

リスクエンジン（集団リスク）

集団における 10 年間の循環器疾患発症危険度を予測するリスク評価ツールの開発

研究分担者	村上義孝	東邦大学医学部社会医学講座医療統計学分野	教授
研究分担者	三浦克之	滋賀医科大学社会医学講座公衆衛生学部門	教授
研究分担者	上島弘嗣	滋賀医科大学アジア疫学研究センター	特任教授
研究代表者	岡村智教	慶應義塾大学医学部衛生学公衆衛生学	教授

研究要旨

EPOCH-JAPAN循環器データベースを用いて死亡予測モデルの妥当性検討と、一般集団に適用した場合の死亡者数予測・妥当性の検討を実施した。死亡予測モデルの妥当性検討の結果、実測値と予測値との関連は良好であり、相関係数も一部を除き0.95を超えていること、男女別の傾向に違いはないこと、疾患別では女性CHDでイベントが少ない影響から相関図での関連が不明瞭であることが示された。死亡予測モデルを一般集団に適用した場合の死亡者数予測（5年間(2010-2014年)）では、男性では実データと予測モデルの比がCVD:83.7%, 脳卒中:92.7%, CHD:71.4%と良好である一方、女性ではCVD:54.7%、脳卒中:54.8%と予測モデルの結果が観察値の半数程度、CHDでは19.8%と過小評価であった。

EPOCH-JAPAN循環器データベースを用いた死亡予測モデルのモデル妥当性は良好であった。また一般集団に適用した場合の死亡者数予測の実施可能性が示され、男性での妥当性が確認された。

A. 研究目的

循環器コホートを中心に構成された大規模コホート統合研究プロジェクトである EPOCH-JAPAN のデータを用い、最終年である本年は、1) EPOCH-JAPAN 循環器データベースを用いた死亡予測モデルの妥当性検証と、2) その死亡予測モデルを一般集団に適用した場合の死亡者数予測と妥当性検討を実施したので報告する。

B. 研究方法

(1) EPOCH-JAPAN 循環器データベースを用いた死亡予測モデルの妥当性検証

EPOCH-JAPAN を構成するコホートのうち 1990 年付近にベースライン調査を実施したコホートデータを用い、10 年後の死亡予測モデルを作成した。またこの死亡予測モデルの妥当性検証を行うため、1995 年

付近にベースライン調査を実施した EPOCH-JAPAN の各コホートデータを用い、比較を実施した。

15 コホートで構成される EPOCH-JAPAN のコホートを、年代別に以下の 3 グループに分けて検討に用いた。なお 1990 年以前に開始した端野・壮瞥研究、NIPPON DATA80、職域コホートの YKK は今回の検討から除外した。

グループ 1 ; 1990 年前後に開始

大迫、小矢部、吹田、放影研、久山、JACC、NIPPON DATA90

(小矢部、JACC は糖尿病の変数がないため除外)

グループ 2 ; 1995 年前後に開始

大崎、大阪、JMS

グループ 3 ; 2000 年以降に開始

愛知県職域、岩手県北

解析方法として上記のグループ 1 のコホート（大迫、吹田、放影研、久山、NIPPON DATA90）を用い、循環器疾患の死亡予測モデルを作成した。エンドポイント(死亡)は全循環器疾患（以下 CVD）、脳卒中、CHD とした。付録に男性の CVD を例として、パラメータ値（ハザード比）、集団平均値、CVD の死亡リスク予測モデルの作成法を示した。付表 1 に死亡予測モデルの回帰パラメータ、ハザード比、10 年間のベースライン生存関数を男女別、疾患別に示し、付表 2 に死亡予測モデルに使用する集団平均値（男女別）を示した。

作成した循環器疾患（CVD,脳卒中、CHD）死亡予測モデルについて、グループ 2（大崎、大阪、JMS）に含まれる 3 つのコホートデータを用い、モデル妥当性を検証した。検討法として 10 年以内の CVD 死亡率を対象に、実測値とモデル予測値を比較した。具体的には各個人の予測死亡率を低い値から順に並べ、10 等分したグループを構成し、実際の死亡割合と予測値の平均とをグラフにより比較した。また上記の 10 グループの実測値と予測値の相関図を各コホートで確認し、相関係数を算出した。

