Modelling Usage and Cost of NSW Hospitals - An Overview

Linc Thurecht
National Centre for Social and Economic Modelling, University of Canberra

International Microsimulation Conference on Population, Ageing and Health: Modelling Our Future

Canberra, December 2003
**Background**

- Three year ARC SPIRT grant
- Industry partners:
  - NSW Health
  - Productivity Commission
  - Health Insurance Commission
- Modelling NSW hospital usage and effect of private health insurance
The Data

- All in-patient separations from public and private NSW hospitals:
  - 1996-97 to 1999-00
  - Patient based
  - Geo-coded down to CD level
  - Cost of each separation
  - Diagnoses details for each separation
Unique Measure of Socioeconomic Status

- Socioeconomic status (SES) imputed at the patient level:
  - Equivalent Family Income quintiles
  - Actual distribution from the 1996 Census for the full NSW population
  - SES imputed by:
    - Census Collection District
    - Sex
    - Age
Hospitalisation per 100 of NSW Population by Age and SES, 1999-00

Public Hospitals and CHD: Average Patient Costs, Public vs Private Patients

Projected Hospital Costs in 2009-10

Modelling Private Health Insurance

A - Assuming 30% PHI Rebate, Medicare Levy Surcharge and Lifetime Health Cover not introduced
B - Assuming 30% PHI Rebate, Medicare Levy Surcharge and Lifetime Health Cover maintained
C - Assuming 30% PHI Rebate removed

## Lifetime Health Cover and Hospital Usage

<table>
<thead>
<tr>
<th></th>
<th>0 - 24</th>
<th>25 - 34</th>
<th>35 - 54</th>
<th>55 - 74</th>
<th>75+</th>
<th>All ages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Public hospitals</strong> %</td>
<td>-3.00</td>
<td>-9.03</td>
<td>-18.03</td>
<td>-21.60</td>
<td>-15.10</td>
<td>-13.60</td>
</tr>
<tr>
<td><strong>Private hospitals</strong> %</td>
<td>9.98</td>
<td>25.42</td>
<td>29.83</td>
<td>37.38</td>
<td>35.99</td>
<td>29.66</td>
</tr>
</tbody>
</table>

**Difference in 2009-10 – Scenario A to Scenario C**

No 30% PHI Rebate, Lifetime Health Cover or Medicare Levy Surcharge

vs

Lifetime Health Cover and the Medicare Levy Surcharge
(but no 30% PHI Rebate)

Average Per Capita Expenditure on Public Hospital - 70 Years and Over, 1998-99

Further Details

  - A Microsimulation Model of Hospital Patients: New South Wales
  - Hospitalisation Rates and Costs by Socioeconomic Status
  - NSW Hospitals: Are They 'Pro-Rich' or 'Pro-Poor'?
  - Projecting the Fiscal Impact of Population Ageing on the Hospital System: A Distributional Analysis
  - Public Policy and Private Health Insurance - Distributional Impact on Public and Private Hospital Usage in NSW
  - Socioeconomic Characteristics of NSW Hospital Users in 1998-99
Modelling Usage and Cost of NSW Hospitals – An Overview

Linc Thurecht

National Centre for Social and Economic Modelling, University of Canberra

ABSTRACT: This paper provides an overview of a socioeconomic model of hospital users in New South Wales developed as part of a three year Australia Research Council grant. The model is based on four years of separations from both public and private hospitals, including New South Wales residents treated outside of the state. A unique patient ID is attached to each separation enabling related hospital separations to be attributed to the same (anonymous) patient. A unique measure of socioeconomic status is then imputed for each patient based on the empirical distribution of equivalent family income on census night of a person living in the same geographic area and of the same sex and age. Costs are also calculated for each separation enabling the total patient cost of treatment to be identified, including *inter alia* average costs of treatment by socioeconomic status, type of hospital or disease. Spatial analysis can be performed at levels as low as an average of 200 households. A model of private health insurance (PHI) is also created to estimate the likelihood that a given individual will purchase PHI, and given the need to attend hospital, whether they will choose a public or private hospital. A facility is also developed that enables projected patient numbers and associated costs to be estimated.
**Objective of the Model**

To understand the socioeconomic characteristics of NSW hospital patients, project future hospital usage and model the effect of private health insurance on both public and private hospital patient numbers.

