

4. We assumed that the distribution of the first listed reason for a visit in the last two weeks in the ANHS survey was a good representation of the most important condition leading to a visit in a fortnight.

5. The probability of each condition category being seen by a doctor was based on data from the last visit in the last two weeks. We assumed that, as 77% of users had just one visit in the last fortnight, that probability gave a good representation of the condition category being seen by the doctor.

6. Only 1.8% of people in the ANHS survey had a 'doctor type' of 'GP and specialist'. We thus assumed that differences in GP visit probabilities derived from combining 'GP only' and 'GP and Specialist' rather than from 'GP only' alone were minimal.

7. Where people in the ANHS survey with 'doctor type' equal to 'GP and specialist' had only one GP visit, we assumed their type had been incorrectly recorded and changed it to 'GP only'.

8. A limitation is that we did not model severity nor co-morbidities as the necessary data were not available.

9. In the matching of doctors, people in the NZHS survey who reported they had not visited the doctor in the last 12 months were assumed to be the same (in terms of which kind of doctor they would likely choose) as those patients in the NPMCS survey who reported they had just one visit to the GP in the last year.

10. NZHS and NPMCS data were matched using non-after-hours visits. The primary diagnosis and doctor action models were also based on non-after-hours data. We assumed that the visits reported in the ANHS survey (information about which we use in the simulation) were also mostly due to non-after-hours visits. This would ensure that profiles of doctors and probabilities of primary diagnosis and doctor actions would be appropriate to use in conjunction with ANHS visit data.

11. The random allocation of doctor profiles for those individuals in the NZHS survey who had zero visits in the last 12 months considered all doctors in NPMCS (that is, not only non-after-hours doctors).

12. Each SAS data step imputed different characteristics in the simulation process (for example, illness in one, and actions in another). We assumed that the results were not affected by having the same seed for the random numbers that drove the imputations in each data step.

7.3.8 Issues re scenario testing

Scenario testing was implemented by re-weighting the population in the base-file according to various individual characteristics. There are underlying assumptions and issues that arise which are listed below:

1. We re-weighted the population according to official projections of demographic composition in 2021 (Statistics NZ - only source of futuristic data).

2. In scenarios, it is assumed that the link between projected demographics and other individual characteristics would remain the same.

3. It assumed values or levels of the other individual characteristics would remain the same (no projected data available).

4. It assumed that ANHS survey data input would remain valid for new people profiles.

5. It assumed the same doctors would be present regardless of changes in each base file record's demographics – otherwise doctors would have to be re-matched to the new people profiles to create a new updated base-file.

6. It assumed that the link between people and patient profiles would remain the same.

7. It assumed that doctor profiles and their link to people profiles would remain the same.

8. It assumed that the relationships between doctor actions and the predictor variables would remain the same, that is, NPMCS prediction models would still hold.

9. It assumed values or levels of the predictors of a doctor action would remain the same (no projected data available on patient or doctor/practice characteristics).

10. We could not reweight doctor nor practice characteristics as the patient was the unit of analysis in the base-file.

11. We could possibly have reweighted NPMCS patient data before matching to NZHS data.

7.3.9 Discussion and Conclusion

The contribution that microsimulation models can make to addressing 'what if?' scenarios and realistic extrapolation into the future is well known *(Gupta and Harding 2007)*. However, a feature of such models that has not been sufficiently emphasised is their potential for drawing together data from different sources. We have attempted to demonstrate this by building a model that can be used to test a range of scenarios on the impact of demographic ageing. Necessarily, the construction of such a model relies heavily on its foundation in empirical data and hence makes reasonably strong assumptions about the plausibility of results arising from their combination.

The major strength of this model lies in its use of existing micro-level data from various relevant sources. There were data available on population occurrence of recent health/illness conditions, GP service utilisation, and GP clinical activity. Statistical matching was employed to create a synthetic base file by combining different data sources. This enhancement enabled a better representation of reality than could be achieved from any one source alone. However, the need to harmonise data sources as to, for example, their intrinsic classification categories and time scales, meant that there was potential information that was lost to the model. Another requirement was that a variable needed to be present in both the external data source (either ANHS or NPMCS) and the base file derived from NZHS. Furthermore, as longitudinal data were not available to derive transition probabilities, the model is a static one, using static ageing methods, so that any extrapolation into the future is hedged with assumptions.

As much as possible, data sources were selected that related to New Zealand and a specific time period circa 2002. The obvious exception on both counts was our use of the ANHS 1995 to obtain recent illness information which was otherwise not available. The rationale, borne out by evidence, was that the two countries shared much social similarity and particularly (pertinent here) in their demographic structure and primary health care system. A limitation of the illness and service utilisation data (from Australia) was that it was based on a 2-week reporting period so that the simulation of multiple fortnights to create a year resulted in over-estimation due to the fact that actual fortnights are not independent of one another. We also felt justified in aligning our simulated results, for example, the number of GP visits, to New Zealand 2002 because of identified and quantifiable differences to Australian data.

There were other limitations of the ANHS tabulated data employed for the simulation of morbidity experience and GP use. Proportions were broken down by age group, gender, and household type where numbers allowed; other ANHS variables could not be used as there were no sufficiently compatible counterparts in our New Zealand base file. Ethnicity was not considered as the ethnic composition of Australia and New Zealand were very different, and so there was no common variable that could be used; however, ethnicity was used in the simulation of GP activity which relied entirely on New Zealand data.

We were also limited by the data available for our measure of family support and have used the proxy of 'being partnered or not' (a re-grouping of household type), the rationale being that adults who are partnered receive more support and care and may not have to visit the doctor as much as the unpartnered *(Ostberg & Lennartsson 2007; Prior & Hayes 2003; Van Houtven & Norton 2004)*. Initial exploration showed that partnership had a moderate effect; however as it was a key component in our scenario testing, it was retained.

We operationalised GP repertoire by using the age of the GP as a proxy for their cohort, and assuming that older GPs may have, for example, a higher prescribing rate because of the training

their cohort received. Underpinning this is the rationale that, once the early pattern has been set, GPs tend to persist in a particular practice style *(Davis et al 2000)*.

We attempted to design the simulation process so that it followed as much as possible the pathway to care. Effectively, a health history was created for each person in our synthetic population. We attempted to reproduce the realistic linkage from one health event to another embodied by the pathway to care, both logically and chronologically. However, this was not always possible depending on the availability and nature of the data. Therefore, the modelling process has necessarily been an iterative one of continuous verification and validation, with incremental progress, and a balance between the criteria of underlying logical sense and producing stable plausible results. We have attempted to generate more robust estimates by averaging the results of 100 simulation runs.

In testing various scenarios by manipulating specific factors of interest, the model must assume that everything else remains the same including inherent structural relationships. This is a limit on the realism that can be achieved. Our provisional results showed little relative change proportionally though that translated to not inconsiderable change in absolute numbers of GP visits and doctor activities. The relative effect of demographic ageing may be moderated by healthier living and improved medical care (*Knickman & Snell 2002; Rechel et al 2009*) though we did not account for these factors explicitly in the model.

A novel microsimulation approach was successfully applied to a synthetic sample of individuals made by combining existing data sources *(Davis et al 2010; Pearson et al 2011).* A working prototype model of primary medical care in New Zealand for 2002 was constructed which produced plausible results for key parameters. Furthermore, the model was able to be used to test the impact of various scenarios involving demographic ageing via a projection to 2021. Model projections suggest limited change in system demand. There is potential to improve, to increase the complexity of, and to extend the model. This will enhance its usefulness as a scenario-testing tool for policy purposes.

8. References

Abello A, Lymer S, Brown L, Harding A, Phillips B. 2008. Enhancing the Australian National Health Survey data for use in a microsimulation model of pharmaceutical drug usage and cost. Journal of Artificial Societies and Social Simulation; 11(3)2. <http://jasss.soc.surrey.ac.uk/11/3/2.html>.

Aboderin I. 2004. Modernisation and ageing theory revisited: Current explanations of recent developing world and historical Western shifts in material family support for older people. Ageing and Society; 24:29-50.

Afgani M, Sinanovic C, Haas, H. 2008. Anomaly detection using the Kullback-Leibler divergence metric. First International Symposium on Applied Sciences on Biomedical and Communication Technologies, pp. 1–5, 25–28.

Alegre J, Arcarons J, Calonge S, Manresa A. 2000. Statistical matching between different datasets: An application to the Spanish Household Survey (EPF90) and the Income Tax File (IRPF90). In: The workshop on fighting poverty and inequality through tax benefit reform: Empirical approaches. Barcelona.

ASCAPE <http://ascape.sourceforge.net>.

Australian Bureau of Statistics. 1996. National Health Survey 1995: Users' guide. Canberra: Australian Bureau of Statistics. <http://www.abs.gov.au>.

Australian Bureau of Statistics. 1997. National Health Survey 1995: Summary results. Canberra: Australian Bureau of Statistics. <http://www.abs.gov.au>.

Australian Bureau of Statistics. 2004. Statistical matching of the HES and NHS: An exploration of issues in the use of unconstrained and constrained approaches in creating a basefile for a microsimulation model of the pharmaceutical benefits scheme. Canberra: Australian Bureau of Statistics.

Brown L, Harding A. 2002. Social modelling and public policy: Application of microsimulation modelling in Australia. Journal of Artificial Societies and Social Simulation; 5(4).

Bryant J, Teasdale A, Tobias M, Cheung J, McHugh M. 2004. Population ageing and government health expenditures in New Zealand, 1951-2001. The Treasury, Wellington.

Cohen M. 1991. Statistical matching and microsimulation models. In: The uses of microsimulation modelling, review and recommendations, vol. 1, eds CF Citro, EA Hanushek. Washington: National Academy Press.

Complex Systems Modellling Group. 2010. Modellling in Healthcare. American Mathematical Society. Providence, Rhode Island, USA.

Davis P, Gribben B, Scott A, Lay-Yee R. 2000. Do physician styles persist over time? Continuities in patterns of clinical decision-making in a stable cohort of general practitioners. Journal of Health Services Research and Policy; 5:200-207.

Davis P, Gribben B, Lay-Yee R, Scott A. 2002. How much variation in clinical activity is there between general practitioners? A multi-level analysis of decision-making in primary care. Journal of Health Services Research & Policy; 7: 202-208.

Davis P, Lay-Yee R, Pearson J. 2010. Using microsimulation to create a synthesised data set and test policy options: the case of health service effects under demographic ageing. Health Policy; 97:267-27. doi:10.1016/j.healthpol.2010.05.014.

Dickey D, Donaghy S, Smith JM. 2003. Tips for Running Simulations in SAS. North Carolina State University, Raleigh, NC.

Edwards KL, Clarke GP. 2009. The design and validation of a spatial microsimulation model of obesogenic environments for children in Leeds, UK: SimObesity. Social Science and Medicine; 69(7): 1127–1134.

Fone D, Hollinghurst S, Temple M, Round A, Lester N, Weightman A, Roberts K, Coyle E, Bevan G, Palmer S. 2003. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. Journal of Public Health Medicine; 25: 325-335.

