Coding Response to a Case-Mix Measurement System Based on Multiple Diagnoses

Colin Preyra

Objective. To examine the hospital coding response to a payment model using a case-mix measurement system based on multiple diagnoses and the resulting impact on a hospital cost model.

Data Sources. Financial, clinical, and supplementary data for all Ontario short stay hospitals from years 1997 to 2002.

Study Design. Disaggregated trends in hospital case-mix growth are examined for five years following the adoption of an inpatient classification system making extensive use of combinations of secondary diagnoses. Hospital case mix is decomposed into base and complexity components. The longitudinal effects of coding variation on a standard hospital payment model are examined in terms of payment accuracy and impact on adjustment factors.

Principal Findings. Introduction of the refined case-mix system provided incentives for hospitals to increase reporting of secondary diagnoses and resulted in growth in highest complexity cases that were not matched by increased resource use over time. Despite a pronounced coding response on the part of hospitals, the increase in measured complexity and case mix did not reduce the unexplained variation in hospital unit cost nor did it reduce the reliance on the teaching adjustment factor, a potential proxy for case mix. The main implication was changes in the size and distribution of predicted hospital operating costs.

Conclusions. Jurisdictions introducing extensive refinements to standard diagnostic related group (DRG)-type payment systems should consider the effects of induced changes to hospital coding practices. Assessing model performance should include analysis of the robustness of classification systems to hospital-level variation in coding practices. Unanticipated coding effects imply that case-mix models hypothesized to perform well *ex ante* may not meet expectations *ex post*.

Key Words. Prospective payment, case mix, hospital costs

Case-mix based reimbursement of hospitals is internationally pervasive, with the most common¹ patient classification systems in use being those related to diagnostic related groups (DRG). Since the introduction of the Prospective Payment System (PPS) in the United States, a contentious aspect of the DRG system has been the presence of cost variation within DRGs that could

potentially be explained by measurable differences in patient severity. The literature generally concludes that residual intra-DRG variation for a given patient sample can be reduced by refining classification groups to take into account the presence of combinations of specific complications and comorbid conditions (Horn et al. 1985; Averill et al. 1992; Schwartz et al. 1996; Lynk 2001).

These findings have been the basis for policy changes and recommendations. Based on their recent reviews of the DRG system, the Medicare Payment Advisory Commission (MedPac 2000) found that modifying the DRG case-mix measure using complication and comorbid diagnoses to roughly triple the number of DRGs significantly reduced residual variance in cost per discharge in historical samples of typical Medicare discharges. There were at least two key policy conclusions from this work: the first was that "PPS payment rates based on the refined DRGs would reflect more accurately providers' production costs than those currently in use" (MedPac 2000, p. 72); the second was that the more refined case-mix measure could "substantially diminish the role of the Indirect Medical Payment adjustment [a proxy for case mix] and improve payment equity among hospitals" (Medpac 2000, p. 76). MedPac subsequently recommended that the Center for Medicare and Medicaid Services (CMS) adopt a system that allowed estimated complexity to vary within base DRGs, footnoting that "although the possibility that some hospitals may overstate their patients' severity cannot be ruled out, it seems unlikely that this would be a major source of error in the payment estimates" (2000, p. 77). As reported by Baker (2002), CMS ultimately did not accept the MedPac recommendation in part because there was "no ability to predict with accuracy how such a change might affect coding behavior among hospitals" (2002, p. 13); although Baker suggests that the decision will likely be revisited.

Much of the concern that hospitals will alter their coding practices in response to changes in the patient classification system stems from experience following the introduction of the PPS in the United States (for example, Hsiah et al. 1988; Steinwald and Dummit 1989; Ginsburg and Carter 1986). This study is meant to inform the current debate by describing the Canadian experience of incorporating modifications to a case-mix measurement system similar in structure to those under consideration in the United States and other jurisdictions. In doing so, it describes hospital coding responses as well as the

resultant impact on adjustment factors and predictive accuracy in a hospital payment model.

THE POLICY CHANGE

In Ontario, hospital services are provided to approximately 13 million residents by almost exclusively nonprofit private hospital corporations. Roughly 85 percent of overall hospital operating expenditures are funded by the Ministry of Health and Long Term Care. Following the establishment of global budgets in 1969, hospital base adjustments had been determined by inflation, program changes, and demographic change. Since 1988, additional adjustments have been made based on hospitals' relative cost per adjusted discharge. Using a model similar to that used by CMS under the PPS, hospital expected cost per discharge is calculated adjusting for factors such as case mix, teaching intensity, geography, and size. Hospitals for which expected costs exceed actual costs are targeted for "equity-based" funding increases in an amount proportional to the gap between actual and expected. In addition to direct funding, the model is extensively used for forecasting, planning, and financial analysis, and is reported in the financial quadrant of the annual Ontario hospital report card.

