Author's response to reviews

Title: ICD-10 Coding Algorithms for Defining Comorbidities of Acute Myocardial Infarction

Authors:

Lawrence So (lso@ucalgary.ca)
Dewey Evans (devans@phsa.ca)
Hude Quan (hquan@ucalgary.ca)

Version: 2 **Date:** 27 November 2006

Author's response to reviews: see over

CHAPS
CENTRE FOR HEALTH AND POLICY STUDIES

FACULTY OF MEDICINE

Centre for Health and Policy Studies
Health Sciences Centre G230
3330 Hospital Drive N.W.
Calgary, Alberta, Canada T2N 4N1
Telephone: (403) 210-8617
Fax: (403) 210-3818
Email: hquan@ucalgary.ca

Nov 27, 2006

Mr. Iratxe Puebla
Senior Assistant Editor, BMC-series Editor

Dear Mr. Puebla;

RE: Manuscript Number: MS 2141822252114551, " ICD-10 Coding Algorithms for Defining Comorbidities of Acute Myocardial Infarction"

Thank you for giving us the opportunity to resubmit the above manuscript. We greatly appreciate the comments provided by the three reviewers and have revised the manuscript in response to the comments that you provided us in your letter dated November 6, 2006. Below, we provide an itemized summary of the changes made to the paper. Reviewers' comments are shown in bold, followed by our responses. Additions to the manuscript itself are bolded.

Reviewer #1

1.pp.7. How is length of stay defined: (discharge date- admission date + 1)? Are patients dying the day of the event excluded? - Please clarify

Length of stay was defined as: Discharge date – admission date. This means that all AMI patients discharged on the day of the admission or following day are excluded from the administrative dataset regardless of whether they're discharged alive or dead. We excluded patients dying ≤ 1 day because our discharge data included emergency room deaths as in-hospital deaths, which is inconsistent with many administrative data used in the literature. If we included these patients with LOS≤1 day, it would be difficult to distinguish between emergency room deaths from in-hospital deaths. We also excluded patients discharged alive ≤ 1 day in order to improve the accuracy of our AMI diagnosis. The following comment was added on page 7 to clarify the exclusion:

"Patients who are discharged on the same day of the admission or the following day…"

2.pp 7. If patients dying the day of the event are not included- how would an inclusion affects the prediction of mortality?

To address this important question, we re-analyzed our data after including deaths within 1 day. Including the 3,049 patients died within one day produced a decrease in the C-statistics for all our models, as shown in the below table. This decrease is expected because the inclusion of patients died within one day reduces each model's ability to distinguish between deceased and alive patients by using patients' comorbidities. For these patients, the severity of AMI is more of a risk factor than the patients' comorbidities. Nevertheless, we can still infer that there is little difference in predictive ability between the ICD-9 and ICD-10 models and our conclusions remain. The table shows that the ICD-9 and ICD-10 C-statistics are about the same.

	ICD-9 Data In-hospital	**ICD-10 Data** In-hospital	**ICD-9 Data** 30-Day	**ICD-10 Data** 30-Day	**ICD-9 Data** 1 Year	**ICD-10 Data** 1 Year
C-Statistic (Original)	0.841	0.826	0.816	0.809	0.815	0.812
C-Statistic (with AMI deaths within one day)	0.799	0.799	0.787	0.791	0.796	0.804

We added the following comments on page 14 about this limitation:

"We excluded patients who were discharged on the day of admission or the following day. The limitation of this exclusion is that some true AMI cases were excluded. We re-analyzed our data after including deaths within 1 day. Including the 3,049 deaths within one day produced a decrease in the C-statistics for all our models. This suggests that the C-statistics reported in our study are slightly overestimated. Nevertheless, there remains a small difference between the predictive ability of ICD-9 and ICD-10 models.**"**

3. pp 12. Could better treatment of diabetes be related to the non-significant effect of diabetes with complications related to mortality?
4.pp.12 Has the definition of AMI changed in the two study periods e.g. introduction of biochemical markers for AMI and could this influence the associations between mortality and some of the factors in the index? - in the last paragraph in the Discussion section occurrence of AMI is mentioned -please be more specific?!

We analyzed data from 1994 to 2004. Over our study period, advancements in medicine has resulted in increased precision in diagnoses and effectiveness in treatments for diabetes, and other diseases, such as coronary artery disease. These advancements may explain the slight differences between ICD-9 and ICD-10.

On the page 14, we add the following statements:

'In the 10 year period, the criteria for AMI diagnosis were revised and efficient treatment to chronic disease was produced. For example, biomarkers of cardiac troponin and creatinine kinase-MB mass were included in diagnosis of myocardial infarction to increase sensitivity and specificity. The advances in disease treatment and management can partly explain the slight variation between ICD-9 and ICD-10 models that can not be accounted for in our analysis."

