数理解析研究所講究録
第 1910 巻 2014 年 1-19

変量効果モデルを用いた複数の読影者による画像診断法の精度の推定

佐伯　浩之[1,3]、丹後　俊郎[2]、汪 金芳[1]
[1] 千葉大学大学院 理学研究科
[2] 医学統計学研究センター
[3] 富士フイルムＲＩファーマ株式会社

Hiroyuki Saeki[1,3], Toshiro Tango[2] and Jinfang Wang[3]
[1]Graduate School of Science, Chiba University
[2]Center for Medical Statistics
[3]Development Department, FUJIFILM RI Pharma Co., Ltd.

1　はじめに

ある疾患を検査する診断法の性能を示す指標として、真に疾患である被験者において診断法が陽性判定を示す確率である「感度 (sensitivity)」と、真に疾患ではない被験者において診断法が陰性を示す確率である「特異度 (Specificity)」がある。画像診断法を考えた場合は、診断法から得られた画像によって疾患の有無を評価するため、読影者が必要となる。特に、画像診断法の臨床試験では読影結果の再現性を確認することを目的として、複数の読影者が同一の画像を独立に評価する。(Lehr and Kashanian, 2009)。

従来、複数の読影者から得られた結果については、合議や多数決を行うことにより単独の読影者による結果として取り扱われていたが、合議は非独立の評価であるためにバイアスを生じる可能性があり、多数決は複数の読影者による結果のバラツキを考慮できない

〒104-0031 東京都中央区京橋 2-14-1

ため、これらの方法は主要な評価に対して推奨されていない (FDA; 2004、Obuchowski and Liber; 2008、CHMP; 2009)。このような課題に対して、Saeki and Tango (2011) は複数の読影者に基く相関のある割合の差の推論のための非劣性検定、信頼区間及び標本サイズの推定方法を提案している。しかしながら、複数の読影者から得られた感度と特異度を要約するための適切な方法については、これまでに提案されていない。そこで、本研究では複数の読影者により評価された診断方法の感度及び特異度の要約、並びに 95% 信頼区間の推定方法を導出することを目的とする。

2 データ構造と統計モデル

ある種の疾患 D に適用した診断法 M の性能を示す指標として、感度 p と特異度 q は以下のように定義される。

$$p = P(+|D) \tag{1}$$

$$q = P(-|\bar{D}) \tag{2}$$

ここで、診断法 M を想定される母集団 $\mathcal{P}$ から無作為に抽出された被験者に適用し、得られたデータを O とする。画像診断法を考えた場合には、O は無作為に抽出された被験者に画像診断法 M を実施して得られた画像データと解釈できる。複数の読影者の母集団 $\mathcal{R}$ から、無作為に K 名の読影者を抽出し、$R_1, \cdots, R_K$ とする。そして、読影者 R_k が画像データ O を読影し、陽性と判定した場合には $Y_k = +$、陰性と判定した場合には $Y_k = -$ を出力する。したがって、Y_k はベルヌーイ分布に従い、陽性及び陰性判定の確率変数として以下のように表せる。

$$p_k = P(Y_k = +) = P(+|D, R_k), \quad k = 1, \cdots, K \tag{3}$$

$$q_k = P(Y_k = -) = P(-|\bar{D}, R_k), \quad k = 1, \cdots, K \tag{4}$$

以降では、最も単純な状況として読影者 2 名での判定結果を対象とする。読影者 2 名での感度及び特異度の構造を表 1 に示す。表 1 より、読影者 1 及び 2 の感度と特異度は以

表 1 読影者 2 名における感度と特異度の構造.

