Severity variations within DRGs: Measurement of hospital effects by use of data on significant secondary diagnoses and procedures

DON HINDLE, PIETER DEGELING AND ONO VAN DER WEL

Don Hindle is Visiting Professor and Pieter Degeling is Associate Professor in the Medical Faculty of the University of New South Wales. Ono van der Wel is Director Health Costing and Evaluation in the South Australian Health Commission.

Abstract

The Diagnosis Related Group classification has provided an excellent basis for enhancing the equity of resource allocation between public acute hospitals. However, it underestimates the higher levels of severity and consequent costliness of referral hospitals.

This paper describes a practical way of measuring within-DRG variations in severity, which can be used to increase the precision of casemix-based funding. It involves the regression of length of stay against the numbers of significant diagnoses and procedures, and hence the prediction of additional justified costs. An example is given of its application to data from South Australian public hospitals.

The context: DRG-based funding and the issue of severity

Since 1994 the South Australian Health Commission has allocated funds among public acute care hospitals in proportion to their casemix-weighted patient volumes (Commonwealth Department of Health and Family Services 1997). The key measure of casemix is the Australian variant of the Diagnosis Related Group (DRG) classification (AN-DRGs). While it is effective in many respects, there are legitimate concerns about its ability to measure variations in severity.

This would be of little consequence if patients were randomly assigned to hospitals. However, a sophisticated referral process has developed over time, whereby the more complicated patients are directed to a subset of hospitals higher in the referral chain. It is therefore reasonable to hypothesise that there are systematic variations: that some hospitals tend to receive more than the average number of relatively costly patients in most DRGs, and might therefore be unfairly funded at the margins.

This hypothesis has some support in the scientific literature. For example, Sharkey et al. (1991) studied the costs of admitted patients with burns in a sample of United States hospitals. They found that DRGs explained 24 per cent of the variation in total costs. By using the Severity of Illness Index, they were able to define sub-categories, which increased variance explanation to 44 per cent. They concluded that '… hospital funding which relies exclusively on categorisation by DRG is inequitable and, in some circumstances, discourages increased understanding of treatments which improve outcomes'.

MacKenzie et al. (1991) analysed trauma patients in Maryland by use of the Injury Severity Score. They argued that, while trauma care had been regionalised for sound clinical and economic reasons, DRG-based funding would provide perverse incentives to abandon the progress that had been made. Similar results of within-DRG severity variations have been reported by Young, Macioce and Young for trauma patients (1990), Froehlich and Jarvis (1991) for paediatric patients, Stoskopf and Horn (1992) for schizophrenia, Hammond and Ward (1991) for elderly patients with burns, Rosenthal and Landefeld (1993) for elderly patients, and Rapoport et al. (1990) for patients requiring intensive care. Studies of all types of patients have been reported by Bradbury, Stearns and Steen (1991), Horn et al. (1991) and Averill et al. (1992).

Several kinds of approaches have been suggested whereby account may be taken of within-DRG severity variations in the funding context. They may be described under four partially overlapping headings.

First, the *DRG assignment logic* might be refined. One option involves the use of more sophisticated ways of defining the diagnosis and procedure clusters. For example, there has been support for use of the Patient Management Categories approach, which permits the formation of multiple clusters of related conditions for a single episode (Young 1985). The Disease Staging classification has also attracted interest, mainly because of the way in which conditions are grouped according to disease process, and then ranked according to models of their sequential manifestation (Young 1988). Other options include revision of methods of splitting the diagnosis and procedure clusters. For example, experiments have been undertaken whereby the clusters are split by score ranges of the Computerised Severity Index (Horn & Horn 1986).

The second type of approach involves *sub-classification of DRGs through funding rules.* For example, higher payment rates per DRG might be made to teaching hospitals on the grounds that they attract more severely ill patients. In effect, another variable (type of hospital) is being used as an indicator of casemix. Similarly, the long-stay outlier payments in several United States and Australian funding models represent an attempt to use length of stay as an indicator of severity of condition. Less powerful and direct adjustments include the socioeconomic disadvantage component of the United States Medicare's prospective payment system, and the research, teaching and acute transfer payment increments in several Australian models.

