Diagnosis-based risk adjustment and Australian health system policy

Ronald Donato and Jeffrey Richardson

Abstract

Diagnosis-based risk adjustment is increasingly seen as an important tool for establishing capitation payments and evaluating appropriateness and efficiency of services provided and has become an important area of research for many countries contemplating health system reform.

This paper examines the application of a risk-adjustment method, extensively validated in the United States, known as diagnostic cost groups (DCG), to a large Australian hospital inpatient data set.

The data set encompassed hospital inpatient diagnoses and inpatient expenditure for the entire metropolitan population residing in the state of New South Wales. The DCG model was able to explain 34% of individual-level variation in concurrent expenditure and 5.2% in subsequent year expenditure, which is comparable to US studies using inpatient-only data. The degree of stability and internal consistency of the parameter estimates for both the concurrent and prospective models indicate the DCG methodology has face validity in its application to NSW health data sets. Modelling and simulations were conducted which demonstrate the policy applications and significance of risk adjustment model(s) in the Australian context.

This study demonstrates the feasibility of using large individual-level data sets for diagnosis-based risk adjustment research in Australia. The results suggest that a research agenda should be established to broaden the options for health system reform.

Aust Health Rev 2006: 30(1): 83–99

Ronald Donato, MEc, Lecturer (Economics)

School of Commerce, University of South Australia, Adelaide, SA.

Jeffrey Richardson, PhD, Foundation Director

Centre for Health Economics, Monash University, Melbourne, VIC.

Correspondence: Mr Ronald Donato, School of Commerce, University of South Australia, City West Campus, North Terrace Adelaide, Adelaide, SA 5000.
ronald.donato@unisa.edu.au

What is known about the topic?

Since the mid 1990s, many European countries have engaged in extensive research into health-related risk adjustment methods utilising diagnostic information. Risk adjustment is seen as an important mechanism for establishing weighted prospective capitation payments for both competitive and non-competitive health care systems and for evaluating the appropriateness and efficiency of services provided.

What does this paper add?

This is the first study using a large hospital data set for risk adjustment research in Australia. It shows the superiority of diagnosis-based risk adjustment methods over conventional approaches. The paper documents an improved method of assessing the efficiency of area health services, analysing capitated funding in a competitive environment and identifying high cost users for case management purposes.

What are the implications?

Diagnosis-based risk adjustment should be considered by policy makers at the federal level (for broadening health system reform options) and at a state level (for comparing area health services). ◆

Risk adjustment and capitation

POLICY MAKERS THROUGHOUT most of the industrialised world are grappling with ways to improve the efficiency and equity of their health care systems, to contain macro-level expenditure and to make their health systems more innovative and responsive to consumer preferences. To this end, risk adjustment methods are increasingly being seen as an important policy tool for both competitive and non-competitive population-based health care systems.[1] Risk adjustment accounts for differences in individual health risk and, hence, the expected health care resource requirements of patients. Risk adjustment can play an important role in establishing appropriate funding levels and ensuring resource allocation is made more equitable by being matched to the health care needs of the population.[1-4] It can also

assist in the assessment of provider performance. If the casemix adjustment is insensitive to differences in health care status of patients, providers might be penalised for treating sicker patients or be incorrectly assessed in terms of their relative efficiency in treating such patients.[5-7]

Various methods have been used for predicting resource requirements based on individual-level health risk. To date, most attention has been given to the development of models in which risk adjustment is based upon diagnostic information.[8] These methods use diagnostic information obtained from administrative data generated from patient encounters with the health system. These data are used to infer the patient's medical problem and, from this, their likely need for health services.[4] A substantial research agenda in the development and implementation of diagnosis-based risk adjustment methods has been under way in the United States since the early 1980s and in European countries since the mid 1990s.[3,9]

Although the purpose of risk adjustment is to pursue efficiency and equity objectives, the aims differ slightly depending on whether third-party purchasers operate in a competitive or non-competitive environment. In a competitive environment where health plans can compete for enrolees, such as those in the Netherlands, Israel, Belgium and the US Medicare system for example, efficiency in the operation of the insurance markets is of critical importance.[8,10] A fundamental consideration here is for risk adjustment to be used as the basis of individual capitation rates, and accordingly it must be able to explain variances in *predictable* expenditures of the individual. A good risk adjuster need not explain variances in expenditure associated with unpredictable or random events, as conventional insurance and risk-sharing arrangements are applied in such circumstances. Inadequate risk-adjusted capitation payments create incentives for health funds to exploit unpriced heterogeneity by avoiding those individuals where expected costs exceed what is paid and, conversely, selecting those enrolees where payments exceed expected costs. That is, there are perverse incentives to select good risk and avoid bad risk (ie, risk selection).

There is also growing interest in the use of individual-level data in non-competitive health systems, and research into the potential use of risk adjustment is now taking place in countries such as the UK, Canada and Spain.[5-7] In these countries, individual-level risk adjustment methods are seen as useful for devolving budgets down to smaller and less well defined geographical populations, such as primary care groups and large general practices in the UK and regional health areas in Canada, and for assessing efficiency of primary health care clinics in Spain. Consequently, methods of funding that accommodate variations in the health characteristics of individual patients are increasingly needed for equitable resource allocation across population groups and for valid measurement of performance across health care providers.

Australia operates both a predominantly publicly funded universal health insurance system through Medicare, as well as a significantly subsidised voluntary private insurance sector covering 45% of the population for private hospital care and ancillary services. Risk adjustment methods have potential relevance in the public health system for establishing capitation funding (such as in New South Wales area health services [AHS]) as well as for monitoring the performance of provider networks who provide health services for large public purchasers such as the Department of Veterans' Affairs. Importantly also, risk adjustment methods have significant relevance in the private health insurance sector for establishing prospective reinsurance arrangements in order to promote incentives for efficiency while maintaining the principles of community rating. The Australian mix of financing and service provision has made managed competition, a longer-term market-based solution aimed at integrating private health insurance into the universal framework of Medicare, a viable option for reform and, consequently, risk adjustment is of key importance.[10-11]

This paper reports on an exploratory inquiry into the Australian application of a particular risk adjustment method that has been extensively validated in the US, known as diagnostic cost groups (DCGs), to a large Australian hospital inpatient

data set. The study applies the risk adjustment system software, known as DxCG, to individual-level hospital inpatient diagnosis and expenditure data. The database is the inpatient records for 1996–97 and 1997–98 for the metropolitan regions of New South Wales.

