Author(s): Reinikainen, Jaakko; Laatikainen, Tiina; Karvanen, Juha; Tolonen, Hanna

Title: Lifetime cumulative risk factors predict cardiovascular disease mortality in a 50-year follow-up study in Finland

Year: 2015

Version:

Please cite the original version:

Reinikainen, J., Laatikainen, T., Karvanen, J., & Tolonen, H. (2015). Lifetime cumulative risk factors predict cardiovascular disease mortality in a 50-year follow-up study in Finland. International Journal of Epidemiology, 44(1), 108-116. https://doi.org/10.1093/ije/dyu235

Lifetime cumulative risk factors predict cardiovascular disease mortality in a 50-year follow-up study

Jaakko Reinikainen[1], Tiina Laatikainen[2,3,4], Juha Karvanen[1] and Hanna Tolonen[2]

[1] Department of Mathematics and Statistics, University of Jyvaskyla, Jyväskylä, Finland

[2] Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland

[3] Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland

[4] Hospital District of North Karelia, Joensuu, Finland

Corresponding author:

Jaakko Reinikainen

Department of Mathematics and Statistics, University of Jyväskylä

P.O. Box 35 (MaD), FI-40014 University of Jyväskylä, Finland

Email: jaakko.o.reinikainen@jyu.

Word Count:

Abstract: 247 words, Article: 2856 words

Summary

Background Systolic blood pressure, total cholesterol and smoking are known predictors of cardiovascular disease (CVD) mortality. Less is known about the effect of lifetime accumulation and changes of risk factors over time as predictors of CVD mortality, especially in very long follow-up studies.

Methods Data from the Finnish cohorts of the Seven Countries Study were used. The baseline examination was in 1959 and seven re-examinations were carried out approximately in five-year intervals. Cohorts were followed up for mortality until the end of 2011. Time-dependent Cox models with regular time-updated risk factors, time-dependent averages of risk factors and latest changes in risk factors, using smoothing splines to discover nonlinear effects were used to analyse the predictive effect of risk factors for CVD mortality.

Results A model using cumulative risk factors, modelled as the individual-level averages of several risk factor measurements over time, predicted CVD mortality better than a model using the most recent measurement information. This difference seemed to be most prominent for systolic blood pressure. U-shaped effects of the original predictors can be explained by partitioning a risk factor effect between the recent level and the change trajectory. The change in body mass index predicted the risk although body mass index itself did not.

Conclusions The lifetime accumulation of risk factors and the observed changes in risk factor levels over time are strong predictors of CVD mortality. It is important to investigate different ways of using the longitudinal risk factor measurements to take full advantage of them.

Keywords Cardiovascular diseases, risk factors, longitudinal study, mortality

Key Messages

- Both long-term exposure to increased risk factors on individual-level and changes in classical risk factors predict the risk of cardiovascular disease mortality.
- Especially the cumulative value of systolic blood pressure is a stronger predictor than the most recent value.
- Simplistic use of longitudinal risk factor measurements in modelling may underestimate the importance of them.

Introduction

Cardiovascular disease (CVD) risk factors have been studied widely, but there is still a demand for advice on the utilization of longitudinal risk factor measurements in risk prediction. Most cohorts include only baseline measurements with follow-up of outcomes using registers or other means of data collection. Thus using only the baseline measurements has been the primary approach in building risk prediction models (**1, 2**) resulting in challenges with unknown risk factor changes in long follow-ups and inverse associations immediately prior to death because of severe diseases. Less attention is so far paid to taking advantage of longitudinal risk factor measurements.

Recently, research has been done on utilizing longitudinal risk factor measurements in risk prediction. The effect of visit-to-visit variability in blood pressure has been shown to be associated with all-cause mortality and stroke (**3, 4**). In connection with a patient's hospital death, it has been found out that using time-dependent risk factors improves the predictive ability of a model compared to a time-fixed (time-independent) model (**5**). Longitudinal measurements have also been used to describe the relation of population-level changes in risk factors to the risk of coronary death (**6**).

There have been some earlier efforts to utilize longitudinal measurements by calculating individual-level changes and averages of risk factors (**7-9**). Nevertheless, using only linear effects of predictors may prevent researchers from finding true, possibly nonlinear, effects of risk factors. Categorisation of continuous variables has sometimes been used in order to overcome this problem (**10**), but it may be an inefficient way in finding nonlinearities and therefore other methods, e.g., splines, should be considered (**11**).

Our aim is to predict CVD mortality using longitudinal risk factor measurements and two individual-level variables derived from them, namely, changes between the latest two measurements and time-

dependent averages. We also compare the predictive ability of models that use longitudinal measurement information differently. Data from the Finnish cohorts of the Seven Countries Study are used in the analyses. These data suit our purposes well, because the cohorts have seven re-examinations carried out approximately in five-year intervals after the baseline measurements in 1959 and are followed up for over 50 years.