**Target Audience**

Public policy makers in the health field, and in particular, those interested in distributional issues relating to hospital usage and private health insurance.

**Method and Platform**

Starting with a database of all in-patient separations from NSW public and private hospitals over the four year period 1996-97 to 1999-00, a patient identifier is added, patient records are geo-coded and the estimated cost of each separation is calculated. Socioeconomic status is then imputed onto each patient record based on the empirical distribution of the socioeconomic status of the NSW population on census night. Separate models are then built to project hospital usage and the effect of private health insurance on the choice between public and private hospitals.

The model was constructed using SAS on a Windows PC.

**Policy Environment**

Total expenditure on health in Australia in 2001-02 was estimated to be $66.6 billion, or 9.3% of GDP, with real growth averaging 4.6% over the past decade (AIHW 2003). Hospitals comprise nearly a third of this with estimated expenditures of $21.5 billion in 2001-02.

With increasing pressure on the hospitals sector and declining private health insurance membership, the government has introduced a number of policies designed to increase the take-up of private health insurance. One of these initiatives, the 30% rebate on private health insurance cost the government $2.3 billion in 2002-03 (DHA 2003).

**References**

See references at the end of this document.
Introduction

This project is being conducted under an Australian Research Council (ARC) Strategic Partnerships with Industry Research and Training (SPIRT) grant entitled ‘Health Policy and Socioeconomic Status in Australia: Improving Decision Support Tools’. The industry partners in this jointly sponsored grant are the New South Wales Department of Health, the Productivity Commission and the Health Insurance Commission.

The major aims of the project are to impute socioeconomic status and patient-level expenditures to all people who attend a New South Wales (NSW) hospital as an in-patient and, following on from this, to carry out a series of distributional analyses of the NSW hospital patient population. By imputing socioeconomic status onto individual-based administrative data — rich in information on services provided and associated costs — we expected that distributional issues could be analysed in considerably greater detail and in more complex ways than have previously been possible.

Other aims are to develop a model to project hospital usage and to construct a private health insurance model the purpose of which is to estimate private health insurance coverage and the split of usage and expenditures between private and public hospitals under current and possible future policy parameters.

Data Overview

The base dataset includes all in-patient separations from NSW hospitals, both public and private, over the four year period from 1996-97 to 1999-00. Data items include demographic and administrative information, and details of diagnoses and procedures associated with each separation. A full description of the datasets is provided in Thurecht et al (2003a). Certain important aspect of the datasets are briefly discussed below.

Each record has a patient identifier attached to it enabling separations from the same patient to identified. The matching is performed using AutoMatch which is a generalised record linkage system based on probabilistic record linkage methods (refer to MatchWare Technologies (1997)). A useful overview of the probabilistic linking of data and the AutoMatch software is provided by Jaro (1995).

Each patient record is also geo-coded internally by NSW Health at both the Statistical Local Area (SLA) and Census Collection District (CD) level.1 Geo-coded data permits both spatial comparisons of interest (eg rates of hospitalisation in urban areas compared to rural or remote areas), and enabled a more accurate measure of socioeconomic status to imputed onto the patient records.

The cost of each separation is calculated internally by the NSW Department of Health based on the NSW Hospital Cost Data Collection and NSW Private Hospital Cost

1 A CD is a spatial area comprising on average around 200 households. In NSW an SLA contains on average around 60 CDs.
Collection. An adjustment is then made to reflect any revenue the government may have received for treating the patient. This may arise, for example, if a patient was treated for an injury that was covered by workers’ compensation.

