



HAL
open science

Ten-year risk of all-cause mortality: assessment of a risk prediction algorithm in a French general population

Emilie Bérard, Vanina Bongard, Dominique Arveiler, Philippe Amouyel, Aline Wagner, Jean Dallongeville, Bernadette Haas, Dominique Cottel, Jean-Bernard Ruidavets, Jean Ferrières

► **To cite this version:**

Emilie Bérard, Vanina Bongard, Dominique Arveiler, Philippe Amouyel, Aline Wagner, et al.. Ten-year risk of all-cause mortality: assessment of a risk prediction algorithm in a French general population. *European Journal of Epidemiology*, 2010, 26 (5), pp.359-368. <10.1007/s10654-010-9541-6>. <hal-00655377>

HAL Id: hal-00655377

<https://hal.science/hal-00655377v1>

Submitted on 28 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Ten-year risk of all-cause mortality: Assessment of a risk prediction algorithm in a French general population

Emilie Bérard^a, Vanina Bongard^{a*}, Dominique Arveiler^b, Philippe Amouyel^c, Aline Wagner^b, Jean Dallongeville^c, Bernadette Haas^b, Dominique Cottel^c, Jean-Bernard Ruidavets^a, Jean Ferrières^{a,d}

^aDepartment of Epidemiology, Health Economics and Public Health, UMR 558 INSERM - Université de Toulouse, Centre Hospitalier Universitaire de Toulouse, Toulouse, France.

^bDepartment of Epidemiology and Public Health, EA 3430, Université de Strasbourg, Faculté de Médecine, Strasbourg, France.

^cDepartment of Epidemiology and Public Health, Institut Pasteur de Lille - INSERM U774, Université Lille Nord de France, F-59000 Lille, France.

^dDepartment of Cardiology B, Centre Hospitalier Universitaire de Toulouse, Toulouse, France.

Conflict of interest: none declared

* Corresponding author:

Professor Jean Ferrières, MD
Department of Epidemiology
Toulouse University School of Medicine
37 allées Jules Guesde
31073, Toulouse Cedex
France
Phone: +33 561 521 870 ; Fax: +33 561 145 627
jean.ferrieres@cict.fr

Wordcount : 3706

Abstract

Background While assessment of global cardiovascular risk is uniformly recommended for risk factor management, prediction of all-cause death has seldom been considered in available charts. We established an updated algorithm to predict absolute 10-year risk of all-cause mortality in apparently healthy subjects living in France, a country with high life expectancy.

Methods Analyses were based on the Third French MONICA Survey on cardiovascular risk factors (1995-1996) carried out in 3208 participants from the general population aged 35-64. Vital status was obtained 10 years after inclusion and assessment of determinants of mortality was based on multivariable Cox modelling.

Results One-hundred-fifty-six deaths were recorded. Independent determinants of mortality were living area (Northern France), older age, male gender, no high-school completion, smoking, systolic blood pressure ≥ 160 mmHg, LDL-cholesterol ≥ 5.2 mmol/l, and diabetes. Score sheets were developed to easily estimate 10-year risk of death. For example, a non diabetic, heavy smoker, 46-year old man, living in South-Western France, who did not complete high-school, with LDL-cholesterol ≥ 5.2 mmol/l and systolic blood pressure < 160 mmHg, has a 17% probability of death in the ten coming years. The C-statistic of the prediction model was 0.76 [95% CI: 0.72-0.80] with a degree of overoptimism estimated at 0.0058 in a bootstrap sample. Calibration was satisfying: *P*-value for Hosmer-Lemeshow χ^2 test was 0.483.

Conclusion This prediction algorithm is a simple tool for guiding practitioners towards a more or less aggressive management of risk factors in apparently healthy subjects.

Abstract word count: 242

Keywords

All-cause mortality; Low-risk population; Risk estimation; Risk factors; Risk prediction algorithm.

Abbreviations

CI: Confidence intervals

HDL-cholesterol: High density lipoprotein cholesterol

HRs: Hazard ratios

LDL-cholesterol: Low density lipoprotein cholesterol

Introduction

Prediction algorithms estimating a probability to develop a disease over a given period of time according to baseline characteristics have emerged in clinical practice [1,2]. A limitation of such algorithms lies in the fact that predictions are highly dependent on the average level of risk of the cohort used to build the formula, and thus may not be perfectly applicable to other populations [3,4]. Nevertheless, the use of prediction algorithms is uniformly recommended as a decision-support for therapies aimed at preventing cardiovascular diseases [5-7]. In France, Health Agencies also recommend the use of cardiovascular risk prediction algorithms for the management of risk factors, but no consensus has clearly been reached on the formula to use, as none is perfectly adapted to the level of risk of the French population. Prediction algorithms adapted to low risk populations are particularly warranted. Besides, to our knowledge, none of the available charts allows the prediction of all-cause death, which is however the most ultimate endpoint for evaluating the impact of prevention.

