

# **Research Article**

# A Comparison of Generalized Additive Models to Other Common Modeling Strategies for Continuous Covariates: Implications for Risk Adjustment

Lynne Moore<sup>1,2\*</sup>, James A Hanley<sup>2</sup>, Alexis F Turgeon<sup>1,4</sup> and André Lavoie<sup>1,3</sup>

<sup>1</sup>Department of Epidemiology and Biostatistics. McGill University, Montreal, Quebec, Canada

<sup>2</sup>Unité de traumatologie-urgence-soins intensifs, Centre de Recherche du CHA (Hôpital de l'Enfant-Jésus), Université Laval, Quebec City, Quebec, Canada
<sup>3</sup>Département d'anesthésiologie, Division de soins intensifs, Hôpital de l'Enfant-Jésus, Université Laval, Quebec City, Quebec, Canada
<sup>4</sup>Département de médecine sociale et préventive, Université Laval, Quebec City, Quebec, Canada

#### Abstract

Common modeling strategies for quantitative covariates include single linear terms, dummy variables on categories, Fractional Polynomials (FP) and cubic smoothing splines in Generalized Additive Models (GAM). The goal of this study was to evaluate the impact of using GAM over other common covariate modeling strategies on risk adjustment. Analyses were based on inter-hospital mortality comparisons in a Canadian provincial trauma system (n=123,732; 59 hospitals). Parameter estimates describing the increase in log odds of mortality for one hospital compared to the reference were adjusted with five quantitative covariates modeled using 1) single linear terms, 2) dummy variables on 2, 3, 4, 5 categories, 3) FP, and 4) GAM. The parameter estimates generated by the first three modeling strategies were compared to that generated by the GAM using mean standardized difference. Mean standardized difference (95% CI) was 71.69 (51.7-91.7) for single linear terms, 21.1 (14.3-28.9); 23.4 (15.6-31.2); 49.6 (28.1-71.1); and 48.5 (28.8-68.2) for dummy variables on 2, 3, 4, and 5 categories, respectively and 12.7 (10.0-15.4) for FP. Results suggest that GAM, FP and at least 4 risk-homogeneous categories provide equivalent risk adjustment to smoothing splines in GAM while single linear terms and less than 4 categories may induce residual confounding.

**Keywords:** Multivariable regression; Quantitative covariates; Generalized additive models; Fractional polynomials; Risk adjustment; Residual confounding

# Introduction

When evaluating the effect of an exposure using observational data, adjustment for potential confounders is usually inevitable and is generally performed using statistical modeling. Risk adjustment models often include quantitative variables that may not have a linear relationship with the independent variable. As such, if they are modelled without transformation, the postulate of linearity, essential for parametric modeling strategies, is violated leading to bias and coverage problems [1]. If these variables are of primary interest, violating the postulate of linearity can lead to effect estimates that are too close to the null [2]. If these variables are used to control confounding, violation of the postulate of linearity can lead to residual confounding. Consequent influence on the measure of exposure effect is difficult to predict.

The literature abounds with advice on how to model quantitative covariates [3-6]. Some have suggested that dummy variables on a minimum of five categories are sufficient [7]. Others have suggested that parametric transformation with fractional polynomials (FP) [8] or non-parametric functions in Generalized Additive Models (GAM) [9] can improve control of confounding. Most studies that have evaluated the impact of modeling strategy on risk adjustment have been based on simulated data and limited to null associations between exposure and disease [6,10,11]. Therefore, while GAM theoretically offer the most flexible modeling strategy, [12] there is a lack of evidence based on real data that they lead to better control of confounding.

The objective of this study was to evaluate the impact of using GAM over other commonly used covariate modeling strategies on risk adjustment.

# Methods

## Study setting

The study was based on the comparison of adjusted hospital

mortality estimates across trauma centers in the province of Quebec, Canada. The context of trauma provides a good test bed for the evaluation of risk adjustment as there is great heterogeneity in the baseline risk of injured patients across centers, most of the indicators used to describe baseline risk are quantitative in nature, and the relationship between many risk factor indicators and mortality is non-linear. We were primarily interested in how parameter estimates describing the log odds of mortality in one trauma center compared to another trauma center, adjusted for quantitative risk factors, would vary according to the modeling strategy applied to risk factors. Of secondary interest was the effect of modeling strategy on statistical significance.