5.Table 3: Consider to expand Table 3 with the figures from Table 2 and delete table 2.

We agree with this revision, but having the confidence intervals suggested by Reviewer 2 and frequencies from Table 2 reduces the readability of Table 3. Hence, we did not expand Table 3 with the figures from Table 2.

6. Table 5 Consider to have a footnote that it is patients surviving the day of AMI event (+1 day? – See earlier comment) if this is true?

The following note was added below Table 5:

"Note: Both ICD-9 and ICD-10 data excluded patients who were discharged on the same day of admission or the following day."

7. Table 6 ---do--- / 8. Table 6 AIM (sic)
We fixed the spelling error, i.e., from "AIM" to "AMI" in the title of Table 6.

Reviewer #2

1. The abstract should contain information on the comorbidities that were studied, as well as on the numbers of patients that were included.

The abstract has specified the comorbidities studied and total number of patients in the BC administrative dataset. That is, we added the comorbidities in the objective section.

"The nine comorbidities that we examined were shock, diabetes with complications, congestive heart failure, cancer, cerebrovascular disease, pulmonary edema, acute renal failure, chronic renal failure, and cardiac dysrhythmia."

In the results section we added the number of cases: "49,861 AMI patients in a Canadian province during 1994 - 2004."

2. The authors applied logistic regression to study the relation between ICD-coded comorbidities and 1-year mortality. Why didn't they choose Cox' regression, which accounts for events occurring over a longer period of time, as well as for decreasing numbers of patients at risk?

We could use a Cox regression to account for events occurring over a longer period of time and decreasing risk of mortality. However, we used logistic regressions to predict mortality for fix periods, i.e., in-hospital, 30-day, and one-year mortality to increase the comparability of our findings to past studies, which commonly use this modeling approach. Further, the focus of this study is to assess the effect of a series of conditions in their entirety rather than analyzing each condition specifically.

We added the following comment on page 8:

"An alternative modeling approach is to use a Cox regression to account for the risk of mortality over time. However, logistic regression was used to predict mortality for fixed periods. This approach has been commonly used in previous studies. Therefore, comparability of our study to previous studies could be increased by using the logistic regression approach."

3. It is unclear if (and which) model-building strategy was applied. How should the ORs in table 4 be interpreted? Are these adjusted for all comorbidities?

The objective of this study is to not develop a new prediction rule for the ICD-10 data, but to determine whether ICD-10 coding algorithms can be matched to ICD-9 algorithms. Given that the nine comorbidities used in this study were validated by Tu et al's study and referenced by other studies, this prediction rule is expected to generate similar estimates. Hence, no model-building strategies such as backward or forward stepwise regression were used.

We added the following comments on page 8:

"Each OR should be interpreted as controlling for all nine comorbidities, age and sex …No model building strategy was employed because this 11 variable prediction rule had been validated by Tu et al's study and referenced subsequently by studies using this prediction rule."

4. The authors study the predictive models in terms of discrimination. They should also study the performance of their models in terms of calibration: how well were mortality rates predicted?

We added statistics showing predictive ability and goodness-of-fit i.e., Brier Score and R-square in Table 6. Both statistics show that ICD-9 data is similar to ICD-10 data. The following comments were added on page 8, 9 and 11:

"Another measure we used to assess the models' predicted probabilities was the Brier score. The Brier score measures the mean squared difference between expected probability of dying and its actual occurrence. Low Brier scores indicate that there is a small difference, which shows that the model predicted well. The R-square summary statistic is used to assess how well the model fits with the data. Particularly, R-square measures the proportion of the total variation in mortality explained by the logistic regression model… The same result was observed when comparing the Brier scores between ICD-9 and ICD-10 regressions. The lower Brier scores for ICD-10 regressions relative to ICD-9 regression indicated that it performed slightly better than the ICD-9 regressions. Both ICD-9 and ICD-10 models explained approximately 20 percent of the variation in mortality, as shown by the R-squares in Table 6. The small difference between ICD-9 and ICD-10 models' R-squares shows again the similarity between the two coding systems."

5. The reader should be informed on the completeness of follow-up.

The proportion of patients who can't be linked between the inpatient data and registry file are lost to follow up. This information is unavailable to us, so we cannot inform the reader of this proportion. However, according to Li et al. (2006), the highest correct linkage rate between these two dataset is 98.9 percent, so excluding this proportion should not affect our results significantly. For the proportion of patients who move out of the BC province without notice of population registry and consequently are lost in the follow-up, we expect to be also small given that the largest period that mortality is studied is one year. Again, we lack the information in our dataset to inform the reader of this proportion.