Developing an updated algorithm for predicting all-cause mortality would be complementary to the assessment of cardiovascular risk with the available charts, and thus would lead to a more global screening of subjects at risk, in order to get a better prevention of mortality. The aim of this analysis was to assess determinants of all-cause mortality in a French general population aged 35 to 64, and to establish an updated algorithm for predicting the probability of 10-year all-cause mortality according to baseline risk factors.

Methods

Study population

We studied a sample of 3403 subjects randomly selected from the general population to participate in the Third French MONICA Survey on the prevalence of cardiovascular risk factors [8-11]. The study was approved by an institutional ethics committee in agreement with the French law on human biomedical research and the Declaration of Helsinki. The population-based sample was randomly recruited using polling lists, between December 1994 and July 1997 among middle-aged men and women (35-64 years old), living in Northern France (Lille centre), North-Eastern France (Strasbourg centre) or South-Western France (Toulouse centre). The informed consent to participate in the study was obtained from each subject. The participation rate was 66% [11].

Since we aimed to build an algorithm for prediction of risk of mortality in primary prevention of cardiovascular disease and in apparently healthy people, subjects with the following medical histories were excluded from the analyses: clinical or subclinical ischemic heart disease (International Classification of Disease, 9th revision, codes 410.0 to 414.9), clinical or subclinical atherosclerotic cerebrovascular disease (433.0 to 438.9, except codes 437.3 to 437.7), documented atherosclerosis in other arteries such as aorta, renal or lower limb arteries (440.0 to 440.9), chronic renal failure (585.0 to 585.9), chronic respiratory insufficiency (496.0 to 496.9), chronic heart failure (428.0 to 428.9), chronic liver disease or cirrhosis (571.0 to 571.9) and cancer, excluding benign neoplasms and in-situ carcinoma (140.0 to 209.9 and 235.0 to 239.9). Overall, 179 participants were excluded for the above reason. Besides 16 other could not be kept in analyses because of missing data on potential determinants of mortality (7 regarding smoking habits and 9 regarding blood pressure measurement).

Questionnaires and measurement of clinical parameters

Extensive questionnaires were administered by a trained and certified medical staff. Data on socio-economic level, previous medical history, drug intake, cardiovascular risk factors and life style habits were recorded. Educational level was assessed by a report of level of graduation or school drop-out. Alcohol consumption was quantified with a 7-day recall method referring to a typical week. Height, weight and arterial blood pressure (mean of two measurements performed with a standard sphygmomanometer in a sitting position after a 5-minute rest, at least) were measured according to standardised protocols by the medical staff.

Laboratory methods

Blood samples were taken after at least 10 hours of overnight fasting. Serum total cholesterol and triglycerides were measured by enzymatic assays (Boehringer, Mannheim, Germany). High density lipoprotein cholesterol (HDL-cholesterol) measurement was done after sodium phosphotungstate-magnesium chloride precipitation of apo B-containing lipoproteins. Low density lipoprotein cholesterol (LDL-cholesterol) was determined by the Friedewald [12] formula when triglycerides were below 4.6 mmol/l (400 mg/dl). Glucose levels were measured using a conventional enzymatic method based on hexokinase-glucose-6-phosphate deshydrogenase. Diabetes was assessed for people with fasting blood glucose ≥ 7 mmol/l (126 mg/dl) or under hypoglycaemic drug treatment.

Endpoint Determination

Vital status (cause and date of death when applicable) was obtained on December 31th, 2006 for each participant through the French national database recording all deaths occurring in French people [13]. Causes of death were coded according to the International Classification of Diseases, 9th revision (for death occurring before 2000) or 10th revision (for death occurring thereafter). Our primary endpoint was all-cause mortality. Secondly we studied cardiovascular deaths included deaths from ischemic heart diseases, atherosclerotic cerebrovascular diseases, or atherosclerosis and instantaneous deaths. Authorizations to use data were obtained as requested by the French law.

Statistical analysis

Statistical analysis was performed on STATA statistical software, release 9.2 (STATA Corporation, College station, TX, USA). In univariate analysis, qualitative variables were compared with χ^2 -test (or Fisher's exact test when necessary). Student's t-test was used to compare the distribution of quantitative data according to qualitative variables (or Mann-Whitney's test when the distribution of the continuous variable was skewed or when homoscedasticity was rejected).

Survival analysis:

A Kaplan-Meier survival analysis was performed for various potential explanatory variables and differences in survival functions were tested between exposed and unexposed subjects using the Log-Rank test. When the log-linearity hypothesis was not fully respected, continuous variables were transformed into qualitative data. Hazard

ratios (HRs) for mortality and 95% confidence intervals (CI) were assessed using a Cox proportional hazard regression model. All variables listed in *Table 1* have been tested in univariate analysis. Only those with P -value < 0.20 were introduced in the multivariate Cox model (*Table 3 footnote*). A backward procedure was applied to assess variables that were significantly and independently associated with mortality (P -value < 0.05). The proportional-hazard assumption was tested for each covariate by the “log-log” method plotting ($-\ln\{-\ln(\text{survival})\}$) for each category of a nominal covariate, versus $\ln(\text{analysis time})$. None of the assumptions could be rejected.