Methods

Cohorts

The analyses were conducted using the Finnish cohorts of the Seven Countries Study. These cohorts consisted of all men who were born between 1900 and 1919 in two geographically defined rural areas located in Eastern and South-Western Finland (**12**). The baseline survey was conducted in 1959 and the re-examinations in 1964, 1969, 1974, 1984, 1989, 1994 and 1999. Information on individuals' lifetime and the cause of death is available until the end of 2011.

These two cohorts consist of 1711 men aged 40 to 59 at the baseline. Characteristics of the cohorts are presented in detail in Table 1. At the end of the follow-up period, 16 men were still alive and 850 had died from cardiovascular disease (CVD). Mortality data was obtained from the National Causes of Death Register through record linkage. CVD deaths were defined as ICD-8 codes 390-458, ICD-9 codes 390-459 and ICD-10 codes I00-I99. The median follow-up time from baseline to death or the end of the follow-up was 23.1 years. For the purposes of our study, we restricted the analyses to men who were examined at least twice, which resulted in data with 1540 individuals.

(Table 1 here)

Statistical analysis

In the modelling of the follow-up data with longitudinal risk factor measurements, we applied time-dependent Cox models (**13**), where we used age as the time-scale (**14, 15**). For continuous predictor variables, smoothing splines (**16**) were used to identify possibly nonlinear effects. Four degrees of freedom (three with change in cholesterol) were used for splines to control the amount of smoothing. The proportional hazards assumption of the Cox models was checked using Schoenfeld residuals (**16**).

As predictor variables we used classical chronic disease risk factors and two variables derived from them: 1) latest change and 2) time-dependent average. Assume we have a risk factor x and longitudinal measurements on it at time points $t(i)$, $i = 0, 1, 2, \ldots$. In a regular time-dependent Cox model, the baseline measurement $x_{t(0)}$ is used in the time interval $(t(0), t(1)]$, the second measurement $x_{t(1)}$ is used in the interval $(t(1), t(2)]$, and so on.

The latest change was calculated on individual-level as the difference of the latest two measurements. That is, the value $x_{t(i)} - x_{t(i-1)}$ is used in the interval $(t(i), t(i+1)]$. The time-dependent average was calculated on individual-level as a mean of the most recent and all previous measurements. In other words, the value mean($x_{t(i)}, x_{t(i-1)}, \ldots, x_{t(0)}$) is used in the interval $(t(i), t(i+1)]$. Both of these derived variables are time-dependent.

CVD death was the end-point in all models. Four different approaches to model the longitudinal risk factor information were considered:

(a) Traditional time-dependent model in which the risk factors are updated approximately every five years to use only the most recent risk factor measurements.
(b) Model (a) with individual-level risk factor specific latest changes.
(c) Time-dependent model in which the risk factors are updated approximately every five years to use the individual-level risk factor specific averages.
(d) Model (c) with individual-level risk factor specific latest changes.

In each model, systolic blood pressure (SBP), body mass index (BMI, kg/m^2), total cholesterol, resting heart rate, smoking status and physical activity were considered. Smoking status (current smoking) is dichotomous (yes/no), physical activity is categorical with three levels (sedentary or invalid / light work / hard work) and all other variables are continuous. Physical activity, which was measured only in 1959, 1964 and 1969, was used as an adjustment variable rather than a variable of direct interest.

The time-dependent average was also calculated for smoking, which resulted in a continuous variable and is interpreted as the percentage of the follow-up time the individual has been smoking. If a risk factor measurement was missing, the previous observed value was used in the modelling. The participation rate was high in all examinations but declined over time when members of the cohorts grew very old (Table 1).

The model selection for Models (a)-(d) was carried out so that in the beginning all appropriate variables were in the model and smoothing splines were used with all continuous variables. Then, the variable or nonlinear component of a spline with the largest p-value was dropped one by one until each variable had a *p*-value less than 0.05. When a nonlinear component of a spline was dropped, the variable was treated as a regular continuous variable. All analyses were carried out with the R statistical software (**17**). Cox models were fitted using the `coxph` function from the `survival` package (**18**).

When comparing the predictive ability of different models, the continuous version of net reclassification improvement (NRI) (**19**) and integrated discrimination improvement (IDI) (**20**) indices were used. These two indices have become popular alternatives for using the area under the receiver

operating characteristic curve (AUC) since it has been observed that the improvement in AUC between two models is not so good in detecting changes in model discrimination as expected (**21, 22**).

The NRI and IDI indices use predicted event probabilities in their estimation. Here, ten-year predicted probabilities for CVD death were used. They were obtained using the appropriate Cox models, and as the interest was in the prediction of CVD mortality, we excluded the individuals died for other causes during the prediction interval from the calculations of the model comparison indices. The NRI measures the amount of correct reclassification when the predicted risk of CVD death is compared between the 'new' and the 'old' model. A reclassification is considered to be correct if an event obtains a higher predicted risk by the new model than by the old model or if a non-event obtains a lower predicted risk. The NRI is estimated as the proportion of correct minus incorrect reclassifications among events, plus the proportion of correct minus incorrect reclassifications among non-events. The IDI is the difference in discrimination slopes between the two models. The discrimination slope of a model is defined as the average predicted risk of events minus the average risk of non-events. The IDI can also be seen as a difference in average sensitivity minus average (1 - specificity) between the models.