Gabe J, Calnan M, Bury M. 1991. The sociology of the health service. London: Routledge.

Garces J, Rodenas F, Sanjose V. 2003. Towards a new welfare state: The social sustainability principle and health care strategies. Health Policy; 65:201-215.

Gilbert N, Troitzsch K. 2005. Simulation for the social scientist. Maidenhead: Open University Press.

Goldstein H. 1987. Multilevel models in educational and social research. New York: Oxford University Press.

Graham P, Blakely T, Davis P, Sporle A, Pearce N. 2004. Compression, expansion or dynamic equilibrium? The evolution of health expectancy in New Zealand. Journal of Epidemiology and Community Health; 58: 659-666.

Gupta A, Harding A. 2007. (eds) Modelling our future: Population ageing, health and aged care. Amsterdam: Elsevier.

Jagger C, Matthews R, Spiers N, Brayne C, Comas-Herrera A, Robinson T, Lindesay J, Croft P. 2006. Compression or expansion of disability? Forecasting future disability levels under changing patterns of diseases. London: King's Fund.

JAVA <http://java.com/en>.

Knickman JR, Snell EK. 2002. The 2030 problem: Caring for aging baby boomers. Health Services Research; 37(4): 849-884.

Lloyd-Sherlock P. 2000. Population ageing in developed and developing regions: implications for health policy. Social Science and Medicine; 51: 887-895.

Lymer S et al. Undated. Constructing a basefile for the microsimulation modelling of the use and costs of medical services in Australia. National Centre for Social and Economic Modelling, University of Canberra, Technical Paper, Table 13.

Lymer S, Brown L. 2006. The methodological issues of using cross-sectional data for pharmaceutical health outcome analysis. National Centre for Social and Economic Modelling, University of Canberra, Technical Paper. <http://www.natsem.canberra.edu.au>.

Lymer S, Brown L, Harding A. 2008. HealthMod: A microsimulation model of the Australian Medical Benefits System in 2001. National Centre for Social and Economic Modelling, University of Canberra, Technical Paper. <http://www.natsem.canberra.edu.au>.

Mannion O, Lay-Yee R, Wrapson W, Davis P, Pearson J. In Press 2011. JAMSIM: A microsimulation modellling policy tool. Journal of Artificial Societies and Social Simulation.

Martini A, Trivellato U. 1997. The role of survey data in microsimulation models for social policy analysis. Labour; 11(1): 83–112.

McGrail K, Green B, Barrer ML, Evans RG, Hertzman C, Normand C. 2002. Age, costs of acute and long-term care and proximity to death: evidence for 1987-88 and 1994-95 in British Columbia. Age and Ageing; 29:249-253.

Ministry of Health. 1999. Taking the pulse: The 1996/97 New Zealand Health Survey. Wellington: Ministry of Health. <http://www.moh.govt.nz>.

Ministry of Health. 2004a. Population ageing and health expenditure: New Zealand 2002-2051. Wellington: Ministry of Health. <http://www.moh.govt.nz>.

Ministry of Health. 2004b. Longer life, better health? Trends in health expectancy, New Zealand 1996-2001. Wellington: Ministry of Health. <http://www.moh.govt.nz>.

Ministry of Health. 2004c. A portrait of health: Key results of the 2002/03 New Zealand Health Survey. Wellington: Ministry of Health. <http://www.moh.govt.nz>.

Moore G, Showstack J. 2003. Primary care medicine in crisis: towards reconstruction and renewal. Annals of Internal Medicine; 138:244-247.

Morrissey K, Clarke G, Ballas D, Hynes S, O'Donoghue C. 2008. Examining access to GP services in rural Ireland using microsimulation analysis. Area; 40(3): 354–364.

OECD. 2006. Projecting OECD health and long-term care expenditures: What are the main drivers? OECD Economics Department. Working Paper No. 477. <http://www.oecd.org>.

OECD. 2009. Policies for healthy ageing: An overview. OECD Directorate for Employment, Labour and Social Affairs. Health Working Paper No. 42. <http://www.oecd.org>.

Ostberg V, Lennartsson C. 2007. Getting by with a little help: The importance of various types of social support for health problems. Scandinavian Journal of Public Health; 35(2): 197-204.

Pearson J, Lay-Yee R, Davis P, O'Sullivan D, von Randow M, Kerse N, Pradhan S. 2011. Primary care in an ageing society: Building and testing a microsimulation model for policy purposes. Social Science Computer Review; 29(1):21-36. doi:10.1177/0894439310370087. < http://ssc.sagepub.com/content/29/1/21>.

Prior PM, Hayes BC. 2003. The relationship between marital status and health: An empirical investigation of differences in bed occupancy within health and social care facilities in Britain, 1921-1991. Journal of Family Issues; 24(1): 124-148.

R Development Core Team. 2009. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing <http://www.R-project.org>.

Rässler S. 2002. Statistical matching: A frequentist theory, practical applications, and alternative Bayesian approaches. New York: Springer.

Raudenbush S, Bryk A. 1986. A hierarchical model for studying school effects. Sociology of Education.

Raymont A, Lay-Yee R, Davis P, Scott A. 2004. Family doctors: Methodology and description of the activity of private GPs: The National Primary Medical Care Survey (NatMedCa): 2001/02. Wellington: Ministry of Health. <http://www.moh.govt.nz/natmedca>.

Rechel B, Doyle Y, Grundy E, McKee M. 2009. How can health systems respond to population ageing? Copenhagen: World Health Organisation.

Rodgers WL. 1984. An evaluation of statistical matching. Journal of Business and Economic Statistics; 2(1): 91–102.

Salmond C, Crampton P, Sutton F. 1998. NZDep96 Index of Deprivation. Wellington: Health Services Research Centre, Victoria University.

SAS 9.2, SAS Institute Inc, Cary, North Carolina, USA <http://www.sas.com>.

Schabenberger O. 2005. Introducing the GLIMMIX procedure for generalised linear mixed models. SAS Institute Inc., Cary, NC.

Smith DM, Clarke GP, Harland K. 2009. Improving the synthetic data generation process in spatial microsimulation models. Environment and Planning A; 41: 1251–1268.

Snijders T, Bosker R. 1999. Multilevel Analysis: An introduction to basic and advanced multilevel modelling. Sage Publications.

Sox HC. 2003. The future of primary care. Annals of Internal Medicine; 138: 230-231.

Sutherland H, Taylor R, Gomulka J. 2002. Combining household income and expenditure data in policy simulations. Review of Income and Wealth; 48(4): 517–536.

Swedish National Institute of Public Health. 2006. Healthy ageing: A challenge for Europe. Stockholm: Swedish National Institute of Public Health. <http://www.healthyageing.nu>.

Van Houtven CH, Norton EC. 2004. Informal care and health care use of older adults. Journal of Health Economics; 23(6): 1159-1180.

von Randow M, Davis P, Lay-Yee R, Pearson J. In Press 2011. Data matching to allocate doctors to patients in a microsimulation model of the primary care process in New Zealand. Social Science Computer Review.

Wolfson MC. 1994. POHEM - A framework for understanding and modelling the health of human populations. World Health Statistics Quarterly; 47(3-4): 157-176.

Zaidi A, Scott A. 2001. Base dataset for the SAGE model. SAGE Technical Note 1. London: ESRC-SAGE Research Group.

9. Appendices

9.1 Abbreviations

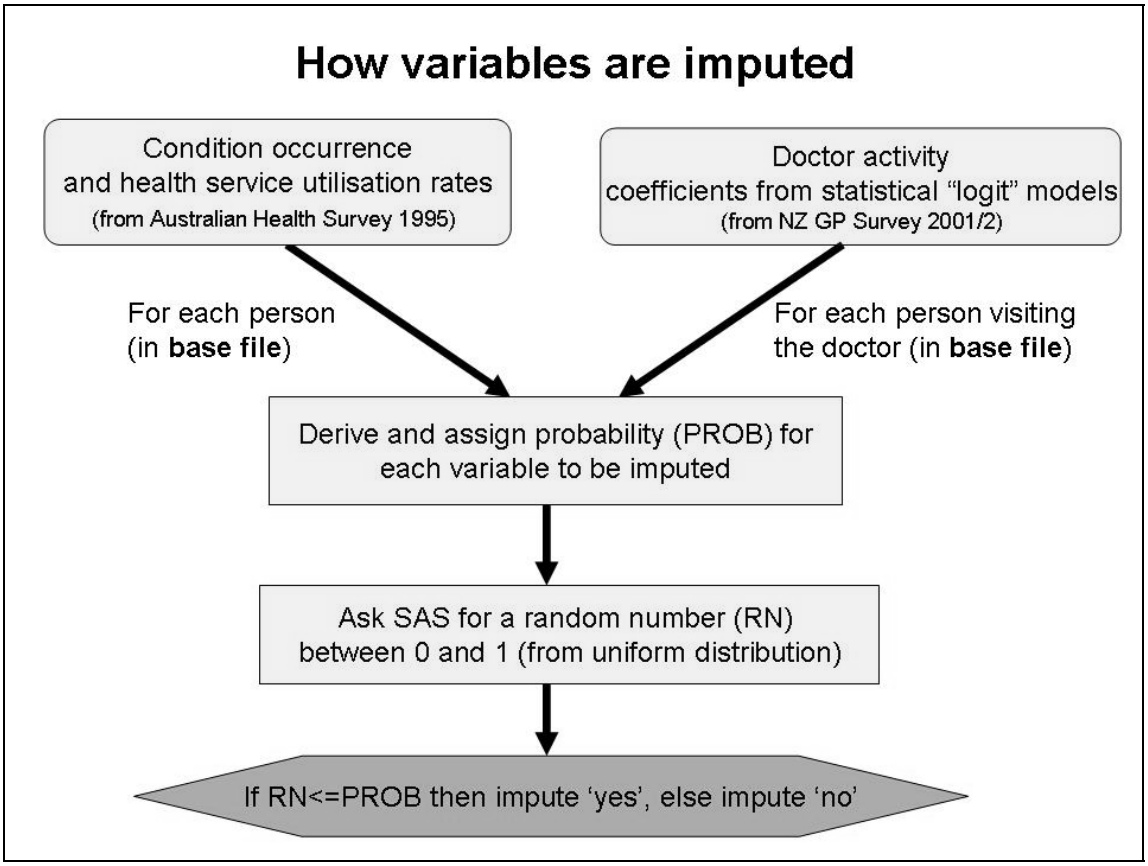
AGEGRP	Age Group
ANHS	Australian National Health Survey
AUS	Australia
CDF	Cumulative Distribution Function
DISAB	Disability
GP	General Practitioner
HHTYPE	Household Type
LT	Long-term illness
MIC	Most Important Condition
NATSEM	National Centre for Social and Economic Modelling
NEC	Not Elsewhere Classified
NI	Not Illness
NPMCS	National Primary Medical Care Survey
NZ	New Zealand
NZDep	New Zealand Deprivation Score
NZHS	New Zealand Health Survey
PCASO	Primary Care in an Ageing Society
PROB	Probability
RN	Random Number
SIM	Simulation
ST	Short-term illness (recent)
STLT	Long-term illness with recent episode
SX	Symptoms (non-specific)

9.2 Random assignment of characteristics

9.2.1 The use of random numbers to implement probabilities

Figure 9.2.1 below shows how random numbers were used in the simulation to convert probabilities, whether derived from tables (ANHS) or statistical models (NPMCS), to assigning characteristics to an individual. Essentially a probability is compared to a random number drawn from a uniform distribution between 0 and 1. If that random number is less than or equal to the probability then the characteristic of interest is determined to be present (otherwise absent).