The third type of approach aims for *improved measures of patient condition*. For example, it would be possible to incorporate new and more specific ICD diagnosis codes for many diseases, in much the same way that low birthweight sub-classes were created in 1989 by incorporation of birthweight ranges in grams (Hindle, Pilla & Scuteri 1991).

A different approach to increased specificity has been proposed by McMahon et al. (1992). They developed an index value by use of the pathology variables in APACHE-2, and found that it explained an additional 14 per cent of the observed variations in length of stay compared with the use of DRGs alone.

Finally, use might be made of *additional variables in the discharge data set.* With minor exceptions, current DRG logic only makes use of four attributes (principal diagnosis, the main procedure, one other diagnosis, and type of discharge).

One possibility is that of using all secondary diagnoses to construct a weighted aggregate score. Iezzoni et al. (1990, 1991) developed a combinatory measure of co-morbidities by use of clinicians' subjective judgements about interactions. A similar model is DRGscale, which consists of a co-morbidity index derived by multiple linear regression (Christofferson, Conklin & Gonella 1988). Other variables which have been considered in the literature include type of admission (emergency or elective), number of procedures, and weighted number of procedures.

South Australia made use of an indirect indicator of within-DRG severity variations in its 1994–95 casemix funding formula. Each hospital was assigned to one of three categories: teaching hospitals (which received a 10 per cent increment to the base DRG payments), other metropolitan and regional (5 per cent increment), and all other (for which no increment was paid). The increments were based on actual costs in the previous year which could not be explained by DRG alone.

The South Australian Health Commission undertook research in 1995 with the objective of replacing hospital category with indicators based more directly on clinical information. In the event, it was replaced by two other measures (Hindle, Cromwell & Halsall 1995).

First, admission to intensive care was used as the basis for separate funding of intensive care units (ICUs). This approach, advocated for some time in the private hospital sector, has progressively gained support elsewhere including New South Wales, Tasmania, Queensland and Western Australia. The second change involved the development of a 'severity adjustment' model whereby the remaining variations in length of stay by DRG were predicted by the number of secondary diagnoses and procedures. Other variables were tested during the study, but rejected for various reasons.

The resultant model was applied to budget allocation for the 1995–96 and 1996–97 financial years. Although there were obvious weaknesses, most people accepted that it represented a better basis for severity adjustment than hospital category alone.

Development of a refined model for severity adjustment in South Australia

The South Australian Health Commission determined that it would attempt further refinement for 1997–98. The study database comprised 332 475 discharges of acute admitted patients from all public hospitals in the State during 1995–96. For each discharge, the available variables included AN-DRG version 3, all DRG assignment fields, and length of stay.

There were 55 851 (17 per cent) records not used for one reason or another. They included 13 054 'unqualified newborns', whose costs are considered by definition to be included in those of the record relating to the mother.

There were 6610 episodes removed because they related to hospitals whose casemix was not considered to be adequately defined by DRG. Also removed were 3245 'nursing home type' and 1156 rehabilitation episodes from all hospitals where they were present (because they are funded through other components of the model).

All ICU days were removed from the analysis. Where they represented only a part of the episode, its length of stay was recalculated by deduction of ICU days. If less than a day remained, the episode was defined to be wholly ICU, and the record was removed in its entirety. There were 8370 records of the latter type.

Finally, 23 456 outlier episodes were removed by use of an interquartile range trimming algorithm, because they constitute a separate payment element in the model. Moreover, there is evidence that some of them represent data errors (and, in particular, failure fully to apply the rules for statistical discharge and type change).