The aims of the study were to assess the validity of the DCG methodology in the Australian context; determine the ability of the diagnostic method to predict concurrent and subsequent year expenditure; and explore the potential application and policy implications of the risk adjustment models. While there have been some previous studies[12-13] in risk adjustment in Australia this is the first study which has applied the method to a large Australian data set. This enabled a more detailed analysis of the robustness and internal consistency of parameter estimates and allowed for particular modelling simulations to be performed in order to demonstrate the policy importance of diagnosis-based risk adjustment methods.

Methods

The DCG method begins by classifying the 15 000 ICD-9-CM (International Classification of Diseases, Ninth Revision, Clinical Modification) codes to one of 545 clinically similar categories known as DxGroups. The 545 DxGroups are in turn clustered into 118 condition categories (CCs) based on clinical similarity and comparable expected costs. The condition categories are organised into hierarchies (HCCs, or hierarchical coexisting conditions) to ensure that only the most expensive of closely related condition categories are assigned to an individual. A full list of the HCC category names is shown in the Appendix. The DCG/HCC system is a multiple condition model and individuals therefore can have more than one HCC allocated to them. The DCG risk adjustment model uses linear, single equation, ordinary least squares (OLS) regression methods to model expected health care expenditures of individuals using demographic and HCC (diagnosis-based) parameter estimates as explanatory variables.

The data used in this analysis were provided by the New South Wales health department (NSW Health). The data set contains hospital inpatient diagnosis and modelled inpatient expenditure data linked at the individual level over 2 years for the fiscal years 1996–97 and 1997–98. It includes the entire population residing in the catchment area of the nine metropolitan area health services in New South Wales. The data set included all episodes of public or private hospitalisation for a population of 4 836 203 and 4 882 934 for the 2 years included in the study. In 1996–97, 831 666 individuals (17.3%) were hospitalised, and in 1997–98 there were 848 454 people (17.4%) hospitalised. Costs for each hospitalisation were estimated using data from the NSW Hospital Comparison Data Book and the NSW Casemix standards for 1999–2000. For acute care, actual cost data were unavailable and costs were estimated using casemix standards based on the NSW Hospital Cost Study for 1997–98 and Version 3.1 (Australian National) Diagnosis Related Groups (DRGs). Costs for sub- and non-acute separations were estimated on the basis of bed-days using data provided in the *NSW hospital comparison data book* — a benchmark costing framework developed by NSW Health. Numbers of non-treated patients residing in the metropolitan AHS were determined by reference to the Australian Bureau of Statistics census data.

A series of multivariable linear regressions were constructed to predict concurrent and subsequent year expenditures. Individual costs were explained by a series of dummy variables for each individual's HCC. To validate the explanatory performance, a split sample method, similar to that used in other research studies, was used to establish model structure and to assess and validate performance.[3,14] The model is first fitted against a sample population (ie, development sample) to establish the model structure and estimate parameter coefficients. A validation process is then undertaken whereby the model structure and coefficient estimates derived from the development sample is applied to the actual data of a second (ie, validated) sample. The explanatory capability of the "out-of-sample" fit in terms of R^2 is assessed against the original development sample. Finally, linear regression is used to determine parameter estimates for a

combined model (ie, development and validated sample sets) which then forms the basis upon which subsequent predictive analysis is undertaken. Various modelling and simulation exercises were undertaken demonstrating particular policy applications of the DCG model(s), and comparisons are made with age–sex and prior cost models to highlight the importance of risk adjustment methods in a policy context.

This study applied Release 5.1 DxCG software: an inpatient multidiagnosis HCC/DCG model. SAS version 8.2 (SAS Institute, Cary, NC, USA) was used to perform all the analysis.

Results

Since the NSW data set only contains hospital expenditures, the concurrent model was confined to developing a casemix profile for the hospitalised-only population. A development sample of 500 000 patients was randomly selected from the hospitalised population of 831 666 with the remaining 331 336 people forming the second, or validated, sample.

Concurrent data model

The explanatory performance of various alternative model structures is summarised in Box 1. Before modifications, the development and second sample were able to explain 31.7% and 29.0% of the variation in concurrent expenditure, respectively. Several modifications were made to the concurrent model. These included, first, trimming expenditure to a per individual threshold of $50 000 to limit the effects of extreme isolated outliers. Secondly, some HCCs were combined when they were unstable (across the two sample sets) due to small population numbers (ie, HCC 040/041; 045/046; 089/090; 101/102; 105/106; 110/111). Thirdly, some HCCs were combined when they had inconsistent cost-rank ordering (HCC 006/007; 072/073; 074/075) Finally, age-interaction terms were retrospectively introduced for a child (ie, less than 18 years old) and for the elderly (ie, greater that 64 years old) when there were more than 1000 people in the cohort and when the variable was statistically significant (at 99%). All other HCCs (including those coefficients that were statistically insignificant) remained in the model unchanged.

Box 1 shows that the cumulative effects of these three modifications led to some minor improvement in model performance with the R^2 for the development and second sample increasing to nearly identical figures of 0.345 and 0.339, respectively. Trimming the per person inpatient expenditure to $500 000 marginally increased the R^2. There were a very small number of individuals

1 Summary of model performance (R^2) in explaining concurrent NSW hospital inpatient expenditures of metropolitan area health services, 1996–97

	Development N=500000	Second sample N=331666	Cross validation	All hospitalised N=831666
Demographic-only model				
Age–sex model*				0.039
DCG/HCC model				
Not combined —no trim	0.317	0.290		
$500k trim + combined HCC + age/interaction	0.345	0.339	0.338	0.342
Alternative expenditure trims				
$250K trim + combined HCC + age/interaction	0.379			
$100K trim + combined HCC + age/interaction	0.442			

* The age–sex model is based on the entire data set of 831 666 hospitalised patients for 1996–97.
DCG = diagnostic cost group. HCC = hierarchical coexisting condition. ◆

whose expenditures exceeded the threshold (32 and 21 people for the development and second sample set, respectively). Combining HCCs and adding age interaction seems to have had a negligible impact upon the explanatory capabilities of the model. For comparative purposes, and as expected, applying alternative expenditure threshold trims of $250 000 and $100 000 improves the R^2 to 0.379 and 0.442, respectively.