Results

Model (a) is a traditional time-dependent model, used as a starting point in the modelling. Table 2 shows the results for the model and Figure 1 illustrates the effects of continuous variables on hazard. BMI seemed to have no effect on CVD mortality, and is therefore not included in the model. Nonlinear effects of SBP and cholesterol suggest that also low risk factor values would be associated with a high risk of CVD death.

Adding the differences of the latest two measurements as predictors to describe individual-level changes in risk factors, we have Model (b). Now SBP is not a nonlinear predictor anymore (Table 2 and Figures 1 and 2), but the effect can be considered linear. The change variables for SBP, cholesterol, BMI, and heart rate in this model indicate that lowering risk factor levels predicts CVD death. It is worth noting that even though BMI itself does not seem to predict the CVD mortality in Models (a) and (b), the change in BMI clearly predicts the risk.

(Table 2 here)

The third model does not only update the risk factor values when the new measurements become available, but it uses individual-level averages of the most recent and all the previous measurements. These averages can be interpreted to model the cumulative effect of risk factors rather than the effect of the current risk factor level. In this Model (c) the change variables are not used, so this is comparable with Model (a). Model (c) shows that when using the time-dependent averages instead of

the regular time-dependent variables, the effects of SBP and cholesterol can be considered linear (Table 2 and Figure 1), unlike in Model (a). Moreover, especially for SBP, the use of the average seems to make it a stronger predictor.

Model (d) uses the averages of risk factors and change variables. The results of Model (d) are presented in Table 2 and Figures 1 and 2. Again, the change in BMI predicts CVD mortality although BMI itself does not. The difference between these models is that in Model (d) the change in heart rate does not predict the risk.

The statistical significance of a risk factor does not necessarily tell much about its epidemiological relevance. To understand better the importance of the predictors, the rank-hazard plots (**23**) were used to compare average and change variables of SBP and cholesterol from Model (d) (Figure 3). Both average variables seem to have steeper lines than the change variables, so they are stronger predictors of CVD mortality and average SBP is stronger than average cholesterol.

In the change in cholesterol values, the entire range from the lower quartile to the upper quartile has virtually the same risk of CVD mortality, but the values below the lower quartile clearly have an increased risk (Figure 3). The strange bend in the right tail of the change in cholesterol can be ignored as it refers only to a few observations and the related confidence interval is large, which is seen in Figure 2.

The differences of the predictive ability of the presented models (b) and (d) were investigated using NRI and IDI indices. All the estimates of indices are positive and their confidence intervals indicate that Model (d) can separate CVD death events and non-events better than Model (b) (Table 3). Thus, the use of averages of risk factors usually improves the model compared to the use of regular time-dependent risk factors, suggesting that the cumulative effects of the risk factors would predict the risk of CVD mortality better than only the most recent information of the risk factor levels.

(Table 3 here)

Discussion

In this study, we presented different ways to use longitudinal risk factor measurements in modelling the risk of CVD mortality in a long-term follow-up study. A model using the individual-level averages of risk factor measurements, representing the lifetime accumulation of the risk, was shown to have a better predictive ability than a model using only the most recent measurement information. It was found out that the partitioning of risk factor effects between the recent level and the change trajectory can explain the U-shaped effects of the original predictors.

The approach employed in this paper belongs to the field of life course epidemiology (**24, 25**), as we consider accumulation and change trajectories of risk factors rather than single measurements. We note that despite the long follow-up, this study is restricted to adults over 40 years of age. For example, Lynch and Davey Smith (**25**) consider the associations of early-life conditions with stroke and coronary heart disease.

The strengths of this study are eight longitudinal risk factor measurements in regular intervals and the 50-year mortality follow-up. These data provide a good overview of the individual level risk factor profiles and their changes in a lifetime. There are some limitations in the present study related to the data as well. The baseline measurement was carried out in 1959, so all the relevant risk factors known nowadays were not measured. Also the development of causes of death diagnostics and the medical treatment of hypertension and high cholesterol may have some effect on the results. However, the medication effects are reflected to risk factor levels, and the role of risk factors in predicting the changes in mortality is greater than the effects of invasive treatments (**26**). Particularly, if the main interest was to obtain probabilities of CVD mortality, competing risks models could be considered. However, in Finland the autopsy rates are high and the quality of causes of death registering has been shown to be good, especially in relation to cardiovascular diseases (**27-29**). Several other diseases have the same risk factors as CVD. It has been shown that the adjustment for competing causes of death will reduce the risk of CVD mortality, especially of older people and groups with less favourable risk-factor profiles. (**30**)

In an earlier study, using a time-dependent model, BMI did not predict a CVD incidence (**2**). In addition to this, it has been observed that weight decrease is associated with high CVD risk (**31**). In this study a concordant result was observed showing that even though BMI itself does not predict a CVD risk, a change in BMI does. Decrease in BMI in late life is usually related to frailty and morbidity leading to higher mortality (**32, 33**).