Figure 9.2.1 How to impute variables in simulation



9.2.2 Using a cumulative distribution function (CDF)

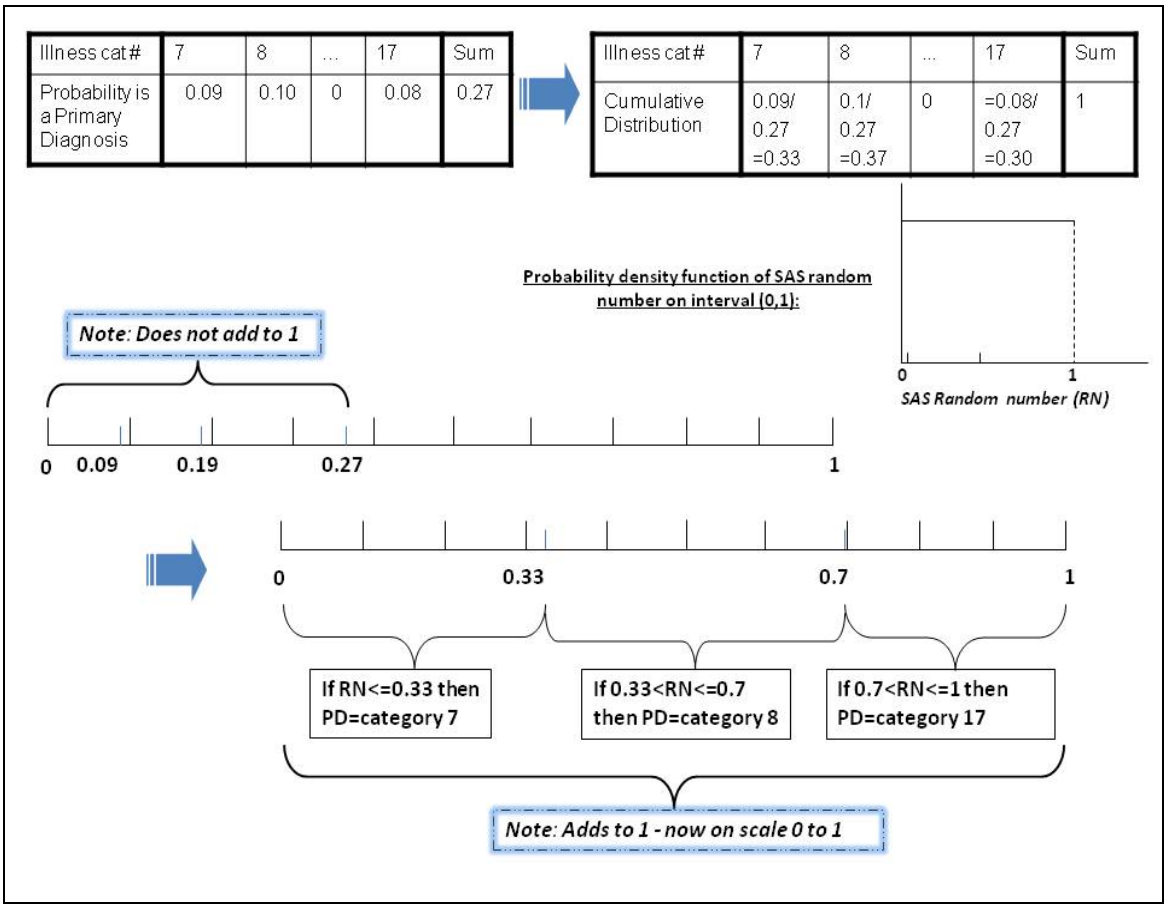
The CDF describes the probability that a random variable is less than or equal to the dependent variable of the function.

An array of probabilities may exist relating to multiple categories of a given characteristic. In order to model such a characteristic, a CDF is created. The aim is to standardise the probabilities

so that they are converted to a scale from 0 to 1. Thus a random uniform number on the same interval (0, 1) can be compared to these standardised probabilities and so used to assign the characteristic.

Figure 9.2.2 below shows an example of how this is done in the case of assigning a primary diagnosis to a visit for someone who has the following characteristics: aged 45-64, male, Maori, with >11 visits to the GP in the last 12 months, and in the lowest deprivation quintile. This person has condition categories 7 ('Cardiovascular/ circulatory diseases'), 8 ('Respiratory diseases') and 17 (Not an Illness/ Non-symptomatic/ Not Stated') present in the fortnight. The crude probability of category 7 being the primary diagnosis is 0.09, for category 8 it is 0.10, and for category 17 it is 0.08. Summing all the crude probabilities together gives 0.27. Dividing each of the crude probabilities (0.09, 0.10, and 0.08) in turn by the sum (0.27) gives a new set of probabilities (0.33, 0.37, and 0.30) for the characteristics that sum to 1.

Figure 9.2.2 How to use a cumulative distribution function (CDF)



These new probabilities can then map to intervals on a random uniform number ranging from 0 to 1. If the random number lies (a) between 0 and (0.09/0.27) then the primary diagnosis would

be assigned as category 7; (b) between (0.09/0.27) and ((0.09/0.27) + (0.1/0.27)) then the primary diagnosis would be assigned as category 8; (c) between ((0.09/0.27) + (0.1/0.27)) and ((0.09/0.27) + (0.1/0.27) + (0.08/0.27)) then the primary diagnosis would be assigned as category 17. The random number is compared until it maps to a primary diagnosis category. Note that ((0.09/0.27) + (0.1/0.27) + (0.08/0.27)) = 1.

9.2.3 Random number generation

We used the SAS RANUNI function to produce non-repeating random draws from a uniform distribution for one run of the simulation (*Dicky, Donaghy & Smith 2003)*. The RANUNI function has a period (the length of the non-repeating pattern) of 2^31 – 1 or 2,147,483,647 numbers which is greater than the estimated number of random numbers needed for one run of the simulation. The seed value used in the RANUNI function specifies the starting point for the set sequence of pseudo-random numbers to be drawn. Once the RANUNI function is used, the seed initially specified in that instance applies to the rest of the SAS data step (no matter what other seed values are requested). A different seed value for random number generation, equal to the run number, was specified for each run of the simulation. Thus each SAS data step, within a call of the macro for a given run, was given the run number as the seed for all random numbers generated within that step. In this way, when conducting multiple runs, we have tried to avoid possible autocorrelation due to overlapping streams of random numbers (each starting from a different seed value).

9.3 Reallocation of 'non-symptomatic' and 'ill-defined' condition categories

The following discussion outlines the investigations carried out on methods of reallocating 'Sx' and 'NI' categories.

9.3.1 Re-allocation – under Option B simulation process (see section 5.4 above)

One way we tried reallocating was by using the visits in NPMCS that only had, for each visit, one listed reason and one listed diagnosis, so that we could be assured that the diagnosis was directly related to the reason for the visit. The distribution of the diagnosis categories for this kind of visit where the reason for the visit was in the 'NI' category was used to reclassify the 'NI' conditions in an investigative simulation (see Table 9.3.1 below). Similarly, we tried to reallocate the 'Sx' category via the distribution of diagnosis categories in this kind of visit where the reason for visit was in the 'Sx' category (see Table 9.3.2 below).

Table 9.3.1 Reallocation of 'Not illness' category: Distribution of diagnosis categories for visits with only one reason, i.e. 'Not illness', and one diagnosis

Condition category	Percentage of visits
Infectious and parasitic diseases	2.8
Neoplasms	3.9
Endocrine/nutritional/metabolic/immunity disorders	2.4
Diseases of blood and blood forming organs	0.2
Mental disorders	3.2
Nervous system/sense organ diseases	6.1
Cardiovascular/Circulatory diseases	6.6
Respiratory system diseases	4.1
Digestive system diseases	3.2
Genitourinary system diseases	2.3
Complications of pregnancy/childbirth/puerperium	0.3
Skin and subcutaneous tissue diseases	7.9
Musculoskeletal and connective tissue diseases	1.9
Congenital anomalies	0.3
Symptoms, signs and ill-defined conditions & Disability NEC	1.3
Injury and poisoning	7.9
Not an Illness, non-symptomatic	45.8

Table 9.3.2 Reallocation of 'Symptoms' category: Distribution of diagnosis categories for visits with only one reason, i.e. 'Symptoms', and one diagnosis

Condition category	Percentage of visits
Infectious and parasitic diseases	10.0
Neoplasms	1.1
Endocrine/nutritional/metabolic/immunity disorders	0.5
Diseases of blood and blood forming organs	0.4
Mental disorders	1.5
Nervous system/sense organ diseases	11.8
Cardiovascular/Circulatory diseases	6.5
Respiratory system diseases	35.9
Digestive system diseases	4.7
Genitourinary system diseases	7.4
Complications of pregnancy/childbirth/puerperium	0.1
Skin and subcutaneous tissue diseases	4.7
Musculoskeletal and connective tissue diseases	1.6
Congenital anomalies	0.01
Symptoms, signs and ill-defined conditions & Disability NEC	6.3
Injury and poisoning	3.2
Not an Illness, non-symptomatic	4.3

We also tried reallocating the 'Sx' category via the distribution of all diagnoses listed for NPMCS visits excluding 'NI'. This was deemed implausible (see Table 9.3.3) and not further pursued.

