

Accounting for measurement error in a two-step procedure using multivariate longitudinal models.

Hendriek C. Boshuizen¹, Mariapaola Lanti², Alessandro Menotti², Joanna Moschandreas³, Hanna Tolonen⁴, Aulikki Nissinen⁴, Srecko Nedeljkovic⁵, Antony Kafatos³, Daan Kromhout¹
¹ National Institute of Public Health and the Environment, Bilthoven, the Netherlands
² Association for Cardiac Research - Associazione per la Ricerca Cardiologica, Rome, Italy
³ Department of Social Medicine, Medical School, University of Crete, Heraklion, Crete, Greece
⁴ Department of Health Promotion and Chronic Disease Prevention, National Public Health Institute, Helsinki, Finland
⁵ Institute for Cardiovascular Diseases, Clinical Center of Serbia, Beograd, Serbia-Montenegro

Example

What is the effect of past changes in blood pressure and cholesterol on mortality due to coronary heart disease (CHD) and stroke in subjects 65 years and older.

For this example individual changes in exposure need to be estimated with a longitudinal mixed model. In that way change is not distorted by regression to the mean, as is the case in the commonly done analyses where crude differences are used between past en baseline values.

First step model:

$$X_{ij} = \beta_0 + \beta_{0i} + \beta_1 a + \beta_2 a^2 + \beta_3 a^3 + \beta_4 y + \beta_5 y^2 + \beta_6 y^3 + \epsilon_i$$

$$\epsilon_i = N(0, \Sigma_\epsilon)$$

$$\begin{pmatrix} \beta_{0i} \\ \beta_{1i} \end{pmatrix} = N(0, \Sigma_\beta)$$

where:
 X_{ij} is the vector of outcomes (SBP and cholesterol) in round j for individual i,
 a is age -50 in measurement round j,
 y is calendar year -70 in measurement round j,
 β_s are vectors of regression coefficients
 Σ_ϵ is a 2x2 covariance matrix with unequal diagonal elements and zero off-diagonal elements
 Σ_β is the (unstructured) 2x2 covariance matrix of the random effects.

Second step model:

Cox Proportional Hazards model as before
 Exposure variable: EBLUP value at current age and the amount of change of exposure over time (β_{11})

Data
 As before



Results

Table 2: Relative risks (95% confidence interval) on Coronary Heart Disease (CHD) and stroke mortality in those over 65 years of age when simultaneously entering current exposure and change in exposure in the model. Seven Countries Study 35-year follow-up (1958-1999).

	# CHD/stroke in analysis*	Per 20 mmHg SBP		Per mmol/l serum cholesterol	
		Change (RR per 10 years)	RR at current age	Change (RR per 10 years)	RR at current age
Coronary Heart Disease Mortality					
- age and cohort	863	0.33 (0.19,0.60)	1.70 (1.53,1.90)	0.50 (0.34,0.73)	1.31 (1.21,1.41)
- fully adjusted	738	0.53 (0.26,1.07)	1.56 (1.37,1.77)	0.73 (0.46,1.17)	1.23 (1.13,1.34)
Stroke Mortality					
- age and cohort	500	0.32 (0.15,0.68)	1.84 (1.60,2.12)	0.31 (0.17,0.54)	1.17 (1.06,1.31)
- fully adjusted	445	0.31 (0.13,0.74)	1.93 (1.63,2.29)	0.57 (0.30,1.14)	1.17 (1.04,1.32)

* Only events after age 65 are included. Also, analyses using exposure 25 years ago use only events from year 25-35
[†] The RR for SBP declines with age due to the presence of an interaction term between age and SBP. Here the RR at age 75 is presented

Conclusions example

- past SBP is more predictive of stroke mortality than more recent SBP
- a similar trend is seen for SBP and CHD mortality, and for cholesterol
- method of Tsiatis is a flexible and feasible method that is more valid than the usually applied methods to analyze health effects of changes in risk factors