The original case-mix methodology used in the acute payment model was Case Mix Groups (CMG),¹ a Canadian adaptation of DRGs. The methodology uses ICD-9 diagnoses and Canadian Classification of Procedure (CCP) codes to assign each patient to one of roughly 600 resource groups within 25 Major Clinical Categories (MCC) based on the Most Responsible Diagnosis determined at discharge. Like Medicare DRG, the CMG methodology may use the presence or absence of secondary diagnoses and patient age when determining assignment to a resource group. As experienced by other jurisdictions, a perceived shortcoming of the hospital funding methodology had been the existence of intra-CMG variation in resource use that could potentially be explained by more complete use of diagnostic information on patient charts. In the beginning of fiscal year 1997/1998, following several years of development work led by the Canadian Institute for Health Information (CIHI), the province adopted a fundamental change to the CMG methodology meant to achieve greater precision in cost estimates through extensive use of secondary diagnostic information. This "complexity" methodology incorporates diagnosis typing information to identify chronic conditions, complications, and multisystem failure concomitant with the acute

care episode (Benoit, Skea, and Mitchell 2000). The hospital assigns a type to each coded diagnosis. Type M indicates the most responsible diagnosis; types 1 and 2 indicate comorbidities and complications requiring increased length of stay or resources; type 3 indicates additional diagnoses not affecting the patient's course of treatment. All diagnoses except type 3 diagnoses can potentially combine to affect the case weight assigned to each abstract. With the exception of day surgery and cases within several MCCs (those relating mainly to obstetrics, neonates, mental disease and disorders, and HIV infections), cases within a base CMG are potentially assigned to one of four complexity levels and one of three age groups. Instead of making exclusive use of diagnoses and procedures for classification, the CMG complexity methodology additionally uses age, which "tends to limit the impact of subjectivity in coding practice" (Canadian Institute for Health Information 2000b). Age and complexity levels within CMGs are not applied where they do not improve homogeneity in terms of length of stay. Hence there are up to 12 possible subgroups within applicable CMGs, resulting in roughly 3,400 resource groups overall. Expected resources generally increase with age and complexity level. Complexity level 1 indicates no complexity; level 2 indicates complexity related to chronic conditions; level 3 is complexity related to "serious conditions"; and level 4 relates to "potentially life threatening conditions." For example, in 1997 under the CMG methodology, a typical patient older than age 70 assigned to CMG 146, Simple Pneumonia and Pleurisy with Complications, had an expected cost of \$4,234. Under the CMG with complexity methodology, that patient's expected cost could range from \$3,016 to \$11,861, depending on how the secondary diagnoses combined to determine the patient's complexity level. This case-mix methodology was used by the province for hospital funding purposes for five years until 2002.

METHODS

There are three main aspects to the analytical approach used here to assess the impact of the policy change. The first analysis is meant to address whether there was a discernible coding response in the years following the introduction of the complexity methodology. The primary variable of interest is complexity adjusted case mix, controlling for changes in demographics, practice patterns, patient acuity, and CMG mix. Previous studies (Steinwald and Dummit 1989; Ginsburg and Carter 1986) have shown that coding responses to changes in

the patient classification system can vary according to hospital type. In addition, there is a common perception that unmeasured complexity in DRG-type systems is more significant for hospitals that act as referral centers, such as those with a teaching mission (MedPac 2000). Hence, trends are reported separately for major teaching hospitals (identified as members of the Ontario Council of Teaching Hospitals). In order to account for shifts in population age structure over time, trends were examined standardizing to the 1996/1997 population using the direct method.

Variation in case mix over time can also result from changes in CMG mix and the incidence of atypical cases, namely: deaths, transfers, long-stay outlier cases, and patients that are voluntarily discharged against medical advice. To account for differences in CMG mix, trends are disaggregated and reported for high-volume CMGs. The impact of atypical cases is controlled for by separately examining trends for typical cases. The time trends of measured acuity and actual resource consumption and morbidity are also of interest. Without a longitudinal coding response, changes in average inpatient case weight would be correlated with changes in observable indicators such as average length of stay; an increased incidence of cases categorized as "life threatening" would be associated with an increase in mortality rates. Overall, the trend analysis is meant to establish whether hospitals altered their reported incidence of secondary diagnoses given that the complexity methodology provided incentives to do so.

The remaining econometric analyses assess the impact of the policy change on a hospital payment model. The first econometric analysis assesses whether the introduction of the complexity instrument and subsequent changes in hospital coding practices resulted in a reduction in the weight of the teaching factor in the cost regression. The regression model specified is similar to those used for developing the original PPS indirect medical cost adjustments (Pettengill and Vertrees 1982; Dalton, Norton, and Kilpatrick 2001),

$$\begin{aligned} \ln(\text{oper cost/case})_{it} = & \alpha + \beta_1 \ln(1 + \text{IRB})_{it} + \beta_2 \ln(\text{case-mix index})_{it} \\ & + \beta_3 \ln(\text{patient days})_{it} \\ & + \beta_4 \ln(\text{distance})_{it} + \sum \delta_t \text{year}_t \\ & + \sum \delta_t^* \text{year}_t \times \ln(1 + \text{IRB})_{it} + v_{it} \end{aligned}$$

The literature originally regarded the Intern and Resident to Bed (IRB) ratio as an indicator of the indirect costs of medical education. There is an emerging consensus that the teaching variable also serves to account for patient severity and complexity not accounted for by the case-mix variable. Therefore,

improvements in coding accuracy and comprehensiveness would be expected to reduce the importance of the teaching variable as a determinant of costs. In other words, the coefficient on IRB would be expected to decline as the case-mix measure accounted for an increased portion of the variation in hospital costs. In order to identify the magnitude of this effect, dummy variables for each year are interacted with the IRB ratio to allow the estimated teaching coefficient to vary over time. Examining δ_t^* , the coefficients on the interacted year and teaching variables, indicates the extent to which a more comprehensive case-mix measure reduced the importance of the teaching factor in the payment model.² Patient days are included as an independent variable to incorporate potential hospital scale effects. In Ontario, geographic dispersion has been found to be an important determinant of hospitals costs (Joint Policy and Planning Committee 2001), and is measured using distance to nearest facility. Factor price indices are not used for hospital payment purposes since variation is accepted to be unsystematic (Preyra 1998; Joint Policy and Planning Committee 2001); for example, most healthcare workers belong to unions that negotiate wages through provincial collective bargaining. Year-specific dummy variables are used to allow the regression intercept to vary, controlling for cost variation related to time, such as inflation.