### Study data

The study sample was drawn from the Quebec Trauma Registry, which contains information on all trauma patients admitted to any of the 59 trauma centres in the province of Quebec, Canada according to the following uniform inclusion criteria: death, intensive care unit admission, hospital length of stay of more than two days, or transfer from another hospital. Data is extracted from patient files in each trauma center, according to standardized coding protocols, and then centralized at the Quebec Ministry of Health where it is screened and

\*Corresponding author: Lynne Moore, Unité de traumatologie - urgence - soins intensifs, Centre de recherche du CHA (Hôpital de l'Enfant-Jésus), 1401, 18eme rue, Quebec City, Quebec, Canada, G1J 1Z4, Tel: 01-418-649-0252; Fax: 01-418-649-5733; E-mail: lynne.moore.cha@ssss.gouv.qc.ca

Received November 09, 2010; Accepted January 11, 2011; Published January 14, 2011

**Citation:** Moore L, Hanley JA, Turgeon AF, Lavoie A (2011) A Comparison of Generalized Additive Models to Other Common Modeling Strategies for Continuous Covariates: Implications for Risk Adjustment. J Biomet Biostat 2:109. doi:10.4172/2155-6180.1000109

**Copyright:** © 2011 Moore L, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Page 2 of 5

added to the Quebec Trauma registry. Periodic validation, supervision by a data coordinator, yearly on-going training for data coders, an electronic forum of coding queries, and thrice-yearly meetings with key stakeholders (e.g. trauma physicians, researchers, administrators), all help improve data quality. Trauma centers with less than 100 patients (n=7) were excluded from analyses.

## Statistical analysis

The outcome variable of interest was hospital mortality. Risk adjustment was based on five variables commonly used for case-mix control in trauma research: 1) age; 2) the Injury Severity Score, an ordinal variable varying between 1 and 75 that measures increasing anatomical injury severity [13]; 3) the Glasgow Coma Score, an ordinal variable varying between 3 and 15 that measures increasing level of consciousness [14]; 4) respiratory rate, an interval-scale variable varying from 0 to around 100; and 5) systolic blood pressure, an interval-scale variable varying from 0 to around 300. The Injury Severity Score is based on clinical evaluations performed throughout the hospital stay whereas the Glasgow Coma Score, respiratory rate, and systolic blood pressure are measured on arrival at the emergency department of the trauma center.

Risk adjustment variables were entered in a logistic regression model using each of the four modeling strategies: 1) a single linear term, 2) dummy variables on either 5, 4, 3, or 2 categories, 3) FP, and 4) GAM. Categories were based on clinically plausible cut-points that are widely used in the literature. However, alternative cut-points were evaluated in sensitivity analyses.

## **Fractional polynomials**

Introduced by Royston and Altman in 1994 [14], fractional polynomials (FP) are an extension of traditional quadratic and cubic polynomials whereby a wider range of powers is used. Royston suggests that the following group of powers gives all the flexibility needed: -2, -1, -0.5, 0, 0.5, 1, 2, and 3, where 0 represents the logarithm. Models can be first degree (with only one polynomial term) or can include more terms. According to Royston, two terms (second degree model) offer sufficient flexibility in most situations. Optimal FP were based on minimizing model deviance using STATA software (Stata Corporation, College Station, TX, Version 7.0). Clinical plausibility was verified by inspecting curves of the functional relationship between risk factors and the risk of mortality.