（２）一般集団の死亡率予測と妥当性の検討

（１）で作成した死亡予測モデルを集団の健診結果が実際に存在する市町村に適用し、5 年間の累積死亡数を算出した。また予測された累積死亡数を実際の死亡者数と比較し、死亡予測モデルの妥当性を検討した。対象となる市（以下、当該市）の 2010 年の特定健診データから検査値および問診票の情報（集団平均値および割合）を入手した。これら情報を前年度報告書で検討した修正係数(SMR)を使用した予測モ

デルに投入し、集団における累積死亡率を算出した。研究で使用する予測モデルは以下の通りである。

$$\hat{p} = SMR \times (1 - S(t, \bar{x})^z)$$

ここで $S(t, \bar{x})$ はベースライン生存関数、 z は以下に示す共変量の線形結合となる；年齢、収縮期血圧、総コレステロール、糖尿病の有無、喫煙状況。今回の修正係数として 1990 年の EPOCH-JAPAN を基準集団とした標準化死亡比（SMR）を使用した。なお当該市の最新死亡数データが 2015 年であったため、推定期間を 5 年間に変更し、表 1 のように 5 年間のベースライン生存関数をデータより推定した。

解析方法として、予測された 5 年累積死亡率を 2010 年の当該市 40-80 歳人口に乗じることで、当該市における 5 年間の疾患別死亡数を算定した。なお当該市の 40-80 歳の死亡数は統計資料より直接知ることとはできないため、当該市の 2010-2014 年の疾患別死亡数（全年齢）に当該市が含まれる県の 40-80 歳死亡の占有割合を乗じることで算定した。

（倫理面への配慮）

本研究では匿名化されたデータを用いるため、個人情報保護に関する問題は生じない。「人を対象とした医学系研究に関する倫理指針」に基づいて実施し、資料の利用や管理などその倫理指針の原則を遵守した。また全体の研究計画については慶應義塾大学、データ管理についてはデータ管理機関である滋賀医科大学の倫理委員会の承認を得ている。

C. 研究結果

（１）EPOCH-JAPAN 循環器データベースを用いた死亡予測モデルの妥当性検証

図 1 から図 3 に、各コホートにおける

10年以内の死亡率と予測値との関連を性別、疾患別に示した（便宜上、各コホート名はA,B,Cとした）。この結果から実測値と予測値との関連は良好であり、相関係数も一部を除き0.95を超えていた。男女別の傾向の違いはないものの、疾患別では女性CHDではイベントが少ない影響で、相関図での関連が明瞭でないコホートがあった。

（２）一般集団の死亡率予測と妥当性の検討

表2に妥当性検討で用いた当該市の40-75歳の2010年の検査値等の平均値、割合を示した。表3の当該市の死亡予測モデルに必要な修正係数（SMR）をみると男性CVDを除きSMRの値は100%以下であり、特にCHDで低い値を示した（男性41.6%、女性16.7%）。

表4に当該市の5年間(2010-2014年)の疾患別死亡数と、40-80歳の推計死亡数を示した。40-80歳の死亡%を見ると、男性では約半数（CVD;53.8%、脳卒中;51.3%、CHD;64.4%）であるのに対し、女性では約2-3割（CVD;23.4%、脳卒中;23.9%、CHD;31.4%）と全死亡に占める割合が低かった。

表5に表4で算定した当該市の5年間の累積死亡数の実測値と、リスク予測モデルによる予測値との比較を示した。男性では実データと予測モデルの結果の比がCVD:83.7%、脳卒中:92.7%、CHD:71.4%と良好である一方、女性ではCVD:54.7%、脳卒中:54.8%と予測モデルの結果が観察値の半数程度となり、CHDでは19.8%と過小評価となっていた。

D. 考察

今回、EPOCH-JAPAN循環器データベースを用いた死亡予測モデルの妥当性の検証

を行い、その予測能が良好であることを示した。またその死亡予測モデルを一般集団に適用した場合の死者数予測を実施し、実施可能性を確認した。実測データと比較した結果、男性での死者数での妥当性が確認された。

