Original Article

Improving the Diagnosis Related Grouping Model's Ability to Explain Length of Stay of Elderly Medical Inpatients by Incorporating Function-linked Variables†

S Sahadevan,[1] *FAMS, DGM (Lond), FRCP (Edin)*, A Earnest,[2] *MSc*, YL Koh,[3] *MBBS*, KM Lee,[4] *MBBS*, CH Soh,[5] *MSc*, YY Ding,[1] *FAMS, DGM (Glas), FRCP (Edin)*

Abstract

Introduction: This study first aimed to determine the adequacy of the Diagnosis Related Grouping (DRG) model's ability to explain (1) the variance in the actual length of stay (LOS) of elderly medical inpatients and (2) the LOS difference in the same cohort between the departments of Geriatric Medicine (GRM) and General Medicine (GM). We then looked at how these explanatory abilities of the DRG changed when patients' function-linked variables (ignored by DRG) were incorporated into the model. Materials and Methods: Basic demographic data of a consecutively hospitalised cohort of elderly medical inpatients from GRM and GM, as well as their actual LOS, discharge DRG codes [with their corresponding trimmed average length of stay (ALOS)] and selected function-linked variables (including premorbid functional status, change in functional profile during hospitalisation and number of therapists seen) were recorded. Beginning with ALOS, function-linked variables that were significantly associated with LOS were then added into two multiple liner regression models so as to quantify how the functional dimension improved the DRGs' abilities to explain LOS variances and interdepartmental LOS differences. Forward selection procedure was employed to determine the final models. For the interdepartmental analysis, the study sample was restricted to patients who shared common DRG codes. Results: 114 GRM and 118 GM patients were studied. Trimmed ALOS alone explained 8% of the actual LOS variance. With the addition of function-linked variables, the adjusted R^2 of the final model increased to 28%. Due to common code restrictions, the data of 79 GRM and 78 GM patients were available for the analysis of interdepartmental LOS differences. At the unadjusted stage, the median stay of GRM patients was 4.3 days longer than GM's and with adjustments made for the DRGs, this difference was reduced to 3.9 days. Additionally adjusting for the patients' functional features diminished the interdepartmental LOS discrepancy even further, to 2.1 days. Conclusion: This study demonstrates that for elderly medical inpatients, the incorporation of patients' functional status significantly improves the DRG model's ability to predict the patients' actual LOS as well as to explain interdepartmental LOS differences between GRM and GM.

Ann Acad Med Singapore 2004;33:614-22

Key words: Aged, Prospective payment system, Resource allocation

Introduction

Casemix refers to the numbers and types of patients within a healthcare setting and Diagnosis Related Groupings (DRGs) represent one mode of classifying casemix.[1] In essence, DRGs are categories of clinically meaningful patient conditions which require similar levels of hospital resources for their treatment.[2] The original aim of DRGs at its inception in 1967 was to enable the measurement and evaluation of the outcomes of health services, but it soon became apparent to the Health Care Financing

[1] Department of Geriatric Medicine
[2] Department of Clinical Epidemiology
Tan Tock Seng Hospital, Singapore
[3] Department of Medicine
National University Hospital, Singapore
[4] Department of Forensic Psychiatry
Institute of Mental Health, Singapore
[5] Department of Biostatistics
Harvard University, USA
Address for Reprints: Dr Suresh Sahadevan, Department of Geriatric Medicine, Tan Tock Seng Hospital, 11, Jalan Tan Tock Seng, Singapore 308433.
Email: suresh_sahadevan@ttsh.com.sg

† A commentary on this topic is printed on page 660.

Administration in the United States that a prospective funding mechanism for Medicare patients – which was then being increasingly regarded as a vital strategy to curb the nation's spiralling healthcare costs – could also be built on the DRG classification system.[3,4]

In the United States, the DRG-based subvention system attracted controversy from its very beginnings.[5] Its relative effectiveness as a cost-containment tool was increasingly appreciated; also perceived positively was the DRG system's ability to provide relevant patient-centred information that enhanced effective management of hospitals as well as enabled healthcare organisations to better plan for future resource utilisation. On the other hand, concerns were voiced about the limitations of the DRG approach, especially its inability to adequately incorporate the severity dimension of illnesses.[6]

Despite these controversies, the relative success of the DRG-based prospective payment system in controlling Medicare costs in the United States[7] led to its adoption in many other countries over the last two decades. Casemix funding of inpatient and daycare surgery services in public sector hospitals was introduced to Singapore in October 1999 and it was based on the Australian National DRG or AN-DRG system. The primary impulse motivating this change was the need for a more effective healthcare cost-containment strategy than the existing per diem funding, which would also be fair and, most importantly, be capable of motivating the healthcare providers themselves to reduce, or at least minimise the escalation of, their operational costs.

