Reviewer’s report

Title: Spatial variation in the management and outcomes of acute coronary syndrome

Version: 2 Date: 4 May 2005

Reviewer: Joanna Moschandreas

Reviewer’s report:

General comments.
A combination of “traditional” statistical methods (linear regression, logistic regression, cluster analysis) and techniques used in the field of geomatics (the geographic information system ArcGIS, geographically weighted regression, cartographic representations) are employed for data analysis and presentation purposes in this interesting article. Outcomes are hospital death (HD, yes/no), length of stay in hospital (LoS, in days), invasive cardiac procedures (ICP, yes/no) and early hospital readmission for those patients surviving hospitalization for acute coronary syndrome (EHR, readmission for heart disease in the first 30 days following discharge). Covariates considered are distance to specialized cardiology centre, geographic location, age and gender.

As it stands, the statistical analysis appears inadequate in certain respects, and revision is required. It is essential that the three points provided below are addressed:

1) The regression models (logistic and linear regression). In the Discussion (pg 15) and the Conclusion (pg18-19) it is stated that the main findings of the study are essentially those presented in Table 2. In the corresponding statistical models, however, no 2nd order interactions were considered (see also Specific comment 9) below). The authors themselves note that a ‘paradoxical finding’ between LoS, ICP and proximity may be explained by interaction between the covariates. These models need to be refitted, including interaction terms. If interactions are present, then the main effects cannot be interpreted so the conclusions resulting from main effects analyses (presented in Table 2) may be misleading. In addition, the residuals will obviously change if the models change. These analyses need to be rerun initially including 2nd order interaction terms in all models.

2) The hierarchical cluster analysis (HCA) undertaken. In the application of HCA to group the standardized ratios in the 16 Quebec regions it is not clear which proximity measure was used (for the distance matrices), why the choice of four clusters was made for each of the ICP, HD & EHR ratios or if, in fact, there was any statistical basis to the approach. The last sentence of paragraph 1 on page 8 (“This method was used instead of…..”) needs to be clarified. It would appear preferable not to perform the HCA, and simply to present the ratios in each region in Figure 2.

3) GWR and binary response. For the logistic regression models (i.e. those with ICP, HD and EHR as dependent variables), it was not possible to statistically assess spatial variation using the package GWR (as mentioned on pg 9). The ad-hoc approach used here cannot lead to any conclusions about spatial variability or the lack of it (pg 12). The title does not adequately reflect the contents of the manuscript (as spatial heterogeneity is only statistically assessed for LoS).

Specific comments & minor corrections
1) pg 7 It is not clear why the variable representing distance to the cardiology centre requires categorization. In addition, it is not certain that the categories chosen are indeed the most appropriate. It is stated that the categories chosen (<32km, 32-64 km, 64-105 km & >105km) were based on aerial transportation times of 60, 90 & 120 minutes [Ref 11 & Abstract Ref 18]. Reference
11 provides the same cut-offs but the transportation times were based on ground ambulance data and round-trip helicopter distances provided by the University of Michigan (unpublished data).

2) pg 7 The distance was “defined by the centroid of the postal code area”. How is “centroid” defined (presumably not in the mathematical sense). For example, is it the largest concentration of dwellings?

3) pg 8 1st sentence of paragraph 2. Replace “Multiple log-normal analysis was…” by “Multiple linear (log-normal) regression analysis was….”

4) pg 8 State which package was used for the cluster analysis.

5) pg 9 Global estimates and local estimates. Please explain this terminology. Are the global estimates those obtained by the ordinary least squares approach?

6) pg 9 State the version of ArcGis used and the producers of the package. pg 9 State the versions of GWR, SAS & SPSS used (and producers).

7) pg 9 State which package was used for the regression analyses, and therefore the computation of residuals.

8) pg 11. 2nd paragraph. There is an unjustified etiological emphasis on the interpretation of the results. eg “The presence of an ICP reduces HD rates” - could it not be that subjects who are too ill are not subjected to ICP?

9) pg 11 2nd paragraph. It is noted that women had decreased odds of receiving ICP. Possible interactions between gender and ICP on the odds of HD and EHR should be considered. See General comment 1.

10) pg 11 Were any statistical tests performed to confirm the postulated relationships between LoS, ICP and distance (Table 3)? If not, the last section of paragraph 2 and Table 3 should be removed. See also 13).

11) pg 18 The word ‘robust’ should be avoided as there is no indication that the models are indeed ‘robust’ in the statistical sense.