The counting of diagnoses and procedures

The previous model took account of all secondary diagnoses and procedures listed in the discharge records. The current study involved four refinements, as follows:

- only diagnoses listed as co-morbidities and complications (CCs) in the DRG logic were counted
- only procedures used in DRG assignment were counted
- DRG exclusion logic was applied, whereby some CCs were not counted if they were in combination with a related principal diagnosis
- CCs were weighted according to a four-point scale of severity (1 for minor, 2 for moderate, 3 for major, and 4 for catastrophic) used by the refined DRG classification.

Regression analysis to compute adjustment factors

The mean length of stay for any DRG varies between hospitals for a variety of reasons, but two are of particular relevance here. First, hospitals may differ with respect to their efficiency. This is of no interest in terms of casemix measurement, since the aim is to reduce the difference to factors under the control of hospital managers.

Second, length of stay at one hospital may be higher or lower, as a consequence of casemix differences not reflected in the DRG classification. The aim of casemix-based funding is to take account of all patient variations, and therefore this cause of differences in length of stay (and consequent costs) is undesirable. One such hypothesised cause of within-DRG variation is the need to provide care in respect of additional conditions and required procedures which are not taken into account in the DRG logic.

This study was directed at the latter factor, and it was therefore necessary to control for other causes. Several approaches were considered. In the event, a relatively simple method was used, which involved five analytical steps.

In Step 1, the lengths of stay for every patient were adjusted so that the DRG means were the same for every hospital. This involved use of the following formula:

$$DL_{ijk} = AL_{ijk} * (\sum_{i,j} AL_{ijk}/N_k)/(\sum_{i} AL_{ijk}/n_{jk})$$

where

ijk	patient i in hospital j, in DRG k
N_k	total number of cases in DRG k
n_{jk}	total number of cases in DRG k at hospital j
AL_{ijk}	actual length of stay of patient i in hospital j, DRG k
DL_{ijk}	adjusted length of stay of patient i in hospital j, DRG k

The effect of this adjustment is elimination of the differences in means between hospitals, but retention of the variations between patients in the same hospital. It is therefore assumed that the effect of within-DRG variations in severity, as reflected in the presence of significant diagnoses and procedures, applies to a similar extent across all observed lengths of stay.

In Step 2, linear regression models were constructed for each DRG. In each case the dependent variable was adjusted length of stay in days, and the predictor variables were number of significant procedures and weighted number of significant diagnoses as defined above. The standard least squares model was applied, with no constraint on the constant term, to give a model for each DRG with the following form:

$$DL_{ik} = m_1 wd_{ik} + m_2 p_{ik} + b$$

where

DL_{ik}	adjusted length of stay, patient i in DRG k
m_1	regression coefficient, weighted number of significant diagnoses
wd_{ik}	weighted number of significant diagnoses, patient i in DRG k
m_2	regression coefficient, number of significant procedures
p_{ik}	number of significant procedures, patient i in DRG k
b	constant

Regression models were constructed for 400 DRGs. The others were excluded for one of three main reasons.

First, no attempt was made to estimate effects for DRGs with fewer than 20 cases across all hospitals (after the reductions in scope as described above). This was a simple expedient, reflecting both the low likelihood of significant effects and the pressure of time to undertake the analyses.

Second, a small number of DRGs were excluded because they wholly or almost entirely comprised same-day cases. This is a matter of logic: regression is irrelevant since there is no observable within-class variation in length of stay.

Finally, 124 DRGs were excluded because the results were not significant, as determined by the F statistic at $p = 0.05$. Other checks were made on the remainder, in terms of R^2 (the coefficient of multiple determination) and the coefficients, but they led to no further exclusions.