EPOCH-JAPANを用いた死亡予測モデルの構築には1990年代スタートのコホートデータを用い、その妥当性検討には1995年スタートの3つのコホートを使用した。このようにモデル構築以外の外部コホートデータによる妥当性検証がなされたことは有意義であり、リスク予測モデルの信ぴょう性を高めたといえる。今回は10年以上の観察期間をもつEPOCH-JAPANの3コホート研究のデータを用いた。これは2000年代以降にスタートしたコホートデータでは観察期間が10年未満と短く、今回の検討に不適と判断したからである。EPOCH-JAPAN内のコホートデータの更新により、現在に近い時期の10年観察データでの妥当性が検証可能となると思われる、今後期待される。

今回の検討では、昨年度提案した県間の地域間差、疾患の時代的変遷等の要因を修正する、乗数を含んだ死亡予測モデルを実際の市町村データに用いた。人口動態統計とコホートデータとの比較から算出した係数に基づく修正に際しては、昨年度使用した直接法による比較法であるCMF(Comparative Mortality Figure)に代えて、間接法による較法のSMR(Standardized mortality ratio)を採用した。当該市では日本の全市町村と同様に、疾患別の年齢階級別死亡数の資料はない。そのため全年齢の死亡数を用いて修正係数を計算する方法としてSMRを使用した。CMFとSMRは基準集団と対象集団の年齢別死亡率で等比性が成り立つ場合、同様の数値になることが知られている。このこと

から今回 SMR を使用したことの問題は少ないと思われる。

死亡予測モデルを一般集団に適用した場合の死亡者の予測数を実際のデータと比較した結果、男性の CVD や脳卒中では 8,9 割の一致が見られたのに対し、CHD 特に女性では約 2 割と過小評価の結果となった。CHD の診断精度は市町村によりバラツクこと、リスク評価モデル作成時の投薬・治療状況と現在では変化があるなど他因子の影響は否定できない。また当該市の人口規模での CHD 死亡数は少ないため、その中での死亡数予測の限界が示されたといえる。今回の死亡予測モデルは県単位の集団におけるリスク評価を目的として開発されてきた。今後は県レベルの一般集団での妥当性評価も合わせて必要となると思われる。

E. 結論

EPOCH-JAPAN 循環器データベースを用いた死亡予測モデルのモデル妥当性は良好であり、一般集団に適用した場合の死亡者数予測の結果、実施可能性が示され、一部の妥当性が確認された。

F. 健康危機情報

なし

G. 研究発表

1. 論文発表

なし

2. 学会発表

Murakami Y, Okamura T, Miura K, Ueshima H, on behalf of the EPOCH-JAPAN research group. Development of a cardiovascular disease risk prediction model for population health planning in JAPAN: EPOCH-JAPAN STUDY. Society for Social Medicine and Population Health and International Epidemiology Association European Congress Annual Scientific Meeting (Cork, Ireland) 2019.

H. 知的財産権の出願・登録状況（予定を含む）

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

付録 本研究のリスク予測モデルの説明

1990 年前後に追跡を開始した大迫、吹田、放影研、久山の 4 つのコホートデータを用い、エンドポイント(死亡)を全循環器疾患（以下 CVD）、脳卒中、CHD とした、循環器疾患の死亡予測モデルを作成した。以下に男性 CVD を例に、死亡リスク予測モデルの作成法を説明する。

CVD男性		パラメータ	ハザード比	集団平均値
年齢		0.111	(1.12)	55.88
SBP		0.011	(1.01)	134.03
総コレステロール		0.001	(1.00)	199.02
糖尿病の有無		0.359	(1.43)	0.08
喫煙	禁煙	0.109	(1.12)	0.27
	現在喫煙	0.416	(1.52)	0.53

Cox モデルにより算出されたパラメータ値およびベースライン生存関数を用い、死亡確率 $\hat{p}$ は以下の式によって求められる。

$$\hat{p} = 1 - S(t, \bar{x})^Z$$

ここで $S(t, \bar{x})$ はベースライン生存関数、パラメータと共変量の線形結合の和を示す Z は以下のよ

$$z = \exp(\beta_1(\text{age} - \text{mean}(\text{age})) + \beta_2(\text{sbp} - \text{mean}(\text{sbp})) - \beta_3(\text{tchol} - \text{mean}(\text{tchl})) \\ + \beta_4(\text{dm} - \text{mean}(\text{dm})) + \beta_5(\text{quit} - \text{mean}(\text{quit})) + \beta_6(\text{smk} - \text{mean}(\text{smk})))$$