A similar Cox analysis was conducted to assess independent determinants of cardiovascular mortality. Participants who died from a non-cardiovascular cause during follow-up were then censored.

Development of a scoring scheme:

The Cox model was used to provide a formula predicting probability of all-cause death according to baseline characteristics and risk factors. To convert results from the Cox model into absolute risk estimates, we calculated survival within our population by using Kaplan-Meier statistics, as previously described by Wilson and colleagues [1]. To provide a tool easy to use in clinical practice, score sheets for prediction of absolute risk of 10-year all-cause mortality were developed, based on β -coefficients of the final Cox proportional hazard model.

Model performance:

The C-statistic was calculated to evaluate the discriminatory ability of the prediction model. We performed an internal validation of the model using a bootstrapping method.

We produced 500 random bootstrap samplings with replacement. Each sample comprised 3208 subjects (same size as the original data set). The model was applied to each bootstrap sample and the C-statistic was estimated each time. The mean difference between the C-statistic estimated from using the original data set and the C-statistic estimated from using each bootstrap sample was calculated. This mean difference can be considered as a measure of the degree of overoptimism introduced in the validation of our model. Finally, the C-statistic could be corrected by subtracting the degree of overoptimism from the original C-statistic [14].

To assess the calibration of our prediction model, the Hosmer-Lemeshow χ^2 statistic was calculated for 8 degrees of freedom [15]. Besides, we plotted the observed 10-year incidence of mortality (Kaplan-Meier estimation) and the mean predicted risk in each deciles of predicted risk.

Results

Description of the population sample

Table 1 describes the population sample by centre. There were 3208 apparently healthy participants (1056 in Lille, 1010 in Strasbourg and 1142 in Toulouse centres) with age ranging from 35 to 64 at inclusion. A total of 156 deaths were recorded during the follow-up period (32 (21%) cardiovascular deaths, 75 (48%) cancer deaths, and 49 (31%) deaths from other causes). *Table 2* shows the main characteristics of deceased and non-deceased participants. Centre, older age, male gender, low educational level, smoking, alcohol consumption, diabetes, hypertension, high triglycerides and LDL-cholesterol and low HDL-cholesterol were significantly associated with death.

Survival analysis

The median follow up time was 10.8 years. *Table 3* shows variables independently and significantly associated with 10-year all-cause mortality after multivariate analysis, according to the Cox proportional hazard regression model: living area, age, gender, educational level, smoking, diabetes, systolic blood pressure and LDL-cholesterol. All first-order interactions between mortality and these independent variables were considered (in particular interactions with gender). None exceeded the significance threshold of 0.05. Among the 156 deaths that occurred during the 10-year follow-up, 32 (21%) were due to a cardiovascular cause (ischemic heart disease, atherosclerotic cerebrovascular disease, atherosclerosis or instantaneous death). Significant determinants of cardiovascular mortality were very similar to determinants of all-cause mortality: age older than 54, male gender, smoking ≥ 15 pack-years, diabetes, and systolic blood pressure ≥ 160 mmHg. High LDL-cholesterol tended to be associated with cardiovascular mortality but did not reach the significance level.

Model performance

The C-statistic of the prediction model was 0.7585 [95% CI: 0.7182-0.7989]. The overoptimism correction based on the bootstrap simulation was estimated at 0.0058, leading to a corrected C-statistic equal to 0.7527. The calibration Hosmer-Lemeshow χ^2 statistic was 7.50 with 8 degrees of freedom (P -value for lack of fit = 0.483). Figure 2 compares the observed 10-year incidence of all-cause mortality and the mean predicted risk by deciles of predicted risk.

Scoring Scheme

Using 10-year Kaplan-Meier statistic ($S(t_{10})_{KM} = 0.96$) and β -coefficients of the final Cox proportional hazard model for all-cause mortality (*Table 3*), as previously described by Wilson and colleagues [1], we estimated absolute 10-year risk of all-cause mortality according to baseline characteristics and risk factors (*Appendix*).

To provide a tool easy to use in clinical practice, score sheets for prediction of absolute risk of all-cause mortality were developed from β -coefficients of the final Cox proportional hazard model (*Figure 1*).