With the advent of a DRG-based prospective payment system in Singapore, departments of geriatric medicine primarily focusing on frail older patients have experienced poor financial performance – as reflected by their longer length of inpatient stay – compared to departments of general medicine. Such has been the case with the Department of Geriatric Medicine (GRM) at Tan Tock Seng Hospital, which typically admits elderly patients with certain core geriatric syndromes (such as immobility, falls, dementia and incontinence). These syndromes are characterised by functional declines and generally, the additional consideration for admission to GRM is that the patients' functional declines are judged to be amenable to rehabilitation. Those elderly patients who do not require a significant functional assessment (either because they are still independent in spite of their illnesses or they are already chronically and totally disabled with no prospects for rehabilitative recovery) are more straightforwardly managed in the General Medicine Department (GM).

It had become increasingly clear to us that while our department's admission policy (selectively targeting which elderly medical inpatients require the services of GRM) promoted efficiency at 1 level of the healthcare system, it simultaneously generated inefficiency at the level of our DRG-based financial performance. We recognised that the nature of geriatric syndromes was such that they would generally prolong acute hospital stays (even when the major part of the attendant rehabilitation was conducted off-site). We thus hypothesised that the main reason for the adverse DRG-related reports stemmed from the systematic inability of the instrument's organ-based diagnostic categories to adequately embrace the functional dimension of medical illnesses in the elderly. The study reported herein was conducted to specifically to verify this hypothesis.

Materials and Methods

One hundred and fourteen GRM patients (all 65 years of age and above) and 118 elderly general medicine (GM) patients (in the same age bracket) admitted consecutively over a 2-month period in 1999 from our acute-care hospital were prospectively studied, regardless of their admitting diagnoses. The following variables were collected from each patient: (a) age (young-old = 65 to 74 years; old-old = 75 years and above); (b) gender; (c) functional status [independent = independent in both ambulation and basic activities of daily living (ADL); impaired = requiring assistance in either ambulation or in 1 or more of the basic ADL]. This functional status variable was measured at 3 different points in time in relation to the hospitalisation episode: the pre-morbid period (defined as being before the onset of the latest round of illness), at admission time and at discharge time; (d) change in functional status at discharge time compared to the pre-morbid period (no change; worsened); (e) total number of referrals to therapists (physiotherapists, occupational therapists, speech therapists) during hospitalisation; (f) referrals to medical social worker (MSW) during hospitalisation (yes; no); (g) the actual length of stay (LOS) during the hospitalisation period and (h) the DRG-specified trimmed average LOS (ALOS) pertaining to that same hospitalisation. LOS was used as a proxy for resource consumption because the data on patients' actual cost of hospitalisation (which, unlike LOS, also reflects intensity of care) was unavailable at the time of the study.

Apart from age, gender, trimmed ALOS and actual LOS, all the other variables were selected because of their ability to reflect various aspects of the functional dimension of the patients (i.e. their inabilities to walk, eat, dress, bathe or use the toilet), a dimension which the largely organ-based DRG model tends to ignore. For ease of description in this paper, these other variables (together with age and gender) are sometimes referred to as a group of function-linked variables.

This non-interventional study was done with the approval of our hospital's ethics committee.

Statistical Analysis

This study involved 2 main multiple linear regression analyses and in both, the patients' actual LOS was the dependent variable. As the actual LOS was skewed and the residuals in the final models were not normally distributed, we analysed this variable on the natural logarithmic scale. The first of the multivariable analyses looked at how well the variances in the actual lengths of stay of all patients in the study was explained initially by the DRG's trimmed ALOS variable alone and subsequently, with the addition of function-linked variables. (Prior to the multivariable modelling, the overall distributions of the variables as well as their associations with the patients' actual lengths of stay were described. Only the statistically significant variables obtained through the bivariable analyses were considered as co-variates for the multivariable analysis.) The model-building was done sequentially, beginning with the DRG's trimmed ALOS as the first independent variable to be entered into the equation so as to highlight the impact of how the subsequent addition of other variables changed the model's overall ability to explain the patients' actual LOS beyond what the DRG variable itself could do.

Interdepartmental LOS differences (between GRM and GM) were the focus of the second multiple regression analysis and for this part of the study only those patients from the 2 departments who showed common DRG codes were included. Common codes ensured that the same DRG-specified trimmed ALOS applied to both departments. The patients' actual LOS (on the log scale) again was the dependent variable and model-building was likewise done sequentially. This time, however, the modelling began with the department variable (GRM or GM) to first determine its sole explanatory power (in addressing the variance of the patients' actual LOS) as well as to quantify the initial LOS difference between the 2 departments. Subsequently, the DRG's trimmed ALOS was added to the equation followed by, at the last stage, other significant function-linked variables, so as to note the incremental changes in the model's ability to explain the patient's actual LOS as well as to determine how the interdepartmental LOS difference progressively altered when statistical adjustments were made for these additional variables.