	Judgment of (Rater1, Rater2)				
	(+,+)	(+,−)	(−,+)	(−,−)	Total
D	p_{++}	p_{+-}	p_{-+}	p_{--}	1
$\bar{D}$	q_{++}	q_{+-}	q_{-+}	q_{--}	1

下のように示される。

$$
\begin{aligned}
p_1 &= p_{++} + p_{+-} \\
p_2 &= p_{++} + p_{-+} \\
q_1 &= q_{--} + q_{-+} \\
q_2 &= q_{--} + q_{+-}
\end{aligned}
$$

読影者 1 及び 2 が同一の画像データ O を読影した結果得られる観測感度を $\hat{p}_1$ と $\hat{p}_2$、観測特異度を $\hat{q}_1$ と $\hat{q}_2$ とし、これらをロジット変換したものを各々 y_{a1}、y_{a2} 及び y_{b1}、y_{b2} として検討を進める。

まず最初に、fixed effects model (FEM) を考える。ここでは、2 名の読影者が同一の被験者集団 D 又は $\bar{D}$ から得られた画像データ O を評価するため、被験者集団内での相関を考慮する必要がある。従って、感度の推定を目的とした以下の FEM を定義する。

$$
\begin{aligned}
&\begin{pmatrix} y_{a1} \\ y_{a2} \end{pmatrix} = \begin{pmatrix} \mu_a \\ \mu_a \end{pmatrix} + \begin{pmatrix} \epsilon_{a1} \\ \epsilon_{a2} \end{pmatrix} \\
&\begin{pmatrix} \epsilon_{a1} \\ \epsilon_{a2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} , \begin{pmatrix} \sigma_{a1}^2 & \sigma_{a12} \\ \sigma_{a12} & \sigma_{a2}^2 \end{pmatrix} \right)
\end{aligned}
\tag{5}
$$

ここで、μ_a はロジット変換された真の感度、ϵ_{a1}、ϵ_{a2} は、平均が 0、分散が σ_{a1}^2 と σ_{a2}^2、共分散が σ_{a12} の多変量正規分布に従うと仮定する。なお、特異度の FEM は、式 (5) の y_{a1}、y_{a2}、μ_a、ϵ_{a1}、ϵ_{a2}、σ_{a1}^2、σ_{a2}^2、σ_{a12} の各々を、y_{b1}、y_{b2}、μ_b、ϵ_{b1}、ϵ_{b2}、σ_{b1}^2、σ_{b2}^2、σ_{b12} で置き換えることにより定義する。

次に、univariate random effects model (UVRM) を考える。θ_{a1} 及び θ_{a2} を読影者 1 及び読影者 2 が読影した場合のロジット変換された真の感度として以下のモデルを定義する。

$$
\begin{aligned}
&\begin{pmatrix} y_{a1} \\ y_{a2} \end{pmatrix} = \begin{pmatrix} \theta_{a1} \\ \theta_{a2} \end{pmatrix} + \begin{pmatrix} \epsilon_{a1} \\ \epsilon_{a2} \end{pmatrix} \\
&\begin{pmatrix} \epsilon_{a1} \\ \epsilon_{a2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} , \begin{pmatrix} \sigma_{a1}^2 & \sigma_{a12} \\ \sigma_{a12} & \sigma_{a2}^2 \end{pmatrix} \right)
\end{aligned}
\tag{6}
$$

さらに、θ_{a1}、θ_{a2} に対して、読影者間変動を考慮した以下のモデルを定義する。

$$
\begin{aligned}
&\begin{pmatrix} \theta_{a1} \\ \theta_{a2} \end{pmatrix} = \begin{pmatrix} \mu_a \\ \mu_a \end{pmatrix} + \begin{pmatrix} e_{a1} \\ e_{a2} \end{pmatrix} \\
&\begin{pmatrix} e_{a1} \\ e_{a2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} , \begin{pmatrix} \tau_a^2 & 0 \\ 0 & \tau_a^2 \end{pmatrix} \right)
\end{aligned}
\tag{7}
$$

ここで、e_{a1} 及び e_{a2} は分散が τ_a^2 の多変量正規分布に従うと仮定する。以上で定義した式 (6) 及び (7) に基づき、以下の UVRM を定義する。

$$\begin{pmatrix} y_{a1} \\ y_{a2} \end{pmatrix} = \begin{pmatrix} \mu_a \\ \mu_a \end{pmatrix} + \begin{pmatrix} \epsilon_{a1}^* \\ \epsilon_{a2}^* \end{pmatrix} \tag{8}$$
$$\begin{pmatrix} \epsilon_{a1}^* \\ \epsilon_{a2}^* \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} , \begin{pmatrix} \sigma_{a1}^2 + \tau_a^2 & \sigma_{a12} \\ \sigma_{a12} & \sigma_{a2}^2 + \tau_a^2 \end{pmatrix} \right)$$