#### Generalized additive models

Generalized Additive Models are an extension of Generalized

Linear Models that are based on the specification of a parametric link function (the logit is used for logistic regression) and allow independent variables to be modeled with parametric or non parametric functions. Non-parametric functions (cubic splines are used in this study) allow the expression of relations to mortality that are non-linear. While the location of knots in a cubic smoothing spline has little influence on results, the number of knots (equivalent to the smoothing parameter usually expressed as the number of degrees of freedom) is important. Too many degrees of freedom can lead to overfitting (describing random error rather than a true underlying relationship) whereas choosing too few can lead the analyst to miss important trends (the linear model is an extreme example of this with a smoothing parameter equivalent to 1 degree of freedom). Unfortunately, the algorithms available for choosing the optimal smoothing parameter (e.g. maximum of the generalizedcross validation function) are not yet well developed and can generate misleading results. For this reason, a fixed smoothing parameter equivalent to four degrees of freedom was used for each variable. Data simulations performed by Hastie and Tibshirani have demonstrated that four degrees of freedom is suitable in most circumstances [12]. However, other smoothing parameters were evaluated in sensitivity analyses. Again, clinical plausibility of functional relationships was verified graphically. GAM were fit using the PROC GAM procedure in the Statistical Analysis System (SAS Institute, Cary, NC, version 9.2).

# Model fit

The model fit offered by each of the four modeling strategies was evaluated by fitting the logistic models described in Table 1.

The fit of each model was described using Akaike's Information Criteria (AIC). The AIC measures goodness of fit and is calculated as a function of the maximum of the likelihood function and a correction for the number of parameters in the model. The discrimination of each model was also evaluated described by the Area Under the receiver Operating Characteristic curve (AUC). The AUC varies between 0 and 1 where a model with an AUC=1 discriminates deaths from survivors perfectly and a model with an AUC=0.5 discriminates no better than chance alone. AUC were compared using the non-parametric method for dependant samples described by Delong et al. [15]. Finally, curves of *observed* and *predicted* mortality were compared for univariate models. Observed mortality proportions were calculated for as many small intervals of quantitative variables as possible based on the constraint of at least 10 events per interval.

# **Risk adjustment**

The influence of modeling strategy on risk adjustment was assessed

Model	Description	Specification
I	Single linear term	$Logit(P) = \alpha + \beta_1 AGE + \beta_2 ISS + \beta_3 GCS + \beta_4 RR + \beta_5 SBP$
lla	Dummy variables on 5 categories	$ \begin{array}{l} \text{Logit}(P) = \alpha + \beta_{\text{A}}\text{GE}_{55:64} + \beta_{2}\text{AGE}_{65:74} + \beta_{3}\text{AGE}_{75:84} + \beta_{4}\text{AGE}_{364} + \beta_{5}\text{ISS}_{9:15} + \beta_{6}\text{ISS}_{16:24} + \beta_{7}\text{ISS}_{25:40} + \beta_{6}\text{ISS}_{360} + \beta_{9}\text{GCS}_{9:12} + \beta_{10}\text{GCS}_{6:8} + \beta_{11}\text{GCS}_{4:5} + \beta_{12}\text{GCS}_{3} + \beta_{13}\text{RR}_{0} + \beta_{14}\text{RR}_{1:5} + \beta_{15}\text{RR}_{6:9} + \beta_{16}\text{RR}_{329} + \beta_{17}\text{SBP}_{0} + \beta_{18}\text{SBP}_{1:49} + \beta_{13}\text{SBP}_{50:75} + \beta_{20}\text{SBP}_{76:89} \end{array} $
IIb	Dummy variables on 4 categories	$ \begin{array}{l} \text{Logit}(\text{P}) = \alpha + \beta_{1}\text{AGE}_{_{65}74} + \beta_{3}\text{AGE}_{_{75:84}} + \beta_{4}\text{AGE}_{_{84}} + \beta_{5}\text{ISS}_{_{9:15}} + \beta_{6}\text{ISS}_{_{16:24}} + \beta_{8}\text{ISS}_{_{24}} + \beta_{9}\text{GCS}_{_{9:12}} + \beta_{_{10}}\text{GCS}_{_{6:8}} + \\ \beta_{11}\text{GCS}_{_{3:5}} + \beta_{_{13}}\text{RR}_{_{0:5}} + \beta_{_{15}}\text{RR}_{_{6:9}} + \beta_{_{16}}\text{RR}_{_{29}} + \beta_{_{17}}\text{SBP}_{_{0:49}} + \beta_{_{19}}\text{SBP}_{_{50:75}} + \beta_{_{20}}\text{SBP}_{_{76:89}} \\ \end{array} $
llc	Dummy variables on 3 categories	$ \begin{array}{l} \text{Logit}(P) = \alpha + \beta_2 \text{AGE}_{_{65.74}} + \beta_4 \text{AGE}_{_{774}} + \beta_5 \text{ISS}_{_{9.15}} + \beta_8 \text{ISS}_{_{>15}} + \beta_9 \text{GCS}_{_{9.12}} + \beta_{_{12}} \text{GCS}_{_{3.8}} + \beta_{_{13}} \text{RR}_{_{0.9}} + \beta_{_{16}} \text{RR}_{_{>29}} + \beta_{_{17}} \text{SBP}_{_{0.75}} + \beta_{_{20}} \text{SBP}_{_{76.89}} \end{array} $
IId	Dummy variable on 2 categories	$Logit(P) = \alpha + \beta_{4}AGE_{>65} + \beta_{5}ISS_{>15} + \beta_{9}GCS_{>8} + \beta_{13}RR_{0.9 >29} + \beta_{17}SBP_{0.89}$
111	FP	$ \begin{array}{l} \text{Logit}(P)\text{=} \ \alpha + \beta_1 \text{AGE}^3 + \beta_2 (\text{AGE}^3 \text{ X log}(\text{AGE}))_1 + \beta_3 \text{ISS}^{-1} + \beta_4 \text{log}(\text{ISS})_1 + \beta_5 \text{log}(\text{GCS}) + \beta_6 \text{GCS}^3 + \beta_7 \text{log}(\text{RR}) + \beta_8 (\text{log}(\text{RR}) + \beta_1 \text{log}(\text{RR}))_1 + \beta_9 \text{SBP} + \beta_{10} \text{log}(\text{SBP})_1 \end{array} $
IV	GAM	$Logit(P)= \alpha + s(AGE) + s(ISS) + s(GCS) + s(RR) + s(SBP)$