After examining any shifts in the relative importance of the teaching and case-mix variables in the model, the final econometric analysis assesses whether, overall, more extensive use of secondary diagnoses for case-mix adjustment substantively improved payment accuracy. Testing the hypothesis that the measured complexity improved predictive power involves assessing the incremental contribution of the secondary diagnoses to the overall explanatory power of the cost regression. For any hospital, the case-mix index incorporating complexity can be decomposed into two parts, $CMI_{PLX} = CMI_{BASE} \times \Delta$. CMI_{BASE} represents the portion of a hospital's case-mix index that would arise without consideration of multiple diagnoses (i.e., is based on age, principal procedures, and most responsible diagnoses only) Δ represents the portion of a hospital's measured case mix that is due only to the reported secondary diagnoses. The complexity adjustment allows hospitals with higher values of Δ to have higher estimated resources per case conditional on CMG mix. To operationalize the calculation, the method of Preyra and Sandor (2002) is used to recalibrate all hospital case-mix indices ignoring all secondary diagnostic chart information.³ Applying this methodology and forming the ratio of each hospital's complexity adjusted case-mix index to their noncomplexity-adjusted index provides a hospital-specific estimate of Δ .

Therefore, in a given year t , the cost model may be written as

$$\begin{aligned}\ln(\text{oper cost/case})_{it} = & \alpha_t + \beta_{1t} \ln(1 + \text{IRB})_{it} + \beta_{2t} \ln(\text{CMI}_{\text{BASE}})_{it} \\ & + \beta_{3t} \ln \Delta_{it} + \beta_{4t} \ln(\text{patient days})_{it} \\ & + \beta_{5t} \ln(\text{distance})_{it} + \varepsilon_{it}\end{aligned}$$

The incremental contribution of the complexity variable to the overall predictive power of the regression is the concept of interest. In theory, the coefficients β_2 and β_3 should be identical. The above specification allows the marginal contribution of the complexity and noncomplexity components of case mix to differ. Even if the true coefficients were equal, the estimated coefficients could still differ substantially. For example, if there was significant variation in hospital coding of complications and comorbidities not related to resource use, then the complexity component would be measured with error. Attenuation bias could drive the estimate of β_3 to zero if the measurement error “noise” dominated the complexity “signal.”

The models were estimated by case weighted ordinary least squares with Huber-White robust standard errors with clustering by hospital. All models underwent examination of residuals to ensure the regression assumptions were satisfied and robustness of conclusions was verified by comparing the effects of the regression with and without statistical outliers and high-leverage observations.

DATA

Financial data at the hospital level and clinical data at the patient level covering years “1996/1997 to 2001/2002” were obtained from the Ministry of Health and Long Term Care for all Ontario acute care hospitals. Hospital-specific case-mix indices were calculated by forming the ratio of total case weights to total discharges. Since CIHI annually reestimates the case weighting methodology as new case costed data become available, longitudinal case-mix comparisons could confound methodological changes with coding changes. Therefore, annual improvements in the underlying costing methods were controlled for using the 2002 “regrouped” version of the complexity grouper in all years of the study. The ministry annually collects data on student days by facility. Student days include third and fourth year undergraduates, interns and residents involved in the provision of medical care. For the purposes of the hospital funding formula, teaching intensity is defined as the ratio of medical student days to occupied bed day days (IRB). The student days of three hospitals were known to be reported incorrectly in

Table 1: Descriptive Statistics for 6-Year Study Sample (Total observations = 786)

<i>Weighted Mean</i>	<i>Year</i>					
	<i>1996/1997</i>	<i>1997/1998</i>	<i>1998/1999</i>	<i>1999/2000</i>	<i>2000/2001</i>	<i>2001/2002</i>
Cost per discharge (\$)	2,311	2,282	2,360	2,412	2,517	2,765
Census days	163,730	161,593	159,269	156,605	159,090	155,055
Case mix index	0.82	0.81	0.82	0.82	0.82	0.86
Teaching intensity (IRB)	0.116	0.121	0.112	0.113	0.108	0.113
Distance (km)	22	20	21	21	21	21
<i>Total discharges</i>	1,939,864	1,899,876	1,898,925	1,991,542	2,035,212	2,014,541
<i>Total number of hospitals</i>	134	129	128	131	131	133

“1996/1997 or 1997/1998” and subsequently corrected. For any erroneous year, we used the following year of data as a proxy, resulting in stable teaching factors for these hospitals across all six years of the sample.