For example, the score is 18 points (*Figure 1*) for a man living in South-Western France, aged 46, who did not complete high school, has smoked more than 15 pack-years, is non diabetic and present with LDL-cholesterol ≥ 5.2 mmol/l (200 mg/dl) and systolic blood pressure < 160 mmHg. This corresponds to a 17% risk of death in the ten coming years, according to the scoring scheme (*Figure 1*), the exact probability given by the formula being 16% (*Appendix*). This 10-year all-cause mortality risk is 3 times higher than the average 10-year all-cause mortality risk estimate of same age men (5.3%) (*Figure 1*) and almost 5 times higher than the risk estimate of a same age man without modifiable risk factors (3.6%) (i.e. a man living in South-Western France, who did not complete high school, has never smoked, is non diabetic and present with LDL-c < 5.2 mmol/l (200 mg/dl) and blood pressure < 160 mmHg).

Discussion

Over the past three decades, life expectancy has greatly increased in most European countries consequently to a dramatic decrease in cardiovascular mortality [16]. Because of such changes, predictions made with available risk charts may not be still accurate as they are based on rather old cohorts. Updated risk algorithms, adapted to current European populations, are therefore warranted. The present study assesses

independent determinants of all-cause mortality in a recent cohort (carried out in the mid-nineties) of apparently healthy people. The prediction algorithm we developed is a simple tool for daily clinical practice, requiring solely a clinical examination and a blood test to estimate 10-year absolute risk of mortality for a given subject. The C-statistic of the prediction model reaches 0.76 [95% CI: 0.72 - 0.80], which is very close to the C-statistic of the Framingham prediction algorithms (ranging from 0.68 to 0.77 according to the model chosen [1]) or the SCORE charts (0.71 for the chart using total cholesterol and the low-risk cohort [2]). Thus, our prediction algorithm appears to have a good internal validation but needs to be externally validated before being used in clinical practice, even in a French population. Indeed, the algorithm is based on three French areas only (Northern, North-Eastern and South-Western France), which may not be totally representative of the whole French country. Besides, France is obviously not representative of all European countries, especially regarding cardiovascular mortality, which is well-known to be lower in France than in Northern Europe as described in the so-called French paradox [17]. The latest available statistics from the World Health Organization [18] show that France has one of the highest life expectancy in Europe (81 years at birth, in 2006), similar to other Mediterranean (Spain and Italy) or non-Mediterranean (Switzerland) countries. These areas sharing with France a high life expectancy, close to the one observed in Northern (78 years at birth, in 2005 in Lille area), North-Eastern (80 in Strasbourg area) or South-Western France (82 in Toulouse area), would be particularly appropriate for testing our prediction model.

Obviously, our analysis of the determinants of all-cause mortality is dependent on the data collected during the Third French MONICA population survey. Since this

study was originally intended and designed to analyse cardiovascular risk factors, some potential determinants of all-cause mortality (family history of breast cancer, carcinogenic exposures, viral hepatitis contaminations ...) could not be studied as they were not recorded. The performance of the risk prediction equation would have been enhanced by considering a wider set of non-cardiovascular parameters. However, main causes of death in France and Europe for people aged 35-64 are cancers (lung and upper aero-digestive tract cancers) and cardiovascular diseases [19]. Beyond age, the main risk factors for these cancers are smoking and alcohol consumption which were extensively recorded in our study, alongside with cardiovascular risk factors. Accordingly with the hypothesis that cardiovascular risk factors are also major risk factors for all-cause mortality, a previous European cohort [20] originally intended and designed to analyse risk factors for cardiovascular but also other chronic diseases pointed out that age, smoking, diabetes and high blood pressure are also independent determinants of all-cause mortality in middle-aged subjects.

Beyond cardiovascular risk factors, we found low educational level and living area to be independent determinants of mortality. These two variables are generally considered to be global markers of health condition and are strongly associated with life expectancy [21-23]. Social inequalities, including lower educational levels, have been consistently pointed out as major determinants of health and mortality [21,22]. Suggested explanations include the unequal distribution of health risk behaviours (smoking, alcohol consumption, sedentary lifestyle...), hard occupational exposures, poor health insurance coverage, psychosocial pressures and stress-related factors promoted for instance by financial difficulties and insecure jobs [22]. With higher

mortality rates in Northern as compared to South-Western France, we pointed out another illustration of the North-South gradient in cardiovascular but also all-cause mortality, which has often been related to the French paradox. The concept suggests the interference of living patterns to explain differences in mortality rates [17]. Accordingly, the gradient we observed may be due to differences in life habits such as dietary patterns or wine drinking behaviours, presumably impacting on the distribution of diseases and causes of death [17,23,24]. However, it can also be related to social inequalities, with a lower average socio-economic level in Northern as compared to Southern France [23], and with a lower density of general practitioners, specialists, hospitals and health institutions, and consequently a less easy access to health prevention and care.