上記の説明（全てベースライン時の値）；年齢:age、収縮期血圧:sbp、総コレステロール:tchol、糖尿病の有無:dm、喫煙状況：（禁煙:quit、現在喫煙:smk）

上式から、男性における 10 年間の CVD 死亡の予測確率は次のようになる。

$$p = 1 - 0.9846^Z$$

ここでベースライン生存関数は 0.9846 であり、 Z は以下になる。

$$z = \exp(0.111 \times (\text{age} - 55.88) + 0.011 \times (\text{sbp} - 134.03) - 0.001 \times (\text{tchol} - 199.02) \\ + 0.359 \times (\text{dm} - 0.08) + 0.109 \times (\text{quit} - 0.27) + 0.416 \times (\text{smk} - 0.53))$$

今回使用した各疾患（CVD、脳卒中、CHD）の予測モデルに必要な、男女別パラメータ値、ベースライン生存関数、変数のセンタリングに使用する集団平均値を以下に示す。

付表 1 死亡予測モデルの回帰パラメータ、ハザード比、10 年間のベースライン生存関数

(男女別、CVD・脳卒中・CHD 別)

			男性			女性		
			パラメータ	ハザード比		パラメータ	ハザード比	
CVD	年齢		0.111	(1.12)		0.132	(1.14)	
	SBP		0.011	(1.01)		0.010	(1.01)	
	総コレステロール		0.001	(1.00)		-0.001	(1.00)	
	糖尿病の有無		0.359	(1.43)		0.445	(1.56)	
	喫煙	禁煙	0.109	(1.12)		0.393	(1.48)	
		現在喫煙	0.416	(1.52)		0.491	(1.63)	
	ベースライン生存関数(10年間)		0.9846	—		0.9954	—	
脳卒中	年齢		0.119	(1.13)		0.120	(1.13)	
	SBP		0.012	(1.01)		0.009	(1.01)	
	総コレステロール		-0.002	(1.00)		-0.006	(0.99)	
	糖尿病の有無		0.209	(1.23)		0.087	(1.09)	
	喫煙	禁煙	-0.085	(0.92)		0.321	(1.38)	
		現在喫煙	0.171	(1.19)		0.090	(1.09)	
	ベースライン生存関数(10年間)		0.9951	—		0.9963	—	
CHD	年齢		0.106	(1.11)		0.143	(1.15)	
	SBP		0.014	(1.01)		0.015	(1.01)	
	総コレステロール		0.008	(1.01)		0.005	(1.00)	
	糖尿病の有無		0.557	(1.75)		0.825	(2.28)	
	喫煙	禁煙	0.248	(1.28)		0.632	(1.88)	
		現在喫煙	0.541	(1.72)		0.304	(1.35)	
	ベースライン生存関数(10年間)		0.9954	—		0.9990	—	

付表 2 死亡予測モデルに使用する集団平均値 (男女別)

集団平均値		男性	女性
年齢		55.88	56.54
SBP		134.03	131.58
総コレステロール		199.02	211.42
糖尿病の有無		0.08	0.07
喫煙	禁煙	0.27	0.03
	現在喫煙	0.53	0.10