In Step 3, every discharge record in scope was re-processed to derive a predicted length of stay from the multiple linear regression models. Then the change in days of stay assumed to be a consequence of the presence of the additional diagnoses and procedures was computed for each patient from the difference between predicted and adjusted length of stay, as follows:

$$EL_{ijk} = PL_{ijk} - DL_{ijk}$$

where

- ijk patient i in hospital j, in DRG k
- PL_{ijk} predicted length of stay of patient i in hospital j, patient in DRG k
- DL_{ijk} adjusted length of stay of patient i in hospital j, patient in DRG k
- EL_{ijk} change in days of stay of patient i in hospital j, patient in DRG k

The sum of EL values across all patients in the same DRG is zero, by definition. However, this is not necessarily the case for each hospital. Step 4 involved computing the changes in days of stay for any hospital as:

$$\sum_{ik} EL_{ijk} = \sum_{ik} (PL_{ijk} - DL_{ijk})$$

A second model was developed which took account of the argument that days differ in cost according to the DRG. In this case, the model:

$$\sum_{ik} EL'_{ijk} = \sum_{ik} (PL_{ijk} - DL_{ijk}) * w_j$$

was used, where the w_j terms were the national public hospital DRG cost weights.

The last step involved the estimation of index values for each hospital to reflect their relative levels of change in days predicted by the regression models. Table 1 shows the two models, one with and the other without weighting by the DRG cost relativities. For each model there is a 'crude' adjustment factor which is the sum of actual days in scope and predicted days as a percentage of actual days.

Table 1: Severity adjustment factors by hospital

	Unweighted				Weighted by DRG cost			
Hospital	Actual	Predicted	Adjustment		Actual	Predicted	Adjustment	
code	days	days	Crude	Rebased	days	days	Crude	Rebased
49	2 263	2 598	214.83	114.33	4 558	5 241	214.97	116.53
34	113 159	116 909	203.31	108.20	330 451	339 729	202.81	109.93
516	47 274	48 632	202.87	107.97	94 105	96 575	202.63	109.83
10	80 921	82 924	202.47	107.76	218 246	223 067	202.21	109.61
58	7 468	7 537	200.93	106.94	12 838	13 010	201.34	109.14
32	35 351	35 664	200.89	106.91	86 127	87 154	201.19	109.06
60	3 135	3 165	200.95	106.94	6 195	6 261	201.07	108.99
48	38 809	39 141	200.85	106.89	65 308	65 931	200.95	108.93
6	57 530	58 382	201.48	107.23	168 337	169 144	200.48	108.67
96	3 207	3 232	200.80	106.86	4 383	4 369	199.69	108.24
428	16 073	15 806	198.33	105.55	26 993	26 639	198.69	107.70
25	74 888	75 167	200.37	106.64	179 821	177 288	198.59	107.65
348	7 559	7 459	198.68	105.74	11 656	11 416	197.94	107.29
183	699	687	198.20	105.48	918	898	197.88	107.26
281	13 723	13 423	197.81	105.27	21 543	20 978	197.38	106.99
337	11 760	11 509	197.86	105.30	16 848	16 357	197.09	106.83
53	42 948	41 783	197.29	105.00	77 281	74 996	197.04	106.81
240	2 942	2 868	197.46	105.09	3 950	3 795	196.06	106.28
265	3 733	3 613	196.78	104.72	4 929	4 729	195.95	106.22
234	1 145	1 124	198.10	105.43	1 612	1 542	195.62	106.04
260	1 526	1 480	196.95	104.82	1 871	1 785	195.43	105.94
296	7 960	7 687	196.57	104.62	11 057	10 511	195.07	105.74
204	708	678	195.79	104.20	1 031	980	195.03	105.72
276	6 363	6 118	196.16	104.40	8 494	8 064	194.93	105.66
229	516	494	195.73	104.16	608	574	194.52	105.44
214	1 040	1 001	196.20	104.42	1 513	1 428	194.38	105.36
301	4 090	3 935	196.20	104.41	5 552	5 240	194.37	105.36
384	7 084	6 781	195.72	104.16	10 657	10 040	194.21	105.27
363	4 027	3 864	195.95	104.29	5 032	4 724	193.88	105.09
332	565	534	194.51	103.52	750	702	193.55	104.92
425	4 269	4 075	195.45	104.02	5 722	5 352	193.54	104.91
101	5 631	5 392	195.77	104.19	9 233	8 628	193.44	104.86
420	2 589	2 481	195.84	104.23	3 267	3 042	193.11	104.68
378	782	742	194.79	103.67	1 007	937	193.01	104.62
36	7 557	7 151	194.63	103.58	10 102	9 374	192.80	104.51
394	2 487	2 364	195.03	103.79	3 514	3 257	192.70	104.45

