

### 3. 統合研究リスクエンジン2（個人の10年リスク）

#### 個人の10年以内の循環器疾患の予測リスクの開発

八谷 寛（藤田医科大学医学部公衆衛生学）  
李 媛英（藤田医科大学医学部公衆衛生学）  
田中 佐智子（滋賀医科大学社会医学講座医療統計学）  
村上 義孝（東邦大学医学部社会医学講座医療統計学）  
三浦 克之（滋賀医科大学社会医学講座公衆衛生学）  
辻 一郎（東北大学大学院医学系研究科情報健康医学講座公衆衛生学分野）  
宮本 恵宏（国立循環器病研究センター）  
坂田 清美（岩手医科大学医学部衛生学公衆衛生学講座）  
磯 博康（大阪大学大学院医学系研究科公衆衛生学）  
岡山 明（生活習慣病予防研究センター）  
岡村 智教（慶應義塾大学医学部衛生学公衆衛生学）

#### 背景・意義

個々の人は、リスク因子を種々の組み合わせで有しているため、当該個人の予防介入を行うにあたっては、それらリスク因子の組み合わせから予測される疾病発生リスクの推定と、予測リスクカテゴリー別に定められた標準的方法を医師の判断により用いることが有用である。国内外で既に多くのリスク予測モデル（予測式、リスクエンジン）が開発され、使用されているが、欧米で開発されたモデルはイベント発生率および疾病構成割合が大きく異なるため、日本人へのそのままの適用は困難である。日本国内でも複数のリスクモデルが開発されているが、リスク因子の平均レベルや保有割合、疾病との関連性は時代によって異なる可能性があるため、1990年代後半から2000年代に開始されたコホートを含む最近の集団で開発する意義は大きい。また、職域コホート、都市部や農山村部コホートなど多様なコホートを含む日本を代表する集団でのリスクエンジンが望まれる。そのような広範な対象を含む大規模研究においては、エンドポイントを標準化が容易な死亡とすることが妥当である。さらに、慢性腎臓病（Chronic Kidney Disease: CKD）が循環器疾患の重要なリスク因子であることが広く知られるようになってきているが、CKDを含むリスクエンジンは少なく、特に循環器疾患死亡を標的としたモデルはない。CKDのような重要な新規リスク因子を含むモデルを開発することで、識別能、適合度をより高め、ハイリスク者の正確な同定を行うことが可能となる。また、適切な交互作用項をモデルに含めることもモデルの改善に重要で、そのためには検出力の高い大規模なデータセット（EPOCH-JAPAN 循環器データベース）での解析が有用である。

## 方法

対象：EPOCH-JAPAN 研究の 15 のコホート男性 61,083 人と女性 86,382 人のうち、ベースラインが 1985 年以前、追跡期間の中央値が 10 年未満、さらに本研究で用いるリスク因子の測定が実施されていない 7 コホートを除外した。残りの 8 コホート（大崎・YKK・吹田・久山・ND90・大阪・JMS・愛知職域）の男性 31,214 人と女性 34,552 人のうち、年齢が 40 歳未満あるいは 80 歳以上の者（N=10,579）、現在喫煙の有無・収縮期血圧（SBP）・総コレステロール（TC）・HDL コレステロール（HDL）・尿タンパク・糖尿病の有病者のいずれの変数に欠測がある者（N=8,392）、循環器疾患の既往者（N=1,926）を除外し、最終的に男性 21,491 人、女性 23,378 人を対象とした。