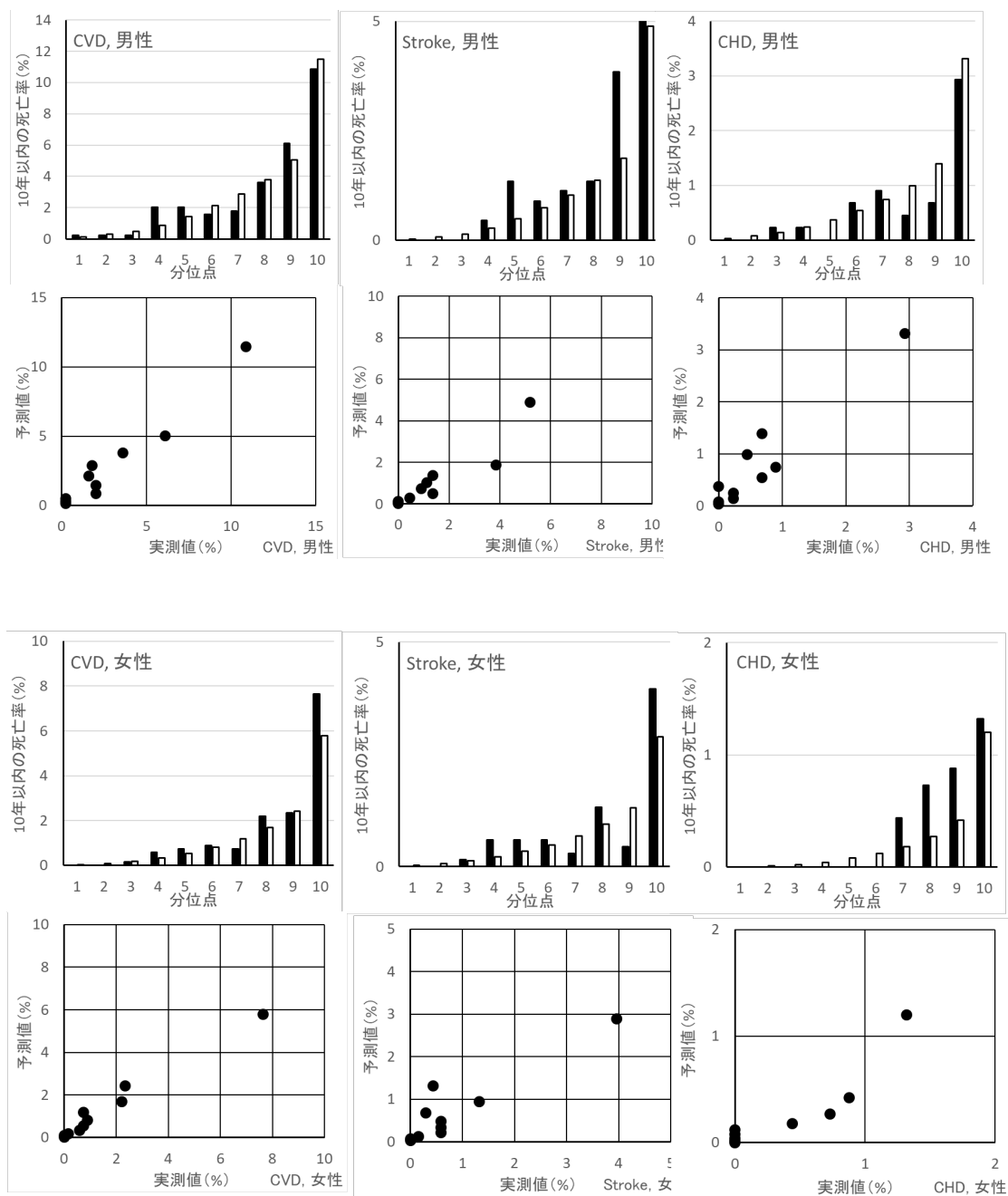


図1 コホート A における 10 年以内死亡率とその予測値との関連

一段：左から順に男性 CVD、脳卒中、CHD の 10 分位の実測値(黒)、予測値(白)の棒グラフ

二段：一段で示した実測値と予測値の間の相関図

三段：左から順に女性 CVD、脳卒中、CHD の 10 分位の実測値(黒)、予測値(白)の棒グラフ

四段：一段で示した実測値と予測値の間の相関図