The calculation of hospital acute operating cost per case is based on financial data that are reported under uniform guidelines. These data are used for hospital funding; they undergo extensive edit checks and are generally felt to be of good quality. As part of the annual recalibration of the hospital funding formula, an expert working group familiar with hospital reporting practices reviews the financial data and identifies data-quality issues that are subsequently addressed with the relevant hospitals. Following a reconciliation process, any hospitals where acute care costs are known to be unreliably measured are excluded from the calibration of the cost model. These six hospitals along with the two specialized children’s hospitals were excluded from the analysis.

Our final longitudinal dataset consists of detailed clinical, financial, and supplementary information for 134 acute care hospitals over six years. These 786 hospital-level observations include patient-level clinical data aggregated from roughly 12 million discharge abstracts. Table 1 provides descriptive statistics for the analytical sample.

RESULTS

Trends in Case Mix

Table 2 displays key trends in the components of hospital case mix to which the complexity methodology is applied, beginning in year 1996/1997 and

Table 2: Growth Rates Since 1997 for Cases with Complexity Adjustments

	Academic Health Sciences Centers						All Hospitals					
	CMI 1997	1998	1999	2000	2001	2002	CMI 1997	1998	1999	2000	2001	2002
A. Average CMI												
Typical	1.33	7%	12%	19%	27%	35%	1.07	5%	9%	14%	18%	25%
Total	1.83	4%	8%	13%	19%	25%	1.45	2%	5%	9%	12%	17%
B. Total Inpatient Cases												
Typical	203,740	-4%	-7%	-8%	-9%	-11%	729,168	-3%	-5%	-5%	-6%	-6%
Total	247,645	-4%	-6%	-8%	-10%	-12%	871,958	-3%	-5%	-5%	-6%	-6%
C. Average Length of Stay												
Typical	5.7	1%	2%	4%	6%	7%	5.3	0%	2%	4%	5%	6%
Total	8.0	-1%	0%	1%	3%	2%	7.3	-2%	-2%	-1%	0%	0%
E. Level 1 Inpatient Cases												
Typical	156,716	-9%	-14%	-20%	-27%	-34%	572,335	-6%	-11%	-14%	-18%	-23%
Total	178,990	-9%	-14%	-21%	-29%	-37%	656,170	-7%	-11%	15%	-19%	-24%
F. Level 2 Inpatient Cases												
Typical	20,872	10%	17%	25%	28%	37%	65,750	10%	19%	28%	34%	48%
Total	28,278	7%	12%	17%	18%	25%	87,370	7%	14%	20%	26%	37%
G. Level 3 Inpatient Cases												
Typical	10,162	21%	34%	59%	86%	111%	26,985	20%	37%	58%	79%	112%
Total	15,126	16%	26%	44%	61%	81%	40,906	14%	27%	43%	57%	82%
H. Level 4 Inpatient Cases												
Typical	6,059	26%	50%	97%	158%	211%	12,609	27%	56%	103%	148%	211%
Total	12,562	20%	35%	69%	103%	139%	26,140	20%	40%	74%	100%	144%
I. Number of Diagnoses per Case												
Comorbidities	1.1	21%	31%	49%	70%	102%	1.0	14%	24%	39%	54%	74%
Complications	0.2	26%	46%	91%	146%	204%	0.1	18%	40%	76%	110%	157%
J. Deaths												
Deaths per 10,000 Cases	427	5%	8%	10%	7%	5%	440	3%	5%	7%	5%	5%

relative to that year through to year 2001/2002. The number of diagnoses per case increased by 24 percent over the sample period, with the annual trend showing no evidence of declining. Secondary diagnoses classified as materially impacting course of treatment, increased steadily in the years following the introduction of the complexity methodology. By 2001/2002, the provincial increase in comorbidities and complication was 74 percent and 157 percent respectively; substantially more in major teaching hospitals.

Over the six years for the study sample, controlling for age, the average inpatient case weight increased by 17 percent for the province as a whole, with the growth in case mix differing substantially by year, by case type, and by facility type. For typical cases, which comprise roughly 85 percent of total inpatient medical and surgical cases, case mix increased by 35 percent in teaching hospitals and 25 percent overall. This phenomenon of differential increase in case mix according to hospital type is similar to that experienced in the United States following the introduction of the PPS (Steinwald and Dummit 1989). The growth rate in average case mix was not driven by a small number of hospitals, although there was considerable variation in case-mix growth at the hospital level, as shown in Figure 1.

In all years, case mix is clustered according to Ministry defined hospital peer groups, although for the community and major teaching hospitals, there was greater dispersion of case mix in 2001/2002 than in 1996/1997.⁴ This increase in variance is striking: for example, among the major teaching

Figure 1: Growth in Case-Mix Index for Typical Inpatients

hospitals, the hospital with the highest inpatient case mix in 2001/2002 experienced a 50 percent increase in case mix over the sample period, while the hospital with the lowest inpatient case mix experienced no change. Detailed analysis of these hospitals' program mix provided no evidence that the CMG mix had changed substantially over time, so changes in measured complexity within CMG drove much of the difference.

A key difference between the CMG methodology with and without complexity was the introduction of complexity grades that take into account patient comorbid conditions that "strongly influence the length of stay" (Canadian Institute for Health Information 2000). Without coding effects, small increases in average length of stay would then suggest concomitant decreases in the share of lower-complexity cases relative to higher complexity cases. As Table 2 shows, the direction of the complexity increase was in the expected direction but of a much larger order of magnitude. Overall, Level 1 cases declined by 24 percent, with Level 2, 3, and 4 cases increasing by 37 percent, 82 percent, and 144 percent respectively. Again, the trends are markedly different according to case type: for example, in Level 4, there was an overall increase of 211 percent in typical cases assigned to the highest complexity level compared to 144 percent for atypical and typical cases combined.