For both multivariable analyses, forward selection procedure was used to determine the final covariates of the best explanatory model (with *P* value of entry and removal at 5% and 10% respectively). Variables that were significant at the 10% level in the bivariable analysis were included in the multivariate model selection procedure. Diagnostic checks for multicollinearity using variance inflation factors, heteroscedasticity and normality using residual plots and outlier/influential observations were also carried out. All statistical tests were conducted using the Statistical Package for Social Sciences, version 9. Level of significance was set at 5%.

Results

Descriptive Features of All Subjects

The demographics and function-linked profile of this study are shown in Table 1 from the perspectives of the respective Departments (GRM and GM) as well as in total. The GRM patients were older and more functionally impaired (premorbidly and at the time of discharge), had more referrals to therapists during their hospitalisations, ultimately had longer lengths of stay and received more referrals to community hospitals (step-down care facilities) or nursing homes at the time of discharge. In brief, GRM patients were older and frailer.

DRG's Ability to Explain LOS (Involving All Subjects)

The bivariable associations between each of the study variables and the patients' lengths of stay are shown in Table 2.

The multivariable analysis was then done sequentially to enhance the explanatory features of the study variables on the patients' LOS. Using the adjusted R^2 of a linear regression model (where the actual LOS constituted the dependent variable), the DRG's trimmed ALOS could only explain 8% of the variance in the actual LOS. In the next stage, all the significant variables shown in Table 2 (including trimmed ALOS) were entered into the regression model, and a forward selection method was used to select the most important variables that could explain the actual LOS. The following variables were then picked up as being statistically significant: functional status at discharge, total number of referrals to therapists and trimmed ALOS. The adjusted R^2 of this model was now 28% (Table 3).

DRG's Ability to Explain Interdepartmental Differences in LOS

This part of the study focused only on patients from both departments who shared common DRG codes. This restriction was necessary so as to ensure that different DRG codes would not be a confounding factor in explaining the differences in LOS between GRM's and GM's study patients. As a result of this restriction, the total number of patients dropped to 157 (79 from GRM and 78 from GM).

It was noted that the overall patterns in the distribution of variables of the common code patients did not differ from the patterns seen in the full group of patients (as shown in Table 1), except for referrals to medical social workers and community hospitals, where narrowing to common DRG codes showed GRM to have significantly more referrals to medical social workers and where the differences between the departments in terms of referrals to community hospitals

Table 1. Descriptive Statistics by Department (Among All Subjects)

	Total	(n = 232)	GRM	(n = 114)	GM	(n = 118)	*P* value #
	n	%	n	%	n	%	
Age (y)							
65-74	88	37.9	30	26.3	58	49.2	
≥75	144	62.1	84	73.7	60	50.8	
Median (Range)	77	(65-98)	81	(65-98)	75	(65-96)	<0.001*
Gender							
Male	99	42.7	49	43.0	50	42.4	0.925
Female	133	57.3	65	57.0	68	57.6	
Overall functional profile (pre-morbid)							
Independent	151	65.1	62	55.3	88	74.6	0.002
Functionally impaired (PD/TD)	81	34.9	52	44.7	30	25.4	
Overall functional profile (admission)							
Independent	89	38.5	37	33.3	52	44.4	0.083
Functionally impaired (PD/TD)	142	61.5	77	66.7	65	55.6	
Overall functional profile (discharge)							
Independent	124	53.4	44	39.5	79	66.9	<0.001
Functionally impaired (PD/TD)	108	46.6	70	60.5	39	33.1	
Change in functional profile at discharge from pre-morbid							
No change	205	88.4	96	84.2	109	92.4	0.053
Worsen	27	11.6	18	15.8	9	7.6	
Referrals during hospitalisation							
Medical Social Worker	26	11.2	17	14.9	9	7.6	0.079
Total no. therapy referrals Median (Range)	0	(0-3)	1	(0-3)	0	(0-2)	<0.001*
Referrals to community hospital or nursing home at time of discharge							
	41	17.7	27	23.7	14	11.9	0.018
Number of outliers	45	19.5	35	31.0	10	8.5	<0.001
Actual length of stay							
Median (Range)	8.0	(2.0-53.0)	11.0	(3.0-53.0)	6.0	(2.0-49.0)	<0.001*

GRM: Department of Geriatric Medicine; GM: Department of General Medicine

Pearson χ^2 test unless specified otherwise

* Mann-Whitney test

or nursing homes became insignificant. Likewise, the bivariable association findings in this restricted cohort were similar to those shown in Table 2, except that the median lengths of stay were no longer significantly different across the categories of age group and changes in functional profile from premorbid to discharge.