エンドポイント：エンドポイントは虚血性心疾患、脳卒中、虚血性心疾患と脳卒中（心血管疾患：CVD）3 種とした。

統計解析：解析は男女合わせて実施した。コホートを層化変数とした層化 COX 比例ハザードモデルを用いて解析した。性別・現在喫煙の有無・尿タンパクの有無・糖尿病の有病者は 2 値変数、年齢・SBP・HDL/TC 比は対数変換をした値を解析に用いた。単変量モデルで計学的に有意な（ $P < 0.10$ ）変数及び性別、年齢それぞれと残りの各変数の間の交互作用項（ $P < 0.10$  のもの）を多変量モデルに投入し、変数減少法（ $P < 0.10$  を基準）で最終モデルに含まれるリスク因子を決定した。なお、エンドポイントが脳卒中、CVD のモデルには HDL/TC 比の予測因子としての多変量モデルでの検討は行っていない。コホートごとの平均 10 年生存率に各コホートの該当エンドポイントの死亡者を重み付けし、集団全体の平均 10 年生存率を算出し、予測式のベースラインハザードとして当てはめた。予測モデルの識別能は解析集団全体における c-statistic を求め、適合度は予測リスクを 10 分位し、その平均予測リスクと観察リスクとの calibration plot で示し、Hosmer-Lemshow 検定で検証した。

## 研究結果

開発に用いたコホートのリスク因子の平均値または保有割合を Table 1 に、各リスク因子の冠動脈疾患、脳卒中、循環器疾患との関連（単変量解析）を Table 2 に示した。多変量モデルで選択された変数は冠動脈疾患モデルで年齢、性別、現喫煙、糖尿病、尿蛋白、収縮期血圧、HDL/TC 比、年齢と収縮期血圧の交互作用、年齢と現喫煙の交互作用であった（Table 3）。脳卒中モデルでは、年齢、現喫煙、糖尿病、尿蛋白、収縮期血圧、年齢と収縮期血圧の交互作用、年齢と現喫煙の交互作用、循環器疾患モデルでは、年齢、性別、現喫煙、糖尿病、尿蛋白、収縮期血圧、年齢と収縮期血圧の交互作用、年齢と現喫煙の交互作用であった。

開発した冠動脈疾患モデルの識別能は 0.84（95%信頼区間：0.73-0.93）、脳卒中モデル

は 0.81 (0.73-0.88)、循環器疾患モデルは 0.82 (0.75-0.88) であった。冠動脈疾患モデルと循環器疾患モデルの適合度は良好であった。脳卒中モデルも概ね適合していたが、Hosmer-Lemshow 検定は有意 (非適合) であった。性年齢階級別の平均リスクを Table 4 に示した。

Table 1. Baseline characteristics of the participating cohorts, means (standard deviation) or percentage.

Cohort	N	Baseline year	Age (year)	Men (%)	Current smoking	SBP (mmHg)	HDLc (mg/dl)	TC (mg/dl)	HDLc/TC ratio	Proteinuria +	Prevalent DM +
Osaki	10898	1995	61.5 (9.4)	47.6	26.0	131.2 (17.5)	51.7 (12.8)	204.0 (35.2)	0.26 (0.07)	1.7	5.5
YKK	2797	1990	47.1 (5.3)	67.4	39.2	119.2 (15.7)	52.9 (13.2)	202.6 (35.0)	0.27 (0.07)	4.0	2.5
Suita	5016	1991	58.6 (10.7)	47.5	29.3	129.8 (22.0)	52.7 (13.8)	210.8 (37.1)	0.26 (0.07)	6.1	4.6
Hisayama	2475	1988	57.5 (10.4)	42.5	25.2	132.4 (20.4)	50.4 (11.7)	207.0 (42.0)	0.25 (0.07)	5.7	8.8
ND 90	5367	1990	56.1 (10.5)	42.8	28.1	138.3 (20.3)	53.8 (15.4)	206.6 (37.9)	0.27 (0.08)	2.8	4.3
Osaka	4767	1996	57.3 (9.6)	36.3	23.7	135.4 (20.5)	59.4 (14.8)	213.4 (36.7)	0.28 (0.08)	2.0	3.5
JMS	9535	1994	57.3 (9.1)	38.7	21.9	130.3 (20.7)	51.2 (12.9)	194.2 (34.7)	0.27 (0.07)	1.8	3.0
Aichi	4014	2002	50.1 (5.6)	81.2	31.1	127.4 (15.7)	58.7 (15.8)	211.7 (34.8)	0.28 (0.08)	2.3	6.7
Overall	44869	-	57.0 (10.1)	47.9	26.7	131.1 (19.8)	53.4 (14.1)	204.8 (36.7)	0.27 (0.08)	2.8	4.6

SBP indicate systolic blood pressure, HDLc, high density lipoprotein cholesterol; TC, total cholesterol, DM, diabetes mellitus.