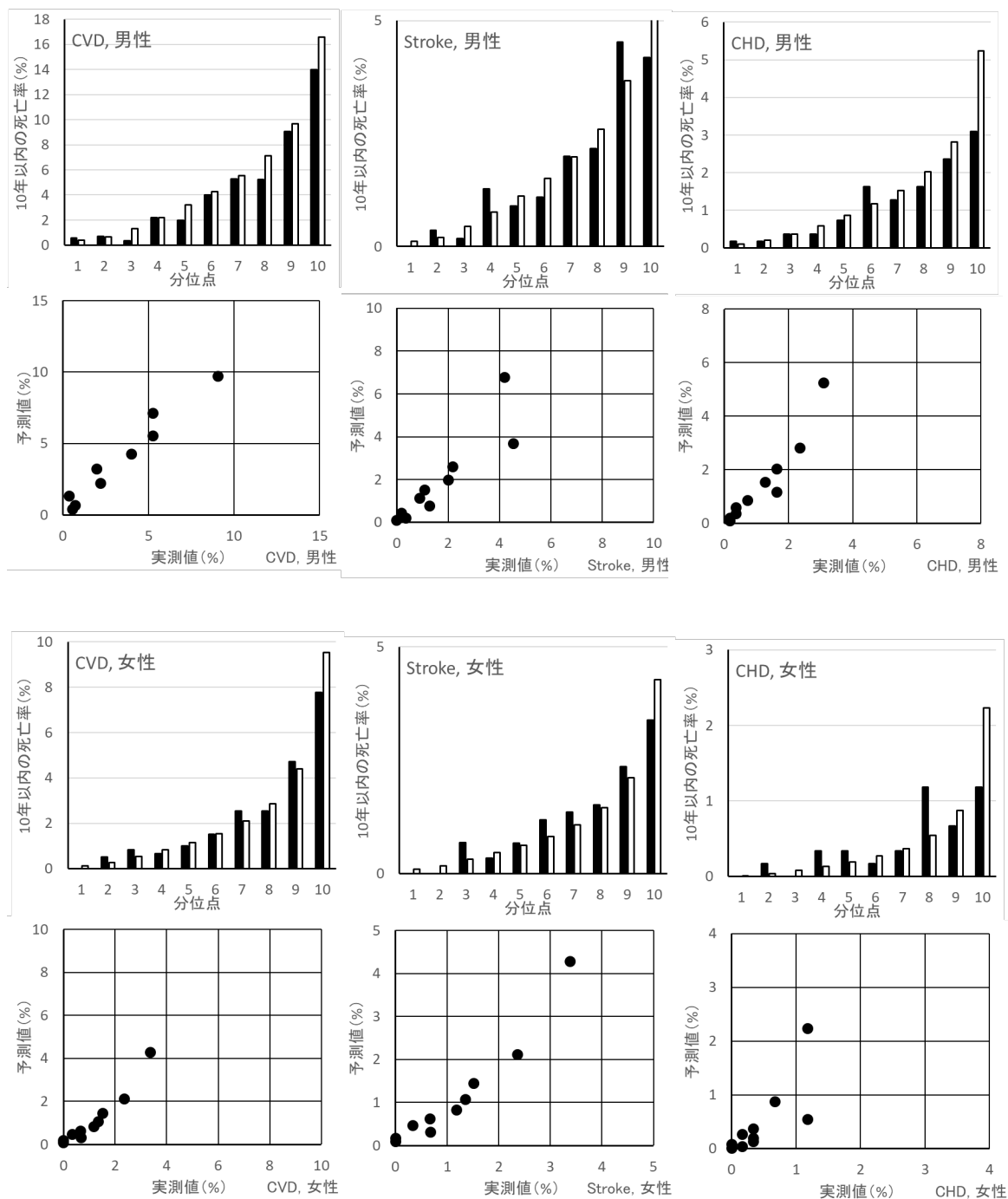


図2 コホート B における 10 年以内死亡率とその予測値との関連

一段：左から順に男性 CVD、脳卒中、CHD の 10 分位の実測値(黒)、予測値(白)の棒グラフ
二段：一段で示した実測値と予測値の間の相関図
三段：左から順に女性 CVD、脳卒中、CHD の 10 分位の実測値(黒)、予測値(白)の棒グラフ
四段：一段で示した実測値と予測値の間の相関図