Logit(P) is the log odds of the probability of hospital mortality, ISS: Injury Severity Score; GCS: Glasgow Coma Score; RR: Respiratory Rate; SBP: Systolic Blood Pressure, FP: Fractional Polynomials, GAM: Generalized Additive Model, S: cubic smoothing spline

Table 1: Models used for analyses.

Citation: Moore L, Hanley JA, Turgeon AF, Lavoie A (2011) A Comparison of Generalized Additive Models to Other Common Modeling Strategies for Continuous Covariates: Implications for Risk Adjustment. J Biomet Biostat 2:109. doi:10.4172/2155-6180.1000109

by entering a series of dummy variables representing each center compared to a reference (the largest center) into models I to IV. Trauma center effect estimates generated by models I to III were compared to those generated by the GAM (model IV) by calculating standardized difference in parameter estimates, i.e. ( $\beta_{jGAM} - \beta_{jMODEL}$ )/SE( $\beta_{jGAM}$ ) and the proportion of estimates that changed statistical significance. Standardized differences of over 40% were considered important as they are considered to have adverse effects on bias and coverage [16].

Statistical analyses were performed using the Statistical Analysis System (SAS Institute, Cary, NC, version 9.2). The study was approved by the Research Ethics Committee of the 'Centre Hospitalier Affilié Universitaire de Québec' and the 'Commission d'Accès à l'Information du Québec'.

### Results

The study population comprised 123,027 patients of whom 7600 (6.2%) died in hospital. SBP had a Gaussian distribution but all other quantitative variables had asymmetric distributions. There was an important variation in the distribution of risk factors across trauma centers, illustrating the need to adjust for baseline risk in hospital comparisons: Mean age varied between 8.0 (95% CI: 7.9-8.2; paediatric center) and 71.2 (70.5-71.8) years; mean ISS between 5.2 (4.5-5.9) and 14.1 (138-14.4); mean GCS between 12.4 (11.7-13.0) and 14.8 (14.7-14.8); mean RR between 18.0 (17.9-18.1) and 23.1 (22.8-23.3); and mean SBP between 116.3 (115.8-116.9) and 149.8 (148.5-151.1).