To ensure the accuracy of the cumulative incidence of death recorded in our sample, we calculated the expected cumulative incidence that should be observed in our study population if death rates by year, gender and age categories were those found in Northern, North-Eastern and South-Western France between 1995 and 2006. We found an expected cumulative death rate of 7.2% whereas the observed cumulative incidence of death in the study was 5.8% [95% CI: 5.0 - 6.6] (in the whole sample, i.e. before excluding subjects in secondary prevention or those with severe chronic disease). This difference may be related to the possible selection of a study sample in a relatively better health shape than the general population, as already described for participants in research projects [25]. As a consequence, our risk algorithm may slightly underestimate probabilities of death and this should be taken into account when defining risk categories to settle measurements of prevention. Indeed, the aim of prediction algorithms is to discriminate between subjects at low and high risk, to propose

appropriate risk management. Our equation indicates that a man older than 55, with a cluster of risk factors, may have a 10-year risk of death greater than 50%. There is no doubt that this man requires to be referred for an aggressive management of risk factors, but for less extreme situations, the question is where do we have to settle the threshold to define high risk subjects. Given that high cardiovascular risk has been assessed as a 10-year probability of fatal cardiovascular event $\geq 5\%$ according to the SCORE chart [7], and given that cardiovascular mortality represents about 1/4 of all-cause mortality in middle-aged European people, we propose to consider that subjects with a 10-year risk of all-cause mortality $\geq 15\%$, according to our algorithm, are at high risk (i.e. slightly below 20%, to take into account that our equation may underestimate probabilities of death).

In conclusion, various algorithms have been developed to estimate a probability of disease occurrence for a given individual. Such tools are of major importance to screen those people who will benefit from prevention among apparently healthy subjects, and consequently to guide practitioners towards a more or less aggressive management of risk factors. Therefore risk algorithms are key for developing cost-effective disease prevention. Because of a constant improvement of medical research, risk charts need to be updated to provide information in accordance with current medical care. Although we mainly focused on cardiovascular determinants, the tool we developed is complementary to cardiovascular algorithms previously published. Indeed, assessing all-cause mortality enables a more global approach of health prevention and a better identification of priorities for health policies.

Appendix:

The following steps lead to the estimation of the 10-year absolute risk of all-cause mortality for a given subject:

Step 1: Equation A = 0.55 (for people living in Northern France) + 0.04 (for people living in North-Eastern France) + 0.00 (for people living in South-Western France) + 0.00 (for age ranging between 35 and 44) + 0.50 (for age ranging between 45 and 54) + 1.35 (for age ranging between 55 and 64) + 0.00 (for women) + 0.74 (for men) + 0.00 (if high school has been completed) + 0.45 (if high school has not been completed) + 0.00 (for non smokers) + 0.96 (for smokers < 15 pack-years) + 1.04 (for smokers \geq 15 pack-years) + 0.00 (if LDL-cholesterol < 5.2 mmol/l (200 mg/dl)) + 0.49 (if LDL-cholesterol \geq 5.2 mmol/l (200 mg/dl)) + 0.00 (for non diabetic subjects) + 0.50 (for diabetic subjects) + 0.00 (if systolic blood pressure < 160 mmHg) + 0.42 (if systolic blood pressure \geq 160 mmHg).

Step 2: Equation A is calculated for the average level of each explanatory variable:

$$B = 0.55 \times 0.33 + 0.04 \times 0.31 + 0.50 \times 0.35 + 1.35 \times 0.30 + 0.74 \times 0.50 + 0.45 \times 0.66 + 0.96 \times 0.06 + 1.04 \times 0.15 + 0.49 \times 0.08 + 0.50 \times 0.12 + 0.42 \times 0.08 = 1.7873$$

Step 3: $C = A - B$ (where $B = 1.7873$) and $D = e^C$

Step 4: The 10-year survival Kaplan-Meier value $S(t)$ is exponentiated by D and subtracted from 1 to calculate the 10-year risk of all-cause mortality for a given subject, according to his baseline characteristics and risk factors.

$$P = 1 - [S(t)]^D \text{ (where } S(t) = 0.96)$$

Example: The 10-year absolute risk of all-cause mortality is 16% for a man living in South-Western France, aged 46, who did not complete high school, has smoked more than 15 pack-years, is non diabetic, and has LDL-cholesterol \geq 5.2 mmol/l and blood pressure < 160 mmHg.

Step 1: $A = 0.50 \times 1 + 0.74 \times 1 + 0.45 \times 1 + 1.04 \times 1 + 0.49 \times 1 = 3.22$

Step 2: $B = 1.7873$

Step 3: $C = 3.22 - 1.7873 = 1.4327$ and $D = e^{1.4327} = 4.1899$

Step 4: and $P = 1 - 0.96^{4.1899} = 0.1572$

Acknowledgements

We would like to thank the " Institut National de la Santé et de la Recherche Médicale" (INSERM), the "Direction Générale de la Santé (DGS)", the "Institut Pasteur de Lille", the "University Hospital of Lille", the "Fonds d'intervention en Santé Publique", the "Mutuelle Générale de l'Education Nationale", "ONIVINS", the "Fondation de France", the "CPAM of Selestat"; the "Fédération Française de Cardiologie" the "Conseil Régional du Nord-Pas de Calais", Parke-Davis and Bayer pharmaceuticals, and CERIN for their financial supports enabling this work. We did appreciate the collaboration with the INSEE and the health centres in the 3 regions.