We added the following comment on page 14:

"Lastly, the proportion of AMI patients that could not be linked with vital death registry or had moved out of the BC province without notice to the population registry in our study period were lost to follow-up. According to Li et al.(34), the accuracy of linkage between the registry file and inpatient discharge dataset was about 98%. We also suspect that only a few AMI patients

moved out of the BC province given that our longest period of observing mortality was one year. Hence, our results were less likely affected by this small proportion of AMI patients lost to follow-up."

6. Table 3 presents the relation between ICD-9/ICD-10 data and chart review among 193 patients. This is a small sample. Furthermore, several of the comorbidities are rare (e.g. shock N=5; cancer N=10). Therefore, the table will greatly improve if the authors also present 95% CIs around the point estimates of sensitivity, specificity, PPV, NPV. The authors should comment (discussion section) the (high) level of uncertainty of the point estimates. This might be the reason why ICD-9 and ICD-10 datasets poorly matched to chart review data for cerebrovascular disease, pulmonary edema and cardiac dysrhythmia.
We agree with these comments and added confidence intervals in Table 3. We also made the following comments on pages 9, 10 and 12:

"The width of the 95% confidence intervals for the nine comorbidities shows the low prevalence for most of the comorbidities in the chart review data except for cardiac dysrhythmia and congestive heart failure."

"The low positive predictive value for cerebrovascular disease is partly due to the low prevalence.**"**

7. Table 3 presents an indirect comparison between the ICD-9 and ICD-10 coding algorithms. This reviewer would like to see a direct comparison (so: cross-tables for each of the 9 comorbidities).

Please see below for cross-tables for all nine comorbidities.

Reviewer #3

1. Including in the introduction what are the problems and challenges with the transition from the ICD-9 coding system to the ICD-10 coding system. The "so what" question. Also, appropriate references should be added in the "reference" section (that is incorrectly written "REFERNCES")
To emphasize the importance of resolving the problems with transitioning from ICD-9 to ICD-10 coding system, the following comment is added to page 4 and 5:

"Without valid ICD-10 coding algorithms, administrative data can not be used to measure and control for patients' comorbidities when studying their clinical outcomes. Further, future policy and quality decisions using ICD-10 data require the development and validation of ICD-10 coding algorithms for these comorbidities."

The spelling error for "references" has been corrected on page 17.

2. **Justify the choice of the 9 AMI comorbidities. For example, why cancer but no infectious diseases like pneumonia?**
As noted in the introduction, the nine comorbidities were chosen because they were validated by Tu et al.'s paper and several other studies. Further, many other studies and the Canadian government have employed this prediction rule to predict AMI outcomes. Hence, no new comorbidities were added to this prediction rule.

The following comment is added on page 4:

"Risk adjustment is an important tool used in health service research to account for differences in AMI patient's characteristics. To develop such a tool for AMI patients…"

3. **Specify the method of random selection of 4,008 charts for the 4 teaching hospitals.**
On page 6, the first sentence of paragraph two was changed to the following:

"A simple random sample from four teaching hospitals was used to select 4,008 patients' charts to be reviewed."

4. **More detail should be provided on the physician involved in the revision of the comprehensive ICD-10 code list (page 6). Are they clinicians or non clinicians, specialists (which specialty) or family physicians?**

On page 6, we added the following: "The four physicians are specialists in neurology, general internal medicine, and psychiatry"

5. Titles for some tables could be more specific: Table 2 - indication that the table concerns the Alberta random chart review and specification on the total n.
The title of this table has been changed to "Table 2: Prevalence of comorbidity in Alberta chart review data, ICD-9, and ICD-10 data among AMI patients (N = 193)**"**

Table 3 – indication that the table concerns the Alberta random chart review and specification on the total n.
The title of this table has been changed to "Table 3: Agreement between ICD-9/ICD-10 and Alberta chart review data among AMI patients (N = 193)"

Table 5 - indication that the table concerns the BC provincial database and specification on the total n.
The title of this table has been changed to "Table 5: Mortality by comorbidities among AMI patients from the BC provincial database (N = 49,861)"

Again, thank you for providing us the opportunity to revise and resubmit this manuscript. We would like to commend the reviewers for their absolutely outstanding reviews. They obviously took a great deal of time to review the fine details of our paper, and we would certainly appreciate it if you could send them a special thank you on our behalf for their excellent ideas and contribution to a stronger paper. We hope that you will be satisfied with our responses and revisions.

We look forward to hear from you with a decision on our paper.