A number of steps were taken to maximise the integrity of the data supplied by NSW Department of Health, including the attribution of missing values in one separation based on populated values of other separations by the same patient. These steps are fully outlined in Thurecht et al (2003a).

**Imputing a Unique Measure of Socioeconomic Status**

Distributional studies typically use the Australian Bureau of Statistics (ABS) Socio-Economic Indexes for Areas (SEIFAs) to assess differences in the socioeconomic characteristics of the group under examination. These indexes are based on census unit record files and relate to discrete geographic areas. A full description of how the indexes were constructed is provided in ABS (1998).

While the ABS has constructed five SEIFAs, the Index of Relative Socioeconomic Disadvantage (IRSED) is the most commonly used measure of socioeconomic status. However, a major shortcoming of the SEIFAs is that they assign the same socioeconomic status to all people in the same geographic area. This ignores the differences that may be related to the individual (income, education, type of work etc) and to their family circumstances (single, with children etc). This implies that two individuals in the same geographic area with fundamentally different circumstances would be allocated the same index of socioeconomic status without regard to their underlying heterogeneity.

To address this shortcoming, we obtained a custom extract from the 1996 Census unit record files of the equivalent family income (EFI) of the entire NSW population. This extract was taken at the CD-Sex-Age level to provide the actual empirical distribution of socioeconomic status of the entire population.\(^2\) To preserve confidentiality, in addition to rendering a more tractable metric, only EFI quintiles were reported by the ABS. The modified OECD scale was used to calculate EFI (refer to Mejer and Siermann 2000).

While a full description of how socioeconomic status was imputed onto the patient records is provided in Thurecht et al (2003a), a brief description follows.

The EFI was first calculated for each individual on Census night. All EFIs were then ranked and a quintile assigned. The number of people within the same EFI quintile where then counted for each CD, sex and age combination. An EFI quintile was then imputed onto the patient record based on the likelihood of a person from a particular CD and of a given sex and age being in a particular EFI quintile. The imputation was based on a uniformly distributed random number being less than the cumulative probability of a person from the CD-sex-age cell belonging to a particular EFI quintile. The following example is provided:

---

\(^2\) Age was defined in ten year age groups up to age 69 and then 70 years of age and above.
Empirical Distribution from Census Unit Record Files

<table>
<thead>
<tr>
<th>CD</th>
<th>Sex</th>
<th>Age</th>
<th>EFI Quintile</th>
<th>Number of People</th>
<th>Probability^a</th>
<th>Cumulative Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1251405</td>
<td>Male</td>
<td>0–9</td>
<td>1</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1251405</td>
<td>Male</td>
<td>0–9</td>
<td>2</td>
<td>10</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>1251405</td>
<td>Male</td>
<td>0–9</td>
<td>3</td>
<td>0</td>
<td>0.0</td>
<td>0.5</td>
</tr>
<tr>
<td>1251405</td>
<td>Male</td>
<td>0–9</td>
<td>4</td>
<td>0</td>
<td>0.0</td>
<td>0.5</td>
</tr>
<tr>
<td>1251405</td>
<td>Male</td>
<td>0–9</td>
<td>5</td>
<td>10</td>
<td>0.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

^a ‘Probability’ refers to the proportion of people in that CD–sex–age–EFI quintile cell.

Imputation of Socioeconomic Status

<table>
<thead>
<tr>
<th>CD</th>
<th>Person ID</th>
<th>Sex</th>
<th>Age Years</th>
<th>z^b</th>
<th>EFI Quintile Assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>1251405</td>
<td>1</td>
<td>Male</td>
<td>0–9</td>
<td>0.4986</td>
<td>2</td>
</tr>
<tr>
<td>1251405</td>
<td>2</td>
<td>Male</td>
<td>0–9</td>
<td>0.5877</td>
<td>5</td>
</tr>
<tr>
<td>1251405</td>
<td>3</td>
<td>Male</td>
<td>0–9</td>
<td>0.8242</td>
<td>5</td>
</tr>
<tr>
<td>1251405</td>
<td>4</td>
<td>Male</td>
<td>0–9</td>
<td>0.2654</td>
<td>2</td>
</tr>
</tbody>
</table>

^b \(z \sim U[0,1]\).