Table 9.3.3 Reallocation of 'Symptoms' category: Distribution of all diagnoses in NPMCS (excluding NI category)

Condition category	Percentage of visits (excl NI category)
Infectious and parasitic diseases	5.3
Neoplasms	3.0
Endocrine/nutritional/metabolic/immunity disorders	5.1
Diseases of blood and blood forming organs	0.6
Mental disorders	6.2
Nervous system/sense organ diseases	10.1
Cardiovascular/Circulatory diseases	11.5
Respiratory system diseases	18.3
Digestive system diseases	5.4
Genitourinary system diseases	5.7
Complications of pregnancy/childbirth/puerperium	0.4
Skin and subcutaneous tissue diseases	8.3
Musculoskeletal and connective tissue diseases	7.0
Congenital anomalies	0.3
Symptoms, signs and ill-defined conditions & Disability NEC	4.3
Injury and poisoning	8.8

Several investigative simulations were tried: (a) reallocating the 'NI' category via Table 9.3.1; (b) reallocating the 'Sx' category according to Table 9.3.2; (c) reallocating the 'Sx' category using the distribution of all diagnoses in NPMCS (excluding the NI category) according to Table 9.3.3; and (d) reallocating both the 'NI' and the 'Sx' categories using all Tables 9.3.1 – 9.3.3. In each case, the reallocated distributions described above were converted to cumulative distribution functions (CDF). For each condition that needed to be reallocated in the simulation, the relevant CDF was compared to a random number from a uniform distribution on the interval (0, 1). The outcomes of these simulations were compared to the one where no reallocation was made (see Tables 9.3.4 and 9.3.5 below). Redistributing both the 'Sx' and 'NI' categories together gave the best outcomes for validation of conditions and actions. However, as the 'NI' category in the 'no reallocation' simulation was already very close to the bench mark, it was decided that it was inappropriate to reallocate. Reallocating the 'Sx' category on its own, regardless of the method, resulted in too much of the category being redistributed (compared to the benchmark), and while it improved the overall validation error for the conditions, it made it slightly worse for the actions. In addition, it was also felt that the 'reason for visit' variable from NPMCS used to generate the reallocation tables above was too unreliable to be used in this way, as it was known that surveyed doctors inconsistently filled in the 'reason' field, that is, often not completing it at all, or not recording the patient's words and instead writing down a diagnosis. The analysis also had to be limited to one-reason/one-diagnosis visits in order to ensure the link between the reason category and the diagnosis category (as many visits had a different number of reasons to the number of diagnoses, and it was not clear which diagnoses belonged to which reasons). This reduced the numbers from 9272 visits to 1164 visits which when weighted equated to just 18.1 % of the data, possibly too small to give reliable distributions. It was therefore decided not to reallocate either the 'Sx' or 'NI' categories (under the Option B simulation process).

Table 9.3.4 Reallocation of 'Symptoms' and 'Not illness' categories: percentage distribution of population-level conditions for those people assigned in the simulation as having visited the doctor

Condition category	NPMCS (bench mark)	No reallocation	Sx reallocated only (using Table 5.12)	Sx reallocated only (using Table 5.11I)	NI reallocated only (using Table 5.10)	Both NI and Sx reallocated (using tables 5.10 and 5.11)
Infectious and parasitic diseases	4.3	1.9	3.2	2.6	2.5	3.8
Neoplasms	2.4	0.7	0.8	1.1	1.5	1.6
Endocrine/nutritional/ metabolic/immunity disorders	4.1	5.3	5.4	5.9	5.8	6.0
Diseases of blood and blood forming organs	0.5	0.6	0.7	0.7	0.7	0.7
Mental disorders	5.0	2.5	2.7	3.2	3.2	3.5
Nervous system/sense organ diseases	8.2	4.7	6.1	5.9	5.9	7.3
Cardiovascular/ Circulatory diseases	9.3	9.3	10.2	10.6	10.5	11.5
Respiratory system diseases	14.8	13.8	17.6	15.8	14.7	18.5
Digestive system diseases	4.4	7.1	7.8	7.82	7.8	8.5
Genitourinary system diseases	4.6	2.7	3.6	3.5	3.2	4.1
Complications of pregnancy/childbirth/ puerperium	0.3	0.1	0.1	0.1	0.1	0.1
Skin and subcutaneous tissue diseases	6.7	4.9	5.6	5.9	6.7	7.3
Musculoskeletal and connective tissue diseases	5.7	8.7	9.0	9.5	9.1	9.5
Congenital anomalies	0.2	0.1	0.1	0.1	0.2	0.2
Symptoms, signs and ill-defined conditions & Disability NEC	3.5	12.4	0.7	0.5	12.8	1.0
Injury and poisoning	7.1	3.8	4.3	4.9	5.5	6.0
Not an Illness, non-symptomatic	19.0	21.6	22.3	21.9	9.9	10.4
Average error	-	2.2	1.8	1.7	2.2	1.9

Table 9.3.5 Reallocation of 'Symptoms' and 'Not illness' categories: percentage distribution of visits with each kind of doctor action

Doctor action	NPMCS (bench mark)	No reallocation	Sx reallocated only (using Table 5.12)	Sx reallocated only (using Table 5.11I)	NI reallocated only (using Table 5.10)	Both NI and Sx reallocated (using tables 5.10 and 5.11)
Non-drug	62.0	60.1	58.4	59.0	59.1	57.3
Followup	57.2	55.2	54.2	54.9	56.0	55.0
Investigation	24.8	22.1	21.0	21.1	21.8	20.6
Referral	15.8	13.9	12.9	13.5	14.4	13.5
Prescription	66.2	54.0	56.6	56.2	58.9	61.6
Average error	-	4.1	4.6	4.3	3.2	3.6

9.3.2 Re-allocation – under Option C (final) simulation process (see section 5.4)

One possible reason for the discrepancy in the 'Symptoms' outcome was that the subset of conditions from the simulation was not the most comparable to our NPMCS benchmark – we needed a set that matched up as much as possible with conditions that would actually have been presented to the doctor. Earlier in this process, we had had to make a decision regarding which set of conditions we would output.

Showing all conditions for people that had been allocated a visit would include *all* conditions that people had in the fortnight, as they were being modelled using input from *population* level information from the ANHS survey. This would include:

- Conditions brought to the GP
- Population level conditions – not brought to any health practitioner's attention
- Conditions brought to health practitioners other than GPs
- All conditions for all visits in a fortnight.

The outputted conditions table (which shows *all* conditions in a fortnight for people who had at least one visit in a fortnight) is not totally comparable to what would have been presented to the GP for any one visit (see Table 9.3.6 below). Thus the NPMCS column only goes up to 4 distinct condition categories whereas the simulation goes beyond to 8 categories.

Table 9.3.6 Distribution of number of distinct 'all' condition categories: per visit in NPMCS vs per fortnight in simulation

Number of distinct condition categories	Percent in NPMCS – per visit	Percent in simulation (all conditions subset) – per fortnight *
1	60.6	24.8
2	26.9	35.3
3	9.8	24.5
4	2.7	10.9
5		3.5
6		0.9
7		0.2
8		0.0 (to 1 decimal place)

* Final model with Sx redistributed

There are two possible explanations for the over-estimation in the simulation: (a) the 'Sx' category in particular would tend to be the main category for any background conditions at population level not taken to a health professional.; (b) the difference in reporting of actual diagnoses between the GPs filling in the NPMCS survey and GP users reporting their reason for a visit in the ANHS. GPs tend to write down specific diagnoses rather than free text that would have to be coded as 'Sx', while GP users, as lay people, when asked their reason for visiting the doctor, would be more likely to write down symptoms that would be classified as 'Sx'. The concern was that the over-occurrence of the 'Sx' category, in those fortnights and people with visits, would in turn result in too many primary diagnoses being in the 'Sx' category, therefore affecting the accuracy of subsequent doctor action predictions.

Therefore, we needed to: (a) review if the most comparable conditions set from the simulation was being validated against the NPMCS data which would confirm if the 'Sx' category really was being overestimated or if we were just looking at the wrong set; and (b) if there was a genuine overestimate in the 'Sx' category, reallocate this category to the likely diagnoses that would be presented to the GP in order to enable the correct distribution of primary diagnoses to transpire.

An alternative would have been to present the simulated conditions distribution limited just to those conditions that had been deemed as being likely to be seen by the doctor (see Table 9.3.7 below). The spread of unique diagnoses was less like the NPMCS distribution than that including all conditions; the vast majority of fortnights in this alternative set had just one unique condition. This also indicated that perhaps the surveyed person, when asked what they saw the GP for, may have missed out some of the more minor categories and/or that the other visits in the last fortnight that were not surveyed had additional information, and/or indeed the GPs do address all background conditions and not just those the patient intended to bring to a visit.

Table 9.3.7 Distribution of number of distinct 'likely to see the GP' condition categories: per visit in NPMCS vs per fortnight in simulation

Number of distinct condition categories	Percent in NPMCS – per visit	Percent in simulation (conditions deemed 'likely to see GP' subset) – per fortnight *
1	60.6	88.5
2	26.9	9.8
3	9.8	0.7
4	2.7	0.0
5		0.0
6		
7		
8		

* Final model

The decision was made that the more inclusive 'all conditions' was the appropriate set to present as typically a GP would ask about and record all conditions the person had, not just the ones that the patient had in mind to bring. The true distribution was probably somewhere between the two extremes. In fact if we assume that the 'all conditions' set actually does represent all conditions seen by the GP, except that on average one condition was not taken to the GP, the simulated distribution then becomes very close to that of NPMCS (see Table 9.3.8 below).

Table 9.3.8 Distribution of number of distinct 'all' condition categories: per visit in NPMCS vs per fortnight in modified simulation

Number of distinct condition categories	Percent in NPMCS – per visit	Percent in simulation (modified* 'all conditions' subset) – per fortnight *
1	60.6	60.1
2	26.9	24.5
3	9.8	10.9
4	2.7	3.5
5		0.9
6		0.2

*Assuming that fortnights with 2 distinct condition categories really represent fortnights with just 1 condition category that was taken to the GP

To address the simulated over-estimation of the 'Sx' category, we re-visited the method of reallocation. Using previous simulation runs as a guide, it was found that the 'Sx' category was being overestimated by (12.42/3.5) = 3.5 times. Originally, blanket reallocation of *all* 'Sx' conditions resulted in too many being reallocated. Another way was to only reallocate as many as necessary in order to roughly meet the level of 'Sx' conditions that was presented in the benchmark NPMCS. It was decided to reallocate 2 out of every 3 'Sx' conditions for people who had been allocated at least one visit for the fortnight.

We needed to create a second array of conditions for each fortnight that was the same as the original one but also had the reallocation for the 'Sx' category applied. Thus the original array,

with categories aligned to the Australian survey, could be used in conjunction with the Australian data input tables to the simulation; while the second array could be used when needed in conjunction with the NPMCS data, that is, when applying NPMCS-based probabilities to choose a primary diagnosis and when using this chosen primary diagnosis to allocate doctor actions.

Results from the final simulation model were checked to see if the simulated primary diagnosis distribution was roughly comparable to that in NPMCS. The simulated distribution was not considered as an object capable of being manipulated/ reallocated in the simulation. This was because the primary diagnosis probabilities, derived from NPMCS data, had to be applied to the set of conditions likely to be seen by the GP, that is, the 'Sx' category reallocation had to be done at the earlier step.