Validating the model structure by running the estimated coefficients generated from the development sample against the actual data in the validated sample shows an R^2 of 0.338 — a figure very similar to both the development and second sample sets, indicating consistent and valid fit of model structure. As expected, the diagnosis-based model is significantly better than the model with age–sex only. Including the HCC variables and other modifiers increases the R^2 for the combined data set of 831 666 from 0.039 to 0.342. The full model structure showing the parameter estimates is presented in the Appendix. Overall, the explanatory capabilities of the concurrent model are similar to US results when using inpatient data.[15]

Prospective model

For the prospective model, given the absence of other medical costs in the NSW data set, Year 1 inpatient diagnosis and demographic variables are used to predict Year 2 inpatient costs rather than total health costs. One difficulty this presents is the extent of the relationship between Year 1 diagnosis and Year 2 inpatient costs. Specifically, many (Year 1) acute conditions may not necessarily be good predictors of hospital expenditures for the following year. For this reason, a total of 35 HCCs from the original 118 HCC parameters were recommended for exclusion by DxCG system developers, as these categories were considered to have no a priori relationship with subsequent year hospitalisations (see Appendix for the 35 excluded HCCs).

Given potentially small population numbers for some HCCs, a stratified random sample consisting of 400 000 randomly selected hospitalised patients, together with 600 000 randomly selected non-hospitalised patients, from the 1996–97 data set was established for both the development and second sample set. The stratified development sample was then weighted in accordance with the relative proportion of hospitalised to non-hospitalised mix as reflected in the general population.

A summary of the explanatory performance of the various model structures, the R^2, is shown in Box 2. Several modifications were made to the model structure. First, as with the concurrent model, individual-level expenditure was trimmed to a threshold of $500 000 in order to reduce the potential impact of large anomaly outliers. Second, nine HCCs with negative coefficients generated by the development sample (ie, HCCs 003, 009, 021, 056, 058, 059, 074, 093, 113) were

2 Summary of model performance (R^2) in explaining prospective NSW hospital inpatient expenditures of metropolitan area health services for 1996–97 and 1997–98

Weighted	Development sample N=1 000 000	Second sample N=1 000 000	Cross validation	Combined sample N=2 000 000
Age–sex model				0.030
Prior cost				0.040
DCG/HCC model				
Base model: $500K trim	0.052	0.052	0.052	0.052
$250K trim	0.054			
$100K trim	0.062			

DCG = diagnostic cost group. HCC = Hierarchical Coexisting Condition. ◆

constrained to zero. Finally, some of the unstable HCCs with small population sizes and HCCs with inconsistent cost-rank orderings were combined (ie, HCCs 008–010; 013/014; 040/41; 045/046; 050/051; 065/066). Apart from the 35 HCCs that were initially recommended for exclusion, all HCC parameters that were statistically insignificant but had stable and positive coefficients were retained in the model. Weighting the development model, to reflect actual population hospitalisation characteristics, the R^2 for the basic model structure for the development and second sample was 0.052. The R^2 for the cross-validation was 0.052 — a figure which is the same as both the development and second model, demonstrating general soundness in the underlying model structure. For the combined population model of 2 million individuals, the R^2 for DCG/HCC models is also 0.052. The explanatory capability of the prospective model is also comparable with US results using inpatient data. The parameter estimates for the combined prospective model are shown in the Appendix. The combined model structure is used for modelling purposes which are discussed later in the paper.

As expected, reducing the expenditure threshold from $250 000 to $100 000 improves the R^2 for the weighted model from 0.054 to 0.062 respectively. An interesting result is that the DCG/HCC model has higher explanatory capabilities than the prior cost model with the latter weighted model having an R^2 of 0.040. Box 2 also shows that the R^2 for the age–sex model was 0.030.

With regard to the parameter estimates generated by the HCC/DCG model presented in the Appendix, the concurrent model demonstrated a high degree of internal consistency and stability in coefficients across the development and validated sample sets. There was strong internal consistency, with nearly all higher cost HCCs having higher estimated coefficients relative to lower cost HCCs within the same hierarchical category. Differences in parameter estimates between two sample sets are attributable to small population numbers. In both sample sets the estimated coefficients remained highly stable when they were subject to modifications, indicating that the model structure was robust. Applying the Huber–White robust estimator to correct standard errors for heteroscedasticity (to accommodate changing variances in the disturbance term of the dependent variable) only 2 HCCs were statistically insignificant but were retained in the combined model structure.

The estimated coefficients generated by the prospective model are shown in the Appendix. They did not exhibit the same degree of stability and internal consistency as the concurrent model. However, this was not unexpected as the analysis was confined to inpatient-only expenditure. In addition to the a priori exclusion of 35 HCCs, an additional 14 HCCs in the combined model generated negative coefficients and had to be constrained to zero, while a further set of 6 HCCs were combined to maintain stability and cost monotonicity. The parameter estimates across both sample sets were stable when subject to these modifications, again indicating a robust model structure. Correcting standard errors using the Huber–White robust estimator, 16 HCCs were statistically insignificant but were retained in the combined model structure. Despite the limitations of the data set the prospective model exhibited reasonable stability and internal consistency and had demonstrably higher levels of explanatory capabilities relative to both age–sex and prior cost models.

Discussion

The findings from this study suggest that the DCG methodology has validity in its application to a large population health care data set in the Australian context. The policy importance is that the risk adjustment framework can be applied either as a method for monitoring performance of area health services, determining capitation-based funding or for identifying high-cost users for case management purposes.