All non-linear terms for GAM had p-values <0.001, demonstrating that the association between each quantitative variable and hospital mortality had a significant non-linear component. When compared to the *observed* risk of mortality, single linear terms led to an underestimation of risk for patients aged over 80 years (Figure 1a), an overestimation of risk for ISS>40 (Figure 1b), an underestimation of risk for GCS=3, and an overestimation of risk for GCS=4 to 10 (Figure 1c). In addition, single linear terms led to an overestimation of risk for RR>20 (Figure 1d) and an underestimation of risk for SBP<50 and SBP>150 (Figure 1e). Dummy variables on categories led to an underestimation of risk for patients >85 years of age (Figure 1a), an underestimation for ISS>50 (Figure 1b), and they did not pick up the increase in risk for SBP>150 (Figure 1e). The functional relationships between all risk factors and mortality risk described by GAM and FP were very close to observed values. However, for the ISS, RR, and SBP,











GAM and FP functions diverged for extreme values where data was sparse (Figure 1b, 1d, and 1e).

The discrimination and fit of models with a single linear term were inferior to that of the GAM (Table 2). The discrimination and model fit of models based on categories decreased with the number of categories used. Dummy variables on less than 4 categories were not only inferior to the GAM but were also inferior to single linear terms. Discrimination and model fit using FP were equivalent to that of the GAM.

Mean standardized difference in parameter estimates from the GAM model were over 40% when single linear terms or dummy variables on 2 or 3 categories were used but under 40% for FP and dummy variables on 4 or 5 categories (Table 3). Similarly, single linear terms or dummy variables on 2, 3, or 4 categories led to a change in statistical significance over GAM in more than 10% of estimates.

# Sensitivity analyses

Generalized cross validation of deviance functions could not identify 'optimal' smoothing parameters for the GAM model. We therefore increased df by two to a maximum of 20 (or not greater than the number of levels /2) for each variable. Model fit and discrimination







pressure Predicted mortality risk was generated using a univariate logistic model of hospital mortality as a function of the systolic blood pressure.

for the model with maximum df for each variable was similar to that obtained with a fixed smoothing parameter equivalent to 4 degrees of freedom (AIC=34,600; AUC=0.901). When cut-points were changed to reflect alternative clinically plausible categories (ex. GCS<13; ISS>12; age>70), results for models for dummy variables were similar.

# Discussions

The results of this study suggest that the method used for modelling quantitative covariates in a logistic regression model does have an influence on risk adjustment. Of the four modelling strategies used, FP and GAM led to significantly better model fit and discrimination than single linear terms or categories. However, FP, GAM, and dummy variables on at least four categories led to similar parameter estimates. Conversely, single linear terms and less than four categories led to poorer model fit, significantly worse discrimination than GAM, and parameter estimates with a mean standardized difference of more than 40%. If we assume that among the modeling strategies tested, GAM predict outcome most accurately, an assumption supported by the better fit of the GAM model to observed data, results suggest that FP, GAM, or dummy variables on at least 4 categories may offer superior risk adjustment to the other modelling strategies tested.

The potential for residual confounding when quantitative covariates are modeled with single linear terms in spite of non-linear relations to response has been widely documented [6,7,17]. In addition, several studies have shown evidence that using dummy variables on categories for risk adjustment can lead to residual confounding [10,18]. In the 1960s, Cochrane studied the effect of crude versus fine categorization of confounding variables on exposure effect estimates and found that bias decreased as the number of categories used increased [18]. Royston showed that FP fit data better than single linear terms and dummy variables on categories [8]. More recent literature has suggested that GAM can improve risk adjustment compared to single linear terms or parametric transformations of linear terms [9]. In a study comparing spline regression to dummy variables on five categories, little difference in exposure OR effect estimates was observed [17]. The results of this study therefore confirm previous work and go further by suggesting that a) FP and GAM provide similar model fit and risk adjustment and b) although categories do not fit data as well, using at least four may provide equivalent risk adjustment to FP and GAM.