In earlier studies from the Framingham Heart Study, systolic blood pressure, total cholesterol and smoking have been shown to be significant predictors of CVD (**2**). Our results also demonstrated the predictive power of these same risk factors in traditional time-dependent models. Using the accumulation of individual risk factors (average over all previous measurements) provided new information about the effect of traditional risk factors, demonstrating the impact of the individual-level history of the risk factor levels on the risk of CVD.

In a time-dependent model the use of a measurement carried out just prior death may result in a misleading effect due to frailty and co-morbidity (**34-36**). An illness can lead to the lowering of, e.g., blood pressure and cholesterol, which causes an apparent negative relation between risk factors and mortality. In addition, if we do not take into account that risk factor levels may vary over time and use only the baseline measurements, we encounter the regression dilution problem (**37**). Different

adjustment methods have been proposed to overcome this problem (**38**). This study demonstrated that when changes in risk factors are used as predictors in a time-dependent model, the spurious associations caused by frailty and co-morbidity can be eliminated.

This study confirms the value of longitudinal risk factor information in risk prediction. The results imply that using a simplistic method in handling longitudinal risk factor measurements in a prediction model may produce misleading estimates. The traditional time-dependent Cox model assumes that only the most recent risk factor level affects the risk. This assumption may be too restrictive and prevent researchers from understanding the true importance of the risk factors. We recommend investigators to study different ways to utilize the longitudinal risk factor information. We found out that, especially with SBP, the long-term individual-level average is a stronger predictor of the risk of CVD death than the most recent measurement. However, it remains a topic for further studies to investigate in more detail with which risk factors it is worth using the long-term average and with which the most recent measurement is appropriate.

Conclusions

The risk of CVD mortality was predicted with individual-level changes and long-term averages of classical risk factors using the data from Finnish cohorts of the Seven Countries Study. A model using long-term individual-level averages of risk factors was shown to have a better predictive ability than a standard time-dependent model using only the most recent measurements. The results indicate that it is important to investigate different ways of using the longitudinal risk factor measurements to take full advantage of them.

Funding

The research of the first author was supported by the Emil Aaltonen Foundation.

Conflict of interest: none declared.

References

1. Conroy RM, Pyorala K, Fitzgerald AP, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003 Jun;**24**(11): 987-1003.
2. Pencina MJ, D'Agostino RB, Sr., Larson MG, Massaro JM, Vasan RS. Predicting the 30-year risk of cardiovascular disease: the Framingham heart study. *Circulation* 2009 Jun 23;**119**(24): 3078-84.
3. Muntner P, Shimbo D, Tonelli M, Reynolds K, Arnett DK, Oparil S. The relationship between visit-to-visit variability in systolic blood pressure and all-cause mortality in the general population: findings from NHANES III, 1988 to 1994. *Hypertension* 2011 Feb;**57**(2): 160-6.

4. Rothwell PM, Howard SC, Dolan E, et al. Prognostic significance of visit-to-visit variability, maximum systolic blood pressure, and episodic hypertension. *Lancet* 2010 Mar 13;**375**(9718): 895-905.
5. Wong J, Taljaard M, Forster AJ, Escobar GJ, van Walraven C. Addition of time-dependent covariates to a survival model significantly improved predictions for daily risk of hospital death. *J Eval Clin Pract* 2012 Apr;**19**(2): 351-7.
6. Menotti A, Lanti M, Kromhout D, et al. Forty-year coronary mortality trends and changes in major risk factors in the first 10 years of follow-up in the seven countries study. *Eur J Epidemiol* 2007;**22**(11): 747-54.
7. Farchi G, Capocaccia R, Verdecchia A, Menotti A, Keys A. Risk factors changes and coronary heart disease in an observational study. *Int J Epidemiol* 1981 Mar;**10**(1): 31-40.
8. Kahn HA, Dawber TR. The development of coronary heart disease in relation to sequential biennial measures of cholesterol in the Framingham study. *J Chronic Dis* 1966 May;**19**(5): 611-20.
9. Wilson PW, Hoeg JM, D'Agostino RB, et al. Cumulative effects of high cholesterol levels, high blood pressure, and cigarette smoking on carotid stenosis. *N Engl J Med* 1997 Aug 21;**337**(8): 516-22.
10. Sesso HD, Stampfer MJ, Rosner B, Gaziano JM, Hennekens CH. Two-year changes in blood pressure and subsequent risk of cardiovascular disease in men. *Circulation* 2000 Jul 18;**102**(3): 307-12.
11. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Statistics in Medicine* 2006 Jan 15;**25**(1): 127-41.
12. Karvonen MJ, Blonqvist G, Kallio V, et al. C4. Men in Rural East and West Finland. *Acta Medicine Scandinavica* 1966;**180**(460): 169-90.
13. Fisher LD, Lin DY. Time-dependent covariates in the Cox proportional-hazards regression model. *Annu Rev Public Health* 1999;**20**: 145-57.
14. Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. *Am J Epidemiol* 1997 Jan 1;**145**(1): 72-80.
15. Thiebaut ACM, Benichou J. Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study. *Statistics in Medicine* 2004 Dec 30;**23**(24): 3803-20.
16. Therneau TM, Grambsch PM. *Modeling survival data : extending the Cox model*. New York: Springer; 2000.
17. R Core Team. R: A language and environment for statistical computing. R Foundation for Statatistical Computing, Vienna, Austria.; 2013. URL:http://www.R-project.org.
18. Therneau T. A Package for Survival Analysis in S. R package version 2.37-4. 2013. URL:http://CRAN.R-project.org/package=survival.
19. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011 Jan 15;**30**(1): 11-21.
20. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008 Jan 30;**27**(2): 157-72; discussion 207-12.
21. Spitz MR, Amos CI, D'Amelio A, Jr., Dong Q, Etzel C. Re: Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. *J Natl Cancer Inst* 2009 Dec 16;**101**(24): 1731-2; author reply 2.
22. Wang TJ, Gona P, Larson MG, et al. Multiple biomarkers for the prediction of first major cardiovascular events and death. *N Engl J Med* 2006 Dec 21;**355**(25): 2631-9.
23. Karvanen J, Harrell FE. Visualizing covariates in proportional hazards model. *Statistics in Medicine* 2009 Jun 30;**28**(14): 1957-66.
24. Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int J Epidemiol* 2002 Apr;**31**(2): 285-93.