Monitoring efficiency of area health services

An important application of the concurrent model is to compare and explain differences in utilisation

patterns across area health services once adjustments have been made for health status. Box 3 shows a comparison between the mean actual expenditures for hospitalised patients and mean expected hospital expenditures based on the DCG/HCC and age–sex concurrent model grouped by the nine AHS regions in metropolitan New South Wales. The relative index shown in Box 3 represents each region's actual expenditure as a proportion of the average state expenditure. For example, the average actual expenditure for the hospitalised population for AHS-A was 6% above the state mean expenditure. The expected index shows the expected expenditure in the respective models as a proportion of the actual state average expenditure. From this information, a regional efficiency index can be calculated as the ratio of the actual to expected expenditure. For example, based on the DCG/HCC model the expected index for AHS-A was 1.12, which implies that, given the utilisation and age–sex profile of its population, the mean hospital expenditure for that region was expected to be 12% above the state average. However, since the actual mean expenditure was only 6% above the state mean expenditure, this implies that AHS-A has an efficiency index of 0.95. That is, AHS-A is 5 percentage points more efficient than the state average after adjusting for (diagnosis-based) health status of its hospitalised population. The efficiency indices for all the area health services based on both the DCG model and an age–sex model are shown in Box 3.

An interesting outcome shown in Box 3 is the difference between the efficiency indices generated by the DCG/HCC and age–sex demographic models. In particular, AHS-A has an efficiency index of 0.95 according to the DCG/HCC model and 1.05 according to the age–sex model predictions. There is also a significant difference in the efficiency index for AHS-G with the DCG/HCC model indicating an efficiency index of 1.11 compared with an index of 1.04 for the demographic model. There are also some noticeable differences for several other AHS regions which would translate into non-trivial differences in the budget allocation using the two algorithms.

Ostensibly, Box 3 represents the efficiency of a region to manage its hospital costs for its hospitalised population. It would of course be desirable to have medical costs incorporated into the analysis so that the performance is based on the regional health authorities' ability to manage overall patient costs including the effects of services substitution.

3 Concurrent model: relative efficiency of metropolitan area health services (AHS) for hospitalised population for 1996–97 year: DCG/HCC model and age–sex model

				DCG			Age–sex only		
AHS	N	Actual $	Relative index	Predicted $	Expected index	Efficiency index	Predicted $	Expected index	Efficiency index
AHS-A	81 823	4341	1.06	4582	1.12	0.95	4147	1.02	1.05
AHS-B	129 504	4120	1.01	4199	1.03	0.98	4289	1.05	0.96
AHS-C	109 660	3792	0.93	3853	0.94	0.98	3824	0.94	0.99
AHS-D	50 264	3716	0.91	3608	0.88	1.03	3721	0.91	1.00
AHS-E	124 206	3618	0.89	3665	0.90	0.99	3788	0.93	0.96
AHS-F	52 417	4403	1.08	4626	1.13	0.95	4447	1.09	0.99
AHS-G	92 976	4295	1.05	3875	0.95	1.11	4133	1.01	1.04
AHS-H	58 138	4253	1.04	4230	1.04	1.01	4158	1.02	1.02
AHS-I	132 678	4346	1.06	4282	1.05	1.01	4251	1.04	1.02
All NSW	831 666	4081	1.00	4081	1.00	1.00	4081	1.00	1.00

DCG = diagnostic cost group. HCC = hierarchical coexisting condition. ◆

The important feature of the above analysis is that variations in health status measures, even when aggregated at a regional level, may not be adequately captured by age–sex demographic characteristics. Diagnosis-based risk adjustment offers the potential for more refined measures of casemix adjustment of population groups. The policy implication here is that if the federal and state governments engaged in systematic reform and devolved a single stream of funding for health care services down to area service boards, then developing diagnosis-based risk adjustment methods would be a more appropriate policy tool for establishing the funding base and assessing the relative efficiency of area health services. Reliance on age–sex demographic methods results in misleading assessment of the efficiency of area health services.

Competing health insurers and risk selection

The importance of establishing adequate capitation payments so as to minimise the risk-selection incentives within a competitive environment, such as in a managed competition framework or for prospective reinsurance arrangements, is demonstrated below. A simulation exercise was conducted which treated the 9 AHS regions as competing health plans which received two types of capitated payments; one based on age–sex model predictions, and the other on DCG/HCC model predictions. Each health plan is assumed to risk-select patients based on prior cost modelling information. More specifically, each health plan is assumed to risk-select patients when the capitated payment received for an individual is greater than their predicted costs based on the prior cost information; and dump patients when losses are predicted. Profit and losses are then calculated by subtracting the actual costs incurred by individuals after risk selection from the capitation payments received by the plans.

Results reported in Box 4 show that when health plans do not engage in risk selection total profit across all 9 health plans sum to zero as total capitated payments for age–sex and DCG/HCC methods are both modelled to equal total actual costs across the entire state. However if age–sex is used as the basis for establishing capitated payments, and given zero transaction costs, health plans on average will dump around 15.6% of patients and profits totalled across the 9 health plans will increase from zero to $130.5 million (when extrapolated to the entire population of 4.8 million persons). Thus considerable incentives for risk selection exist under age–sex capitation payments (given our assumption that health plans will utilise prior cost information, which is superior to age–sex model predictions). However, if capitation payments are based on DCG/HCC model predictions, and given the relatively poorer explanatory capabilities of the prior cost model, the 9 health plans would risk incurring aggregate losses totalling $28.8 million if they attempt to risk-select based on the prior cost model. Risk selection is thus a loss making strategy and health plans would therefore be expected not to engage in such activity and instead earn zero (ie, normal) economic profits.

Box 4 also shows the extent of, and incentives for, risk selection where there are alternative transaction costs of $50 and $100 per person. These scenarios recognise that person-dumping is not costless. As expected, the extent of risk selection reduces with greater transaction costs of dumping, but the incentives to risk-select are still

4 Simulated profits/losses based on age-sex and DCG/HCC capitated payment methods utilising prior cost modelling information to risk select

Transaction costs per patient dumped	Age–sex capitation		HCC/DCG capitation	
	Risk-selected	Total profit ($millions)	Risk-selected	Total profit ($ millions)
No risk selection	0	$0	0	$0
$0	15.6%	$130.5	64.1%	–$28.8
$50	11.9%	$133.0	12.2%	–$126.4
$100	8.8%	$146.2	9.2%	–$109.0

DCG = diagnostic cost group. HCC = hierarchical coexisting condition. ◆

considerable. In the case of age–sex capitation, total absolute profits increase with greater transaction costs since fewer people are dumped (ie, more people are enrolled in total). If transaction costs rise to sufficient levels risk selection will eventually reduce to zero and normal economic profits will be earned.