9.4 ANHS 1995 data: input tables

Table 9.4.1 Probability of outcomes by condition category and age group (ANHS)

Condition category	Agegrp	ST condition occurrence rate	STLT condition occurrence rate	Probability of condition being seen by a GP
1	1	0.02949	0.00356	0.27476
1	2	0.0311	0.00495	0.13452
1	3	0.024	0.00372	0.13314
1	4	0.01504	0.00325	0.30042
2	1	0.00047	0.0005	0.40751
2	2	0.00152	0.00256	0.46147
2	3	0.00559	0.00833	0.27992
2	4	0.01449	0.02167	0.41206
3	1	0.00356	0.00405	0.13773
3	2	0.01769	0.02375	0.0729
3	3	0.08625	0.09831	0.0836
3	4	0.06559	0.1565	0.0915
4	1	0.00202	0.00302	0.05768
4	2	0.00451	0.00607	0.03954
4	3	0.00383	0.00612	0.07672
4	4	0.00691	0.01318	0.16402
5	1	0.0093	0.00929	0.12655
5	2	0.0236	0.01876	0.17093
5	3	0.0305	0.02406	0.1418
5	4	0.04241	0.015	0.07302
6	1	0.03812	0.01327	0.33899
6	2	0.04431	0.02598	0.15496
6	3	0.04136	0.03654	0.12427
6	4	0.06523	0.07827	0.12811
7	1	0.0014	0.00281	0.31749
7	2	0.00857	0.02614	0.15869
7	3	0.05568	0.1852	0.12208
7	4	0.21627	0.43639	0.1126
8	1	0.15893	0.11	0.24694
8	2	0.12228	0.10252	0.20242
8	3	0.09382	0.10443	0.17896
8	4	0.08216	0.11057	0.21808
9	1	0.07786	0.00521	0.10568
9	2	0.06754	0.01974	0.09116
9	3	0.08733	0.04739	0.09328
9	4	0.11842	0.08747	0.07709

Condition category	Agegrp	ST condition occurrence rate	STLT condition occurrence rate	Probability of condition being seen by a GP
10	1	0.01699	0.00449	0.23501
10	2	0.03239	0.00985	0.17015
10	3	0.05326	0.01633	0.12564
10	4	0.02289	0.01942	0.21312
11	1	0.00056	0	0.01485
11	2	0.00225	0.00029	0.21964
11	3	0	0	.
11	4	0	0	.
12	1	0.06234	0.01946	0.17483
12	2	0.06188	0.01821	0.12254
12	3	0.06373	0.01529	0.1766
12	4	0.06264	0.01534	0.18318
13	1	0.0273	0.01344	0.17211
13	2	0.06381	0.05597	0.15433
13	3	0.08014	0.14549	0.14494
13	4	0.09694	0.23626	0.12424
14	1	0.00051	0.00272	0.03668
14	2	0.00028	0.00043	0.13814
14	3	0	0.00056	0
14	4	0.00102	0.00009	0.20178
15	1	0.14274	0.01313	0.08821
15	2	0.24687	0.01387	0.05918
15	3	0.19163	0.01512	0.07595
15	4	0.16077	0.01396	0.07861
16	1	0.06874	0.00241	0.17776
16	2	0.05728	0.0038	0.2304
16	3	0.04665	0.00423	0.19681
16	4	0.04984	0.00452	0.1862
17	1	0.26511	0.00006	0.13199
17	2	0.35122	0	0.15843
17	3	0.38329	0.00052	0.17858
17	4	0.44439	0.0003	0.32592

Table 9.4.2 Probability of outcomes by condition category and gender (ANHS)

Condition category	Gender	ST condition occurrence rate	STLT condition occurrence rate	Probability of condition being seen by a GP
1	1	0.02423	0.00339	0.21088
1	2	0.03	0.00457	0.19211
2	1	0.00273	0.00532	0.37683
2	2	0.00432	0.00526	0.37504
3	1	0.01208	0.0475	0.10461
3	2	0.05265	0.04819	0.07547
4	1	0.00169	0.00178	0.06941
4	2	0.00579	0.00982	0.08502
5	1	0.01689	0.01544	0.13701
5	2	0.02719	0.01647	0.13573
6	1	0.03727	0.02786	0.19047
6	2	0.05056	0.03167	0.18887
7	1	0.03642	0.09045	0.1236
7	2	0.04453	0.10851	0.12189
8	1	0.11997	0.09949	0.20731
8	2	0.12991	0.11366	0.22895
9	1	0.07471	0.0283	0.10163
9	2	0.08818	0.0282	0.08499
10	1	0.00533	0.00615	0.25409
10	2	0.0544	0.01458	0.15909
11	1	0.00005	0	0
11	2	0.00174	0.00018	0.18242
12	1	0.05484	0.01649	0.16698
12	2	0.07014	0.01893	0.15461
13	1	0.05289	0.06845	0.15678
13	2	0.06274	0.09247	0.13504
14	1	0.00022	0.00158	0.08745
14	2	0.00057	0.00092	0.02925
15	1	0.15718	0.01114	0.07141
15	2	0.21699	0.01658	0.07412
16	1	0.06354	0.00333	0.21931
16	2	0.05325	0.00361	0.17298
17	1	0.28695	0.00025	0.16875
17	2	0.38781	0.00008	0.1914

Table 9.4.3 Probability of outcomes by condition category and original household type (ANHS)

Condition category	Original Hhtype*	ST condition occurrence rate	STLT condition occurrence rate	Probability of condition being seen by a GP
1	1	0.02674	0.00404	0.13707
1	2	0.02564	0.0047	0.16591
1	3	0.0289	0.00317	0.25316
2	1	0.00792	0.00875	0.44644
2	2	0.00443	0.00763	0.35595
2	3	0.00134	0.00173	0.3616
3	1	0.04973	0.08424	0.09717
3	2	0.04807	0.06777	0.07943
3	3	0.01026	0.01572	0.10628
4	1	0.00577	0.01061	0.153
4	2	0.00471	0.00617	0.06109
4	3	0.00214	0.00414	0.07425
5	1	0.04779	0.02927	0.15812
5	2	0.02358	0.01601	0.12954
5	3	0.0135	0.01235	0.13077
6	1	0.05575	0.04749	0.14401
6	2	0.04503	0.03437	0.13979
6	3	0.03956	0.01987	0.28549
7	1	0.09259	0.2014	0.12596
7	2	0.05296	0.13996	0.11758
7	3	0.01259	0.02685	0.14447
8	1	0.10117	0.11268	0.18607
8	2	0.10398	0.1033	0.20057
8	3	0.15491	0.10871	0.24195
9	1	0.09426	0.05041	0.07676
9	2	0.08135	0.03774	0.09245
9	3	0.0782	0.01166	0.09991
10	1	0.03346	0.01471	0.17487
10	2	0.03914	0.0133	0.1659
10	3	0.0187	0.00595	0.18682
11	1	0.00035	0	0
11	2	0.00155	0.0001	0.14671
11	3	0.00031	0.0001	0.35502
12	1	0.07433	0.01686	0.18194
12	2	0.06319	0.01731	0.1504
12	3	0.05863	0.01841	0.16465
13	1	0.07933	0.15383	0.14776
13	2	0.0731	0.10782	0.14038
13	3	0.03493	0.03025	0.15503

Condition category	Original Hhtype*	ST condition occurrence rate	STLT condition occurrence rate	Probability of condition being seen by a GP
14	1	0	0.0002	0
14	2	0.00033	0.00046	0.15184
14	3	0.00057	0.00241	0.03465
15	1	0.21733	0.01664	0.08487
15	2	0.21107	0.01345	0.06389
15	3	0.15233	0.01361	0.08242
16	1	0.06639	0.00544	0.18761
16	2	0.0498	0.00353	0.20725
16	3	0.06588	0.00287	0.19262
17	1	0.44354	0.00012	0.2141
17	2	0.36435	0.00023	0.20208
17	3	0.27925	0.00011	0.13824

* Household type: 1=live *without* someone >=15yrs age; 2=live *with* someone >=15yrs age and partnered (husband/wife or de facto, boyfriend or girlfriend); 3=live *with* someone >=15yrs age and *not* partnered, where 'someone >=15yrs age' is the definition of an adult.

Table 9.4.4 Probability of outcomes by condition category and derived household type (ANHS)

Condition category	Derived Hhtype*	ST condition occurrence rate	STLT condition occurrence rate	Probability of condition being seen by a GP
1	1	0.03023	0.0042	0.20766
1	2	0.02161	0.00348	0.1922
1	3	0.01858	0.00373	0.14392
2	1	0.00095	0.00145	0.44967
2	2	0.00764	0.01323	0.35109
2	3	0.01176	0.01317	0.37591
3	1	0.01005	0.01309	0.08462
3	2	0.08126	0.11449	0.08009
3	3	0.07248	0.13192	0.10237
4	1	0.00316	0.00442	0.04608
4	2	0.0049	0.00685	0.07318
4	3	0.0051	0.01317	0.2022
5	1	0.01586	0.01364	0.15568
5	2	0.02925	0.01852	0.12006
5	3	0.0484	0.02614	0.10805
6	1	0.04096	0.0191	0.2405
6	2	0.04752	0.04707	0.11843
6	3	0.0563	0.06322	0.14108
7	1	0.00469	0.01351	0.17886
7	2	0.10029	0.25877	0.11536
7	3	0.14853	0.32123	0.11868

Condition category	Derived Hhtype*	ST condition occurrence rate	STLT condition occurrence rate	Probability of condition being seen by a GP
8	1	0.14211	0.10657	0.2285
8	2	0.08683	0.10046	0.18936
8	3	0.09617	0.12173	0.20057
9	1	0.07313	0.01188	0.09893
9	2	0.09467	0.05841	0.09519
9	3	0.10845	0.07082	0.06601
10	1	0.02406	0.00695	0.19458
10	2	0.04549	0.01759	0.14178
10	3	0.03408	0.01715	0.16755
11	1	0.00133	0.00013	0.17758
11	2	.	.	.
11	3	.	0	.
12	1	0.06213	0.01888	0.15125
12	2	0.06191	0.01544	0.16865
12	3	0.06676	0.01498	0.20271
13	1	0.04405	0.03296	0.15951
13	2	0.08249	0.16361	0.13181
13	3	0.09547	0.2151	0.14251
14	1	0.00041	0.00167	0.05257
14	2	0.00053	0.00044	0.11962
14	3	0	0.00025	0
15	1	0.19053	0.01347	0.07122
15	2	0.17208	0.01311	0.07267
15	3	0.20034	0.01856	0.08528
16	1	0.06348	0.00305	0.19989
16	2	0.04367	0.00364	0.19265
16	3	0.05784	0.00602	0.19293
17	1	0.30463	0.00003	0.14593
17	2	0.38645	0.00047	0.2317

* Household type: 1=child, 2=partnered adult, 3=unpartnered adult.

