大規模データベース研究におけるベイズ流潜在クラスモデルを用いた疾患アウトカム定義の情報バイアスの分析方法

宇野　慧　　総合研究大学院大学　統計科学専攻　博士課程（3年次編入学）4年

1. 本研究の背景

近年、診療報酬の請求情報（レセプト）や、電子カルテなどの医療情報をもとにした大規模データベース（DB）を用いた疫学研究が増えている。これらのDB研究から得られたエビデンスは、公衆衛生や日常診療、政策立案などに広く用いられている。一方で、多くのDBは詳細な臨床や病理の情報を含んでおらず、そこから定義した疾患の発生や薬剤の使用、疾患の重症度などの情報に誤分類が生じる場合があり、疫学研究を立案・実施するうえでの大きな懸念となる。このような懸念に対処するために、疾患の発生等を定義するためのアルゴリズムの妥当性を評価する取り組みがなされており、一般的にバリデーション研究と呼ばれている。典型的なバリデーション研究は以下のように実施される。まず、診断、手術、臨床検査、治療およびそれらの組み合わせなど、DBから疾患アウトカムを定義するために、複数の定義の候補を作成する。次に、ゴールドスタンダード（GS）と呼ばれる、確度の高い定義を特定して参照に用いる。次に、前述のDBに基づく複数の疾患定義とGSとを比較し、感度、特異度、陽性的中度などの指標をもとに妥当性を定量的に評価する。 こうした一連のプロセスにより、最善と判断された疾患定義を疫学研究に用いることになる。

バリデーション研究の多くは、参照に用いるGSとして、カルテ情報や疾患レジストリの登録情報（登録されていれば当該疾患を発症したと考える）を用いているが、これらの情報に誤分類が生じている可能性は考慮していないことが多い。参照元となる情報に誤分類がある場合、これに基づき評価した疾患定義の妥当性にはバイアスが生じ、さらにこの疾患定義を用いた疫学研究の結果にもバイアスをもたらす可能性がある。こうした問題に対処するため、本研究ではバリデーション研究における疾患定義の妥当性を定量的に評価するための新しいアプローチとして、潜在クラスモデル（Latent Class Model）をもとにした手法を検討している。潜在クラスモデルは真の状態が未知のクラス変数（本研究では疾患あり/なしの状態をあらわす2値変数）をモデル化するための手法であり、様々な発展的手法が提案されている。しかしながら、本研究で対象としているようなバリデーション研究にこうした手法を適用した事例は稀であり、応用面における性質などは十分に検討されていない。そのため、本研究では、Dendukuri & Josephの研究 [1] で提案されたベイズ流の潜在クラスモデルを、バリデーション研究に応用する新しい取り組みを行った。具体的には、Satoらの研究論文[2]の集計表から3つの疾患定義の頻度データを再現し、これらのデータに対してベイズ流潜在クラスモデルを適用した。なお、本研究は丹後俊郎氏（医学統計学研究センター）との共同研究である。

2. 潜在クラスモデルの概要

以下では、参照用のGSとして疾患レジストリの登録情報（レジストリ定義）を用い、DB情報に基づく疾患定義（DB定義）と比較する状況を検討する。レジストリ定義、およびDB定義のいずれも疾患有無の判定結果は＋（陽性）、－（陰性）の2値で評価し、結果をT_jであらわす。ただし $j(=1,\dots,J)$ は疾患有無を判定するための各種の定義jをあらわす添え字である。$T_j=1$ が陽性、$T_j=0$ が陰性を表す。次に、真の疾患有無の状態（観測不能）をあらわすための変数Dを導入する。$D=1$ は疾患あり、$D=0$は疾患なしをあらわす。真の疾患有無は直接観察できず、利用可能なデータ等から推測することになるため、母集団での当該疾患の有病割合の変数$\pi=\mathrm{P}(D=1)$も導入する。以上の変数をもとに、DB定義の妥当性を判断するための定量的指標を以下の通り定義することができる。