While dummy variables on categories, FP, and GAM are useful strategies for modelling non-linear associations to outcome, each method has its disadvantages. As shown in our data, dummy variables on categories are sensitive to the number of categories used. In addition, the choice of cut-points is crucial; multiple risk-homogeneous categories are ideal but may lead to categories with low sample sizes resulting in unstable regression coefficients. Furthermore, dummy variables lead to step functions that are clinically implausible and do not use all the information in the quantitative variable, which can lead to loss of statistical power [5].

FP and GAM allow for smooth non-linear relations to the dependant variable and model quantitative variables over the whole range of values. However, they are mathematically complex and both require careful implementation. In the case of FP, the correct choice of powers is critical and relying solely on empirical model fit (deviance statistics are used in STATA) may lead to clinically implausible associations. In the case of cubic spline functions used in GAM, the

Model	AUC	AIC
I) Single linear term	0.886 (0.882-0.889)*	37 944
IIa) Dummy variables on 5 categories	0.894 (0.891-0.897)*	35 364
IIb) Dummy variables on 4 categories	0.887(0.884-0.890)*	35 799
IIc) Dummy variables on 3 categories	0.876 (0.872-0.880)*	37 729
IId) Binary variables	0.854 (0.850-0.858)*	39 594
III) FP	0.901 (0.898-0.904)	35 054
IV) GAM	0.901 (0.898-0.905)	34 948

FP: Fractional Polynomials, GAM: smoothing splines in a Generalized Additive Model

\*p<0.0001 when compared to the generalized additive model

Table 2: Discrimination and model fit.

Model	Mean standardized difference (95% CI)	% changed statistical significance (95% CI)
No adjustment	172.3 (138-207)	51.9 (0.38-0.66)
<ol> <li>Single linear term</li> </ol>	71.69 (51.7-91.7)	15.4 (6.9-28.1)
II) Categories (5)	21.1 (14.3-28.9)	5.8 (1.2-12.1)
IIb) Categories (4)	23.4 (15.6-31.2)	11.5 (4.4-23.4)
IIc) Categories (3)	49.6 (28.1-71.1)	15.4 (6.9-28.1)
IId) Binary variables	48.5 (28.8-68.2)	13.5 (5.6-25.8)
FP	12.7 (10.0-15.4)	3.9 (0.5-13.2)
GAM	0	0

FP: Fractional Polynomials, GAM: smoothing splines in a Generalized Additive Model

Table 3: Mean standardized difference in parameter estimates, and percent change in statistical significance (p<0.05) compared to the GAM model.

Page 4 of 5

Citation: Moore L, Hanley JA, Turgeon AF, Lavoie A (2011) A Comparison of Generalized Additive Models to Other Common Modeling Strategies for Continuous Covariates: Implications for Risk Adjustment. J Biomet Biostat 2:109. doi:10.4172/2155-6180.1000109

number of knots (equivalent to the smoothing parameter usually expressed as the number of degrees of freedom) is important. Too many degrees of freedom can lead to over fitting whereas important trends can be missed if too few are used. Unfortunately, the algorithms available for choosing the optimal smoothing parameter, such as generalized cross validation of deviance functions, are not yet well developed and can generate misleading results [12]. Identification of the minimum of such functions is not simple as they may have no or multiple minima. Furthermore, use of such criteria can often lead to over-fitting and clinical implausible associations. The choice of power functions for FP and the choice of smoothing parameters for smoothing splines in GAM should therefore always be accompanied by a graphical verification of functional associations with outcome to verify clinical plausibility. Splines can also have unpredictable behaviour at extremes of variable ranges where data is sparse. This will have little influence on confounding control as estimates are very imprecise in these areas but can be disarming when presenting functional relationships graphically. Accompanying fitted functions with confidence bounds will convey information on precision. Like a single linear term, FPs and smoothing splines are sensitive to extreme observations. Therefore, performing influence analysis and checking residuals is important when using these methods. Finally, due to the inherent complexity of non-parametric functions, GAM with smoothing splines cannot be written explicitly and are therefore difficult to reproduce. Models based on FP may therefore represent the most advantageous modeling strategy as they are comparatively intuitive and can be fully expressed in written form.