Again using the adjusted R^2 of the linear regression model (with LOS – analysed on the log scale – being the dependent variable) on this restricted cohort, the influence of the Department factor (GRM or GM) on the variance of the LOS was first studied. It was found that the Department factor could explain 17% of the actual LOS's variance and this equation also showed that, on average, the GRM patients stayed 4.3 extra days compared to GM. When the DRG's trimmed ALOS was next added into the multivariable analysis, the model's adjusted R^2 increased to 23% and the interdepartmental LOS difference decreased to 3.9 days.

In the final model of this restricted cohort, all of the above-mentioned study variables (including trimmed ALOS and Department) were subjected to a forward selection procedure. The following variables were then selected as being statistically significant in explaining patients' actual LOS: functional profile at discharge, total number of referrals to therapists, trimmed ALOS and department. GRM patients, however, were now staying, on average, only 2.1 days more than GM patients (Table 4).

In terms of diagnostics, we found that the residuals for the models were reasonably normally distributed, with a homogeneous variance. We also did not find any substantial multi-collinearity between the final independent variables in the model.

Table 2. Comparison of Average Length of Stay by Various Covariates (Among All Subjects)

Covariates	Actual length of stay (days)	*P* value*
	Median (Range)	
Age (y)		0.011
65-74	7.0 (2.0-53.0)	
≥75	9.0 (2.0-49.0)	
Gender		0.377
Male	8.0 (2.0-39.0)	
Female	8.0 (2.0-53.0)	
Overall functional profile (pre-morbid)		<0.001
Independent	6.5 (2.0-49.0)	
Functionally impaired (PD/TD)	11.0 (3.0-53.0)	
Overall functional profile (admission)		<0.001
Independent	6.0 (2.0-35.0)	
Functionally impaired (PD/TD)	10.0 (3.0-53.0)	
Overall functional profile (discharge)		<0.001
Independent	6.0 (2.0-35.0)	
Functionally impaired (PD/TD)	11.0 (3.0-53.0)	
Change in functional profile (indep/impaired) at discharge from pre-morbid		0.006
No change	8.0 (2.0-53.0)	
Worsen	11.0 (3.0-49.0)	
Referrals to medical social worker		0.009
No	8.0 (2.0-53.0)	
Yes	11.0 (4.0-33.0)	
Referrals to community hospital or nursing home at time of discharge		0.002
No	7.0 (2.0-49.0)	
Yes	11.0 (4.0-53.0)	
Total number of therapy referrals	r = 0.436	<0.001

* Mann-Whitney test
r: Spearman's Correlation

Table 3. Predictors of Length of Stay Among All Subjects (Final Model)

	All patients		
	β-coefficient	95% CI	*P* value
(Constant)	1.65 (0.08)	(1.50, 1.81)	<0.001
Overall functional profile at discharge (1: functionally impaired 0: independent)*	0.22 (0.08)	(0.06, 0.39)	0.008
Total number of therapy referrals	0.24 (0.04)	(0.15, 0.32)	<0.001
Trimmed ALOS	0.03 (0.01)	(0.007, 0.057)	0.011
Adjusted r-square	0.282		
r-square	0.292		

ALOS: average length of stay; Standard error in parentheses.
* Calculated on the dichotomous scale (indept vs functionally impaired)
Length of stay calculated on the natural logarithmic scale

Analysis Without Outliers

Given the fact that GRM had a high percentage of outliers, we also wanted to ascertain whether, and how, the major patterns noted above, would change if we were to remove outliers from the analyses altogether. We thus repeated the earlier multiple regression analyses, but this time focusing only on the inliers.

The final model selected to explain the actual LOS included the following significant variables: overall functional profile at admission, total number of therapy referrals and trimmed ALOS. Trimmed ALOS alone could explain 23.8% (adjusted R-square) of the variation in the actual LOS. Once adjusted for the above-mentioned function-related variables, the final model could explain 31.4% of the variation.

As for explaining the diminishing department differences among patients on common DRG codes, the final model selected overall functional profile at admission, trimmed ALOS, referral to MSW and department as the significant covariates. Department alone presently explained 10.1% of the variation in the actual LOS, and GRM patients stayed on average only 2.5 days longer than GRM patients. After adjusting for trimmed ALOS, the adjusted R^2 went up to 28.1% and the interdepartmental difference narrowed to 1.9 days. Finally, after adjusting for the function-related variables mentioned above, the adjusted R-square went up to 34.5% and the interdepartmental difference was now 1.7 days.