Total inpatient cases fell over the sample period so, for example, the share of total cases that were assigned the highest complexity level increased by 160 percent. Without coding effects, this growth in high complexity cases, termed "life threatening conditions" should have been associated with significantly longer lengths of stay as well as increased morbidity, in particular, deaths in hospitals. However, despite the large increases in these highest-complexity cases, the age standardized mortality rate in hospitals increased by only 5 percent overall and the average length of stay for hospital inpatients was unchanged.

The growth in case mix was not uniform across peer groups and case type. Table 3 provides a more disaggregated view of the growth in typical average case weight for the twenty highest volume medical inpatient CMGs. Overall, the average case weight assigned to the highest-volume CMG, simple pneumonia and pleurisy, experienced an increase of 27 percent, with a 48 percent growth rate in teaching hospitals. These growth rates are even more striking considering that there was only a 1.5 percent increase in average length of stay in this CMG over the six-year period in teaching hospitals as well as in the province overall.

Other case types that have also been the subject of scrutiny by auditors in many jurisdictions, such as heart failure and chronic bronchitis, experienced

Table 3: Case-Mix Trends for Highest Volume Inpatient Medical/Surgical CMGs

	Academic Health Sciences Centers						All Hospitals					
	CMH1997	1998	1999	2000	2001	2002	CMH1997	1998	1999	2000	2001	2002
<i>Case-Mix Group</i>												
143 SIMPLE PNEUMONIA AND PLEURISY	1.18	7%	12%	19%	30%	48%	1.04	5%	8%	15%	19%	27%
253 MAJOR INTESTINAL AND RECTAL PROCEDURES	2.37	1%	5%	9%	15%	20%	2.35	1%	3%	7%	9%	14%
222 HEART FAILURE	1.18	8%	15%	22%	36%	46%	1.11	5%	9%	14%	20%	28%
579 MAJOR UTERINE AND ADNEXAL PROCEDURES WITHOUT MALIGNANCY	0.92	0%	1%	1%	3%	5%	0.90	0%	1%	1%	2%	3%
354 KNEE REPLACEMENT	2.00	0%	0%	2%	4%	5%	1.99	0%	0%	1%	2%	3%
294 ESOPHAGITIS, GASTROENTERITIS, AND MISCELLANEOUS DIGESTIVE DISEASE	0.58	3%	8%	11%	21%	29%	0.54	2%	4%	8%	9%	14%
013 SPECIFIC CEREBROVASCULAR DISORDERS EXCEPT TRANSIENT ISCHEMIC ATTACKS	1.58	3%	9%	17%	23%	29%	1.51	3%	7%	12%	16%	22%
352 HIP REPLACEMENT	2.07	1%	1%	3%	5%	7%	2.07	0%	1%	2%	3%	4%
179 CORONARY BYPASS WITH HEART PUMP WITHOUT CARDIAC CATH	3.05	3%	6%	14%	20%	30%	3.03	3%	5%	13%	19%	29%
251 GASTROSTOMY AND COLOSTOMY PROCEDURES	3.78	3%	4%	11%	14%	20%	3.77	4%	8%	11%	14%	20%
662 FEMUR OR PELVIC PROCEDURES FOR TRAUMA	2.21	3%	6%	10%	16%	19%	2.18	2%	4%	8%	10%	14%
189 PERCUTANEOUS TRANSLUMINAL CORONARY ANGIOPLASTY W/O CARDIAC CC	1.72	2%	3%	7%	8%	15%	1.72	1%	2%	6%	7%	15%
208 AMI WITHOUT CARDIAC CATH WITHOUT SPECIFIED CARDIAC CONDITIONS	1.45	0%	2%	4%	8%	12%	1.41	1%	2%	3%	4%	6%
237 ARRHYTHMIA	0.67	13%	19%	18%	31%	38%	0.64	8%	11%	13%	18%	25%
001 CRANIOTOMY PROCEDURES	2.65	9%	14%	22%	30%	38%	2.72	7%	11%	18%	24%	34%
205 AMI WITHOUT CARDIAC CATH WITH CONGESTIVE HEART FAILURE	2.15	5%	7%	14%	17%	22%	2.09	3%	5%	8%	11%	15%
140 CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD)	1.10	5%	10%	16%	16%	15%	1.05	4%	7%	10%	11%	14%
186 PERMANENT PACEMAKER IMPLANT WITHOUT SPECIFIED CARDIAC CONDITIONS	2.85	3%	6%	8%	9%	13%	2.89	2%	3%	5%	6%	8%
040 TRACHEOSTOMY AND GASTROSTOMY PROCEDURES	13.90	- 1%	- 2%	14%	15%	14%	13.20	- 2%	- 3%	10%	11%	11%
142 CHRONIC BRONCHITIS	1.00	10%	18%	23%	30%	35%	0.95	7%	14%	17%	21%	25%

large increases over the period; the average CMI increase in teaching hospitals for these two CMGs was 46 percent and 35 percent respectively. The average length of stay of these same cases increased by 7 percent for heart failure and actually declined by 3 percent for chronic bronchitis cases. Case-mix groups related to cardiac care also experienced large increases in measured acuity. In addition, CMI growth seems to have increased faster in medical than surgical cases, although this is not always the case.