25. Lynch J, Davey Smith G. A life course approach to chronic disease epidemiology. *Annu Rev Public Health* 2005;**26**: 1-35.
26. Laatikainen T, Critchley J, Vartiainen E, Salomaa V, Ketonen M, Capewell S. Explaining the decline in coronary heart disease mortality in Finland between 1982 and 1997. *Am J Epidemiol* 2005 Oct 15;**162**(8): 764-73.
27. Mahonen M, Salomaa V, Torppa J, et al. The validity of the routine mortality statistics on coronary heart disease in Finland: comparison with the FINMONICA MI register data for the years 1983-1992. Finnish multinational MONItoring of trends and determinants in CArdiovascular disease. *J Clin Epidemiol* 1999 Feb;**52**(2): 157-66.
28. Pajunen P, Koukkunen H, Ketonen M, et al. The validity of the Finnish Hospital Discharge Register and Causes of Death Register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil* 2005 Apr;**12**(2): 132-7.
29. Tolonen H, Salomaa V, Torppa J, Sivenius J, Immonen-Raiha P, Lehtonen A. The validation of the Finnish Hospital Discharge Register and Causes of Death Register data on stroke diagnoses. *Eur J Cardiovasc Prev Rehabil* 2007 Jun;**14**(3): 380-5.
30. Berry JD, Dyer A, Cai X, et al. Lifetime risks of cardiovascular disease. *N Engl J Med* 2012 Jan 26;**366**(4): 321-9.
31. Strandberg TE, Strandberg AY, Salomaa VV, et al. Explaining the obesity paradox: cardiovascular risk, weight change, and mortality during long-term follow-up in men. *Eur Heart J* 2009 Jul;**30**(14): 1720-7.
32. Corrada MM, Kawas CH, Mozaffar F, Paganini-Hill A. Association of body mass index and weight change with all-cause mortality in the elderly. *Am J Epidemiol* 2006 May 15;**163**(10): 938-49.
33. Dahl AK, Reynolds CA, Fall T, Magnusson PK, Pedersen NL. Multifactorial analysis of changes in body mass index across the adult life course: a study with 65 years of follow-up. *Int J Obes (Lond)* 2014 Aug;**38**(8): 1133-41.
34. Glynn RJ, Field TS, Rosner B, Hebert PR, Taylor JO, Hennekens CH. Evidence for a positive linear relation between blood pressure and mortality in elderly people. *Lancet* 1995 Apr 1;**345**(8953): 825-9.
35. Schatz IJ, Masaki K, Yano K, Chen R, Rodriguez BL, Curb JD. Cholesterol and all-cause mortality in elderly people from the Honolulu Heart Program: a cohort study. *Lancet* 2001 Aug 4;**358**(9279): 351-5.
36. Tilvis RS, Valvanne JN, Strandberg TE, Miettinen TA. Prognostic significance of serum cholesterol, lathosterol, and sitosterol in old age; a 17-year population study. *Ann Med* 2011 Jun;**43**(4): 292-301.
37. Clarke R, Shipley M, Lewington S, et al. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol* 1999 Aug 15;**150**(4): 341-53.
38. Frost C, Thompson SG. Correcting for regression dilution bias: comparison of methods for a single predictor variable. *Journal of the Royal Statistical Society Series a-Statistics in Society* 2000;**163**: 173-89.