In essence, compared to the analysis with outliers, the parallel analyses with inliers reveal that DRG's trimmed ALOS has a better explanatory status for both the variances in LOS as well as interdepartmental LOS differences. While function-linked variables still increased the adjusted R^2 (beyond that of the DRG variable) in the final models that looked at LOS variances and interdepartmental differences, overall, this contribution from the function-linked variables was less in magnitude compared to the outlier-included models. In sum, when compared to the inliers, the outliers were significantly characterised by many more function-related difficulties.

Since outliers are methodologically induced by the DRG model itself and GRM's outlier frequency was significantly higher and more problematic than GM's, the ensuing discussion revolves around the results that have included outliers.

Discussion

The importance of the functional dimension in both the diagnostic evaluation and the appropriate management of the clinical problems of frail elderly patients has been

Table 4. Explaining Diminishing Department Differences Among Common Code Subjects (Final Model)

	All patients		
	β-coefficient	95% CI	*P* value
(Constant)	1.55 (0.11)	(1.34, 1.76)	<0.001
Overall functional profile at discharge (1: functionally impaired 0: independent)*	0.28 (0.10)	(0.09, 0.47)	0.004
Trimmed ALOS	0.04 (0.02)	(0.004, 0.07)	0.025
Total number of therapy referrals	0.13 (0.05)	(0.03, 0.24)	0.015
Department (GRM)	0.30 (0.10)	(0.10, 0.50)	0.004
Adjusted r-square	0.314		
r-square	0.332		

ALOS: average length of stay; GRM: Department of Geriatric Medicine; Standard error in parentheses.

* Calculated on the dichotomous scale (indept vs functionally impaired)

Length of stay calculated on the natural logarithmic scale

Note: Interdepartmental difference was calculated as 2.1 days, after adjusting for the other covariates. We used the median length of stay for the GM patients (i.e. 6 days) to calculate the interdepartmental difference. We obtained the ratio of the length of stay between GRM and GM, by taking the exponential of the model coefficient (i.e. exp(0.30)). This gave us a value of 1.35. The departmental difference was then calculated as 1.35*6 - 6 = 2.1 days.

receiving increasing attention in medical literature. The traditional organ-based approaches of general medicine are notoriously inadequate for this cohort of patients. Although organ-defined medical diagnoses are important, function is central to the health and well-being of older people. Medical illness may present with functional decline – often as geriatric syndromes such as impaired cognition (delirium and dementia), impaired mobility, instability (falls), incontinence and impaired feeding. Functional decline has been shown to be associated with adverse outcomes such as increased mortality,[8-10] hospitalisation,[11] delayed discharge[12] and institutionalisation.[13] Available evidence indicates that appropriate geriatric intervention known as geriatric evaluation and management (GEM) can improve several of these outcomes.[14] It is becoming clearer that good care of frail older patients not only requires good medical management but also an adequate period of rehabilitation in the many instances where functional decline has resulted from acute illness or hospitalisation itself.

In essence, this study shows how the addition of function-linked variables to the DRG model significantly improves the latter's ability to explain firstly, the variances in the LOS of a cohort of hospitalised elderly patients and secondly, the interdepartmental LOS differences between GM and GRM patients within this same cohort. While it is true that, in absolute terms, the explanatory power of LOS variance in the final model is modest (28%), the value of the study becomes clearer as we reflect that if the existing DRG-only model was effective, it should have adequately explained the variance in the LOS of our study cohort (far better than our observed 8%). Moreover, a good DRG model should hardly result in significant interdepartmental LOS differences for the same DRGs. The study thus highlights the inadequacy of our current DRG tool and suggests that for elderly patients at least – and in particular, frail elderly patients – the incorporation of functional status can be 1 way of improving the legitimate predictive functions of the DRG.

Differences in LOS between GRM and GM can simply be a reflection of different case protocols and practices between the departments. One part of such interdepartmental variance can therefore legitimately arise from differences in specialty-specific practices that are good and appropriate for patients while another part for the variance may arise from differences in departmental practice efficiencies. It can thus be argued that what the observed interdepartmental differences in this study is showing is the inefficiency of the medical approaches taken by GRM compared to GM, for the same diagnostic conditions. This conclusion however is unlikely because when the function-linked variables were selected and adjusted for by multivariable analysis, not only did the adjusted R^2 of the statistical model consistently increase beyond what the DRG variable itself could contribute, but more importantly, the interdepartmental LOS difference diminished considerably. If the initial problem had been one of inefficient and inappropriate practices, then one would not expect such decreases in LOS differences between GRM and GM merely by adjusting for the patients' functional characteristics. In fact, the findings of the study strongly suggest that the patient severity factor in the older age groups (defined in terms of functional features and changes) is inadequately captured by the existing DRG codes, which are largely organ-based in their definitions.