continued

Table 1: Severity adjustment factors by hospital *continued*

Hospital code	Unweighted: Actual days	Unweighted: Predicted days	Unweighted: Adjustment Crude	Unweighted: Adjustment Rebased	Weighted by DRG cost: Actual days	Weighted by DRG cost: Predicted days	Weighted by DRG cost: Adjustment Crude	Weighted by DRG cost: Adjustment Rebased
409	13	12	195.41	104.00	14	13	192.69	104.45
426	172	164	195.29	103.93	153	142	192.59	104.40
322	1 008	945	193.82	103.15	1 310	1 212	192.48	104.34
121	2 248	2 115	194.08	103.29	2 669	2 462	192.24	104.21
250	2 137	1 995	193.35	102.90	3 013	2 777	192.17	104.17
358	977	914	193.51	102.99	1 406	1 295	192.07	104.11
147	1 715	1 618	194.31	103.41	2 135	1 954	191.53	103.82
404	2 053	1 911	193.08	102.75	2 671	2 444	191.48	103.79
198	2 077	1 936	193.22	102.83	2 811	2 567	191.32	103.70
353	9 743	9 098	193.38	102.91	13 723	12 528	191.29	103.69
116	1 255	1 170	193.24	102.84	1 741	1 585	191.03	103.55
286	1 834	1 727	194.19	103.35	2 328	2 114	190.78	103.42
209	534	495	192.85	102.64	665	603	190.75	103.40
137	3 240	3 016	193.07	102.75	4 345	3 943	190.74	103.39
162	953	877	192.03	102.20	1 148	1 036	190.24	103.12
193	2 168	2 005	192.49	102.44	3 102	2 797	190.16	103.07
106	15	14	192.18	102.27	16	15	190.11	103.05
224	710	653	191.98	102.17	919	828	190.08	103.03
291	2 514	2 333	192.78	102.60	3 007	2 708	190.05	103.02
317	746	685	191.85	102.10	980	881	189.96	102.97
389	2 657	2 429	191.45	101.89	3 841	3 434	189.42	102.68
157	1 944	1 774	191.27	101.79	2 743	2 449	189.27	102.59
188	1 481	1 376	192.90	102.66	2 140	1 909	189.20	102.56
126	1 092	1 002	191.80	102.08	1 483	1 322	189.16	102.53
312	83	77	191.76	102.05	80	71	188.66	102.27
173	1 019	923	190.51	101.39	1 286	1 140	188.60	102.23
245	2 149	1 958	191.12	101.71	2 610	2 308	188.41	102.13
219	539	492	191.14	101.72	705	622	188.30	102.07
142	1 093	1 000	191.47	101.90	1 699	1 498	188.18	102.00
399	734	670	191.30	101.81	911	802	188.02	101.92
132	509	456	189.57	100.89	689	606	187.99	101.90
342	1 360	1 230	190.44	101.35	1 696	1 491	187.90	101.85
414	1 506	1 371	191.03	101.67	2 118	1 861	187.87	101.84
85	2 984	2 683	189.89	101.06	4 199	3 664	187.26	101.51
90	2 016	1 828	190.71	101.50	3 165	2 755	187.05	101.39
368	695	625	189.89	101.06	1 052	915	187.01	101.37
178	315	284	190.13	101.19	431	374	186.80	101.25
168	382	340	189.06	100.62	613	531	186.69	101.20
255	12	10	188.41	100.27	19	16	186.09	100.87
327	789	707	189.60	100.90	1 018	875	185.97	100.80
152	510	449	187.90	100.00	890	752	184.48	100.00
All	**681 762**	**681 762**			**1 497 053**	**1 497 053**		

Also shown are 'rebased' adjustment factors. They are derived by calculating a rebasing factor RF as:

$$RF = 100 / C_{min}$$

where C_{min} is the crude factor for the lowest hospital. Then the rebased values for every hospital are:

$$C_j = R_j * RF$$

Other approaches might arguably be preferred, as a matter of logic. However, this is consistent with the approach applied in the previous study, and gives a similar degree of adjustment to that applied in all models since 1994–95.