The above situation is of policy relevance for the Commonwealth government which is currently developing age–sex related risk adjusters for determining reinsurance arrangements for health insurers (scheduled for introduction in mid 2006). In order to underpin the community rating principle while also trying to create incentives for efficiency, health funds with "sicker" population cohorts will be prospectively reimbursed on a capitated basis through reinsurance arrangements based on the age–sex profile of their enrolled population. However if the age–sex capitated payments are inadequate to compensate health funds for having a sicker patient cohort then incentives exist to risk-select healthier patients at the expense of the sicker population. The simulation exercise presented shows quite clearly the superiority of diagnosis-based risk-adjusted capitated payments over age–sex related casemix with respect to reducing the incentives for risk selection and promoting efficient third-party purchasing.

Similarly, if a managed competition model were contemplated in Australia, developing adequate risk adjusters is an essential prerequisite in order to establish capitated payments and minimise incentives for risk selection. As long as risk adjusters can explain the predictable component of the variation in individual-level expenditure at least as well as methods utilised by health funds then the incentives for risk selection are significantly reduced.

Identifying high cost users

Another interesting application of the prospective model is the ability to identify high cost users. The task performed in this study was to use the prospective DCG model, the prior cost model and age–sex model to generate the 20 000 highest cost users (ie, the highest 1%). Similar to other studies,[16] a comparison is then made between the three models by observing the degree of overlap between the individuals predicted to be high cost users by the respective models with the individuals who actually were high cost users in Year 2. Of the 20 000 individuals identified and predicted by the DCG/HCC model to be the highest cost users, 3004 people (15%) matched the *actual* 20 000 highest cost individuals. The prior cost and age–sex models were able to identify 2472 individuals (12.4%) and 1318 individuals (6.6%), respectively, who were in the highest 20 000 actual users.

5 Proportion of overall costs explained by the highest 20 000 individuals predicted by prior cost and DCG/HCC models, weighted for stratification

Model	No. of people	High cost users	Actual Year 2 costs
DCG/HCC only	13 980	0.70%	2.2%
Both	6020	0.30%	2.0%
Prior cost only	13 980	0.70%	1.7%

DCG = diagnostic cost group. HCC = hierarchical coexisting condition. ◆

Box 5 reports another result for the highest 20 000 users predicted by both the prior cost and DCG/HCC model. Interestingly, 6020 individuals were predicted in common by the two models. This represented about 30% of each model's 20 000 highest cost users. However, this result could be used to account for more of the overall actual costs than the remaining 70% in both models.

The combined use of prior cost and diagnostic information has greater capabilities for identifying people who will be expensive in the following year. This conclusion is consistent with those reported in other similar analyses.[16]

The policy implication is that diagnostic information becomes an important tool for identifying potentially high cost patients. Accordingly, early identification and intervention through timely

and coordinated care offers the potential to improve health outcomes and also reduce health expenditure. Reliance on only prior cost information is a less robust method for determining high cost patients than diagnosis-based methods.

Conclusion

Research into and the application of diagnosis-based risk adjustment methods has been under way in many European countries since the mid 1990s as part of the broader approach to health care reform. Risk adjustment is increasingly being seen as an important policy tool for establishing prospective capitation payments for both competitive and non-competitive third-party purchasing health systems and to evaluate appropriateness and efficiency of services provided.[1-4]

The degree of stability and internal consistency of the estimated coefficients in the regression results for both the concurrent and prospective DCG/HCC models as well as the general explanatory performance comparable to similar overseas studies indicate DCG methodology has face validity in its application to NSW Health data sets. Findings from this exploratory study suggest that further research into the validity and potential usefulness of diagnosis-based risk adjustment methods ought to be pursued. However, further research into these methods, in the Australian context, should involve data sets with linked hospital and medical expenditure data. This will permit a full exploration of the potential capabilities associated with the use of individual-level information. To this end, the federal government has engaged in an exploratory initiative linking state hospital clinical and expenditure data with federal medical and pharmaceutical expenditure data at the individual level. Importantly also, information on prescription drugs has now become available at the individual level. Combining hospital diagnostic information with pharmaceutical information in developing risk adjustment models will permit research similar to the work undertaken in Europe and the US.[17-19] Models incorporating both inpatient and pharmaceutical information perform significantly better than inpatient-only models and represent state-of-the-art risk adjustment methodology in circumstances where out-of-hospital diagnostic information is unavailable. However, notwithstanding recent federal government initiatives, the establishment of linked patient-level data is still underdeveloped and priority needs to be given to the area. A recently published paper details how data linkage protocols in Australia can be established and pursued in a more comprehensive way.[20]

Australia appears to be committed to a mixed public–private federal–state system for the provision of health services, with the present system reflecting a series of ad hoc policies over time. This has led to a fragmented and uncoordinated health system characterised by perverse incentives such as cost-shifting between federal and state government jurisdictions. Further, inherent financial tensions between policies supporting the voluntary private insurance sector and those supporting the universal health system have not been addressed since the introduction of Medicare, leading to perennial structural instability.[21] It is also important to note that the Productivity Commission in its recent review of the National Competition Policy (NCP) specifically highlighted health care as a "prime candidate" for a nationally coordinated approach to health system reform and has called for a national independent public inquiry into Australia's health care system akin to the Hilmer inquiry that preceded the NCP.[22]

The development and use of individual-level risk adjustment methods make possible the consideration of reform options such as capitated single fund-holding to area health authorities or fund-holders; the development of prospective reinsurance arrangements for health insurers; as well as more encompassing competition-based reforms, such as managed competition. Such models could offer possible solutions to the inherent problems currently confronting the Australian health care system.[1,11,23] This study demonstrates the feasibility of utilising very large individual-level linked data sets for risk adjustment research in Australia, and the results pre-

sented suggest that such a research agenda should now be undertaken to underpin the options for system reform.