Because these results are age standardized, all of the increase in case mix within CMGs comes from shifts in the distribution of measured complexity within CMGs. We have observed a consistent shift within most CMGs toward the highest complexity levels. For example, in 1996/1997, 2 percent of provincial typical simple pneumonia and pleurisy cases were assigned to complexity level four. By 2001/2002, there had been a 400 percent increase in the share of these cases provincially, 500 percent in teaching hospitals. Since complexity effects are a function of hospital type, mix of CMGs and the nature of hospital coding practices, hospital-level case mix varies longitudinally, by year and hospital. Consistent with above results, the coefficient of variation of Δ , the pure complexity component, increased steadily from 208 in the first year to 660 in the final year of the sample.

Econometric Analysis

Table 4 shows the regression results for the cost model using the six years of data.⁵ Overall, the patient days variable is a statistically insignificant determinant of unit costs in Ontario. As expected, isolated hospitals tend to have higher unit costs than similar nonisolated hospitals. This variable likely combines scale and scope effects, as small hospitals in Ontario tend to be isolated, and isolated hospitals cannot generally specialize because they are often sole community providers. The case-mix variable is a fundamental determinant of hospital costs.

Interacting the fixed-year effects with the hospital-specific teaching intensity factors allows for a test of the hypothesis that the teaching coefficient in each year varies relative to the base year. Teaching intensity in all years is a significant determinant of unit operating costs but there is no evidence that the observed changes in hospital coding practices resulted in reductions in the teaching coefficient in the five years following the introduction of the complexity methodology.

Table 5 contains the necessary information to determine if the recognition of reported secondary diagnoses in the classification system

Table 4: Regression Results using Weighted Least Squares with Huber-White Correction

<i>Dependent Variable</i>	<i>log (oper cost per case)</i>
Intercept	7.8
<i>(reference year: 1996/1997)</i>	
YR 1997/1998	− 0.012 (.008)*
YR 1998/1999	0.008 (.011)
YR 1999/2000	0.032 (.013)*
YR 2000/2001	0.087 (.014)***
YR 2001/2002	0.133 (.012)***
<i>Teaching Intensity</i>	
ln (teaching)	0.794 (.10)***
ln (1+IRB) × YR 1997/1998	0.006 (.078)
ln (1+IRB) × YR 1998/1999	0.045 (.084)
ln (1+IRB) × YR 1999/2000	0.100 (.08)
ln (1+IRB) × YR 2000/2001	0.001 (.111)
ln (1+IRB) × YR 2001/2002	− 0.004 (.11)
distance	0.004 (.001)**
log (days)	0.004 (.004)
log (case-mix index)	1.043 (.025)***
R ²	0.934

Robust standard errors are in parentheses:

- * $p < .05$;
- ** $p < .01$;
- *** $p < .0001$.

Table 5: Cross Sectional Regression Results Testing for Complexity Impact

<i>Year</i>	<i>Base Component of CMI</i>		<i>Complexity Component of CMI</i>	
	<i>Coefficient</i>	<i>Standard Error</i>	<i>Coefficient</i>	<i>Standard Error</i>
1996/1997	1.13	(0.07)***	0.001	(0.104)
1997/1998	1.17	(0.06)***	0.51	(0.53)
1998/1999	1.15	(0.07)***	− 0.06	(0.34)
1999/2000	1.15	(0.07)***	0.06	(0.3)
2000/2001	1.13	(0.08)***	− 0.16	(0.3)
2001/2002	1.05	(0.08)***	0.197	(0.195)

Robust standard errors are in parentheses:

- * $p < .05$;
- ** $p < .01$;
- *** $p < .0001$.

improved the accuracy of hospital payments. Overall, there is a strong relationship between the base CMI and hospital unit cost variation. However, in no year is there statistical evidence that the complexity component of

hospitals' case mix accounted for a significant amount of variation in hospital unit costs.

DISCUSSION

There are three main results from the preceding empirical analyses. The first is that, since the introduction of the complexity methodology, the mean and variance of hospital case mix has increased significantly without concomitant changes in morbidity or resource use. The second result is that the substantial changes in hospital measured case mix did not significantly alter the teaching coefficient in a standard payment regression. The third result is that the changes in coding practices did not translate to increased accuracy in the estimation of hospital specific costs.

As in other jurisdictions, stakeholders in Ontario anticipated that hospitals with high incidence of complications and comorbidities would become better differentiated through use of the complexity methodology. Since hospitals acting as referral centers, such as academic health sciences centers, would tend to have a higher incidence of complexity than other hospitals, the average case-mix indices for these hospitals was expected to rise disproportionately following the introduction of the complexity adjustment.

In a key contribution, Dalton, Norton, and Kilpatrick (2001) showed that, although cost regressions typically identify a statistical relationship between teaching intensity and hospital costs, there is little evidence for a causal link. Rather, their work supports the recent conclusions of MedPac that teaching intensity is a major proxy for omitted case-mix effects. Based on this reasoning, more comprehensive reporting and recognition of secondary diagnoses would tend to reduce the weight of the teaching-adjustment factor in hospital payment models. However, the Ontario experience suggests that this theoretical prediction may not always obtain in practice. One reason for this finding could be that the teaching factor does indeed relate mainly to indirect costs of teaching that are uncorrelated with case mix. Another reason, that is consistent with the Dalton, Norton, and Kilpatrick interpretation is that, overall, the complexity methodology introduced more noise than signal in the measurement of patient acuity. The data cannot reject this hypothesis, since the complexity component of case mix did little to reduce unexplained variation in hospital cost per case—the very purpose for which it was introduced.