It should be noted that there are certain limitations associated with this imputation method. The first is that while this approach will distribute EFI quintiles according to the empirical distribution on census night, there may be systematic links between socioeconomic status and hospital admissions that are not fully captured by the uniformly distributed assignment of socioeconomic status. The second is that there is no direct link between the socioeconomic status of a given patient from year to year. This is because in the NSW hospitals datasets patients can only be traced within financial years but not across financial years.

Projecting Hospital Usage

One of the objectives of the project was to develop a facility to project future hospital usage. The methodology developed was based on two factors – forecast population growth and change in the propensity to utilise particular hospital services. Usage was projected from the base year of 1999-00 out to 2009-10. While a more detailed description of the methodology is provided in Thurecht et al (2003b), an overview of the approach is provided below.

To calculate the effect of population growth on usage, weights were calculated at the SLA, sex and five year age group level. The weights were based on the ratio of the forecast population in 2009-10 and estimated resident population in 1999-00. For example, if the population in a particular SLA-sex-age cell was projected to double, then the population weight for that patient record in 2009-10 would be two.

To calculate the propensity to utilise particular services, the patients were grouped into related service types as shown in Thurecht et al (2003 a and b). This was done

^3 Forecast population growth was taken from ABS (2000) and the estimated resident population from ABS (2001).
separately for each of the four years from 1996-97 to 1999-00. The propensity for service usage was then taken as the number of patients at the SLA-sex-age level for each service, divided by the underlying estimated resident population for that cell in that year. This produced four observations for each cell from which a linear trend was then projected to estimate the propensity for service usage in each cell in 2009-10. Figure 2 demonstrates this approach. This estimated 2009-10 propensity was then divided by the actual propensity for service usage within each cell in 1999-00 to produce a service propensity weight for each cell in 2009-10. For example, if the estimated propensity for service usage for a particular SLA-sex-age-service type in 2009-10 was 0.75 and the actual rate in 1999-00 was 0.5, then a service trend weight of 1.5 would be assigned to that record.

Figure 2: Projecting SRG propensity

Note: Similar estimates of SRG propensity were taken for 1996-97 and 1998-99 as shown here for 1997-98 and 1999-00.

The final weight for each patient record in 2009-10 was then taken as the product of the 2009-10 population and service propensity weights. Using the above examples, the final weight would be three (two multiplied by 1.5). The aggregate weights were then calibrated against the projected number of separations forecast from the NSW Health Activity Projections Plus Interventions (APPI) model. APPI is an activity projection tool utilised by the NSW Department of Health in which age and sex standardised trends in acute inpatient activity are used to project activity for 2006-07 and 2011-12. Activity projections for 2009-10 were then interpolated based on the annual compound growth in the APPI projections between the two years.

By projecting the trend propensity we make some important implicit assumptions. First, that trends in patterns of disease do not change over the projection period. Note that we do not hold the pattern of disease static, but merely assume that the changes that occurred over the four year sample period will continue over the projection period. Second, that there is no change in health policies that impacts on the hospital sector.
One area in which this assumption is problematic relates to the introduction of Lifetime Health Cover for private health insurance in July 2000. However, the impact of the various government policies introduced in recent years relating to PHI — and the flow-on effect to hospitals — is addressed in the private health insurance model. Third, trends in the supply of hospital services over the sample period are assumed to be the same over the projection period. Finally, trends in the move towards the provision of out-patient treatments that have occurred over the four years to 1999-00 are assumed to continue over the projection period.