Table 1 Characteristics of the follow-up data by the examination years[a]

	Year							
	1959	1964	1969	1974	1984	1989	1994	1999
Participation rate for individuals alive (%)	98	97	97	96	92	86	87	68
Age for individuals alive	49.8 (5.5)	54.7 (5.5)	59.4 (5.5)	64.0 (5.4)	73.0 (5.1)	76.9 (4.7)	80.9 (4.1)	84.1 (3.9)
Age for participants	49.9 (5.5)	54.7 (5.5)	59.4 (5.5)	64.0 (5.4)	73.0 (5.1)	76.8 (4.8)	80.9 (4.2)	83.5 (3.5)
Cumulative all-cause mortality	0	117	282	486	945	1186	1393	1522
Cumulative CVD mortality	0	41	110	209	480	616	713	777
SBP (mmHg)	143.8 (20.7)	139.4 (21.3)	147.6 (23.6)	151.3 (22.3)	154.0 (22.7)	155.1 (22.9)	151.1 (21.0)	142.2 (21.2)
BMI (kg/m^2)	23.7 (3.2)	24.3 (3.6)	24.8 (3.9)	25.0 (3.8)	25.7 (4.1)	26.3 (3.9)	26.4 (4.1)	26.2 (3.3)
Total cholesterol (mmol/l)	6.7 (1.3)	6.8 (1.2)	6.9 (1.3)	6.6 (1.2)	6.1 (1.2)	5.7 (1.1)	5.5 (1.0)	5.7 (1.0)
Heart rate	67.7 (13.0)	68.5 (12.6)	67.3 (13.0)	70.7 (14.3)	67.9 (12.1)	68.3 (12.3)	70.0 (14.8)	69.7 (13.3)
Current smoker (%)	63	54	49	36	18	13	11	2
Physical activity (%) (sedentary or invalid / light work / hard work)	10/16/74	9/15/76	2/38/61	-	-	-	-	-

[a] Values for continuous variables are mean (SD)

Table 2 Model (a): time-dependent model using only the most recent risk factor measurements, Model (b): time-dependent model using only the most recent risk factor measurements and the changes in continuous variables, Model (c): time-dependent model using the averages of the risk factor measurements and Model (d): time-dependent model using the averages of the risk factor measurements and the changes in continuous variables

	Model (a)			**Model (b)**			**Model (c)**			**Model (d)**		
Risk factor	Hazard ratio	95% CI	*p*-value	Hazard ratio	95% CI	*p*-value	Hazard ratio	95% CI	*p*-value	Hazard ratio	95% CI	*p*-value
SBP (per 10 mmHg), linear part of the spline[a]	1.07	(1.04, 1.10)	< 0.01	1.11	(1.08, 1.16)	< 0.01	1.19	(1.15, 1.24)	< 0.01	1.18	(1.13, 1.23)	< 0.01
SBP (per 10 mmHg), nonlinear part of the spline			0.03									
Chol. (mmol/l), linear part of the spline[a]	1.10	(1.04, 1.16)	< 0.01	1.14	(1.07, 1.21)	< 0.01	1.16	(1.09, 1.24)	< 0.01	1.15	(1.08, 1.24)	< 0.01
Chol. (mmol/l), nonlinear part of the spline			0.01			0.02						
Heart rate	1.01	(1.00, 1.01)	< 0.01	1.01	(1.01, 1.02)	< 0.01	1.01	(1.00, 1.02)	0.01	1.01	(1.00, 1.02)	0.01
Smoking, yes	1.42	(1.21, 1.67)	< 0.01	1.44	(1.22, 1.69)	< 0.01	1.59	(1.33, 1.90)	< 0.01	1.59	(1.33, 1.90)	< 0.01
Physical activity (sedentary, invalid)[b]	1.81	(1.18, 2.79)	0.01	1.76	(1.14, 2.72)	0.01	1.69	(1.10, 2.60)	0.02	1.67	(1.08, 2.57)	0.02
Physical activity (light work)[b]	1.31	(1.13, 1.53)	< 0.01	1.29	(1.11, 1.51)	< 0.01	1.26	(1.08, 1.47)	< 0.01	1.25	(1.08, 1.46)	< 0.01
Change in SBP (per 10 mmHg), linear part of the spline				0.89	(0.86, 0.93)	< 0.01				0.96	(0.93, 0.99)	0.01
Change in SBP (per 10 mmHg), nonlinear part of the spline						0.01						0.05
Change in BMI				0.96	(0.92, 0.99)	0.02				0.96	(0.92, 0.99)	0.02
Change in chol. (mmol/l), linear part of the spline[a]				0.91	(0.84, 0.98)	0.01				0.98	(0.91, 1.05)	0.55
Change in chol. (mmol/l), nonlinear part of the spline												0.04
Change in heart rate				0.99	(0.99, 1.00)	0.03						

[a] If the model does not use a spline for this variable, these are the results for the usual linear predictor.

[b] Reference level for physical activity is heavy work.