What would drive an increase in the variation of case mix without any associated increase in predictive power? One possibility is that hospitals

systematically engaged in upcoding, although without the “gold standard” of chart abstraction, this cannot be concluded. Still, the experience of other jurisdictions using case-based payment systems suggests that the incidence of upcoding is likely nontrivial. More generally, however, increased case-mix variation could be the result of systematic or nonsystematic differences in diagnostic coding. Sources of error include missing and misinterpreting information on the chart and misinterpreting coding guidelines. Variation can arise during documentation of the medical record, because of errors or inherent subjectivity in diagnoses; during the process of translating the physician narrative to discrete classification groups; from differences in physician charting practices; or depending on the use of manual versus electronic charts. General concerns about the use of secondary diagnoses for provider comparisons continue to be raised in the literature (Iezzoni 1997; Geraci 2002; Romano et al. 2002). Although there have been few recent studies of the quality of Ontario’s administrative data, the Institute for Clinical Evaluative Sciences (Williams and Young 1996) reviewed studies of the quality of diagnostic information in Ontario hospital administrative data and concluded that “clinical data on secondary diagnosis, comorbidities, and complications are less likely to be recorded accurately and comprehensively in hospital discharge abstracts, and the rates of agreement on case-mix may be accordingly low” (1996, p. 343).

The ability to draw conclusions and to extrapolate these results to other regions is limited. This study has been undertaken with administrative data, and has not used the method of chart abstraction to assess the quality of the reported diagnostic information. The population of Ontario nonprofit hospitals may not generalize to other jurisdictions. If nonprofit hospitals do tend to respond less aggressively than proprietary hospitals to incentives to code diagnoses, then it would be expected that regions where there is a different ownership mix would experience different case-mix effects. Canada does not have formal data quality monitoring mechanisms, such as Quality Improvement Organizations in the United States, nor are Canadian physicians required to attest to the accuracy of diagnoses, although Hsieh and Ahern (1992) demonstrates that these are important determinants of coding variation. This study has examined the impact of a particular grouping methodology. An inpatient grouper that incorporates secondary diagnostic information in a substantially different manner than the complexity methodology considered here would have produced different results. Even if a complexity methodology does not increase accuracy in estimating costs at the hospital level, it still does not follow that the instrument is not preferred. Within organizations, complexity adjustment could still be useful for patient

management or, if risk selection is a major concern, this method could be justified despite modest statistical performance.

The complexity adjustment adopted by Ontario is likely the most substantial revision to a DRG type system considered in the literature thus far. Despite this study's limitations, the results have implications for policy. In Ontario, the variation in measured case mix, comorbidities, and complications has resulted in considerable discussion among stakeholders. In year 2002, following consideration of this paper's results along with supporting analyses, the Ontario Ministry of Health and Long Term Care suspended the use of the complexity methodology in the hospital funding model. Other jurisdictions may derive several lessons from this experience. First, predictions regarding the effects of changes to patient classification systems should consider the endogeneity of hospital coding practices. That is, coding should be recognized as a function of reimbursement policy. Second, the usefulness of secondary diagnoses to improve payment accuracy is limited by the accuracy and homogeneity of reporting across hospitals. Therefore, data quality should be carefully assessed before and following any policy change. Third, coding variation can result for a variety of reasons, and measures should be undertaken to ensure coding standards are clear and enforced. In Ontario, anecdotally, a great deal of coding variation occurred because of varying interpretations of coding guidelines on the part of hospitals. Fourth, changes to the classification system invariably alter the sum and distribution of case-based payments. For example, the value of the average case-weight increase of 17 percent since the introduction of complexity was roughly \$690 million CAD 2002, or roughly 12 percent of acute expenditures. Because the growth in case mix varied considerably by peer group and by hospital, the distribution of expected costs was significantly altered over the period.

The coding responses examined in this paper occurred in the presence of a funding model that rewarded hospitals *at the margin* only. Since Ontario is following a gradual approach to implementation of its prospective payment funding model, the potential distortions were mitigated. Importantly, even though the financial implications for hospitals of changing their coding practices were relatively small, there was a significant change in behavior. Payment organizations, such as the CMS, considering the full adoption of classification systems that rely heavily on secondary diagnoses might expect an even greater coding response than that experienced in Ontario. Other jurisdictions might also consider gradual approaches to implementation, using local pilot studies and incorporating international experience when considering major adjustments to case-mix measurement systems.

ACKNOWLEDGMENTS

I would like to thank George Pink, Paul Sandor, the *HSR* referees, and the JPPC Technical Working Group for their helpful comments.