Table 9.4.5 Mean number of GP visits by age group (ANHS)

Agegrp	*Mean GP visits*
1: 0-24	1.2961824
2: 25-44 yrs	1.3769319
3: 45-64 yrs	1.3736907
4: 65 + yrs	1.3614768

Table 9.4.6 Mean number of GP visits by gender (ANHS)

Gender	*Mean GP visits*
Male	1.3360876
Female	1.3587936

Table 9.4.7 Mean number of visits by original household type (ANHS)

Original Hhtype	*Mean GP visits*
1: Live without other >=15	1.3933138
2: Live with other >=15 & partnered	1.3609225
3: Live with other >=15 & not partnered	1.3137644

Table 9.4.8 Mean number of visits by derived household type (ANHS)

Derived Hhtype	*Mean GP visits*
1: Child	1.2860546
2: Partnered adult	1.3609225
3: Unpartnered adult	1.3647659

9.5 Verification tables: simulated results compared to the Australian Health Survey (ANHS 1995)

Table 9.5.1 Occurrence rates of 'Short-Term and Long-Term' ('STLT') conditions per fortnight

Condition category	Aus Health Survey (%)	Simulation 1st run (%)	Absolute error
Infectious & parasitic diseases	0.40	0.41	0.006
Neoplasms	0.53	0.53	0.003
Endocrine/nutritional/metabolic/immunity disorders	4.78	4.89	0.11
Diseases of blood & blood forming organs	0.58	0.60	0.01
Mental disorders	1.60	1.63	0.03
Nervous system/sense organ diseases	3.00	3.10	0.12
Cardiovascular/circulatory diseases	10.00	10.17	0.22
Respiratory system diseases	10.66	10.63	0.03
Digestive system diseases	2.83	2.88	0.05
Genitourinary system diseases	1.04	1.05	0.02
Complications of pregnancy/childbirth/puerperium	0.01	0.01	0.002
Skin & subcutaneous tissue diseases	1.77	1.73	0.05
Musculoskeletal & connective tissue diseases	8.05	8.17	0.12
Congenital anomalies	0.13	0.13	0.002
Symptoms, signs, ill-defined conditions & disab nec	1.39	1.33	0.06
Injury & poisoning	0.35	0.34	0.01
Not illness/unspecified	0.02	0.02	0.001
		Average error	0.05

Table 9.5.2: Occurrence rates of 'Short-Term' ('ST') conditions per fortnight

Condition category	Aus Health Survey (%)	Simulation 1st run (%)	Absolute error
Infectious & parasitic diseases	2.71	2.69	0.02
Neoplasms	0.35	0.35	0.002
Endocrine/nutritional/metabolic/immunity disorders	3.25	3.42	0.17
Diseases of blood & blood forming organs	0.38	0.39	0.02
Mental disorders	2.21	2.21	0.001
Nervous system/sense organ diseases	4.39	4.39	0.004
Cardiovascular/circulatory diseases	4.05	4.14	0.09
Respiratory system diseases	12.50	12.50	0.004
Digestive system diseases	8.15	8.20	0.05
Genitourinary system diseases	3.00	3.08	0.08
Complications of pregnancy/childbirth/puerperium	0.09	0.10	0.01
Skin & subcutaneous tissue diseases	6.25	6.20	0.05
Musculoskeletal & connective tissue diseases	5.78	5.74	0.05
Congenital anomalies	0.04	0.04	0.001
Symptoms, signs, ill-defined conditions & disab nec	18.72	18.84	0.12
Injury & poisoning	5.84	5.76	0.08
Not illness/unspecified	33.76	33.60	0.17
		Average error	0.05

Table 9.5.3 Average percentage of population with >=1 recent condition per fortnight

	Aus Health Survey	Simulation 1st run	Absolute error
Average percentage with >=1 recent condition per fortnight	75.9	79.4	3.5

Table 9.6.4 Average percentage of population with >=1 recent condition per fortnight - by age group

Age group (years)	Aus Health Survey	Simulation 1st run	Absolute error
0 - 24	65.7	69.2	4.5
25-44	76.2	79.5	3.3
45-64	83.6	87.6	4.0
65+	93.0	95.1	2.1
		Average error	3.5

Table 9.5.5 Average percentage of population with >=1 recent condition per fortnight - by gender

Gender	Aus Health Survey	Simulation 1st run	Absolute error
female	80.5	83.4	2.9
male	71.3	75.2	3.9
		Average error	3.4

Table 9.5.6 Average percentage of population with >=1 recent condition per fortnight – by household type

Household type	Aus Health Survey	Simulation 1st run	Absolute error
Live W ITHOUT someone >=15yrs	85.9	90.1	4.2
Live WITH someone >=15yrs and partnered	81.1	84.1	3.0
Live WITH someone >=15yrs and NOT partnered	67.4	71.8	4.4
		Average error	3.9

Table 9.5.7 Distribution of most important condition (leading to a visit)

Condition category	Aus Health Survey (%)	Simulation 1st run (%)	Absolute error
Infectious & parasitic diseases	2.7	2.44	0.3
Neoplasms	1.4	1.1	0.3
Endocrine/nutritional/metabolic/immunity disorders	2.5	2.7	0.2
Diseases of blood & blood forming organs	0.3	0.2	0.1
Mental disorders	1.9	2.0	0.02
Nervous system/sense organ diseases	5.7	5.7	0.06
Cardiovascular/circulatory diseases	6.2	6.3	0.1
Respiratory system diseases	20.7	21.1	0.3
Digestive system diseases	4.2	4.0	0.2
Genitourinary system diseases	2.9	2.7	0.3
Complications of pregnancy/childbirth/puerperium	0.08	0.1	0.02
Skin & subcutaneous tissue diseases	5.6	5.0	0.6
Musculoskeletal & connective tissue diseases	8.0	7.9	0.07
Congenital anomalies	0.04	0.04	0
Symptoms, signs, ill-defined conditions & disab nec	6.1	5.9	0.2
Injury & poisoning	5.6	4.9	0.7
Not illness/unspecified	26.1	28.0	1.8
		Average error	0.3

Table 9.5.8 Distribution of number of GP visits by 'most important condition'

	Conditionc ategory	Number of GP visits									
		1	2	3	4	5	6	7	8	9	10
ANHS	1	0.7204	0.2049	0.0258	0.0316	0.0172	-.	-	-	-	-
Simulation	1	0.70981	0.22049	0.02082	0.032367	0.01651	-.	-	-	-	-
ANHS	2	0.5746	0.3168	0.0862	0.0134	0.0057	0.0032	-	-	-	-
Simulation	2	0.57509	0.31565	0.07873	0.013999	0.011502	0.00502	-	-	-	-.
ANHS	3	0.7551	0.1621	0.0649	0.0063	0.0035	.-	0.0031	.	0.0049	.-
Simulation	3	0.78691	0.14462	0.04707	0.007967	0.005023	-	0.003962	.	0.004453	.-
ANHS	4	0.7729	0.1403	0.0266	-.	0.0603	-	-.	-.	-.	-.
Simulation	4	0.76181	0.16353	0.00779	-.	0.066867	.-	-.	- .	-	.-
ANHS	5	0.6691	0.2087	0.062	0.046	0.0091	-	.-	-.	-.	0.0051
Simulation	5	0.65764	0.22007	0.07712	0.032337	0.003721	.-	. -	. -	. -	0.009115781
ANHS	6	0.6934	0.2302	0.0451	0.0187	0.0034	.-	0.0073	-.	-.	0.0019
Simulation	6	0.68904	0.23154	0.04725	0.020399	0.003832	.-	0.006815	- .	-.	0.001121033
ANHS	7	0.7769	0.17	0.0308	0.0128	0.0058	.-	0.0013	.-	-.	0.0024
Simulation	7	0.77444	0.17001	0.03409	0.014996	0.003021	.-	0.001543	. -	. -	0.001897017
ANHS	8	0.7981	0.1566	0.0285	0.0115	0.0033	0.0015	-.	-.	-.	0.0004
Simulation	8	0.79501	0.15916	0.02673	0.013461	0.003752	0.00162	-.	.-	-.	0.00268123
ANHS	9	0.7153	0.2006	0.0573	0.0111	.-	0.0063	-.	.-	-.	0.0094
Simulation	9	0.71514	0.20994	0.05401	0.010085	. -	0.00487	. -	. -	. -	0.005955143
ANHS	10	0.6858	0.2287	0.0587	0.0165	0.0036	0.0067	-.	.-	-.	-.
Simulation	10	0.69278	0.22644	0.05689	0.018011	0.001442	0.00444	-.	- .	-.	.-
ANHS	11	0.2087	0.4682	0.0954	.-	0.0201	0.2076	-.	.-	-.	-.
Simulation	11	0.15011	0.41067	0.15205	-.	0.001733	0.28545	-.	- .	-.	.-
ANHS	12	0.7911	0.1455	0.0402	0.0073	0.0083	0.0023	0.0053	.-	-.	-.
Simulation	12	0.77975	0.15265	0.04775	0.00501	0.00637	0.00282	0.005654	- .	-.	.-
ANHS	13	0.6451	0.2508	0.056	0.0322	0.0074	0.0024	0.0009	.-	0.0011	0.0042
Simulation	13	0.639	0.26148	0.05631	0.029562	0.006995	0.00039	0.000401	.-	0.001127	0.00474247
ANHS	14	0.3576	0.6424	.-	-.	-.	-	.-	-.	-.	-
Simulation	14	0.39928	0.60072	- .	-.	.-	-	- .	-.	.-	-
ANHS	15	0.7243	0.1925	0.0501	0.0313	.-	-.	-.	-.	-.	0.0018
Simulation	15	0.72395	0.19277	0.05019	0.031042	- .	-.	.-	. -	. -	0.00204547
ANHS	16	0.637	0.2632	0.0568	0.0277	0.0092	0.0053	-.	-.	-.	0.0008
Simulation	16	0.63114	0.26413	0.05639	0.028796	0.011894	0.00764	-.	- .	-.	0.000007663
ANHS	17	0.8336	0.1396	0.0163	0.0066	0.0013	0.0007	0.0019	-.	-.	.-
Simulation	17	0.8377	0.13825	0.01482	0.005598	0.00098	0.00031	0.002339	.-	. -	.-

Table 9.5.9 Average percentage of population with >=1 GP visit per fortnight

	Aus Health Survey	Simulation 1st run	Absolute error
Overall	21.0	21.2	0.2
Age group			
0-24	17.3	17.4	0.1
25-44	18.8	18.7	0.1
45-64	22.4	23.2	0.9
65+	35.6	34.6	1.0
Gender			
Female	23.5	23.3	0.2
Male	18.5	18.9	0.4
Household type			
Live W ITHOUT someone >=15yrs	27.6	28.0	0.4
Live WITH someone >=15yrs and partnered	22.2	22.2	0.1
Live WITH someone >=15yrs and NOT partnered	17.9	18.3	0.4

Table 9.5.10 Average percentage of population with no GP visit per fortnight

	Aus Health Survey	Simulation 1st run	Absolute error
Overall	79.0		
Age group			
0-24	82.7	82.6	0.1
25-44	81.2	81.3	0.1
45-64	77.7	76.8	0.9
65+	64.4	65.4	1.0
Gender			
Female	76.5	76.7	0.2
Male	81.5	81.1	0.4
Household type			
Live W ITHOUT someone >=15yrs	72.4	72.0	0.4
Live WITH someone >=15yrs and partnered	77.8	77.8	0.01
Live WITH someone >=15yrs and NOT partnered	82.1	81.7	0.4

Table 9.5.11 Average percentage of population with a condition but no GP visit per fortnight

	Aus Health Survey	Simulation 1st run	Absolute error
Overall	72.3		
Age group			
0-24	73.6	74.9	1.3
25-44	75.3	76.5	1.2
45-64	73.3	73.5	0.2
65+	61.8	63.7	1.9
Gender			
Female	70.8	72.0	1.2
Male	74.1	74.9	0.8
Household type			
Live W ITHOUT someone >=15yrs	67.9	68.9	1.0
Live WITH someone >=15yrs and partnered	72.7	73.6	0.9
Live WITH someone >=15yrs and NOT partnered	73.4	74.5	1.1