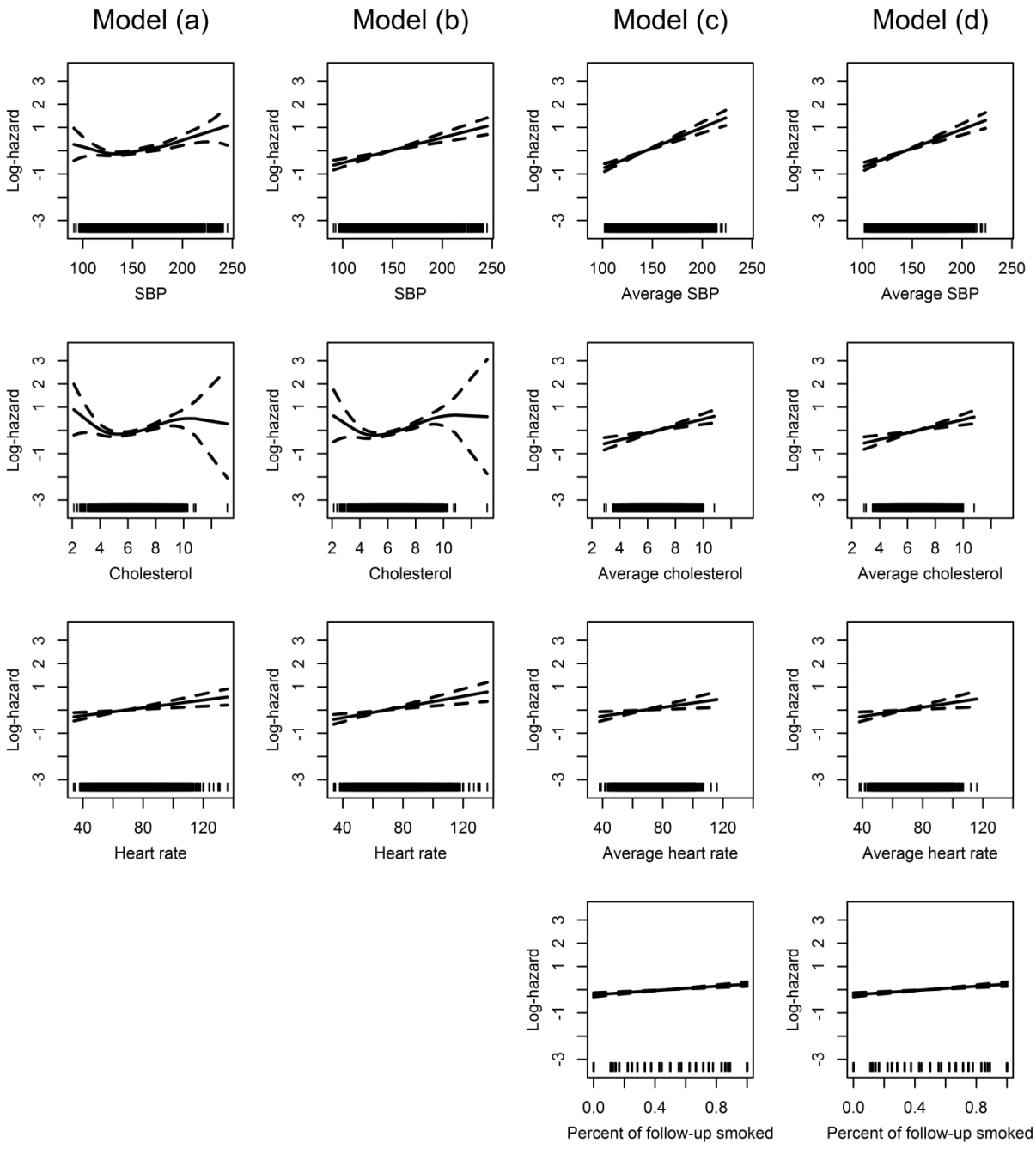

Figure 1 Models (a) - (d): log-hazard ratios and 95% confidence intervals for continuous variables except for the change variables, which are presented in Figure 2. Small ticks at the bottom of each panel represent the observed values and help to identify outliers causing wide confidence intervals.