なお、特異度の UVRM は、式 (8) の y_{a1}、y_{a2}、μ_a、ϵ_{a1}^*、ϵ_{a2}^*、σ_{a1}^2、σ_{a2}^2、σ_{a12}、τ_a^2 の各々を、y_{b1}、y_{b2}、μ_b、ϵ_{b1}^*、ϵ_{b2}^*、σ_{b1}^2、σ_{b2}^2、σ_{b12}、τ_b^2 で置き換えることにより定義する。

さらに、画像診断の評価では同一の読影者が感度と特異度の評価を行うため、読影者内における感度と特異度の相関を考慮する必要がある。ここで、τ_{ab} は同一の読影者によって評価された感度と特異度の共分散とする。

$$\begin{pmatrix} \theta_{a1} \\ \theta_{b1} \end{pmatrix} = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} + \begin{pmatrix} e_{a1} \\ e_{b1} \end{pmatrix} \tag{9}$$
$$\begin{pmatrix} e_{a1} \\ e_{b1} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} , \begin{pmatrix} \tau_a^2 & \tau_{ab} \\ \tau_{ab} & \tau_b^2 \end{pmatrix} \right)$$

$$\begin{pmatrix} \theta_{a2} \\ \theta_{b2} \end{pmatrix} = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} + \begin{pmatrix} e_{a2} \\ e_{b2} \end{pmatrix} \tag{10}$$
$$\begin{pmatrix} e_{a2} \\ e_{a2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} , \begin{pmatrix} \tau_a^2 & \tau_{ab} \\ \tau_{ab} & \tau_b^2 \end{pmatrix} \right)$$

以上で定義したモデルの式 (6)、(7)、(9)、(10) に基づき、以下の bivariate random effects model (BVRM) を定義する。

$$\begin{pmatrix} y_{a1} \\ y_{a2} \\ y_{b1} \\ y_{b2} \end{pmatrix} = \begin{pmatrix} \mu_a \\ \mu_a \\ \mu_b \\ \mu_b \end{pmatrix} + \begin{pmatrix} \epsilon_{a1}^{**} \\ \epsilon_{a2}^{**} \\ \epsilon_{b1}^{**} \\ \epsilon_{b2}^{**} \end{pmatrix} \tag{11}$$
$$\begin{pmatrix} \epsilon_{a1}^{**} \\ \epsilon_{a2}^{**} \\ \epsilon_{b1}^{**} \\ \epsilon_{b2}^{**} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} , \begin{pmatrix} \sigma_{a1}^2 + \tau_a^2 & \sigma_{a12} & \tau_{ab} & 0 \\ \sigma_{a12} & \sigma_{a2}^2 + \tau_a^2 & 0 & \tau_{ab} \\ \tau_{ab} & 0 & \sigma_{b1}^2 + \tau_b^2 & \sigma_{b12} \\ 0 & \tau_{ab} & \sigma_{b12} & \sigma_{b2}^2 + \tau_b^2 \end{pmatrix} \right)$$

3 感度と特異度の推定

本章では、前章で定義した FEM、UVRM 及び BVRM に基いた感度と特異度の推定方法を考える。本研究では、医学研究において複数の臨床研究の結果を統合する手法であ

る、メタアナリシスの考えを採用する。ここで、式 (5)、(8) 及び (11) の σ_{a1}^2、σ_{a2}^2、σ_{a12}、及び σ_{b1}^2、σ_{b2}^2、σ_{b12} を既知として、データから求めた推定量 $\hat{s}_{a1}^2$、$\hat{s}_{a2}^2$、$\hat{s}_{a12}$、及び $\hat{s}_{b1}^2$、$\hat{s}_{b2}^2$、$\hat{s}_{b12}$ を仮定する。

$$\begin{pmatrix} y_{a1} \\ y_{a2} \\ y_{b1} \\ y_{b2} \end{pmatrix} = \begin{pmatrix} \mu_a \\ \mu_a \\ \mu_b \\ \mu_b \end{pmatrix} + \begin{pmatrix} \epsilon_{a1} \\ \epsilon_{a2} \\ \epsilon_{b1} \\ \epsilon_{b2} \end{pmatrix} \tag{12}$$