Table 2. Univariable hazard ratios (HRs) and 95% confidence interval (95% CI).

	Coronary heart disease		Stroke		Cardiovascular disease	
	HRs (95% CI)	P	HRs (95% CI)	P	HRs (95% CI)	P
Ln-age (year)	1,469 (533-4,050)	<0.0001	2,486 (1,130-5,468)	<0.0001	2,047 (1,098-3,815)	<0.0001
Men (%)	2.65 (2.06-57.1)	<0.0001	1.43 (1.19-15.3)	<0.0001	1.77 (1.53-60.1)	<0.0001
Current smoking (%)	2.07 (1.63-35.3)	<0.0001	1.40 (1.16-11.9)	0.0006	1.62 (1.40-40.6)	<0.0001
Prevalent DM + (%)	2.69 (1.82-25.1)	<0.0001	2.26 (1.67-27.4)	<0.0001	2.41 (1.90-51.6)	<0.0001
proteinuria + (%)	3.04 (1.95-24.4)	<0.0001	2.72 (1.90-30.1)	<0.0001	2.83 (2.15-54.3)	<0.0001
Ln-SBP (mmHg)	34.9 (16.16-75.2)	<0.0001	28.4 (15.77-51.2)	<0.0001	30.6 (19.2-48.9)	<0.0001
Ln-(HDLc/TC ratio)	0.59 (0.41-0.83)	0.0028	-	-	-	-

DM indicates diabetes mellitus; SBP, systolic blood pressure, HDLc, high density lipoprotein cholesterol; TC, total cholesterol.

Table 3. Multivariate regression coefficients (standard errors).

	Coronary heart disease		Stroke		Cardiovascular disease	
	$\beta$ (standard error)	<i>P</i> value	$\beta$ (standard error)	<i>P</i> value	$\beta$ (standard error)	<i>P</i> value
Ln-age (year)	61.1992 (15.7527)	0.0001	37.4061 (12.3059)	0.0024	45.5499 (9.7502)	<0.0001
Men (%)	0.6587 (0.1450)	<0.0001	-	-	0.3009 (0.0843)	0.0004
Current smoking (%)	15.3639 (4.3414)	0.0004	8.2429 (3.2749)	0.0118	10.6493 (2.5905)	<0.0001
Prevalent DM + (%)	0.5625 (0.1998)	0.0049	0.4568 (0.1570)	0.0036	0.4941 (0.1234)	<0.0001
proteinuria + (%)	0.5824 (0.2297)	0.0112	0.6362 (0.1847)	0.0006	0.6212 (0.1439)	<0.0001
Ln-SBP (mmHg)	46.3700 (13.3193)	0.0005	26.3995 (10.4481)	0.0115	33.3073 (8.2662)	<0.0001
Ln-(HDLc/TC ratio)	-0.3593 (0.1885)	0.0566	-	-	-	-
Ln-age*Ln-SBP	-10.6145 (3.1716)	0.0008	-5.9229 (2.4876)	0.0173	-7.5446 (1.9681)	0.0001
Ln-age*Current smoking	-3.5197 (1.0336)	0.0007	-1.8451 (0.7823)	0.0184	-2.4200 (0.6181)	<0.0001

DM indicates diabetes mellitus; SBP, systolic blood pressure; HDLc, high density lipoprotein cholesterol; TC, total cholesterol.

Table 4. Mean 10-year risk (%) of death from coronary heart disease, stroke, and cardiovascular disease according to 5-year age groups and sex.