## Strengths and limitations

To date, studies that have compared the impact of different modeling strategies on model fit or confounding control been based on single group comparisons [17] or on simulated data that are likely to convey 'unrealistic situations' [5,6,19] and are often limited to quantitative variables with standard normal distributions and null associations between exposure and outcome [6,10,11]. The current study was based four widely-used modeling strategies, on real data, on several covariates with differing probability distributions from normal to extremely skewed, on exposure with varying associations to outcome, and on multiple comparisons of exposure groups.

Certain limitations should be considered however, when interpreting the results. First, the study was based on data collected in one clinical context and only evaluated a dichotomous outcome modelled on a logit scale (logistic model). Results may therefore not generalize to other clinical situations or other model types. However, we believe the use of several risk factors with differing probability distributions, and multiple exposure comparisons add to the external validity of our findings.

Second, other modeling strategies for quantitative independent variables could have been evaluated. For example, higher powers for FPs, different smoothing parameters for GAM, or alternative cut-points for dummy variables could have been used. In addition, nonlinear transformations of quantitative variables could have been evaluated including cusums [20] or negative exponential functions [3]. However, considering the large number of possible combinations, we restricted comparisons to commonly used modeling strategies and cut-points that are widely used in clinical research.

## Conclusions

Results of this study indicate that when quantitative confounders have a non-linear association with outcome, cubic smoothing splines in GAM, FP, or dummy variables on at least 4 categories may offer better control for confounding than single linear terms or dummy variables on less than 4 categories, providing they are implemented with care.

## Funding

Canadian Institutes of Health Research (post-doctoral research award); Fonds de la Recherche en Santé du Québec (research career award); Canadian Health Services Research Foundation (grant # RC2-1460-05), Fonds de la Recherche en Santé du Québec (grant # 015102).

#### References

- Harrell FE, Lee KL, Mark DB (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 15:361-387.
- Rothman KJ, Greenland S, Lash TL(2008) Modern Epidemiology. Philadelphia: Lippincott, Williams, & Wilkins.
- Royston P, Sauerbrei W, Altman DG (2000) Modeling the effects of continuous risk factors. J Clin Epidemiol 53:219-221.
- Greenland S (1995) Avoiding power loss associated with categorization and ordinal scores in dose-response and trend analysis. Epidemiology 6:450-454.
- Greenland S (1995) Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. Epidemiology 6:356-365.
- Brenner H, Blettner M (1997) Controlling for continuous confounders in epidemiologic research. Epidemiology 8:429-434.
- Rothman KJ, Greenland S, Lash TL (2008) Modern Epidemiology. Philadelphia: Lippincott, Williams & Wilkins.
- Royston P, Ambler G, Sauerbrei W (1999) The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol 28:964-974.
- Benedetti A, Abrahamowicz M (2004) Using generalized additive models to reduce residual confounding. Stat Med 23:3781-3801.
- Austin PC, Brunner LJ (2004) Inflation of the type I error rate when a continuous confounding variable is categorized in logistic regression analyses. Stat Med 23:1159-1178.
- Slama R, Werwatz A (2005) Controlling for continuous confounding factors: non- and semiparametric approaches. Rev Epidemiol Sante Publique 53:2S65-80.
- Hastie TJ, Tibshirani H, Friedman JH (1990) Generalized additive models. London: Chapman and Hall.
- Baker SP, O'Neill B (1976) The injury severity score: an update. J Trauma 16: 882-885.
- 14. Jennett B, Teasdale G (1977) Aspects of coma after severe head injury. Lancet 1:878-881.
- DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837-845.
- Collins LM, Schafer JL, Kam CM (2001) A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychol Methods 6:330-351.
- Boucher KM, Slattery ML, Berry TD et al. (1998) Statistical methods in epidemiology: a comparison of statistical methods to analyze dose-response and trend analysis in epidemiologic studies. J Clin Epidemiol 51:1223-1233.
- Cochran WG (1968) The effectiveness of adjustment by subclassification in removing bias in observational studies. Biometrics 24:295-313.
- Maldonado G, Greenland S (1997) The importance of critically interpreting simulation studies. Epidemiology 8:453-456.
- 20. Royston P (1992) The use of cusums and other techniques in modelling continuous covariates in logistic regression. Stat Med 11:1115-1129.