To estimate the likely costs associated with these usage projections, the estimated net cost of each separation in 1999-00 is adjusted for the likely increase in the cost of hospital services over the projection period. This was based on an OLS forecast from the Australian Bureau of Statistics weighted average of eight capital cities health component of the quarterly Consumer Price Index (refer to Table 5 of ABS (2003)).

Modelling Private Health Insurance

The private health insurance (PHI) model predicts the probability that a person with selected characteristics would have private health insurance. This is then applied to the projected number of hospital patients to determine the relative split of patients between public and private hospitals in the projection period. While a full description is given in Walker et al (2003a), a brief overview is provided below.

The model is based on a time series of people covered by PHI from 1983 to 2002, using data from the ABS Private Health Insurance and National Health surveys. Logistic regression was used with age, sex, gross income unit quintiles, a time index and PHI premium costs as the explanatory variables. A spline was fitted to the model around the time of the introduction of Lifetime Health Cover (July 2001) to account for the structural break in the series that occurred at this point in time. Through an iterative process, various transformations and interactions were included in the model specification.

The proportion of people with PHI by age, sex and gross income unit quintile was then estimated and applied to the aggregate projection of hospital usage in 2009-10 to determine the estimated number of patients that would be covered by private health insurance. For example, if 20% of female 25-34 year olds in the population are estimated to hold private health insurance, then 20% of all female 25-34 year old patients are assumed to have private health insurance.

---

4 The gross cost of each hospital separation was based on the NSW Hospital Cost Data Collection. Generally, the cost of each separation was either calculated at the separation level or based on an average cost in that facility for the treatment received by the patient. The average cost was primarily based on the DRG recorded for each separation. Net costs were calculated on the basis of the gross cost less any revenue the government may have received for providing the treatment eg treatment related to workers’ compensation.
The next step was to determine the split of patients between public and private hospitals. This was based on a syndicated survey conducted by TQA Research.\textsuperscript{5} This showed the likelihood that the respondent attended a public or private hospital, conditioned on whether they held private health insurance. Using the number of patients covered by private health insurance estimated above, the survey proportions were then used to determine the number of patients that would attend either a public or private hospital. That is, using survey responses conditioned on whether private health insurance was held or not, the separate populations of insured and uninsured patients were then split between public and private hospitals.

As outlined in Walker \textit{et al} (2003a), this model can then be applied to a number of different policy scenarios and to quantifying their impact across the public and private hospitals sector.

\textbf{Conclusion}

Starting with four years of data that cover all in-patient separations from both public and private hospitals in NSW, a model was built that can assess the distributional characteristics of hospital users, project future hospital usage and model the effect of private health insurance on the patient choice between attending a public or private hospital.

To accomplish this, a patient identifier was first added to each separation record to enable multiple separations by the one person in the same year to be traced. The data was then geo-coded to the CD level and the cost of each separation was estimated. A sharper measure of socioeconomic status was then imputed based on the empirical distribution of equivalent family income for the entire NSW population on census night.

A projection facility estimated hospital usage ten years into the future and showed the split between public and private hospitals. The private health insurance model then enabled these base projections to be adjusted for hypothesised changes to government policy with respect to private health insurance.

The model provides a deeper understanding of the hospital sector in NSW, and in particular the socioeconomic profile of patients. A number of applications have been developed to demonstrate the types of research and policy questions that can be modelled, and to illuminate the distributional and fiscal implications of different policy scenarios.

\textsuperscript{5} This "Survey of Health Care and Insurance, Australia, 2001" involved 5,194 comprehensive telephone interviews with a random sample of insurable unit heads from all areas of Australia.
References


——, 2003, "Consumer Price Index, Australia", catalogue number 6401.0, ABS, Canberra.


MatchWare Technologies 1997, AutoMatch Version 4.01 (software), Burtonsville, Maryland, USA.


——, Walker, A., Pearse, J. and Harding, A., 2003c, "NSW Hospitals: Are They 'Pro-Rich' or 'Pro-Poor'?", Health Outcomes Conference, Canberra, 20-21 August.