$$\begin{pmatrix} \epsilon_{a1} \\ \epsilon_{a2} \\ \epsilon_{b1} \\ \epsilon_{b2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \hat{s}_{a1}^2 & \hat{s}_{a12} & 0 & 0 \\ \hat{s}_{a12} & \hat{s}_{a2}^2 & 0 & 0 \\ 0 & 0 & \hat{s}_{b1}^2 & \hat{s}_{b12} \\ 0 & 0 & \hat{s}_{b12} & \hat{s}_{b2}^2 \end{pmatrix} \right)$$

$$\begin{pmatrix} y_{a1} \\ y_{a2} \\ y_{b1} \\ y_{b2} \end{pmatrix} = \begin{pmatrix} \mu_a \\ \mu_a \\ \mu_b \\ \mu_b \end{pmatrix} + \begin{pmatrix} \epsilon_{a1}^* \\ \epsilon_{a2}^* \\ \epsilon_{b1}^* \\ \epsilon_{b2}^* \end{pmatrix} \tag{13}$$

$$\begin{pmatrix} \epsilon_{a1}^* \\ \epsilon_{a2}^* \\ \epsilon_{b1}^* \\ \epsilon_{b2}^* \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \hat{s}_{a1}^2 + \tau_a^2 & \hat{s}_{a12} & 0 & 0 \\ \hat{s}_{a12} & \hat{s}_{a2}^2 + \tau_a^2 & 0 & 0 \\ 0 & 0 & \hat{s}_{b1}^2 + \tau_b^2 & \hat{s}_{b12} \\ 0 & 0 & \hat{s}_{b12} & \hat{s}_{b2}^2 + \tau_b^2 \end{pmatrix} \right)$$

$$\begin{pmatrix} y_{a1} \\ y_{a2} \\ y_{b1} \\ y_{b2} \end{pmatrix} = \begin{pmatrix} \mu_a \\ \mu_a \\ \mu_b \\ \mu_b \end{pmatrix} + \begin{pmatrix} \epsilon_{a1}^{**} \\ \epsilon_{a2}^{**} \\ \epsilon_{b1}^{**} \\ \epsilon_{b2}^{**} \end{pmatrix} \tag{14}$$

$$\begin{pmatrix} \epsilon_{a1}^{**} \\ \epsilon_{a2}^{**} \\ \epsilon_{b1}^{**} \\ \epsilon_{b2}^{**} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \hat{s}_{a1}^2 + \tau_a^2 & \hat{s}_{a12} & \tau_{ab} & 0 \\ \hat{s}_{a12} & \hat{s}_{a2}^2 + \tau_a^2 & 0 & \tau_{ab} \\ \tau_{ab} & 0 & \hat{s}_{b1}^2 + \tau_b^2 & \hat{s}_{b12} \\ 0 & \tau_{ab} & \hat{s}_{b12} & \hat{s}_{b2}^2 + \tau_b^2 \end{pmatrix} \right)$$

ここで、本研究のアウトカムである観測感度 $\hat{p}_1$ と $\hat{p}_2$ 及び観測特異度 $\hat{q}_1$ と $\hat{q}_2$ は割合であることから、$\hat{s}_{a1}^2$、$\hat{s}_{a2}^2$、$\hat{s}_{a12}$、及び $\hat{s}_{b1}^2$、$\hat{s}_{b2}^2$、$\hat{s}_{b12}$ は、観測感度及び観測特異度をロジッ