We would like to thank all investigators of the MONICA Project for their contribution to the compilation and validation of the data.

References

1. Wilson PW, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;**97**:1837-1847.
2. Conroy RM, Pyörälä K, Fitzgerald AP, et al.; SCORE Project Group. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;**24**:987-1003.
3. Empana JP, Ducimetière P, Arveiler D, et al.; PRIME Study Group. Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations? The PRIME Study. *Eur Heart J* 2003;**24**:1903-1911.
4. Vergnaud AC, Bertrais S, Galan P, et al. Ten-year risk prediction in French men using the Framingham coronary score: results from the national SU.VI.MAX Cohort. *Prev Med* 2008;**47**:61-65.
5. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA* 2001;**285**:2486-2497.
6. Grundy SM, Cleeman JI, Merz CN, et al.; National Heart, Lung, and Blood Institute; American College of Cardiology Foundation; American Heart Association. Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III guidelines. *Circulation* 2004;**110**:227-239.
7. European Society of Cardiology. Fourth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice. European guidelines on cardiovascular disease prevention in clinical practice: executive summary. *Eur Heart J* 2007;**28**:2375-2414.

8. Tunstall-Pedoe H, Kuulasmaa K, Mähönen M, et al. Contribution of trends in survival and coronary-event rates to changes in coronary heart disease mortality: 10-year results from 37 WHO MONICA Project populations. Monitoring trends and determinants in cardiovascular disease. *Lancet* 1999;**353**:1547-1557.
9. Kuulasmaa K, Tunstall-Pedoe H, Dobson A, et al. Estimation of contribution of changes in classic risk factors to trends in coronary-event rates across the WHO MONICA Project populations. *Lancet* 2000;**355**:675-687.
10. Ferrières J, Cambou JP, Ruidavets JB, et al. Trends in acute myocardial infarction prognosis and treatment in southwestern France between 1985 and 1990 (the MONICA Project-Toulouse). *Am J Cardiol* 1995;**75**:1202-1205.
11. Marques-Vidal P, Ruidavets JB, Amouyel P, et al. Change in cardiovascular risk factors in France, 1985-1997. *Eur J Epidemiol* 2004;**19**:25-32.
12. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low density lipoprotein cholesterol in plasma without use of the preparative ultracentrifuge. *Clin Chem* 1978;**18**:499-502.
13. INSERM—IFR69. Mise en oeuvre du décret n_ 98-37 autorisant l'accès aux données relatives au décès des personnes inscrites au Répertoire National d'Identification des Personnes Physiques (RNIPP) dans le cadre des recherches dans le domaine de la santé.
<http://ifr69.vjf.inserm.fr/svcd.html> (2 June 2010)
14. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and, reducing errors. *Stat Med* 1996 ;**15**:361-387.

15. Hosmer DW Jr, Lemeshow S. Applied Logistic Regression. New York: J Wiley; 1989.
16. Kesteloot H, Sans S, Kromhout D. Dynamics of cardiovascular and all-cause mortality in Western and Eastern Europe between 1970 and 2000. *Eur Heart J* 2006;**27**:107-113.
17. Ferrières J. The French paradox: lessons for other countries. *Heart* 2004;**90**:107-111
18. World Health Organization (WHO). http://www.who.int/whosis/database/core/core_select.cfm (2 June 2010)
19. Aouba A, Péquignot F, Le Toullec A, et al. Les causes médicales de décès en France en 2004 et leur évolution 1980-2004. *Bull Epidemiol Hebdo* 2007; **35-36**:308-314.
20. Hoes AW, Grobbee DE, Valkenburg HA, et al. Cardiovascular risk and all-cause mortality; a 12 year follow-up study in The Netherlands. *Eur J Epidemiol* 1993;**9**:285-292.
21. Saurel-Cubizolles MJ, Chastang JF, Menvielle G, et al; EDISC group. Social inequalities in mortality by cause among men and women in France. *J Epidemiol Community Health* 2009;**63**:197-202.
22. Van Oort F, van Lenthe F, Mackenbach JP. Material, psychosocial, and behavioural factors in the explanation of educational inequalities in mortality in the Netherlands. *J Epidemiol Community Health* 2005;**59**:214–220.
23. Cotel D, Dallongeville J, Wagner A, et al. The North-East-South gradient of coronary heart disease mortality and case fatality rates in France is consistent with a similar gradient in risk factor clusters. *Eur J Epidemiol* 2000;**16**:317-322.