Objective

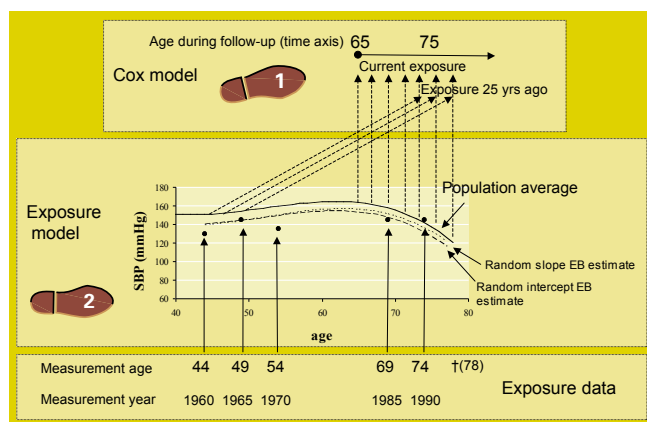
Tsiatis et al. (1995) proposed a relatively simple two-step procedure to adjust regression coefficients for measurement error.

Objectives are:

- extend this method to using a multivariate outcome in the first step of this method
- compare results to results using more commonly applied methods (using attenuation factors/matrixes)
- apply the method in a reanalysis of the 35-year follow-up data of nine cohorts of the Seven Country Study.

Tsiatis AA, DeGruttola V, Wulfsohn MS. Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. JASA 1995;90:27-37.

2-Step method of Tsiatis et al. (1995)



- Step 1:** model individual exposure trajectories using a mixed model for longitudinal data
- Step 2:** use predicted values (EBLUP) from these models in the disease model to estimate relative risks adjusted for measurement error. Predicted values can for instance be: current exposure values, values 25 years ago, average values over a particular period or (with random slope models) change of exposure.

We calculated confidence intervals used an approximate bootstrap method. Tsiatis et al. prescribes refitting the exposure model for each risk set separately. We observed like Andersen and Liestol (2003) little difference in results with fitting the step 1 model only to the entire dataset.

Andersen PK, Liestol K. Attenuation caused by infrequently updated covariates in survival analysis. Biostatistics 2003;4:633-49

Quick bootstrapped confidence intervals

Study subjects were resampled 10 times and 10 exposure models were fitted on these data. EB estimates were calculated from these models for the subjects each the bootstrap sample. For subjects not in a particular bootstrap sample, only the fixed part of the model was available and the random part was randomly drawn from the exposure models on one of the other bootstrap samples. Using these 10 sets of EB estimates, we fitted 10 Cox models. Assuming independence of the errors in the EB estimates and the errors in the coefficients of a single Cox model (β_j), the standard error (SE) of average coefficient of the 10 Cox models (β) can be estimated as:

$$SE_{\beta}^2 = \frac{1}{n} \sum_i SE_{\beta_i}^2 + \left(1 + \frac{1}{n}\right) \frac{1}{(n-1)} \sum_i (\beta_i - \frac{1}{n} \sum_i \beta_i)^2$$

where n is the number of bootstrap samples. The first term in the formula reflects the uncertainty due to the limited information on mortality (<800 cases) in the Cox model and the second term the uncertainty due to the exposure model (approximately 19000 SBP and cholesterol measurements). As the latter is much smaller than the first, 10 bootstrap samples were ample to estimate SE_{β} .

Comparison of methods

We first compared the method with those obtained using a more commonly applied method of adjustment for regression dilution bias, i.e. the use of an attenuation factors or matrixes.

First step model:

$$X_{ij} = \beta_0 + \beta_{0i} + \beta_1 a + \beta_2 a^2 + \beta_3 a^3 + \beta_4 y + \beta_5 y^2 + \beta_6 y^3 + \epsilon_i$$

$$\epsilon_i = N(0, \Sigma_\epsilon)$$

$$\beta_{0i} = N(0, \Sigma_\beta)$$

where:

X_{ij} is the vector of outcomes (SBP and cholesterol) in round j for individual i,
 a is age -50 in measurement round j,
 y is calendar year -70 in measurement round j,
 β_s are vectors of regression coefficients
 Σ_ϵ is a 2x2 covariance matrix with unequal diagonal elements and zero off-diagonal elements
 Σ_β is the (unstructured) 2x2 covariance matrix of the random effects.