Age range (year)	Coronary heart disease		Stroke		Cardiovascular disease	
	men	women	men	women	men	women
40-44	<0.1	<0.1	<0.1	<0.1	0.11	<0.1
45-49	0.12	<0.1	0.13	<0.1	0.24	<0.1
50-54	0.26	<0.1	0.28	0.16	0.51	0.21
55-59	0.50	0.13	0.58	0.35	1.03	0.45
60-64	0.89	0.28	1.09	0.74	1.89	0.94
65-69	1.38	0.54	1.78	1.38	3.03	1.76
70-74	2.29	1.03	3.01	2.54	5.05	3.25
75-80	3.63	1.83	4.83	4.38	8.00	5.60

## Coronary heart disease

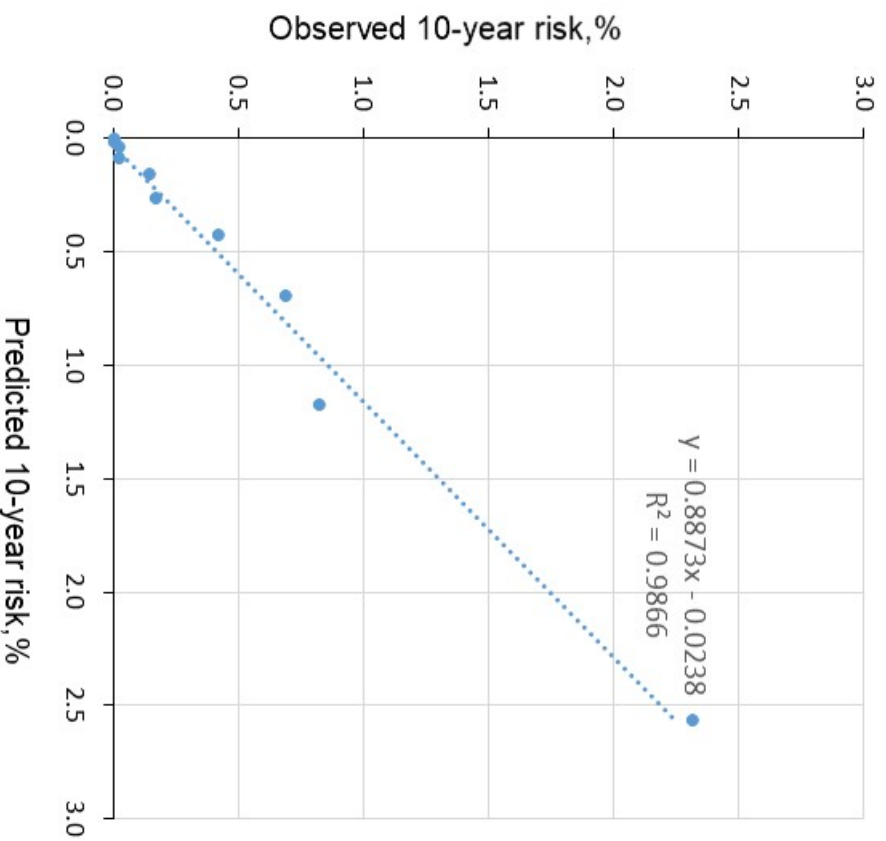


Figure 1. Calibration plots for death from coronary heart disease, showing predicted and observed 10-year risk according to deciles of predicted 10-year risk. Solid point denotes each decile. Coefficient of the regression=0.8725, constant of the regression=-0.0238. Calibration for  $\chi^2=11.39$  P=0.18