24. Perrin AE, Dallongeville J, Ducimetière P, et al. Interactions between traditional regional determinants and socio-economic status on dietary patterns in a sample of French men. *Br J Nutr* 2005;**93**:109-114.
25. Kho ME, Duffett M, Willison DJ, et al. Written informed consent and selection bias in observational studies using medical records: systematic review. *BMJ* 2009;**338**:b866.

Figure legends

Figure 1: Algorithm for 10-year all-cause mortality prediction

Figure 1a: Sum points according to baseline characteristics and risk factors.

Figure 1b: Find the corresponding estimate for 10-year absolute risk of mortality.

Figure 1c: Compare to the average 10-year absolute risk of a same age and gender subject.

* The average 10-year all-cause mortality risk represents the average absolute risk of death for all categories of risk corresponding to a given score.

** The average risk estimates are calculated for a person with average values for baseline characteristics and risk factors.

Figure 2: Observed 10-year incidence of all-cause mortality and mean predicted risk by deciles of predicted risk

Table 1 Main characteristics of the sample, by centre

	Lille	Strasbourg	Toulouse
	Centre	Centre	Centre
	N=1056	N=1010	N=1142
Age (years), mean (\pm SD)	50.2 (\pm 8.0)	50.1 (\pm 8.2)	50.0 (\pm 8.5)
Men, N (%)	513 (49%)	495 (49%)	587 (51%)
High school completed, N (%)	331 (31%)	324 (32%)	430 (38%)
Non-smoking, N (%)	810 (77%)	820 (81%)	897 (79%)
Smoking < 15 pack-years, N (%)	55 (5%)	48 (5%)	80 (7%)
Smoking \geq 15 pack-years, N (%)	191 (18%)	142 (14%)	165 (14%)
No alcohol consumption, N (%)	200 (19%)	246 (24%)	389 (34%)
Drinking \leq 30 g/day for men and \leq 20 g/day for women, N (%)	464 (44%)	506 (50%)	466 (41%)
Drinking > 30 g/day for men and > 20 g/day for women, N (%)	392 (37%)	258 (26%)	287 (25%)
No regular physical activity, N (%)	390 (37%)	182 (18%)	133 (12%)
Light physical activity almost every week, N (%)	537 (51%)	562 (56%)	469 (41%)

Intense physical activity during at least 20 minutes 1 to 2 times a week, N (%)	96 (9%)	193 (19%)	273 (24%)
Intense physical activity during at least 20 minutes \geq 3 times a week, N (%)	31 (3%)	72 (7%)	265 (23%)
Body Mass Index (kg/m ²), mean (\pm SD)	26.5 (\pm 5.0)	26.8 (\pm 4.6)	25.5 (\pm 4.2)
Diabetes, N (%)	165 (16%)	125 (12%)	92 (8%)
Anti-hypertensive drug treatment, N (%)	160 (15%)	143 (14%)	146 (13%)
Systolic blood pressure (mmHg), mean (\pm SD)	133 (\pm 19)	132 (\pm 20)	131 (\pm 18)
Diastolic blood pressure (mmHg), mean (\pm SD)	83 (\pm 12)	85 (\pm 13)	80 (\pm 10)
Lipid-lowering drug treatment, N (%)	108 (10%)	95 (9%)	120 (10%)
Total cholesterol (mmol/l), mean (\pm SD)	5.99 (\pm 1.04)	5.89 (\pm 1.06)	5.78 (\pm 1.02)
LDL-cholesterol (mmol/l), mean (\pm SD)	3.95 (\pm 0.98)	3.79 (\pm 1.02)	3.79 (\pm 0.96)
HDL-cholesterol (mmol/l), mean (\pm SD)	1.48 (\pm 0.42)	1.50 (\pm 0.47)	1.47 (\pm 0.44)
Triglycerides (mmol/l), mean (\pm SD)	1.26 (\pm 0.94)	1.45 (\pm 1.79)	1.18 (\pm 0.75)

SD: Standard deviation.

Lille is located in Northern, Strasbourg in North-Eastern and Toulouse in South-Western France.

Table 2 Main characteristics of deceased and non deceased participants

	Non deceased N=3052	Deceased N=156	<i>P</i> -value
Living area, N (%)			0.003
Lille (Northern France)	985 (32%)	71 (46%)	
Strasbourg (North-Eastern France)	972 (32%)	38 (24%)	
Toulouse (South-Western France)	1095 (36%)	47 (30%)	
Age, N (%)			< 0.001
35-44 years	1112 (36%)	29 (19%)	
45-54 years	1070 (35%)	44 (28%)	
55-64 years	870 (29%)	83 (53%)	
Men, N (%)	1485 (49%)	110 (70%)	< 0.001
High school completed, N (%)	1053 (34%)	32 (20%)	< 0.001
Smoking habits, N (%)			< 0.001
Non-smoking	2435 (80%)	92 (59%)	