Discussion

This model represents a logical approach to the refinement of casemix measures. Significant relationships have been found between length of stay and the additional diagnoses and procedures, and their application has resulted in the identification of differences between hospitals which were consistent with other evidence.

However, the method of analysis has some significant limitations. They need to be noted by users of this approach, and they provide a basis for defining opportunities for further improvement.

First, length of stay has been used as the proxy for cost. This makes sense on practical grounds because it is the only variable correlated with cost which is routinely available with minimal effort. However, there are many possible causes of higher cost in a DRG including multiple procedures, use of capital-intensive facilities and equipment, the need for labour-intensive monitoring and therapy, and use of expensive drugs and prostheses. Not all are associated with longer hospital stays.

It is therefore risky to assume that an increase of i per cent in days of care is equivalent to an increase of i per cent in the costs of care. Indeed, if the model were perfect in every way, this assumption would be unreasonable: it would tend to overestimate the impact on cost of additional days of stay. However, there are features of the model which probably lead to underestimation of costs. For example, the adjustments have not taken account of all DRGs.

Second, the basis for adjustment is each hospital's complexity as indicated by its casemix in the year to which the data related. It may be advisable to take account of changes in the workload in the budget period to which the model is applied.

Third, there is reason to be cautious about the precision of the refined DRG severity class assignments. They have not been reviewed in the Australian context. It might also be argued that the slope of the weights should be less than a factor of four. However, the weighting was applied before the regression models were constructed. It is therefore unlikely that the weighting had a major effect. Indeed, very similar results are obtained if a simple count is used.

Fourth, by standardising mean lengths of stay, the possibility that the effect of severity as measured by this model varies with efficiency has been rejected as a relevant attribute of the model. It would be possible to undertake further analyses to explore the relationship. However, it might not add further validity for the purposes to which the model will be applied.

Fifth, it is extremely difficult to know whether the model is improved by weighting the additional days by DRG cost relativities. Fortunately, Table 1 shows that the results are little different. The best solution may involve both views: additional days in some DRGs are probably more expensive (but perhaps not in proportion to the DRG cost weights) and others are not.

Sixth, there may be significant differences in the accuracy and completeness of reporting of diagnoses and procedures across hospitals. It might therefore be argued that tertiary hospitals do not have greater severity, but simply more careful and thorough coders. The sharing of coding staff among hospitals and the presence of a statewide coding education and audit program reduce the plausibility of this alternative hypothesis, but it cannot be entirely rejected with the available data.

While the model may be criticised on these and other grounds, it is important to appreciate that less sensible assumptions are implied if adjustments of this type are not used. Indeed, there are five arguments in its favour.

First, it has greater clinical validity. Most people would expect that the treatment of additional conditions is systematically associated with increased resource use, and yet the pure DRG model only ever considers a maximum of one procedure and two diagnoses. It considers only one diagnosis in the majority of cases.

Second, elements of the results seem to be extremely robust. For example, all the models used thus far indicate that the larger teaching hospitals have higher levels of complexity. This is a plausible result, and indicates a strong underlying casemix logic.

Third, it is a logical extension of the process of continual refinement of casemix measures, which is a feature of the South Australian strategy. It is consistent with other extensions.

Fourth, it encourages full reporting of significant diagnoses and procedures. Some people might argue that it merely encourages gaming. However, this can be said of any funding model which is affected by clinical measures.

Fifth, it is a low-cost adjustment. It makes use of data which are routinely captured for other purposes, and involves no more than elementary statistical analysis.

In summary, it provides a reasonable solution to immediate needs. Improvements in the AN-DRG classification (including those in version 4 due for release in 1998) may reduce the degree to which external adjustments are required. However, the problems are unlikely to be fully resolved unless major structural changes are made; and these might be undesirable for other reasons.