Acknowledgements

The authors wish to acknowledge the contributions of many individuals whose skills and expertise have been invaluable in making this project possible. We give special thanks to Heather McElroy of the Department of Public Health, University of Adelaide for statistical and computational assistance and to Dr Philip Ryan, Head of Department, for his collegial support of the project. We also appreciate the financial assistance provided by the Commonwealth Department of Health and Aged Care (Hospital Division), which made this project possible. In addition we would like to thank Mr Jim Pearse from NSW Health for developing and making available the comprehensive inpatient data sets used in this project. We also appreciate the assistance provided by Sean Ahern, Rob McGinn and Yang Zhao from DxCG Inc., and Professor Randall Ellis in helping us use the software and interpret the results. The views expressed in this paper are those of the authors, and do not necessarily reflect those of the Centre for Health Economics at Monash University, the Australian Government Department of Health and Ageing or NSW Health.

Competing interests

The authors declare that they have no competing interests.

References

1 Rice N, Smith PC. Capitation and risk adjustment in health care financing: an international progress report. *Milbank Q* 2001; 79: 81-113.

2 Wang MC, Rosen AK, Kazis L, et al. Correlation of risk adjustment measures based on diagnosis and patient self-reported health status. *Health Services and Outcomes Research Methodology* 2000; 1: 351-65.

3 Ash AS, Ellis RP, Pope GC, et al. Using diagnosis to describe populations and predict costs. *Health Care Financ Rev* 2000; 21: 7-28.

4 Iezzoni L, editor. Risk adjustment for measuring health outcomes. Michigan: Health Administration Press, 1997.

5 Majeed A, Bindman AB, Weiner JP. Use of risk adjustment in setting budgets and measuring performance in primary care: how it works. *BMJ* 2001; 323: 604-7.

6 Reid RJ, MacWilliam L, Verhulst L, et al. Performance of the ACG casemix system in two Canadian provinces. *Med Care* 2001; 39: 86-99.

7 Orueta JF, Lopez-De-Munain J, Baez K, et al. Application of the Ambulatory Care Groups in the primary care of a European national health care system: does it work? *Med Care* 1999; 37: 238-48.

8 Van de Ven WPMM, Ellis RP. Risk adjustment in competitive health plan markets. In: Culyer AJ, Newhouse JP, editors. Handbook of health economics. Amsterdam: Elsevier, 2000.

9 Chernichovsky D, van de Ven WPMM. Risk adjustment in Europe. *Health Policy* 2003; 65: 1-3.

10 Segal L, Donato R, Richardson J, Peacock S. Strengths and limitations of competitive versus non-competitive models of integrated captitated fundholding. *J Health Serv Res Policy* 2002; 7 Suppl 1: S56-64.

11 Scotton RB. Managed competition. In: Mooney G, Scotton RB, editors. Economics and Australian health policy. Sydney: Allen and Unwin, 1999.

12 Duckett SJ, Agius PA. Performance of diagnosis-based risk adjustment measures in a population of sick Australians. *Aust N Z J Public Health* 2002; 26: 500-7.

13 Antioch KM, Walsh M. The risk-adjusted vision beyond casemix (DRG) funding in Australia: international lessons in high complexity and capitation. *Eur J Health Econ* 2004; 5: 95-109.

14 Ash AS, Byrne-Logan S. How well do models work? Predicting health care cost. In: Proceedings of the section on statistics in epidemiology. American Statistical Association; 1998: 42-9.

15 Ellis RP, Pope GC, Iezzoni LI, et al. Diagnosis-based risk adjustment for Medicare capitation payment. *Health Care Financ Rev* 1996; 17: 101-28.

16 Ash AS, Zhao Y, Ellis RP. Finding future high-cost cases: comparing prior cost versus diagnosis-based methods. *Health Serv Res* 2001; 36: 194-206.

17 van de Ven WP, van Vliet RC, Leida ML. Health-adjusted premium subsidies in the Netherlands. *Health Aff (Millwood)* 2004; 23: 45-55.

18 Lamers LM, van Vliet RC. The pharmacy-based cost group model: validating and adjusting the classification of medications for chronic conditions to the Dutch situation. *Health Policy* 2004; 68: 113-21.

19 Zhao Y, Ellis RP, Ash AS, et al. Measuring population health risks using in-patient diagnosis and outpatient pharmacy data. *Health Serv Res* 2001; 36: 180-93.

20 Kelman CW, Bass AJ, Holman CD. Research use of linked health data – a best practice protocol. *Aust N Z J Public Health* 2002; 26: 251-5.

21 Donato R, Scotton RB. The Australian health care system. In: Mooney G, Scotton RB, editors. Economics and Australian health policy. Sydney: Allen and Unwin, 1999.

22 Australian Government Productivity Commission. Review of national competition policy reforms. Productivity Commission Inquiry Report no. 33, 28 Feb 2005. Canberra: Commonwealth of Australia, 2005.

23 Richardson J, Segal L. Private health insurance and the Pharmaceutical Benefits Scheme: how effective has recent government policy been? *Aust Health Rev* 2004; 28: 34-47. Available at: <http://www.aushealthreview.com.au/publications/articles/issues/ahr_28_1_300904/ahr_28_1_034-047.asp> (accessed Dec 2005).

(Received 7 Jul 2005, accepted 31 Oct 2005) □

Appendix: Regression results of combined concurrent and prospective DCG/HCC models

			Concurrent model (N=831 666; R^2=0.342)		Prospective model (N=2 000 000; R^2=0.052)	
Parameters			**Estimate ($)**	***P*> |t|**	**Estimate ($)**	***P*> |t|**
Child		Age: less than 18	403.88	<0.0001		
Adult		Age: 18–64	738.81	<0.0001		
Elderly		Age: 65+	826.65	<0.0001		
Female 0–5					226.17	<0.0001
Female 6–12					98.59	<0.0001
Female 13–17					170.25	<0.0001
Female 18–24					349.92	<0.0001
Female 25–34					566.86	<0.0001
Female 35–44					475.53	<0.0001
Female 45–54					531.19	<0.0001
Female 55–64					790.15	<0.0001
Female 65–74					1423.73	<0.0001
Female 75+					2242.95	<0.0001
Male 0–5					306.07	<0.0001
Male 6–12					133.23	<0.0001
Male 13–17					171.86	<0.0001
Male 18–24					258.05	<0.0001
Male 25–34					273.35	<0.0001
Male 35–44					318.83	<0.0001
Male 45–54					515.58	<0.0001
Male 55–64					932.51	<0.0001
Male 65–74					1862.70	<0.0001
Male 75+					2779.37	<0.0001
HCC001		HIV/AIDS	15 785.49	<0.0001	1920.63	<0.0001
HCC002		Septicemia (blood poisoning)/shock	8098.63	<0.0001	384.22	0.0574
HCC003		Central nervous system infections	5858.88	<0.0001	437.87	0.0907
HCC004		Other infectious disease	1245.62	<0.0001	99.92	<0.0001
HCC005		Metastatic cancer	6240.21	<0.0001	518.54	<0.0001
HCC006	C	High cost cancer	6071.67	<0.0001	2858.01	<0.0001
HCC007	C	Moderate cost cancer	6071.67	<0.0001	1800.28	<0.0001
HCC008		Low cost cancers/tumors	2394.45	<0.0001	–	–