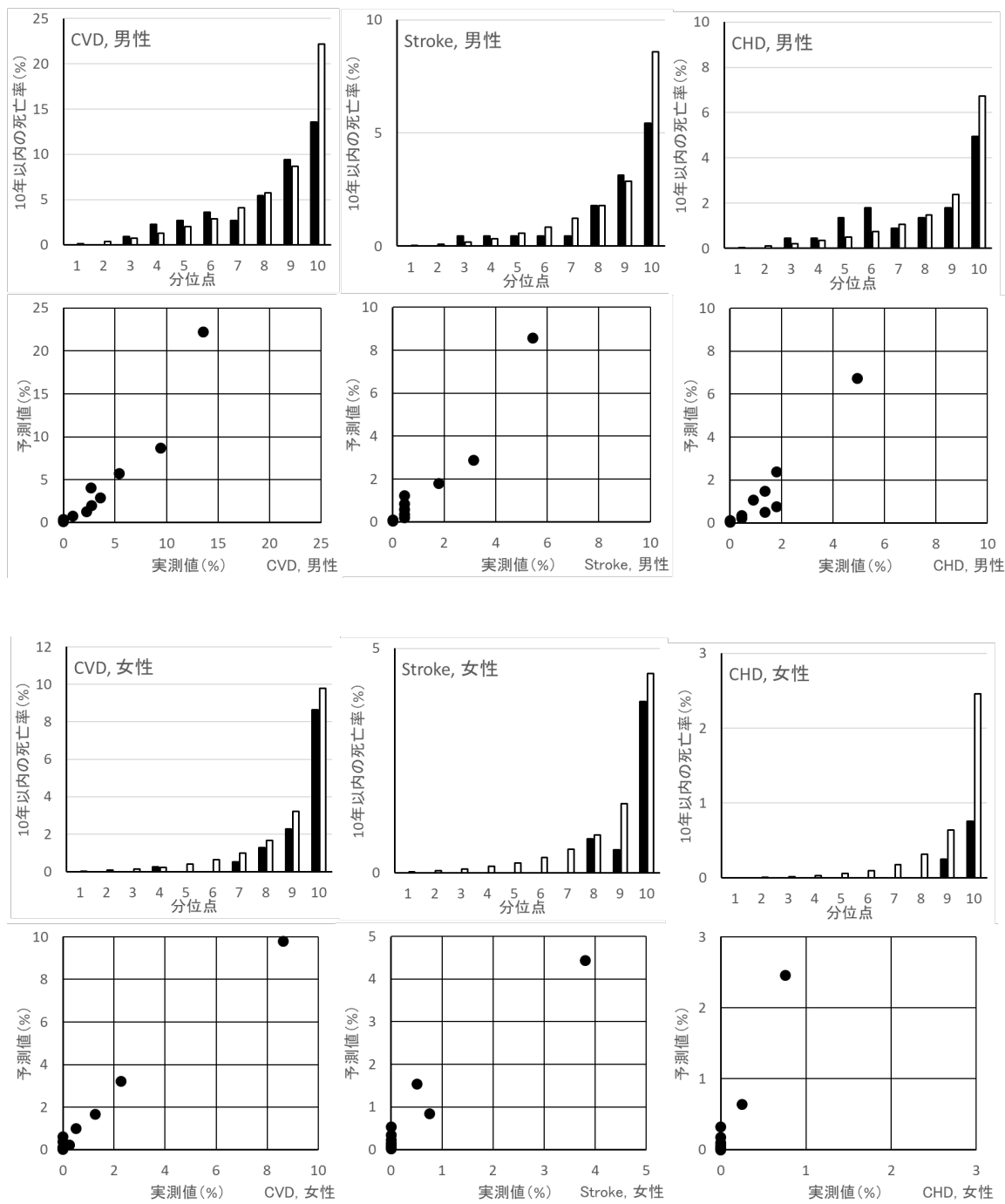


図3 コホート C における 10 年以内死亡率とその予測値との関連

- 一段：左から順に男性 CVD、脳卒中、CHD の 10 分位の実測値(黒)、予測値(白)の棒グラフ
- 二段：一段で示した実測値と予測値の間の相関図
- 三段：左から順に女性 CVD、脳卒中、CHD の 10 分位の実測値(黒)、予測値(白)の棒グラフ
- 四段：一段で示した実測値と予測値の間の相関図

表 1 死亡予測モデルの回帰パラメータ、ハザード比、5年間のベースライン生存関数
(男女別、CVD・脳卒中・CHD 別)