It can be criticised that the above perspectives are only significant when outliers are included in the analysis. While it is true that the DRG's explanatory abilities in determining inliers' LOS was better than with outliers, and function-linked variables additionally contributed less impressively to the explanation of variance in inliers' LOS, we believe it is better to include outliers for several reasons. The outliers – whose funding is manifestly poor in relation to their total accrued costs – are a group defined by the DRG system itself, and every year, the negative financial impact of these outliers, to our Department, is considerable. Moreover, the outlier proportion in GRM has been consistently and markedly higher than GM's, which alludes to an underlying systemic problem. Finally, and most importantly, when function-linked variables are added

onto the DRG model, the explanatory power of the combined model to account for the marked variance in patients' LOS (brought about primarily by the outliers) is very much better. In brief, should the DRG model also incorporate an adequate measure of patients' illness-induced functional difficulties and so adjust for this dimension, the Department of Geriatric Medicine will have a far lower percentage of outliers.

How do our results compare with published findings elsewhere regarding DRGs and their association with resource consumption? The overall poor explanatory power of the DRGs in accounting for LOS variance is consistent with the observation that medical DRGs tend to explain only 5% to 10% of the length of hospital stays.[15] Other investigators have also shown elderly medical inpatients to consistently have unfavourable DRG profiles (in terms of increased hospital stay and/or hospitalisation costs)[11,16-21] and some of these studies have specifically implicated the poor functional status of the patients.[10,17] Most of these papers reported the reality of their ground-level experiences with untrimmed data (that is, data which included outliers). Horn et al in their many papers have continually pointed out the importance of measuring patients' severity of illness (by way of an index which also incorporates the functional dimension of illness) and combining it with DRGs so as to have a model that is markedly better in predicting patients' LOS and hospital costs.[22-27] While, with the use of large administrative data, Jencks et al showed that DRGs required no further adjustments (of patients' age and socio-economic status)[28], other works have pointed out that patterns and averages observed with large levels of aggregation (such as national level data) do not prevail with lower levels of aggregation (such as hospital level data) where variance inevitably increases.[18,20] In fact, the major work that demonstrated the benefits of DRG to the Medicare system in the United States (better value coupled with non-deteriorating standards of clinical care) made use of very large administrative data.[29,30]

The highlighted systematic limitations associated with the use of the DRG model have several practical implications. In collating and studying DRG-related performances at the departmental level, it would thus be inappropriate to conclude that departments focusing on the frail elderly medical patients and having above-average LOS patterns are inefficient in their practices. The inadequate risk-adjustment properties of the DRGs, as evidenced by this study, do not permit such conclusions to be made easily (even when restricting its use to inliers only, the DRG variable explains only around 25% of LOS variances). Similar considerations apply when appraisals are made at the individual physician level.[31] While it has been shown that, compared to unadjusted clinical outcome data, DRG-based casemix adjustment leads to more meaningful physician practice profiles,[32] Horn et al[33] have reported that combining a severity of illness index to the DRG adjustment results in even more accurate practice profiles. As a general principle, performance appraisal with DRG data is best done longitudinally (over time), with a careful consideration of underlying assumptions, rather than cross-sectionally.

DRG-based subvention, because it is working on the principle of averages, thus works best at higher levels of aggregation (e.g. at cluster or hospital level). DRGs are less effective at lower levels of disaggregation (e.g. at departmental or individual doctor level) because the smaller number of cases at these levels mean that 1 or more outliers can potentially skew the averages markedly. While the DRG system therefore is suitable as a means of distributing subvention to the clusters, it has to be used with caution when distributing subvention at departmental level or in determining departmental level financial performance. Thus, DRG-linked departmental performance alone cannot be the guiding principle for hospitals when they redistribute the subventions received to their various departments (while that principle alone may be adequate at the ministry level for allocating subventions to clusters). Hospitals, for example, may choose to centralise the treatment of complex cases in 1 department and fund that department differently to reflect its complexity of care; with such policies, the department treating the complex cases then need not show any significant financial deficits.

However, if these additional considerations are deemed to be rather infeasible at the hospital level, then on the basis of the findings of our study, it becomes important to rectify the systematic bias the current DRG model has against frail elderly medical patients. Several options are available, at least theoretically. One would be to take Geriatric Medicine that focuses singularly on the frail elderly out of the DRG mode of funding altogether. In fact, the AN-DRG excludes geriatric medicine on the basis of its subacute status with subacute care being defined as where the "principal diagnosis (modified for factors such as age and procedures) is not adequate in explaining the need for, or the cost of, the services that she receives".[34] Alternatively, separate funding methods can also be considered for the discipline. Thus while DRG-based funding can cover the acute illness phase of frail elderly patients' hospitalisations, per diem funding can be activated when the patients subsequently receive primarily rehabilitative care either till discharge or till transfer to a step-down care facility. Such methods, however, will increase administrative complexity, requiring also periodic independent audits to minimise gaming or abuse. Outlier funding will also require review.