Appendix, continued

Parameters		Concurrent model (N=831666; R^2=0.342) Estimate ($)	$P>\|t\|$		Prospective model (N=2000000; R^2=0.052) Estimate ($)	$P>\|t\|$
HCC009	Carcinoma in situ	556.01	<0.0001		–	–
HCC010	Uncertain neoplasm	1770.09	<0.0001		–	–
HCC011	Skin cancer, except melanoma	997.36	<0.0001	X	–	–
HCC012	Benign neoplasm	568.70	<0.0001	X	–	–
HCC013	Diabetes with chronic complications	617.69	<0.0001	C	1002.28	<0.0001
HCC014	Diabetes with acute complications	1043.35	<0.0001	C	1002.28	<0.0001
HCC015	Diabetes with no or unspecified complications	221.76	<0.0001		167.79	0.002
HCC016	Protein–calorie malnutrition	1805.42	<0.0001	X	–	–
HCC017	Moderate cost endo/ metab/fluid–electrolyte	2238.47	<0.0001		–	–
HCC018	Other endocrine, metabolic, nutritional	820.24	<0.0001		38.06	0.3019
HCC019	Liver disease	2303.81	<0.0001		1460.04	<0.0001
HCC020	High cost chronic gastrointestinal	1587.37	<0.0001		703.52	<0.0001
HCC021	High cost acute gastrointestinal	2602.76	<0.0001		–	–
HCC022	Moderate cost gastrointestinal	1003.55	<0.0001		–	–
HCC023	Low cost gastrointestinal	303.23	<0.0001	X	–	–
HCC024	Bone/joint infections/ necrosis	4743.52	<0.0001		774.74	0.0002
HCC025	Rheumatoid arthritis/ connective tissue	1627.47	<0.0001		533.61	<0.0001
HCC026	Other musculoskeletal/ connective tissue	1501.04	<0.0001	X	–	–
HCC027	Aplastic and acquired hemolytic anaemias	5993.86	<0.0001		4439.95	<0.0001
HCC028	Blood/immune disorders	3176.81	<0.0001		1066.76	<0.0001
HCC029	Iron deficiency and other anaemias	1125.85	<0.0001		257.09	<0.0001
HCC030	Dementia	890.81	<0.0001	X	–	–
HCC031	Drug/alcohol dependence/psychoses	2813.89	<0.0001		793.93	<0.0001
HCC032	Psychosis/higher cost mental	7461.13	<0.0001		1719.76	<0.0001
HCC033	Depression/moderate cost mental	1689.12	<0.0001		518.84	<0.0001
HCC034	Anxiety	2612.07	<0.0001		542.64	0.0068

Appendix, continued

Parameters			Concurrent model (N=831 666; R^2=0.342) Estimate ($)	$P>\|t\|$		Prospective model (N=2 000 000; R^2=0.052) Estimate ($)	$P>\|t\|$
HCC035		Other mental and substance abuse	492.52	<0.0001		95.94	0.0265
HCC036–039		Mental retardation	1025.76	<0.0001		734.10	0.0464
HCC040	C	Quadraplegia	10 851.01	<0.0001	C	4678.23	<0.0001
HCC041	C	Paraplegia	10 851.01	<0.0001	C	4678.23	<0.0001
HCC042		High cost neurological	3755.33	<0.0001		776.90	<0.0001
HCC043		Moderate cost neurological	897.26	<0.0001		380.25	<0.0001
HCC044		Low cost neurological	549.51	<0.0001	X	–	–
HCC045	C	Respiratory dependent/ tracheostomy	12 033.47	<0.0001	C	1309.89	0.0268
HCC046	C	Respiratory arrest	12 033.47	<0.0001	C	1309.89	0.0268
HCC047		Cardio-respiratory failure and shock	8412.59	<0.0001		513.44	0.0172
HCC048		Congestive heart failure	769.09	<0.0001		258.07	<0.0001
HCC049		Heart arrhythmia	1143.57	<0.0001	X	–	–
HCC050		Acute myocardial infarction	3966.48	<0.0001	C	23.84	0.6347
HCC051		Other acute ischemic heart disease	2417.31	<0.0001	C	23.84	0.6347
HCC052		Chronic ischemic heart disease	697.15	<0.0001	X	–	–
HCC053		Valvular and rheumatic heart disease	1847.98	<0.0001		182.12	0.0211
HCC054		Hypertensive heart disease	1588.51	<0.0001		828.07	0.2240
HCC055		Other heart diagnoses	246.52	0.2002	X	–	–
HCC056		Heart rhythm and conduction disorders	947.55	<0.0001		–	–
HCC057		Hypertension (high blood pressure)	196.16	<0.0001	X	–	–
HCC058		High cost cerebrovascular disease	2965.45	<0.0001		–	–
HCC059		Low cost cerebrovascular disease	1155.48	<0.0001		–	–
HCC060		High cost vascular disease	3388.58	<0.0001		662.12	<0.0001
HCC061		Thromboembolic vascular disease	3654.58	<0.0001		326.84	0.0024
HCC062		Atherosclerosis	1807.90	0.0007	X	–	–
HCC063		Other circulatory disease	1199.18	<0.0001		–	–
HCC064		Chronic obstructive pulmonary disease	1137.34	<0.0001		574.52	<0.0001
HCC065		High cost pneumonia	6756.96	<0.0001	C	131.89	0.4593