Yours sincerely,

Lawrence So
Dewey Evans
Hude Quan

The agreement between Alberta Chart Review Data and BC administrative data

Shock		Chart Data		
		Positive	Negative	Total
ICD-9-CM	Positive	3	2	5
	Negative	2	186	188
	Total	5	188	193
	Sensitivity	60.00		
	Specificity	98.94		
	PPV	60.00		
	NPV	98.94		

Shock		Chart Data		
		Positive	Negative	Total
ICD-10	Positive	3	1	4
	Negative	2	187	189
	Total	5	188	193
	Sensitivity	60.00		
	Specificity	99.47		
	PPV	75.00		
	NPV	98.94		

Diabetes		Chart Data		
		Positive	Negative	Total
ICD-9-CM	Positive	12	3	15
	Negative	3	175	178
	Total	15	178	193
	Sensitivity	80.00		
	Specificity	98.31		
	PPV	80.00		
	NPV	98.31		

Diabetes		Chart Data		
		Positive	Negative	Total
ICD-10	Positive	10	2	12
	Negative	5	176	181
	Total	15	178	193
	Sensitivity	66.67		
	Specificity	98.88		
	PPV	83.33		
	NPV	97.24		

CHF		Chart Data		
		Positive	Negative	Total
ICD-9-CM	Positive	45	5	50
	Negative	10	133	143
	Total	55	138	193
	Sensitivity	81.82		
	Specificity	96.38		
	PPV	90.00		
	NPV	93.01		

CHF		Chart Data		
		Positive	Negative	Total
ICD-10	Positive	44	3	47
	Negative	11	135	146
	Total	55	138	193
	Sensitivity	80.00		
	Specificity	97.83		
	PPV	93.62		
	NPV	92.47		

Cancer		Chart Data		
		Positive	Negative	Total
ICD-9-CM	Positive	9	1	10
	Negative	1	182	183
	Total	10	183	193
	Sensitivity	90.00		
	Specificity	99.45		
	PPV	90.00		
	NPV	99.45		

Cancer		Chart Data		
		Positive	Negative	Total
ICD-10	Positive	9	3	12
	Negative	1	180	181
	Total	10	183	193
	Sensitivity	90.00		
	Specificity	98.36		
	PPV	75.00		
	NPV	99.45		

Cerebrovascular		Chart Data		
		Positive	Negative	Total
ICD-9-CM	Positive	6	12	18
	Negative	0	175	175

Cerebrovascular		Chart Data		
		Positive	Negative	Total
ICD-10	Positive	6	8	14
	Negative	0	179	179

	Total	6	187	193
	Sensitivity	100.00		
	Specificity	93.58		
	PPV	33.33		
	NPV	100.00		

	Total	6	187	193
	Sensitivity	100.00		
	Specificity	95.72		
	PPV	42.86		
	NPV	100.00		

Pulmonary		Chart Data		
		Positive	Negative	Total
ICD-9-CM	Positive	2	0	2
	Negative	26	165	191
	Total	28	165	193
	Sensitivity	7.14		
	Specificity	100.00		
	PPV	100.00		
	NPV	86.39		

Pulmonary		Chart Data		
		Positive	Negative	Total
ICD-10	Positive	2	0	2
	Negative	26	165	191
	Total	28	165	193
	Sensitivity	7.14		
	Specificity	100.00		
	PPV	100.00		
	NPV	86.39		

A Renal		Chart Data		
		Positive	Negative	Total
ICD-9-CM	Positive	12	3	15
	Negative	3	175	178
	Total	15	178	193
	Sensitivity	80.00		
	Specificity	98.31		
	PPV	80.00		
	NPV	98.31		

A Renal		Chart Data		
		Positive	Negative	Total
ICD-10	Positive	12	8	20
	Negative	3	170	173
	Total	15	178	193
	Sensitivity	80.00		
	Specificity	95.51		
	PPV	60.00		
	NPV	98.27		

C Renal		Chart Data		
		Positive	Negative	Total
ICD-9-CM	Positive	15	7	22
	Negative	3	168	171
	Total	18	175	193
	Sensitivity	83.33		
	Specificity	96.00		
	PPV	68.18		
	NPV	98.25		

C Renal		Chart Data		
		Positive	Negative	Total
ICD-10	Positive	13	3	16
	Negative	5	172	177
	Total	18	175	193
	Sensitivity	72.22		
	Specificity	98.29		
	PPV	81.25		
	NPV	97.18		

Dysrhythmia		Chart Data		
		Positive	Negative	Total
ICD-9-CM	Positive	38	0	38
	Negative	68	87	155
	Total	106	87	193
	Sensitivity	35.85		
	Specificity	100.00		
	PPV	100.00		
	NPV	56.13		

Dysrhythmia		Chart Data		
		Positive	Negative	Total
ICD-10	Positive	38	2	40
	Negative	68	85	153
	Total	106	87	193
	Sensitivity	35.85		
	Specificity	97.70		
	PPV	95.00		
	NPV	55.56		