9.6 Validation tables: simulated results compared to GP survey (NPMCS 2001-2)

9.6.1 Conditions - by age group

Table 9.6.1.1 Distribution of conditions for patients aged 0-24 years

Condition category	GP Survey: 0-24 yrs	Simulation 1st run: 0-24 yrs	Absolute error
Infectious & parasitic diseases	8.5	4.5	4.1
Neoplasms	0.8	0.4	0.4
Endocrine/nutritional/metabolic/immunity disorders	0.7	1.1	0.4
Diseases of blood & blood forming organs	0.5	0.4	0.1
Mental disorders	2.8	2.2	0.6
Nervous system/sense organ diseases	13.0	7.9	5.1
Cardiovascular/circulatory diseases	0.7	1.6	0.9
Respiratory system diseases	26.8	26.6	0.3
Digestive system diseases	4.4	7.1	2.7
Genitourinary system diseases	3.2	2.8	0.5
Complications of pregnancy/childbirth/puerperium	0.3	0.1	0.2
Skin & subcutaneous tissue diseases	9.2	7.9	1.3
Musculoskeletal & connective tissue diseases	1.4	4.8	3.4
Congenital anomalies	0.4	0.2	0.2
Symptoms, signs, ill-defined conditions & disab nec	2.8	3.9	1.1
Injury & poisoning	7.6	7.5	0.02
Not illness/unspecified	17.0	21.1	4.1
Total	**100%**	**100%**	
		Average error	1.5

Table 9.6.1.2 Distribution of conditions for patients aged 25-44 years

Condition category	GP Survey: 25-44 yrs	Simulation 1st run: 25-44 yrs	Absolute error
Infectious & parasitic diseases	4.7	3.2	1.5
Neoplasms	2.1	0.8	1.3
Endocrine/nutritional/metabolic/immunity disorders	3.2	3.1	0.1
Diseases of blood & blood forming organs	0.6	0.7	0.2
Mental disorders	7.2	3.8	3.4
Nervous system/sense organ diseases	7.0	6.3	0.8
Cardiovascular/circulatory diseases	3.7	3.6	0.1
Respiratory system diseases	12.6	18.0	5.4
Digestive system diseases	4.3	6.0	1.7
Genitourinary system diseases	6.6	4.4	2.2
Complications of pregnancy/childbirth/puerperium	0.7	0.3	0.4
Skin & subcutaneous tissue diseases	5.8	6.1	0.3
Musculoskeletal & connective tissue diseases	5.1	9.0	3.9
Congenital anomalies	0.2	0.1	0.1
Symptoms, signs, ill-defined conditions & disab nec	4.0	4.9	0.9
Injury & poisoning	7.9	6.1	1.9
Not illness/unspecified	24.4	23.7	0.7
Total	**100%**	**100%**	
		Average error	1.5

Table 9.6.1.3 Distribution of conditions for patients aged 45-64 years

Condition category	GP Survey: 45-64 yrs	Simulation 1st run: 45-64 yrs	Absolute error
Infectious & parasitic diseases	2.4	1.8	0.6
Neoplasms	3.2	1.2	2.1
Endocrine/nutritional/metabolic/immunity disorders	6.5	9.0	2.4
Diseases of blood & blood forming organs	0.2	0.5	0.3
Mental disorders	5.6	3.5	2.1
Nervous system/sense organ diseases	6.3	4.7	1.5
Cardiovascular/circulatory diseases	13.1	12.4	0.7
Respiratory system diseases	9.9	11.7	1.9
Digestive system diseases	4.2	6.7	2.5
Genitourinary system diseases	4.9	4.4	0.5
Complications of pregnancy/childbirth/puerperium	0.2	0.02	0.1
Skin & subcutaneous tissue diseases	4.7	5.2	0.5
Musculoskeletal & connective tissue diseases	7.6	11.6	4.0
Congenital anomalies	0.07	0.03	0.04
Symptoms, signs, ill-defined conditions & disab nec	4.3	3.2	1.1
Injury & poisoning	6.8	3.5	3.3
Not illness/unspecified	20.1	20.6	0.5
Total	**100%**	**100%**	
		Average error	1.4

Table 9.6.1.4 Distribution of conditions for patients aged 65+ years

Condition category	GP Survey: 65+yrs	Simulation 1st run: 65+ yrs	Absolute error
Infectious & parasitic diseases	1.4	1.3	0.1
Neoplasms	3.7	2.1	1.6
Endocrine/nutritional/metabolic/immunity disorders	5.9	8.2	2.3
Diseases of blood & blood forming organs	0.5	0.9	0.4
Mental disorders	4.2	2.3	1.9
Nervous system/sense organ diseases	6.4	5.9	0.4
Cardiovascular/circulatory diseases	19.8	20.2	0.4
Respiratory system diseases	9.2	8.6	0.6
Digestive system diseases	4.9	7.2	2.4
Genitourinary system diseases	4.1	2.1	2.0
Complications of pregnancy/childbirth/puerperium	0	0.02	0.02
Skin & subcutaneous tissue diseases	6.7	3.9	2.8
Musculoskeletal & connective tissue diseases	8.9	11.9	3.1
Congenital anomalies	0.06	0.07	0.01
Symptoms, signs, ill-defined conditions & disab nec	3.1	2.1	1.0
Injury & poisoning	6.2	2.6	3.7
Not illness/unspecified	15.1	20.7	5.6
Total	**100%**	**100%**	
		Average error	1.7

9.6.2 Conditions – by gender

Table 9.6.2.1 Distribution of conditions for female patients

Condition category	GP Survey: female	Simulation 1st run: female	Absolute error
Infectious & parasitic diseases	4.2	2.6	1.6
Neoplasms	2.3	1.1	1.2
Endocrine/nutritional/metabolic/immunity disorders	3.8	5.8	2.0
Diseases of blood & blood forming organs	0.5	0.9	0.4
Mental disorders	5.0	3.0	2.0
Nervous system/sense organ diseases	8.0	5.9	2.1
Cardiovascular/circulatory diseases	8.9	9.4	0.4
Respiratory system diseases	13.1	15.3	2.1
Digestive system diseases	4.3	6.3	2.0
Genitourinary system diseases	5.9	4.7	1.3
Complications of pregnancy/childbirth/puerperium	0.5	0.2	0.3
Skin & subcutaneous tissue diseases	6.5	5.6	0.9
Musculoskeletal & connective tissue diseases	5.9	9.3	3.4
Congenital anomalies	0.15	0.09	0.06
Symptoms, signs, ill-defined conditions & disab nec	3.7	3.7	0.1
Injury & poisoning	6.2	4.2	2.1
Not illness/unspecified	21.2	22.1	0.9
Total	**100%**	**100%**	
		Average error	1.3

Table 9.6.2.2 Distribution of conditions for male patients

Condition category	GP Survey: male	Simulation 1st run: male	Absolute error
Infectious & parasitic diseases	4.5	2.7	1.9
Neoplasms	2.7	1.2	1.5
Endocrine/nutritional/metabolic/immunity disorders	4.4	5.0	0.7
Diseases of blood & blood forming organs	0.4	0.3	0.1
Mental disorders	4.7	2.9	1.8
Nervous system/sense organ diseases	8.7	6.5	2.1
Cardiovascular/circulatory diseases	9.8	10.2	0.4
Respiratory system diseases	17.3	16.9	0.3
Digestive system diseases	4.6	7.4	2.8
Genitourinary system diseases	2.7	1.6	1.1
Complications of pregnancy/childbirth/puerperium	0.03	0.01	0.02
Skin & subcutaneous tissue diseases	7.0	5.9	1.1
Musculoskeletal & connective tissue diseases	5.4	9.7	4.4
Congenital anomalies	0.25	0.11	0.14
Symptoms, signs, ill-defined conditions & disab nec	3.3	3.2	0.1
Injury & poisoning	8.5	5.8	2.7
Not illness/unspecified	15.9	20.6	4.7
Total	**100%**	**100%**	
		Average error	1.5

9.6.3 Visits – by age group

Table 9.6.3.1 Percentage of population with >=1 visit in a year by age group

Age group (years)	NZHS 2002-3 (adults)	Base file (adults + children)	Simulation 1st run
0 - 24	75.9	78.3	98.9
25-44	76.2	76.4	99.2
45-64	83.0	83.0	99.8
65+	94.4	94.5	100.0

Table 9.6.3.2 Average number of visits per year by age group

Age group (years)	GP Survey: GP users	Simulation 1st run - mean visits p yr: population	Simulation 1st run: mean visits p yr: GP users	Aligned simulation: population	Aligned simulation: GP users
0 - 24	5.4	5.9	6.0	4.2	5.4
25-44	5.3	6.7	6.7	4.8	6.0
45-64	6.6	8.2	8.2	5.9	7.4
65+	9.8	12.2	12.2	8.8	11.0

9.6.4 Visits – by gender

Table 9.6.4.1 Average number of visits per year by gender

Gender	GP Survey: GP users	Simulation 1st run - mean visits p yr: population	Simulation 1st run - mean visits p yr: GP users	Aligned simulation: population	Aligned simulation: GP users
Female	7.0	8.2	8.3	5.9	7.4
Male	6.1	6.5	6.5	4.7	5.9

Table 9.6.4.2 Percentage of population with >=1 visit - by gender

Gender	NZHS 2002-3 (adults)	Base file (adults + children)	Simulation 1st run
female	85.0	84.7	99.4
male	76.2	76.6	99.3

9.6.5 Visits – by household type

Table 9.6.5.1 Percentage of population with >=1 visit per year by household type

Household type	NZHS 2002-3 (adults)	Base file (adults + children)	Simulation 1st run
Live W ITHOUT someone >=15yrs	84.6	84.9	99.9
Live WITH someone >=15yrs and partnered	81.4	81.6	99.6
Live WITH someone >=15yrs and NOT partnered	77.3	78.8	98.9

Table 9.6.5.2 Average number of visits per year by household type

Household type	Simulation 1st run - mean visits p yr: population	Simulation 1st run - mean visits p yr: GP users	Aligned simulation - mean visits p yr: population	Aligned simulation - mean visits p yr: GP users
Live W ITHOUT someone >=15yrs	10.0	10.1	7.2	9.1
Live WITH someone >=15yrs and partnered	7.8	7.8	5.6	7.1
Live WITH someone >=15yrs and NOT partnered	6.3	6.3	4.5	5.7

9.6.6 Doctor actions – by primary diagnosis

Table 9.6.6.1 Percentage of visits per year with an investigation by primary diagnosis