Appendix, continued

Parameters			Concurrent model (N=831 666; R^2=0.342) Estimate ($)	$P>\|t\|$		Prospective model (N=2 000 000; R^2=0.052) Estimate ($)	$P>\|t\|$
HCC066		Moderate cost pneumonia	3931.56	<0.0001	C	131.89	0.4593
HCC067		Low cost pneumonia	2828.83	<0.0001		13.92	0.8391
HCC068		Pulmonary fibrosis/other chronic lung disease	1943.90	<0.0001		913.97	<0.0001
HCC069		Pleural effusion/ pneumothorax	3777.18	<0.0001		–	–
HCC070		Asthma	88.81	<0.0001		102.28	<0.0001
HCC071		Other lung disease	2115.80	<0.0001	X	–	–
HCC072	C	High cost eye	1145.41	<0.0001	X	–	–
HCC073	C	Low cost eye	1145.41	<0.0001	X	–	–
HCC074	C	High cost ear, nose and throat	584.27	<0.0001	X	–	–
HCC075	C	Low cost ear, nose and throat	584.27	<0.0001	X	–	–
HCC076		Dialysis status	30515.07	<0.0001		21503.22	<0.0001
HCC077		Kidney transplant status	613.33	0.0394		2324.17	<0.0001
HCC078		Renal failure	2879.25	<0.0001		1102.08	<0.0001
HCC079		Nephritis	541.13	0.0021		439.54	0.0157
HCC080		Other urinary system	1092.71	<0.0001	X	–	–
HCC081		Female infertility	4451.42	<0.0001		628.53	<0.0001
HCC082		Moderate cost genital	987.59	<0.0001	X	–	–
HCC083		Low cost genital	304.68	<0.0001	X	–	–
HCC084		Ectopic pregnancy	858.71	0.0002	X	–	–
HCC085		Miscarriage/abortion	27.73	0.5846	X	–	–
HCC086		Completed pregnancy, major complications	3103.87	<0.0001	X	–	–
HCC087		Completed pregnancy, minor complications	1112.75	<0.0001	X	–	–
HCC088		Normal delivery	1401.31	<0.0001	X	–	–
HCC089	C	High cost pregnancy w/o completion	219.71	0.0450		1188.17	<0.0001
HCC090	C	Low cost pregnancy w/o completion	219.71	<0.0001		930.10	<0.0001
HCC091		Chronic ulcer of skin	4906.14	<0.0001		366.10	<0.0001
HCC092		Other dermatological	965.80	<0.0001	X	–	–
HCC093		Vertebral fractures/spinal cord injuries	3985.82	<0.0001		–	–
HCC094		Hip fracture/dislocation	5417.12	<0.0001	X	–	–
HCC095		Head injuries	5362.25	<0.0001	X	–	–
HCC096		Drug poisoning/internal injury/trauma/ amputation/burns	1705.56	<0.0001		84.07	0.0414

Appendix, continued

Parameters			Concurrent model (N=831 666; R^2=0.342) Estimate ($)	$P>\|t\|$		Prospective model (N=2 000 000; R^2=0.052) Estimate ($)	$P>\|t\|$
HCC097		Other injuries and poisonings	839.63	<0.0001		–	–
HCC098		Complications of care	5856.91	<0.0001		237.57	0.2767
HCC099		Major symptoms	614.37	<0.0001		–	–
HCC100		Minor symptoms, signs, findings	552.07	<0.0001		–	–
HCC101	C	Very high/high cost paediatric	7106.89	<0.0001		4298.97	<0.0001
HCC102	C	High cost paediatric	7106.89	<0.0001		1423.02	<0.0001
HCC103		Moderate cost congenital	1236.75	<0.0001		644.93	<0.0001
HCC104		Low cost congenital	606.51	<0.0001		64.63	0.0473
HCC105	C	Extremely/very low birthweight neonates	37 573.30	<0.0001		490.44	0.4249
HCC106	C	Very low birthweight neonates	37 573.30	<0.0001	X	–	–
HCC107		Serious perinatal problem — newborn	2789.87	<0.0001	X	–	–
HCC108		Other perinatal problem affecting newborn	1209.08	<0.0001	X	–	–
HCC109		Normal, single birth	1175.30	<0.0001	X	–	–
HCC110	C	Major organ transplant status	11 344.06	<0.0001		3807.39	<0.0001
HCC111	C	Other organ transplant status	11 344.06	<0.0001		1185.10	0.0686
HCC112		Artificial opening status/ attention	4597.65	<0.0001		1618.03	<0.0001
HCC113		Elective/aftercare	1784.62	<0.0001		–	–
HCC114		Radiation therapy	4192.68	<0.0001	X		
HCC115		Chemotherapy	5222.20	<0.0001		3436.21	<0.0001
HCC116		Rehabilitation	5001.51	<0.0001	X	–	–
HCC117		Screening/observation/ special exams	950.64	<0.0001		–	–
HCC118		History of disease	143.67	<0.0001	X	–	–
CHCC4			–459.45	<0.0001			
EHCC4			330.18	<0.0001			
CHCC17			–1067.30	<0.0001			
EHCC17			–1938.58	<0.0001			
CHCC23			531.03	<0.0001			
EHCC35			706.57	<0.0001			
EHCC43			–620.71	<0.0001			
CHCC67			–1537.76	<0.0001			

Appendix, continued

Parameters	Concurrent model (N=831666; R^2=0.342) Estimate ($)	$P>\|t\|$	Prospective model (N=2000000; R^2=0.052) Estimate ($)	$P>\|t\|$
EHCC67	−1129.80	<0.0001		
EHCC71	−1057.66	<0.0001		
EHCC80	−610.85	<0.0001		
EHCC98	962.97	<0.0001		
CHCC100	422.77	<0.0001		
CHCC103	2115.11	<0.0001		
EHCC113	−585.39	<0.0001		
EHCC117	−921.64	<0.0001		
CHCC118	662.24	<0.0001		

DCG = diagnostic cost group. HCC = hierarchical coexisting condition.

The letter C denotes HCCs which have been combined

The letter X denotes the 35 HCCs that were originally recommended for exclusion from the prospective model by DxCG system developers. In addition to the excluded HCCs an additional 16 HCCs had negative coefficients which were subsequently constrained to zero.

Age interaction terms in the concurrent model are denoted as CHCC for children under 18 and EHCC for the elderly over 65 years followed by the numeric category consistent with the general HCC categories. ◆