・感度：　真の状態が「疾患あり」の場合に、各種定義jで「陽性」と判定される確率　（$S_j=\mathrm{P}(T_j=1|D=1)$）

・特異度：　真の状態が「疾患なし」の場合に、各種定義jで「陰性」と判定される確率　（$C_j=\mathrm{P}(T_j=0|D=0)$）

・陽性的中度：　各種定義jで「陽性」と判定された場合に、真の状態が「疾患あり」である確率　（$\mathrm{P}(D=1|T_j=1)=\pi S_j/(\pi S_j+(1-\pi)(1-C_j))$）

・陰性的中度：　各種定義jで「陰性」と判定された場合に、真の状態が「疾患なし」である確率　（$\mathrm{P}(D=0|T_j=0)=(1-\pi)C_j/(\pi(1-S_j)+(1-\pi)C_j)$）

以上の設定のもとで、各種の判定定義において「陽性」と判定される確率は、以下の(1)式のとおり有病割合と感度、特異度を用いてあらわすことができる。

$$\mathrm{P}(T_j=1)=\pi\mathrm{P}(T_j=1|D=1)+(1-\pi)\mathrm{P}(T_j=1|D=0) \qquad (1)$$

さらに、本研究では複数の判定定義が相関をもつ状況を想定し、2つの判定の組み合わせ（$l\ (=1,\dots,J)$ および $h(>l,\ =2,\dots,J)$）について感度、特異度それぞれ共分散パラメータ（$\mathrm{Cov}(S_l,S_h)$ および $\mathrm{Cov}(C_l,C_h)$）を導入する。この設定のもとで、2つの判定定義いずれでも「陽性」と判定される（同時）確率は以下の(2)式であらわすことができる。3つ以上の判定定義を用いる場合も同様に拡張でき、本研究では3つの判定定義を扱う状況を想定し、尤度関数（最適化のための目的関数）を構成している。実際の推定においては「判定定義間の相関の有／無」、および「レジストリ定義をGSと見做すか否か」をもとに4パターンのモデルを設定し、各パラメータに無情報事前分布を用いたベイズ推定を行い、結果を比較した。

$$\begin{aligned}\mathrm{P}(T_l=1,T_h=1)=\pi\{&\mathrm{P}(T_l=1|D=1)\mathrm{P}(T_h=1|D=1)+(-1)^{|t_l-t_h|}\mathrm{Cov}(S_l,S_h)\}\\&+(1-\pi)\{\mathrm{P}(T_l=1|D=0)\mathrm{P}(T_h=1|D=0)+(-1)^{|t_l-t_h|}\mathrm{Cov}(C_l,C_h)\}\end{aligned} \qquad (2)$$

3. 事例データと推定結果

本研究では、Satoらが実施した乳がんの疾患アウトカムについてのバリデーション研究[2]の事例を用いた。当該研究は聖路加国際病院の院内がんレジストリ情報をGSと仮定したうえで、DPC情報（DBの一種）から定義した14種類の方法を評価している。本研究ではこのうち、集計結果から頻度データが再現可能な一部の疾患定義（2種類のDB定義、および1種類のレジストリ定義）について検討した。分析の結果は以下の表1に示す。概要としては、判定定義間の相関を考慮し、かつレジストリ定義をGSと見做さない最も柔軟なモデルが、最も良好な予測性能を示した。他のモデルと比べ、有病割合はやや高く、また感度は顕著に低い推定値が得られた。なお、本研究ではシミュレーションによる性能評価も実施しており、適合度や情報量規準の指標において、最も柔軟なモデルが良好な性能を示した。

4. 参考文献

[1]. Dendukuri N, Joseph L. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics* 2001;57(1):158-67.

[2]. Sato I, Yagata H, Ohashi Y. The accuracy of Japanese claims data in identifying breast cancer cases. *Biol Pharm Bull* 2015;38(1):53-7.