ト変換をした上で以下のように計算できる。

$$\begin{pmatrix} y_{a1} \\ y_{a2} \\ y_{b1} \\ y_{b2} \end{pmatrix} = \begin{pmatrix} \mathrm{logit}(\hat{p}_1) \\ \mathrm{logit}(\hat{p}_2) \\ \mathrm{logit}(\hat{q}_1) \\ \mathrm{lotit}(\hat{q}_2) \end{pmatrix} \tag{15}$$

$$\hat{s}_{a1}^2 = \frac{1}{n_D \hat{p}_1(1-\hat{p}_1)} \tag{16}$$

$$\hat{s}_{a2}^2 = \frac{1}{n_D \hat{p}_2(1-\hat{p}_2)} \tag{17}$$

$$\hat{s}_{a12} = \frac{-1}{n_D(1-\hat{p}_1)(1-\hat{p}_2)} \tag{18}$$

$$\hat{s}_{b1}^2 = \frac{1}{n_{\bar{D}} \hat{q}_1(1-\hat{q}_1)} \tag{19}$$

$$\hat{s}_{b2}^2 = \frac{1}{n_{\bar{D}} \hat{q}_2(1-\hat{q}_2)} \tag{20}$$

$$\hat{s}_{b12} = \frac{-1}{n_{\bar{D}}(1-\hat{q}_1)(1-\hat{q}_2)} \tag{21}$$

なお、n_D は疾患ありの被験者数、$n_{\bar{D}}$ は疾患なしの被験者数である。

次に、ロジット変換した感度と特異度の推定を考える。推定量 $\mathbf{m}$ は、FEM、UVRM 及び BVRM に基づき、最尤推定量 $\mathbf{m}_{(FEM)}$、$\mathbf{m}_{(UVRM)}$ 及び $\mathbf{m}_{(BVRM)}$ として求めることができる。

$$\mathbf{m}_{(FEM)} = \left(\mathbf{X}^t\mathbf{R}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{R}^{-1}\mathbf{Y} = \begin{pmatrix} \hat{\mu}_{a(FEM)} \\ \hat{\mu}_{b(FEM)} \end{pmatrix} \tag{22}$$

$$\mathbf{m}_{(UVRM)} = \left(\mathbf{X}^t\mathbf{U}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{U}^{-1}\mathbf{Y} = \begin{pmatrix} \hat{\mu}_{a(UVRM)} \\ \hat{\mu}_{b(UVRM)} \end{pmatrix} \tag{23}$$

$$\mathbf{m}_{(BVRM)} = \left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{Y} = \begin{pmatrix} \hat{\mu}_{a(BVRM)} \\ \hat{\mu}_{b(BVRM)} \end{pmatrix} \tag{24}$$

ここで、

$$\mathbf{X}=\begin{pmatrix}1&0\\1&0\\0&1\\0&1\end{pmatrix}$$

$$\mathbf{R}=\begin{pmatrix}\hat{s}_{a1}^2&\hat{s}_{a12}&0&0\\\hat{s}_{a12}&\hat{s}_{a2}^2&0&0\\0&0&\hat{s}_{b1}^2&\hat{s}_{b12}\\0&0&\hat{s}_{b12}&\hat{s}_{b2}^2\end{pmatrix}$$

$$\mathbf{U}=\begin{pmatrix}\hat{s}_{a1}^2+\hat{\tau}_a^2&\hat{s}_{a12}&0&0\\\hat{s}_{a12}&\hat{s}_{a2}^2+\hat{\tau}_a^2&0&0\\0&0&\hat{s}_{b1}^2+\hat{\tau}_b^2&\hat{s}_{b12}\\0&0&\hat{s}_{b12}&\hat{s}_{b2}^2+\hat{\tau}_b^2\end{pmatrix}$$

$$\mathbf{V}=\begin{pmatrix}\hat{s}_{a1}^2+\hat{\tau}_a^2&\hat{s}_{a12}&\hat{\tau}_{ab}&0\\\hat{s}_{a12}&\hat{s}_{a2}^2+\hat{\tau}_a^2&0&\hat{\tau}_{ab}\\\hat{\tau}_{ab}&0&\hat{s}_{b1}^2+\hat{\tau}_b^2&\hat{s}_{b12}\\0&\hat{\tau}_{ab}&\hat{s}_{b12}&\hat{s}_{b2}^2+\hat{\tau}_b^2\end{pmatrix}$$

$$\mathbf{Y}=\begin{pmatrix}y_{a1}\\y_{a2}\\y_{b1}\\y_{b2}\end{pmatrix}$$

である。

また、各最尤推定量の分散は、以下の式で求められる。

$$\mathrm{Var}(\mathbf{m}_{(FEM)})=\left(\mathbf{X}^t\mathbf{R}^{-1}\mathbf{X}\right)^{-1} \tag{25}$$

$$\mathrm{Var}(\mathbf{m}_{(UVRM)})=\left(\mathbf{X}^t\mathbf{U}^{-1}\mathbf{X}\right)^{-1} \tag{26}$$

$$\mathrm{Var}(\mathbf{m}_{(BVRM)})=\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-1} \tag{27}$$