## stroke

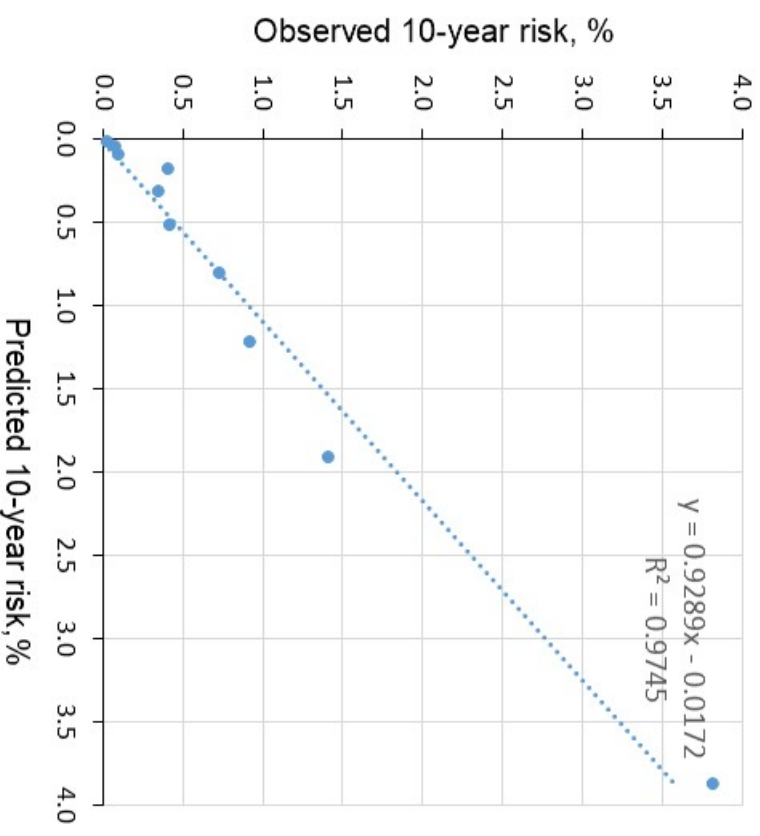


Figure 2. Calibration plots for death from stroke, showing predicted and observed 10-year risk according to deciles of predicted 10-year risk. Solid point denotes each decile. Coefficient of the regression=0.9289, constant of the regression=-0.0172. Calibration for  $\chi^2=23.69$  P=0.003

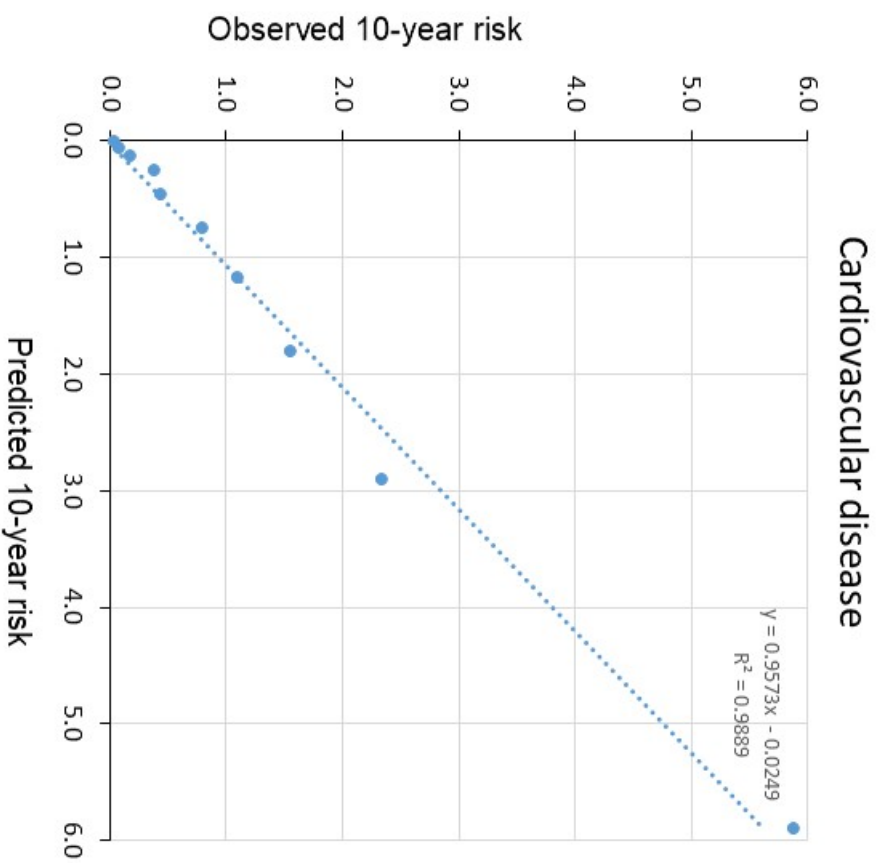


Figure 3. Calibration plots for death from cardiovascular disease, showing predicted and observed 10-year risk according to deciles of predicted 10-year risk.

Solid point denotes each decile.

Coefficient of the regression=0.9573, constant of the regression=-0.0249.

Calibration for  $\chi^2=10.20$  P=0.25