NOTES

1. Case-mix groups (CMG), relative intensity weights (RIW) are referred to throughout this document and are registered trademarks of the Canadian Institute for Health Information (CIHI). Case-mix groups were developed by CIHI; see <http://www.cihi.ca> for more information.
2. Changes in the teaching coefficient over time could be due to changes in teaching intensity as well as case-mix effects, so a version of the model is estimated with the teaching variable fixed at 2001/2002 levels for each hospital (Dalton, Norton, and Kilpatrick 2001).
3. Following review by an expert panel of health services researchers and hospital stakeholders, this methodology was recommended to the Ministry of Health and Long Term Care and accepted as an appropriate case-mix methodology to be used in the absence of reliable information relating to secondary diagnoses.
4. Small hospitals have generally fewer than 3,500 case-mix-adjusted annual discharges, they have a referral population of less than 20,000 people, and tend to be sole community providers. Community hospitals are short-stay acute facilities that are not members of the Council of Teaching Hospitals and are not designated as small (Joint Policy and Planning Committee 2003).
5. The models were estimated allowed teaching intensity to vary over time but the results were similar and excluded for brevity.

REFERENCES

- Averill, R. F., T. E. McGuire, B. E. Manning, D. A. Fowler, S. D. Horn, P. S. Dickson, M. J. Coye, D. L. Knowlton, and J. A. Bender. 1992. "A Study of the Relationship between Severity of Illness and Hospital Cost in New Jersey Hospitals." *Health Services Research* 27 (5): 587–617.
- Baker, J. 2002. "Medicare Payment System for Hospital Inpatients: Diagnosis-Related Groups." *Journal of Healthcare Finance* 28 (3): 1–13.
- Benoit, D., W. Skea, and S. Mitchell. 2000. "Developing Cost Weights with Limited Cost Data-Experiences Using Canadian Case Cost Data." *Case Mix* 2 (3): 88–96.
- Canadian Institute for Health Information. 2000a. *DAD Resource Intensity Weights 2000*. Ottawa, ON: Canadian Institute for Health Information.
- . 2000b. *DAD Expected Length of Stay 2000*. Ottawa, ON: Canadian Institute for Health Information.
- Dalton, K., E. C. Norton, and K. Kilpatrick. 2001. "A Longitudinal Study of the Effects of Graduate Medical Education on Hospital Operating Costs." *Health Services Research* 35 (6): 1267–91.

- Geraci, J. 2002. "The Demise of Comparative Provider Complication Rates Derived from ICD-9_CM Diagnoses." *Medical Care* 40 (10): 847-50.
- Ginsburg, P. B., and G. M. Carter. 1986. "Medicare Case-Mix Index Increase." *Health Care Financing Review* 7 (4): 5-65.
- Horn, S. D., G. Bulkey, P. D. Sharkey, A. F. Chambers, R. A. Horn, and C. Schramm. 1985. "Interhospital Differences in Severity of Illness: Problems for Prospective Payment Based on Diagnosis-Related Groups." *New England Journal of Medicine* 313 (1): 20-4.
- Hsia, D., W. M. Krushat, A. B. Fagan, J. A. Tebutt, and R. P. Kusserow. 1988. "Accuracy of Diagnostic Coding for Medicare Payments Under the Prospective Payment System." *New England Journal of Medicine* 11 (6): 352-5.
- Hsia, D. C., and C. Ahern. 1992. "Medicare Reimbursement Accuracy Under the Prospective Payment System, 1985 to 1988." *American Medical Association Journal* 268 (7): 873-96.
- Iezzoni, L. I. 1997. "Assessing Quality Using Administrative Data." *Annals of Internal Medicine* 15 (127, part 2): 666-74.
- Joint Policy and Planning Committee. 2001. "Integrated Population Based Allocation" [accessed on January 15, 2003]. Available at <http://www.jpcc.org>.
- Lynk, J. 2001. "One DRG, One Price? The Effect of Patient Condition on Price Variation within DRGs and across Hospitals." *International Journal of Health Care Finance and Economics* 1 (2): 111-37.
- MedPac (Medicare Payment Advisory Committee). 2000. *Report to the Congress: Medicare Payment Policy*. Washington, DC: Medicare Payment Advisory Committee.
- Pettengill, J., and J. Vertrees. 1982. "Reliability and Validity in Hospital Case-Mix Measurement." *Health Care Financing Review* 4 (2): 101-27.
- Preyra, C. 1998. "The Econometric Analysis of Hospital Costs with an Application to Scale and Scope Efficiencies in Ontario Hospitals" [dissertation]. Toronto, ON: University of Toronto.
- Preyra, C., and P. Sandor. 2002. *A Prospective Method of Complexity Adjustment to Case Mix Groups*. Ontario: Finance and Information Branch, Ontario Ministry of Health and Long Term Care.
- Romano, P., B. K. Chan, M. Schembri, and J. A. Rainwater. 2002. "Can Administrative Data Be Used to Compare Postoperative Complication Rates Across Hospitals." *Medical Care* 40 (10): 856-67.
- Schwartz, M., L. I. Iezzoni, M. A. Moskowitz, A. Ash, and E. Sawitz. 1996. "The Importance of Comorbidities in Explaining Differences in Patient Costs." *Medical Care* 34 (8): 767-82.
- Steinwald, B., and L. A. Dummit. 1989. "Hospital Case-Mix Change: Sicker Patients or DRG Creep?" *Health Affairs* 8 (2): 35-47.
- Williams, J. I., and W. Young. 1996. "A Summary of Studies in the Quality of Health Care Administrative Databases in Canada." In *Patterns of Health Care in Ontario: The ICES practice Atlas*, edited by V. Goel, J. I. Williams, G. M. Anderson, P. Blackstien-Hirsh, C. Fooks, and C. D. Naylor, pp 339-46. Ottawa, ON: Canadian Medical Association.