Another option would be to improve the DRG model

itself for the type of frail patients that Geriatric Medicine typically manages. This could be by way of adding on new diagnostic terms – that are more appropriate for frail elderly medical inpatients – as principal diagnoses (such as the geriatric syndromes that transcend organ-based classifications) and to then nationally compute their actual resource weights. Alternatively, the current diagnostic entities (based on the International Classification of Diseases system) could be retained, but with validated function-linked variables or severity of illness measures additionally captured so as to more accurately reflect true resource utilisation. It must be recognised, however, that the latter approach requires more unwieldy data collection processes which are perhaps only justifiable if the new system has been empirically demonstrated to be significantly superior to the earlier method.[35] In essence, improving the DRG model for the frail elderly requires major research and administrative efforts. Given also the multiple causes of heterogeneity in patient care, one can reasonably expect that a significant proportion of the problems highlighted in our paper will still remain with all future versions and forms of DRG models.

As the study was conducted 4 years ago, a relevant query would be whether, from then till now, there have been any significant differences in the hospitalisation episodes of the 2 departments since then. The median LOS (in days) between GRM and GM have been, respectively, 10 and 5 (for the year 2001) and 9 and 3 (for the year 2002). The outlier percentages for GRM have been 20% (2001) and 16% (2002) while the corresponding figures for GM were 6% and 7%.[37] In essence, compared to the study period, GRM's median LOS has been decreasing and its outlier figures dropping, but these values have always been markedly higher than GM's. We have no reason to believe that, since the study, more functionally independent elderly patients are being admitted to our department and we thus ascribe the longitudinal improvement in our department's LOS profile to be likely due to the advent of case managers (from mid-2000 onwards) in our hospital. Case managers are trained to look at the functional and social status of patients as well as to proactively arrange for their appropriate discharges or transfers (to downstream facilities). We believe case managers have been actively influential in facilitating the discharge plans of frailer patients. Our department's outlier status, though much less than that seen in 1999, is still markedly high and it is very likely that these outliers are characterised by the same significant functional difficulties as indicated by our study findings.

Conclusion

This study shows the significant extent to which the DRG model is systematically biased against frail elderly medical patients because of its neglect of the functional dimension of illness. This bias results in an inadequate risk adjustment for predicting inpatient costs in this subset of patients. It will not be easy, however, to correct this bias and one can thus expect that future models of DRG will most likely, to varying degrees, retain such limitations. Funding methods for acute hospital-based geriatric medicine will therefore require review and new approaches considered. This paper has also highlighted that DRG is a cost-containment tool whose central strategy of computing average costs of hospitalisation episodes works best at higher levels of aggregation (e.g. nation-wide data and clusters) and whose effectiveness for cross-sectional comparisons progressively diminishes at lower levels of aggregation (e.g. departments and clinicians). It is important for clinicians, healthcare policy makers and administrators to be aware of these matters to enable the wise and careful use of the DRG tool.

Acknowledgements

The authors are grateful to Ms Loh Shu Ching, Manager of Health Information Services Department of Tan Tock Seng Hospital, for always being able to cheerfully and promptly help out – despite her own heavy responsibilities – with the many DRG-related data checks that we continually asked for throughout this period of work.

REFERENCES

1. Cheah J, Chee YC. Case mix – for better or for worse? Singapore Med J 1999;40:9-12.
2. Baker JJ. Medicare payment system for hospital inpatients: diagnosis-related groups. J Health Care Finance 2002;28:1-13.
3. Davis C, Rhodes DJ. The impact of DRGs on the cost and quality of health care in the United States. Health Policy 1988;9:117-31.
4. Siow JK. Casemix in Singapore – a clinician's perspective. Ann Acad Med Singapore 2001;30(Suppl):1-2.
5. Spiegel AD, Kavaler F. The debate over diagnosis related groups. J Community Health 1985;10:81-92.
6. Berenson RA, Pawlson LG. The Medicare prospective payment system and the care of the frail elderly. J Am Geriatr Soc 1984;32:843-8.
7. Wilensky GR. Medicare at 25. Better value and better care. JAMA 1990;264:1996-7.
8. Jarrett PG, Rockwood K, Carver D, Stolee P, Cosway S. Illness presentation in elderly patients. Arch Intern Med 1995;155:1060-4.
9. Inouye SK, Peduzzi PN, Robison JT, Hughes JS, Horwitz RI, Concato J. Importance of functional measures in predicting mortality among older hospitalized patients. JAMA 1998;279:1187-93.
10. Covinsky KE, Justice AC, Rosenthal GE, Palmer RM, Landefeld CS. Measuring prognosis and case mix in hospitalized elders. The importance of functional status. J Gen Intern Med 1997;12:203-8.
11. Ferrucci L, Guralnik JM, Pahor M, Corti MC, Havlik RJ. Hospital diagnoses, Medicare charges, and nursing home admissions in the year when older persons become severely disabled. JAMA 1997;277:728-34.
12. Chin JJ, Sahadevan S, Tan CY, Ho SC, Choo PW. Critical role of functional decline in delayed discharge from an acute geriatric unit. Ann Acad Med Singapore 2001;30:593-9.
13. Fortinsky RH, Covinsky KE, Palmer RM, Landefeld CS. Effects of functional status changes before and during hospitalization on nursing