Primary Diagnosis of visit	GP Survey: % visits investigation	Simulation 1st run: % visits investigation	Absolute error
Infectious & parasitic diseases	27.3	29.6	2.3
Neoplasms	27.4	29.5	2.1
Endocrine/nutritional/metabolic/immunity disorders	47.9	48.4	0.5
Diseases of blood & blood forming organs	59.7	77.4	17.7
Mental disorders	25.1	24.3	0.8
Nervous system/sense organ diseases	11.8	14.7	2.9
Cardiovascular/circulatory diseases	32.6	33.1	0.5
Respiratory system diseases	15.9	18.6	2.7
Digestive system diseases	34.0	49.7	15.7
Genitourinary system diseases	50.1	36.1	14.0
Complications of pregnancy/childbirth/puerperium	35.9	51.4	15.5
Skin & subcutaneous tissue diseases	18.5	22.4	3.9
Musculoskeletal & connective tissue diseases	30.8	34.0	3.2
Congenital anomalies	7.5	10.8	3.3
Symptoms, signs, ill-defined conditions & disab nec	35.6	38.1	2.5
Injury & poisoning	15.6	17.6	2.0
Not illness/unspecified	27.1	29.9	2.8
		Average Error	5.4

Table 9.6.6.2 Percentage of visits per year with a prescription by primary diagnosis

Primary Diagnosis of visit	GP Survey: % visits prescription	Simulation 1st run: % visits prescription	Absolute error
Infectious & parasitic diseases	64.2	67.2009	3.0
Neoplasms	40.9	49.1	8.2
Endocrine/nutritional/metabolic/immunity disorders	81.2	77.2	4.0
Diseases of blood & blood forming organs	39.8	51.6	11.8
Mental disorders	72.8	73.5	0.7
Nervous system/sense organ diseases	70.0	70.8	0.8
Cardiovascular/circulatory diseases	86.1	78.4	7.7
Respiratory system diseases	84.5	81.6	2.9
Digestive system diseases	71.4	63.8	7.6
Genitourinary system diseases	66.5	70.1	3.6
Complications of pregnancy/childbirth/puerperium	35.7	46.3	10.6
Skin & subcutaneous tissue diseases	74.5	73.1	1.4
Musculoskeletal & connective tissue diseases	68.8	65.8	3.0
Congenital anomalies	68.0	72.8	4.8
Symptoms, signs, ill-defined conditions & disab nec	51.3	51.2	0.1
Injury & poisoning	49.4	51.3	1.9
Not illness/unspecified	41.9	44.8	2.9
		Average Error	4.4

Table 9.6.6.3 Percentage of visits per year with a non-drug treatment by primary diagnosis

Primary Diagnosis of visit	GP Survey: % visits non-drug	Simulation 1st run: % visits non-drug	Absolute error
Infectious & parasitic diseases	58.5	61.5	3.0
Neoplasms	81.7	79.9	1.8
Endocrine/nutritional/metabolic/immunity disorders	63.8	63.2	0.6
Diseases of blood & blood forming organs	57.2	74.8	17.6
Mental disorders	71.7	61.0	10.7
Nervous system/sense organ diseases	53.9	57.8	3.9
Cardiovascular/circulatory diseases	53.5	59.1	5.6
Respiratory system diseases	46.9	52.4	5.5
Digestive system diseases	65.5	75.3	9.8
Genitourinary system diseases	77.9	68.7	9.2
Complications of pregnancy/childbirth/puerperium	75.9	87.4	11.5
Skin & subcutaneous tissue diseases	58.0	59.1	1.1
Musculoskeletal & connective tissue diseases	72.3	68.7	3.6
Congenital anomalies	83.3	57.7	25.6
Symptoms, signs, ill-defined conditions & disab nec	70.7	69.1	1.6
Injury & poisoning	74.5	71.1	3.4
Not illness/unspecified	69.3	65.6	3.7
		Average Error	6.9

Table 9.6.6.4 Percentage of visits per year with a follow-up by primary diagnosis

Primary Diagnosis of visit	GP Survey: % visits follow-up	Simulation 1st run: % visits follow-up	Absolute error
Infectious & parasitic diseases	36.5	41.2	4.7
Neoplasms	81.8	83.2	1.4
Endocrine/nutritional/metabolic/immunity disorders	74.8	74.3	0.5
Diseases of blood & blood forming organs	90.7	85.5	5.2
Mental disorders	81.4	78.6	2.8
Nervous system/sense organ diseases	57.6	60.0	2.4
Cardiovascular/circulatory diseases	78.0	79.2	1.2
Respiratory system diseases	41.7	45.7	4.0
Digestive system diseases	61.3	65.0	3.7
Genitourinary system diseases	61.7	67.4	5.7
Complications of pregnancy/childbirth/puerperium	65.6	44.1	21.5
Skin & subcutaneous tissue diseases	52.4	54.4	2.0
Musculoskeletal & connective tissue diseases	65.8	67.8	2.0
Congenital anomalies	80.8	81.1	0.3
Symptoms, signs, ill-defined conditions & disab nec	63.3	65.0	1.7
Injury & poisoning	59.4	60.4	1.0
Not illness/unspecified	51.5	60.3	8.8
		Average Error	4.1

Table 9.6.6.5 Percentage of visits per year with a referral by primary diagnosis

Primary Diagnosis of visit	GP Survey: % visits with referral	Simulation 1st run: % visits referral	Absolute error
Infectious & parasitic diseases	4.2	4.8	0.6
Neoplasms	18.9	21.7	2.8
Endocrine/nutritional/metabolic/immunity disorders	21.3	23.6	2.3
Diseases of blood & blood forming organs	8.1	15.5	7.4
Mental disorders	24.6	25.8	1.2
Nervous system/sense organ diseases	14.0	18.6	4.6
Cardiovascular/circulatory diseases	15.4	19.0	3.6
Respiratory system diseases	5.8	7.1	1.3
Digestive system diseases	18.0	33.1	15.1
Genitourinary system diseases	29.8	25.1	4.7
Complications of pregnancy/childbirth/puerperium	26.5	38.3	11.8
Skin & subcutaneous tissue diseases	10.6	11.7	1.1
Musculoskeletal & connective tissue diseases	33.9	36.6	2.7
Congenital anomalies	39.6	33.8	5.8
Symptoms, signs, ill-defined conditions & disab nec	19.5	25.4	5.9
Injury & poisoning	24.7	28.8	4.1
Not illness/unspecified	14.5	16.8	2.3
		Average Error	4.5

9.6.7 Doctor actions – by age group

Table 9.6.7.1 Percentage of visits per year with an investigation by age group

Age group (years)	GP Survey: % visits investigation	Simulation 1st run: % visits investigation	Absolute error
0 - 24	14.6	16.3	1.7
25-44	31.2	33.4	2.2
45-64	31.7	34.5	2.8
65+	26.4	27.9	1.5
		Average error	2.1

Table 9.6.7.2 Percentage of visits per year with a prescription by age group

Age group (years)	GP Survey: % visits prescription	Simulation 1st run: % visits prescription	Absolute error
0 - 24	65.2	63.7	1.5
25-44	61.1	61.7	0.6
45-64	67.5	66.5	1.0
65+	71.9	67.7	4.2
		Average error	1.8

Table 9.6.7.3 Percentage of visits per year with a non-drug treatment by age group

Age group (years)	GP Survey: % visits non-drug	Simulation 1st run: % visits non-drug	Absolute error
0 - 24	54.9	56.2	1.3
25-44	69.3	65.8	3.5
45-64	65.1	65.4	0.3
65+	62.4	64.4	2.0
		Average error	1.8

Table 9.6.7.4 Percentage of visits per year with a follow-up by age group

Age group (years)	GP Survey: % visits follow-up	Simulation 1st run: % visits follow-up	Absolute error
0 - 24	42.6	47.3	4.7
25-44	55.6	57.4	1.8
45-64	63.7	65.4	1.7
65+	73.8	76.0	2.2
		Average error	2.6

Table 9.6.7.5 Percentage of visits per year with a referral by age group

Age group (years)	GP Survey: % visits referral	Simulation 1st run: % visits referral	Absolute error
0 - 24	9.8	12.9	3.1
25-44	22.2	24.7	2.5
45-64	16.9	19.5	2.6
65+	17.0	17.7	0.7
		Average error	

9.6.8 Doctor actions – by gender

Table 9.6.8.1 Percentage of visits per year with an investigation by gender

Gender	GP Survey: % visits investigation	Simulation 1st run: % visits investigation	Absolute error
female	26.7	29.4	2.7
male	22.1	25.3	3.2
		Average error	3.0

Table 9.6.8.2 Percentage of visits per year with a prescription by gender

Gender	GP Survey: % visits prescription	Simulation 1st run: % visits prescription	Absolute error
female	66.2	65.3	0.9
male	66.5	63.7	2.8
		Average error	1.9

Table 9.6.8.3 Percentage of visits per year with a non-drug treatment by gender

Gender	GP Survey: % visits non-drug	Simulation 1st run: % visits non-drug	Absolute error
female	63.6	64.5	0.9
male	60.1	60.2	0.1
		Average error	0.5

Table 9.6.8.4 Percentage of visits per year with a follow-up by gender

Gender	GP Survey: % visits follow-up	Simulation 1st run: % visits follow-up	Absolute error
female	58.2	61.2	3.0
male	55.7	58.7	3.0
		Average error	3.0

Table 9.6.8.5 Percentage of visits per year with referral by gender

Gender	GP Survey: % visits referral	Simulation 1st run: % visits referral	Absolute error
female	16.3	18.5	2.2
male	15.2	18.8	3.6
		Average error	2.9

9.6.9 Doctor actions – by ethnicity

Table 9.6.9.1 Percentage of visits per year with an investigation by ethnicity

Ethnicity	GP Survey: % visits investigation	Simulation 1st run: % visits investigation	Absolute error
European	25.5	28.4	2.8
Maori	21.0	24.9	3.9
Pacific	17.7	19.9	2.1
Asian	28.9	30.5	1.6
Other	26.7	27.9	1.2
		Average error	2.3

Table 9.6.9.2 Percentage of visits per year with a prescription by ethnicity

Ethnicity	GP Survey: % visits prescription	Simulation 1st run: % visits prescription	Absolute error
European	65.4	64.9	0.5
Maori	69.6	64.0	5.6
Pacific	71.2	69.6	1.6
Asian	68.4	58.8	9.6
Other	65.9	61.8	4.1
		Average error	4.3

Table 9.6.9.3 Percentage of visits per year with a non-drug treatment by ethnicity

Ethnicity	GP Survey: % visits non-drug	Simulation 1st run: % visits non-drug	Absolute error
European	63.8	63.7	0.1
Maori	61.2	60.6	0.6
Pacific	47.5	54.7	7.2
Asian	49.7	60.3	10.6
Other	62.6	58.7	3.9
		Average error	4.5

Table 9.6.9.4 Percentage of visits per year with a follow-up by ethnicity

Ethnicity	GP Survey: % visits follow-up	Simulation 1st run: % visits follow-up	Absolute error
European	57.8	61.3	3.5
Maori	54.6	58.3	3.7
Pacific	52.3	59.8	7.5
Asian	49.6	47.6	2.0
Other	67.3	63.1	4.2
		Average error	4.2

Table 9.6.9.5 Percentage of visits per year with a referral by ethnicity

Ethnicity	GP Survey: % visits referral	Simulation 1st run: % visits referral	Absolute error
European	16.7	19.1	2.4
Maori	14.7	18.6	3.9
Pacific	10.2	15.0	4.8
Asian	12.7	16.8	4.1
Other	15.6	11.5	4.1
		Average error	3.9