References

Averill RF, McGuire TE, Manning BE, Fowler DA et al. 1992, 'A study of the relationship between severity of illness and hospital cost in New Jersey hospitals', *Health Serv Res*, vol 27, no 5, pp 587–606.

Bradbury RC, Stearns FE & Steen PM 1991, 'Interhospital variations in admission severity-adjusted hospital mortality and morbidity', *Health Serv Res*, vol 26, no 4, pp 407–24.

Christofferson J, Conklin JE & Gonella JS 1988, 'The impact of severity of illness on hospital costs', *DRG Monitor*, vol 6.

Commonwealth Department of Health and Family Services 1997, *An Evaluation of Casemix Funding in South Australia, 1994–95*, Commonwealth Department of Health and Family Services, Canberra, January.

Froehlich H & Jarvis WR 1991, 'Economic impact of diagnosis-related groups and severity of illness on reimbursement for central nervous system infections', *J Pediatr*, vol 118, no 5, pp 693–7.

Hammond J & Ward CG 1991, 'Burns in octogenarians', *South Med J*, vol 84, no 11, pp 1316–19.

Hindle D, Cromwell D & Halsall J 1995, 'Estimation of the effects of secondary diagnoses and procedures on within-DRG variations in length of stay', University of Wollongong.

Hindle D, Pilla J & Scuteri JJ 1991, 'Where are DRGs going?' *Australian Health Review*, September.

Horn SD & Horn RA 1986, 'The computerised severity index: A new tool for case mix management', *Journal of Medical Systems*, 10, pp 73–8.

Horn SD, Sharkey PD, Buckle JM, Backofen JE et al. 1991, 'The relationship between severity of illness and hospital length of stay and mortality', *Med Care*, vol 29, no 4, pp 305–17.

Iezzoni LI et al. 1990, 'Illness severity and costs of admissions at teaching and non-teaching hospitals', *JAMA*, vol 264, no 11.

Iezzoni LI, Ash AS, Coffman GA & Moskowitz MA 1991, 'Admission and mid-stay MedisGroups scores as predictors of hospitalisation charges', *Med Care*, vol 29, no 3, pp 210–20.

MacKenzie EJ, Steinwachs DM, Ramzy AI, Ashworth JW & Shankar B 1991, 'Trauma case mix and hospital payment: The potential for refining DRGs', *Health Serv Res*, vol 26, no 1, pp 5–26.

McMahon LF, Hayward RA, Bernard AM et al. 1992, 'APACHE-L: A new severity of illness adjuster for inpatient medical care', *Med Care*, vol 30, no 5, pp 445–52.

Rapoport J, Teres D, Lemeshow S, Avrunin JS et al. 1990, 'Explaining variability of cost using a severity-of-illness measure for ICU patients', *Med Care*, vol 28, no 4, pp 338–48.

Rosenthal GE & Landefeld CS 1993, 'Do older Medicare patients cost hospitals more? Evidence from an academic medical centre', *Arch Intern Med*, vol 153, no 1, pp 89–96.

Sharkey PD, Horn SD, Brigham PA et al. 1991, 'Classifying patients with burns for hospital reimbursement: Diagnosis-related groups and modifications for severity', *J Burn Care Rehabil*, vol 12, no 4, pp 319–29.

Stoskopf C & Horn SD 1992, 'Predicting length of stay for patients with psychoses', *Health Serv Res*, vol 26, no 6, pp 743–66.

Young WW 1985, 'Measuring the cost of care using patient management categories', Final Report, vols 1, 2 and 3, Washington, DC: HCFA Grant no 18-p-97063/3.

Young WW 1988, 'Disease staging and patient management categories', *Med Care*, vol 26.

Young JC, Macioce DP & Young WW 1990, 'Identifying injuries and trauma severity in large databases', *J Trauma*, vol 30, no 10, pp 1220–30.