図1. 疫学研究に用いる医療情報データ、およびバリデーション研究のイメージ

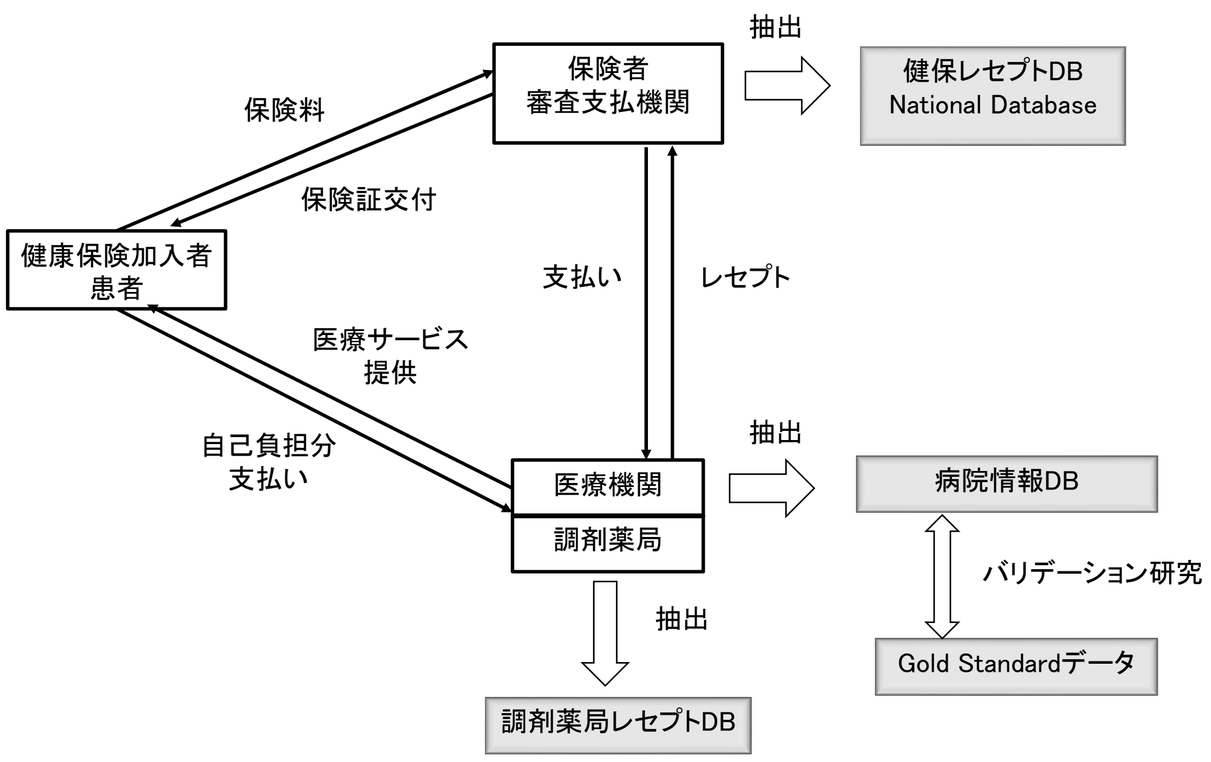

表1. 推定結果の抜粋 (モデル1～4は事後分布の平均値)

		判定定義間の相関なし		判定定義間の相関あり	
		GS	Non-GS	GS	Non-GS
	Sato et al. [2]	モデル 1	モデル 2	モデル 3	モデル 4
有病割合	0.013	0.013	0.014	0.014	0.016
感度$_{DB1}$	0.987	0.986	0.998	0.928	0.925
感度$_{DB2}$	0.904	0.902	0.914	0.855	0.752
感度$_{レジストリ}$	—	—	0.872	—	0.720
特異度$_{DB1}$	0.993	0.993	0.995	0.994	0.996
特異度$_{DB2}$	0.998	0.998	1.000	0.999	0.999
特異度$_{レジストリ}$	—	—	1.000	—	0.999
陽性的中度$_{DB1}$	0.658	0.657	0.753	0.689	0.779
陽性的中度$_{DB2}$	0.873	0.872	0.998	0.933	0.908
陽性的中度$_{レジストリ}$	—	—	0.986	—	0.901
陰性的中度度$_{DB1}$	1.000	1.000	1.000	0.999	0.999
陰性的中度$_{DB2}$	0.999	0.999	0.999	0.998	0.996
陰性的中度$_{レジストリ}$	—	—	0.998	—	0.995
WAIC	—	12454	11536	11618	11536