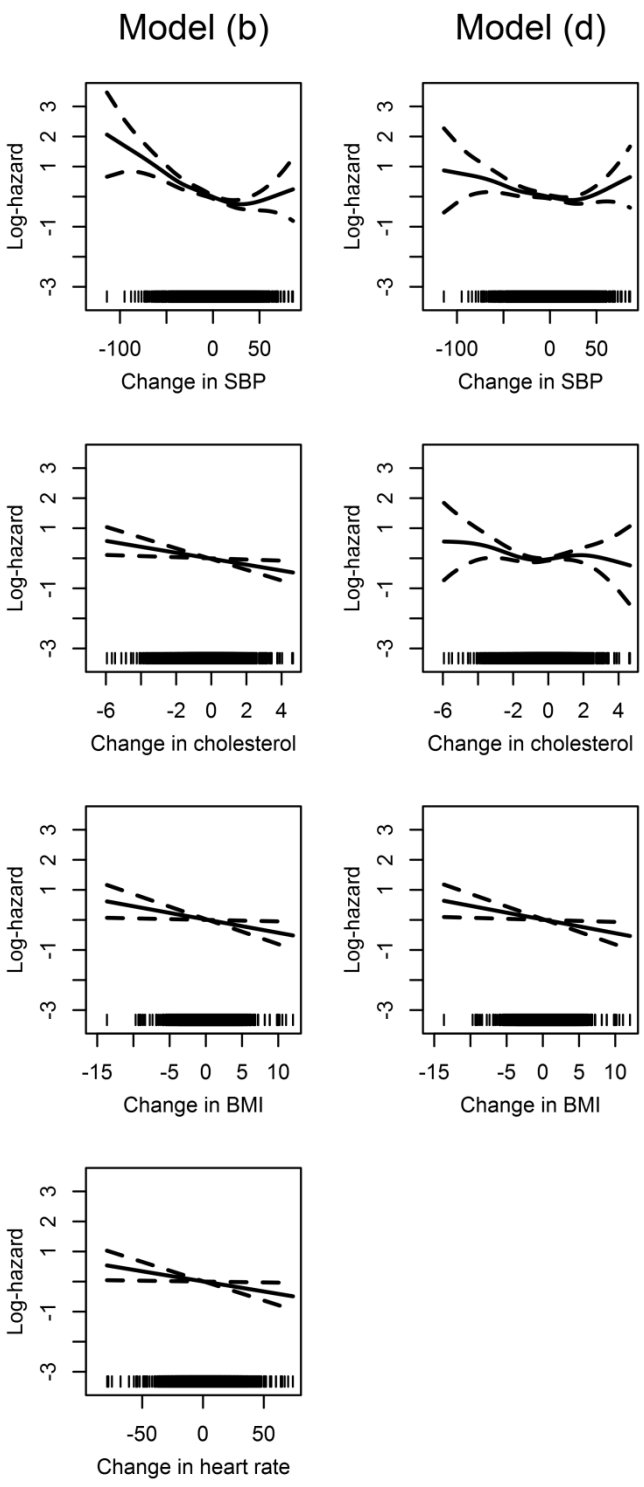

Figure 2 Models (b) and (d): log-hazard ratios and 95% confidence intervals for change variables. Small ticks at the bottom of each panel represent the observed values and help to identify outliers causing wide confidence intervals.

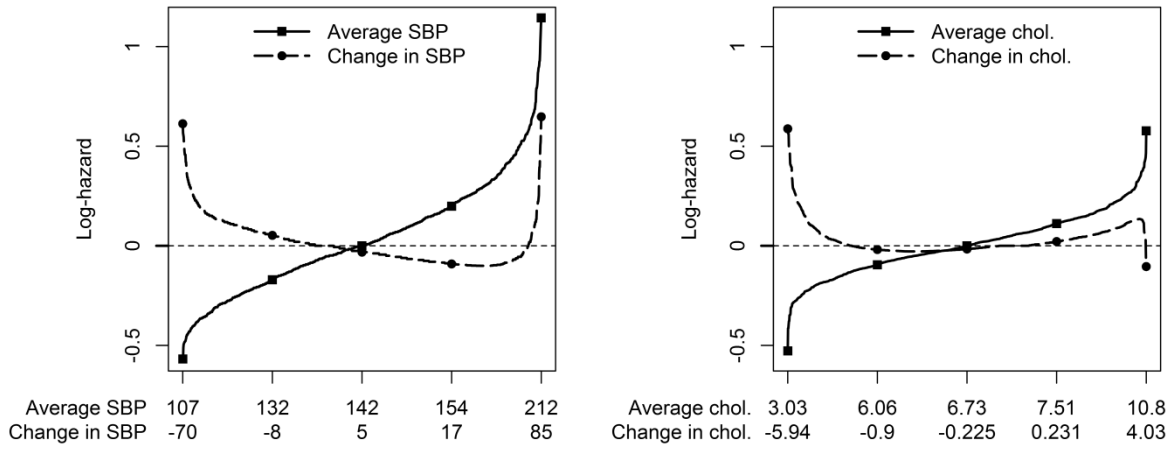

Figure 1 Rank-hazard plots showing the relative importance of average and change variables of SBP and cholesterol in Model (d). The values of each predictor are scaled evenly on the horizontal axis and the values of minimum, first quartile, median, third quartile and maximum are presented. Plots are created using the measurement information of the year 1974 (15 years after the baseline). Reference hazards are the hazards related to median values of average variables and zero values of change variables

Table 3 Net reclassification improvement (NRI) and integrated discrimination improvement (IDI) indices with 95% confidence intervals comparing Model (d) to Model (b). A positive index value indicates that Model (d) has better predictive ability than Model (b)

Prediction interval (years)	**NRI (95% CI)**	**IDI (95% CI)**
1964 – 1973	0.283 (0.117, 0.449)	0.011 (0.006, 0.017)
1974 – 1983	0.105 (-0.035, 0.245)	0.007 (-0.000, 0.015)
1984 – 1993	0.190 (0.019, 0.361)	0.015 (0.005, 0.026)
1994 – 2003	0.087 (-0.207, 0.382)	0.015 (-0.009, 0.038)