

REAL-CAD 患者（冠動脈疾患患者）における心血管イベント発生リスク予測モデルの構築（Developing a risk prediction model for cardiovascular events in REAL-CAD patients）

【はじめに】

前回発表の指摘を受け、REAL-CAD 試験のイベント発生数を予測するために、リスク予測モデルの方法を使うと考えました。今回の抄読会発表はリスク予測モデルの心血管領域の応用状況を紹介します。また、Framingham Heart Study の teaching dataset を使い、いくつかのリスク予測モデルを作成しましたので、その結果を報告させていただきます。

【心血管領域におけるリスク予測モデルの応用状況】

心血管領域におけるリスク予測モデルの開発は、近年盛んな研究分野の 1 つです。Framingham¹ SCORE² ASSIGN³ EUROSCORE⁴ PROCAM⁵ 予測モデルがよく知られているかと思います。日本にも久山町研究の新しい心血管疾患リスク予測モデル⁶、JALS 研究の non-HDL-C、総コレステロールを用いた急性心筋梗塞発症リスクスコア⁷、JPHC 研究⁸の 7つのリスク因子を用いた 10年間の脳卒中発症確率の予測モデルが存在しております。しかし、このリスク予測モデルを含め、いままでの多くの心血管疾患リスク予測モデルは一般集団（general population）に基づいて心筋梗塞や脳卒中の初発リスク(primary risk)を予測するために作られました。REAL-CAD 対象者ようなハイリスク患者（急性心筋梗塞または不安定狭心症の既往のある患者 or PCI または CABG を施行された患者 or 高度狭窄を認める患者）に初発リスク予測モデルは適用できるかどうか疑問です。また、ハイリスク患者を対象し、心血管イベントの再発リスク(subsequent risk)を探索する研究の数は多いとは言えません。その研究の中、初発と再発リスクを両方研究した Framingham⁹ と LIPID¹⁰ スタディで、初発と同じ、年齢、TC、HDL-C、SBP、糖尿病の有無、喫煙は再発のリスク因子として挙げられました。バイオマーカーでは、apoB/apoA-1 は LDL-C と non-HDL-C より強い鬱血性心不全の予測因子と二次予防中の冠動脈疾患患者で示されました¹¹。その他、C 反応性蛋白¹²、左室駆出率¹³も再発の予測因子と報告されました。

【リスク予測モデルの構築指針（一部）】^{14 15}

候補予測因子(Candidate predictors) : theoretically, all potential and not necessarily causal correlates of the outcome of interest. Commonly, however, pre-selection based on subject matter knowledge is recommended. Similar to the outcomes, candidate predictors are clearly defined and measured in a standardized and reproducible way.

欠損値(Missing values) : analysis of individuals with only completely observed data may lead to biased results. Imputation, preferably multiple imputation, of missing values often yields less biased results.

連続予測因子(Continuous predictors) : should not be turned into dichotomies and linearity should not be assumed. Simple predictor transformation can be implemented to detect and model non-linearity, increasing the predictive accuracy of the prediction model.

予測因子の選択(Predictor selection in the multivariable modelling) : selection based on univariable analysis (single predictor-outcome associations) is discouraged. Preferably, if needed, backwards selection or a full model approach should be used, depending on a priori knowledge.

予測パフォーマンスの評価(Model performance measures) : discrimination (eg, c-index), calibration (plots), and (re)classification measures.

内部バリデーション(Internal validation) : bootstrapping techniques can quantify the model's potential for overfitting, its optimism in estimated model performance measures and a shrinkage factor to adjust for this optimism.

外部バリデーション(External validation) : external validation of any type consists of taking the original model or simplified score, with its predictors and assigned weights (eg, regression coefficients), as estimated from the development study; obtaining the measured predictor and outcome values in the new individuals; applying the original model to these data; and quantifying the model's predictive performance.

【Worked sample】

Framingham Heart Study teaching dataset (requested on Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC) website) を用い、first HCHD(Hard coronary heart disease, myocardial infarction or coronary death)発生数を推定するポアソン回帰モデルとロジスティック回帰モデルを作成しました。

ポアソン回帰の場合、i 人 Year 年後イベント発生率 (event rate)の計算式：

$$p(i) = \text{Year} * \exp(\alpha + \beta X_i)$$

ロジスティック回帰の場合、i 人 Year 年後イベント発生確率の計算式：

$$p(i) = \text{Year} * \exp(\alpha + \beta X_i) / (1 + \text{Year} * \exp(\alpha + \beta X_i))$$

Year----追跡期間

β -----回帰係数

X_i -----説明変数

【REAL-CAD への応用課題】

- ü 男・女別 or 患者別（急性心筋梗塞・不安定狭心症・PCI・CABG）の予測モデル？
- ü 予測因子のカテゴリ化
- ü Missing data imputation
- ü Time-term: one year? Six months
- ü And so on