			男性			女性		
			パラメータ	ハザード比		パラメータ	ハザード比	
CVD	年齢		0.111	(1.12)		0.132	(1.14)	
	SBP		0.011	(1.01)		0.010	(1.01)	
	総コレステロール		0.001	(1.00)		-0.001	(1.00)	
	糖尿病の有無		0.359	(1.43)		0.445	(1.56)	
	喫煙	禁煙	0.109	(1.12)		0.393	(1.48)	
		現在喫煙	0.416	(1.52)		0.491	(1.63)	
	ベースライン生存関数(5年間)		0.9945	—		0.9978	—	
脳卒中	年齢		0.119	(1.13)		0.120	(1.13)	
	SBP		0.012	(1.01)		0.009	(1.01)	
	総コレステロール		-0.002	(1.00)		-0.006	(0.99)	
	糖尿病の有無		0.209	(1.23)		0.087	(1.09)	
	喫煙	禁煙	-0.085	(0.92)		0.321	(1.38)	
		現在喫煙	0.171	(1.19)		0.090	(1.09)	
	ベースライン生存関数(5年間)		0.9982	—		0.9986	—	
CHD	年齢		0.106	(1.11)		0.143	(1.15)	
	SBP		0.014	(1.01)		0.015	(1.01)	
	総コレステロール		0.008	(1.01)		0.005	(1.00)	
	糖尿病の有無		0.557	(1.75)		0.825	(2.28)	
	喫煙	禁煙	0.248	(1.28)		0.632	(1.88)	
		現在喫煙	0.541	(1.72)		0.304	(1.35)	
	ベースライン生存関数(5年間)		0.9984	—		0.9997	—	

表 2 妥当性検討に使用した市データの検査値等の平均値・割合 (2010 年、40-75 歳)

		男性	女性
年齢(歳)	平均値	64.7	64.9
収縮期血圧(mmHg)	平均値	131.2	128.0
総コレステロール(mg/dl)	平均値	201.1	210.8
DM(%)	%	14.1	6.9
現在喫煙(%)	%	26.6	7.1

表 3 当該市に死亡予測モデルに必要な修正係数(SMR)

40-79歳	男性			女性		
	SMR	観察値	期待値	SMR	観察値	期待値
CVD	113.0	82.1	72.7	87.7	38.2	43.6
脳卒中	88.9	22.7	25.6	43.6	11.2	25.6
CHD	41.6	7.8	18.8	16.7	3.7	22.0

観察値：当該市の 2009-2011 年の死亡数の平均に、当該市がある県の 40-80 歳の死亡割合を乗じた値。

期待値：EPOCH-JAPAN から計算した 40-49、50-59、60-69、70-79 歳死亡率に、当該市の 2010 年人口を年齢カテゴリ別に乗じた値。

表 4 当該市の 2010-2014 年の疾患別死亡数と 40-80 歳の推計死亡数
(男女別、CVD・脳卒中・CHD 別)

総死亡数(人)	男性			女性		
	CVD	脳卒中	CHD	CVD	脳卒中	CHD
2010年	140	52	17	154	39	11
2011年	160	37	12	188	60	8
2012年	152	49	25	156	52	14
2013年	114	30	12	166	34	10
2014年	153	37	4	166	41	6
5年間の合計	719	205	70	830	226	49
40-80歳の死亡% *	53.8	51.3	64.4	23.4	23.9	31.4
5年間の累積	387.1	105.1	45.1	193.9	54.0	15.4
死亡数(40-80歳)						

*; 2015 年の県全体のデータより推計

表 5 当該市の 5 年間(2010-14 年)の累積死亡数の実測値とモデル予測値との比較
(男女別、CVD・脳卒中・CHD 別)

40-80歳人口	男性(28357人)			女性(31693人)		
(2010年)	CVD	脳卒中	CHD	CVD	脳卒中	CHD
実データ	387.1	105.1	45.1	193.9	54.0	15.4
予測モデルによる	324.2	97.4	32.2	106.4	29.6	3.0
推定死亡者数						
実データとの比 (%)	83.7	92.7	71.4	54.9	54.8	19.8