home admission of older adults. J Gerontol A Biol Sci Med Sci 1999;54:M521-6.
14. Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. Lancet 1993;342:1032-6.
15. Essock SM, Norquist GS. Toward a fairer prospective payment system. Arch Gen Psychiatry 1988;45:1041-4.
16. Sager MA, Rudberg MA. Functional decline associated with hospitalization for acute illness. Clin Geriatr Med 1998;14:669-79.
17. Dunstan EJ, Amar K, Watt A, Seymour DG. First steps in building ACME – an admission case-mix system for the elderly. Age Ageing 1996;25:102-8.
18. Munoz E, Rosner F, Chalfin D, Goldstein J, Margolis IB, Wise L. Financial risk and hospital cost for elderly patients. Age- and non-age-stratified medical diagnosis related groups. Arch Intern Med 1988;148:909-12.
19. Sands GH, Munoz E, Gottesman M, Mulloy K, Wise L. Neurology, age, hospital costs, and DRGs. Neurology 1988;38:655-60.
20. Munoz E, Rosner F, Chalfin D, Goldstein J, Margolis I, Wise L. Age, resource consumption and outcome for medical patients at an academic medical center. Arch Intern Med 1989;149:1946-50.
21. Rosenthal GE, Landefeld CS. Do older Medicare patients cost hospitals more? Evidence from an academic medical center. Arch Intern Med 1993;153:89-96.
22. Horn SD, Sharkey PD, Betram DA. Measuring severity of illness: homogenous case mix groups. Med Care 1983;21:14-30.
23. Horn SD, Bulkley G, Sharkey PD, Chambers AF, Horn RA, Schramm CJ. Interhospital differences in severity of illness. Problems for prospective payment based on diagnosis-related groups (DRGs). N Engl J Med 1985;313:20-4.
24. Horn SD, Sharkey PD, Chambers AF, Horn RA. Severity of illness within DRGs: impact on prospective payment. Am J Pub Health 1985;75: 1195-99.
25. Horn SD, Horn RA, Sharkey PD, Chambers AF. Severity of illness within DRGs. Homogeneity study. Med Care 1986;24:225-35.
26. Horn SD, Horn RA. Reliability and validity of the Severity of Illness Index. Med Care 1986;24:159-78.
27. Horn SD, Sharkey PD, Buckle JM, Backofen JE, Averill RF, Horn RA. The relationship between severity of illness and hospital length of stay and mortality. Med Care 1991;29:305-17.
28. Jencks SF, Kay T. Do frail, disabled, poor, and very old Medicare beneficiaries have higher hospital charges? JAMA 1987;257:198-202.
29. LoGerfo JP. The prospective payment system and quality. No skeletons in the closet. JAMA 1990;264:1995-6.
30. Kahn KL, Rubenstein LV, Draper D, Kosecoff J, Rogers WH, Keeler EB, et al. The effects of the DRG-based prospective payment system on the quality of care for hospitalized Medicare patients: an introduction to the series. JAMA 1990;264:1953-5.
31. Kassirer JP. The use and abuse of practice profiles. N Engl J Med 1994;330:634-6.
32. Salem-Schatz S, Moore G, Rucker M, Pearson SD. The case for case-mix adjustment in practice profiling. When good apples look bad. JAMA 1994;272:871-4.
33. Horn SD, Horn RA, Moses H. Profiles of physician practice and patient severity of illness. Am J Public Health 1986;76:532-5.
34. Lee LA, Eagar KM, Smith MC. Subacute and non-acute casemix in Australia. Med J Aust 1998;169:S22-5.
35. Jencks SF, Dobson A. Refining case-mix adjustment. The research evidence. N Engl J Med 1987;317:679-86.
36. Tan Tock Seng Hospital Data Warehouse. Tan Tock Seng Hospital. Singapore. Accessed October 2003.
37. National Health Group Data Warehouse. Computer Associates. Singapore. Accessed October 2003.