Smoking < 15 pack-years	170 (5%)	13 (8%)	
Smoking \geq 15 pack-years	447 (15%)	51 (33%)	
Alcohol consumption, N (%)			<0.001
No alcohol consumption	805 (26%)	30 (19%)	
Drinking \leq 30 g/day for men and \leq 20 g/day for women	1377 (45%)	59 (38%)	
Drinking > 30 g/day for men and > 20 g/day for women	870 (29%)	67 (43%)	
Physical activity, N (%)			0.125
No regular physical activity	664 (22%)	41 (26%)	
Light physical activity almost every week	1493 (49%)	75 (48%)	
Intense physical activity during at least 20 minutes 1 to 2 times a week	544 (18%)	18 (11%)	
Intense physical activity during at least 20 minutes 3 times a week or more	346 (11%)	22 (14%)	
Body Mass Index (kg/m ²), mean (\pm SD)	26.2 (\pm 4.6)	26.6 (\pm 5.1)	0.212
Diabetes, N (%)	346 (11%)	36 (23%)	< 0.001
Antihypertensive drug treatment, N (%)	417 (14%)	32 (21%)	0.016
Systolic blood pressure (mmHg), mean (\pm SD)	132 (\pm 19)	140 (\pm 21)	< 0.001

Systolic blood pressure \geq 160 mmHg, N (%)	243 (8%)	28 (18%)	< 0.001
Diastolic blood pressure (mmHg), mean (\pm SD)	82 (\pm 12)	85 (\pm 12)	0.033
Lipid-lowering drug treatment, N (%)	307 (10.1%)	16 (10.3%)	0.936
Total cholesterol (mmol/l), mean (\pm SD)	5.88 (\pm 1.03)	5.86 (\pm 1.26)	0.489
Total cholesterol \geq 6.45 mmol/l (250 mg/dl), N (%)	824 (27%)	43 (29%)	0.728
LDL-cholesterol (mmol/l), mean (\pm SD)	3.84 (\pm 0.98)	3.82 (\pm 1.20)	0.461
LDL-cholesterol \geq 5.2 mmol/l (200 mg/dl), N (%)	240 (8%)	20 (13%)	0.008
HDL-cholesterol (mmol/l), mean (\pm SD)	1.49 (\pm 0.44)	1.40 (\pm 0.47)	0.011
HDL-cholesterol < 1 mmol/l (40 mg/dl), N (%)	346 (12%)	31 (21%)	0.001
Triglycerides (mmol/l), mean (\pm SD)	1.28 (\pm 1.23)	1.59 (\pm 1.48)	< 0.001
Triglycerides \geq 1.70 mmol/l (150 mg/dl), N (%)	506 (17%)	34 (23%)	0.063

SD: Standard deviation

Table 3 Independent predictors of all-cause mortality

	β -Coefficient (Cox Model)	Hazard Ratio	95% Confidence Interval	<i>P</i> -value
Toulouse centre (South-Western France)	1.00			
Strasbourg centre (North-Eastern France)	0.04	1.04	[0.67 - 1.61]	0.857
Lille centre (Northern France)	0.55	1.76	[1.21 - 2.58]	0.003
35-44 years old	1.00			
45-54 years old	0.50	1.64	[1.02 - 2.64]	0.040
55-64 years old	1.35	3.89	[2.49 - 6.09]	< 0.001
Women	1.00			
Men	0.74	2.11	[1.48 - 3.01]	< 0.001
Educational level \geq high school completion	1.00			
Educational level < high school completion	0.45	1.57	[1.05 - 2.33]	0.026
Non-smoking	1.00			
Smoking < 15 pack-years	0.96	2.61	[1.45 - 4.70]	0.001

Smoking \geq 15 pack-years	1.04	2.90	[2.01 - 4.17]	< 0.001
No diabetes	1.00			
Diabetes	0.50	1.62	[1.10 - 2.37]	0.014
Systolic blood pressure < 160 mmHg	1.00			
Systolic blood pressure \geq 160 mmHg	0.42	1.57	[1.03 - 2.39]	0.036
LDL-cholesterol < 5.2 mmol/l (200 mg/dl)	1.00			
LDL-cholesterol \geq 5.2 mmol/l (200 mg/dl)	0.49	1.62	[1.01 - 2.61]	0.047
Missing data on LDL-cholesterol level (n=126*)	0.40	1.50	[0.80 - 2.80]	0.207
<hr/>				
Baseline survival Kaplan-Meier function at 10 years: S(t)	0.96			

Variables initially introduced in the multivariate Cox analysis are Living area, Age, Gender, Educational level, Smoking habits, Alcohol consumption, Physical activity, Diabetes, Antihypertensive drug treatment, Systolic and Diastolic blood pressure, LDL-cholesterol, HDL-cholesterol and Triglycerides.

* 126 subjects had missing data regarding LDL-cholesterol level (69 subjects for whom lipid measurement was lacking and 57 for whom LDL-cholesterol could not be estimated with the Friedewald equation as triglycerides were higher than 4.6 mmol/l).