This covariance structure implies that the true values of SBP and cholesterol are correlated, but their measurement errors are independent.

Second step model:

Cox Proportional Hazards model with age as the timescale
 Age of entry: 65 years
 Exposure variable: average EBLUP value over the period baseline-current age (i.e. the age of the risk set)
 All analyses were stratified by cohort to prevent confounding by ecological effects
 All analyses included interaction terms with age when significant.

Method for comparison

- Cox regression using baseline measurements (year 0);
- Regression coefficients from that model, adjusted using an attenuation matrix or attenuation factor based on the measurements in year 0 and year 10.

Data

5238 men in 9 European cohorts of the Seven Country studies (From Finland, Netherlands, Italy, Serbia and Greece). For the first step model all data is used. For the second step model only the follow-up after age 65 is used (average follow-up after age 65 = 12.3 years)

Results

Table 1: Relative risks (95% confidence interval) of systolic blood pressure (SBP) and serum total cholesterol on coronary (CHD) and stroke mortality: a comparison of unadjusted results with results adjusted with different methods for regression dilution bias. Seven Countries Study 35-year follow-up (1958-1999).

	Mortality from Coronary Heart Disease			Mortality from stroke		
	#CHD in analysis*	Per 20 mm Hg SBP at age 75 [†]	Per 1 mmol/l serum total cholesterol	#stroke in analysis*	Per 20 mm Hg SBP at age 75 [†]	Per 1 mmol/l serum total cholesterol
Unadjusted baseline measurement						
- age and cohort	896/886 [‡]	1.32 (1.23,1.41)	1.23 (1.16,1.30)	526/515 [‡]	1.43 (1.32,1.55)	1.12 (1.03,1.21)
- adjusted	747	1.27 (1.18,1.37)	1.18 (1.10,1.25)	450	1.49 (1.37,1.63)	1.08 (0.99,1.18)
- fully adjusted [§]						
Attenuation factor						
- age and cohort	896/886 [‡]	1.83 (1.48,2.28)	1.39 (1.27,1.53)	526/515 [‡]	2.41 (1.84,3.16)	1.20 (1.05,1.37)
- adjusted						
- fully adjusted [§]						
Attenuation matrix						
- fully adjusted [§]	747	1.56 (1.35,1.84)	1.26 (1.14,1.40)	450	2.16 (1.79,2.60)	1.06 (0.93,1.21)
- adjusted						
- fully adjusted [§]						
Method Tsiatis et al. 1995						
- age and cohort	891	1.65 (1.50,1.81)	1.33 (1.23,1.44)	526	1.80 (1.59,2.02)	1.12 (1.00,1.24)
- adjusted						
- fully adjusted [§]	755	1.56 (1.40,1.73)	1.24 (1.14,1.35)	461	1.91 (1.68,2.18)	1.10 (0.98,1.24)

* Only events after age 65 are included
[†] The RR for SBP declines with age due to the presence of an interaction term between age and SBP. RR at age 75 is presented
[‡] The first number applies to SBP, the second to cholesterol
[§] Adjusted for age, BMI, smoking, diabetes, use of anti-hypertensive drugs (all as last measured), SBP or cholesterol, CVD before age 65, family history of CVD (at baseline) and stratified for cohort.

Conclusion on comparison of methods

- Results of both methods were similar, but the two-step method yielded slightly narrower confidence intervals, especially with sparse data
- Univariate method can lead to overestimation of effects
- The two-step method allows answering of more complex questions than the simple method using an attenuation matrix/factor (see example).

P.O. Box 1, 3720 BA Bilthoven, the Netherlands
 www.rivm.nl

hendriek.boshuizen@rivm.nl



National Institute for Public Health and the Environment

Research for man and environment