さらに、変量効果の分散である τ_a^2、τ_b^2 及び τ_{ab} の推定量 $\hat{\tau}_a^2$、$\hat{\tau}_b^2$ 及び $\hat{\tau}_{ab}$ は、DerSimonian and Laird (1986) により提案された方法を多変量で利用できるように拡張した Jackson ら (2010) によるモーメント法に基づいた方法で求めることができる。

$$\hat{\tau}_a=\frac{Q_a-1}{\sum_{k=1}^2 \mathrm{w}_{ak}-\sum_{k=1}^2 \mathrm{w}_{ak}^2/\sum_{k=1}^2 \mathrm{w}_{ak}} \tag{28}$$

$$\hat{\tau}_b = \frac{Q_b - 1}{\sum_{k=1}^{2} \mathrm{w}_{bk} - \sum_{k=1}^{2} \mathrm{w}_{bk}^2 / \sum_{k=1}^{2} \mathrm{w}_{bk}} \tag{29}$$

$$\hat{\tau}_{ab} = \frac{Q_{ab}}{\sum_{k=1}^{2} \sqrt{\mathrm{w}_{ak}\mathrm{w}_{bk}} - \sum_{k=1}^{2} \mathrm{w}_{ak}\mathrm{w}_{bk} / \sum_{k=1}^{2} \sqrt{\mathrm{w}_{ak}\mathrm{w}_{bk}}} \tag{30}$$

ここで、

$$\mathbf{W} = \mathbf{X}^t\mathbf{R}^{-1}$$
$$\mathbf{W}_{(11)} = \mathrm{w}_{a1}\ , \quad \mathbf{W}_{(12)} = \mathrm{w}_{a2}$$
$$\mathbf{W}_{(23)} = \mathrm{w}_{b1}\ , \quad \mathbf{W}_{(24)} = \mathrm{w}_{b2}$$

$$Q_a = \sum_{k=1}^{2} \mathrm{w}_{ak} \left(y_{ak} - \hat{\mu}_{a(FEM)}\right)^2$$

$$Q_b = \sum_{k=1}^{2} \mathrm{w}_{bk} \left(y_{bk} - \hat{\mu}_{b(FEM)}\right)^2$$

$$Q_{ab} = \sum_{k=1}^{2} \sqrt{\mathrm{w}_{ak}\mathrm{w}_{bk}} \left(y_{ak} - \hat{\mu}_{a(COV)}\right)\left(y_{bk} - \hat{\mu}_{b(COV)}\right)$$

$$\hat{\mu}_{a(COV)} = \frac{\sum_{k=1}^{2} \sqrt{\mathrm{w}_{ak}\mathrm{w}_{bk}}\, y_{ak}}{\sum_{k=1}^{2} \sqrt{\mathrm{w}_{ak}\mathrm{w}_{bk}}}\ , \quad \hat{\mu}_{b(COV)} = \frac{\sum_{k=1}^{2} \sqrt{\mathrm{w}_{ak}\mathrm{w}_{bk}}\, y_{bk}}{\sum_{k=1}^{2} \sqrt{\mathrm{w}_{ak}\mathrm{w}_{bk}}}$$

である。

以上の方法で得られた推定値に基づき、読影者 2 名の場合の読影者間変動を表す分散共分散行列は以下のように示される。

$$\Sigma_{DL} = \begin{pmatrix} \hat{\tau}_a^2 & 0 & \hat{\tau}_{ab} & 0 \\ 0 & \hat{\tau}_a^2 & 0 & \hat{\tau}_{ab} \\ \hat{\tau}_{ab} & 0 & \hat{\tau}_b^2 & 0 \\ 0 & \hat{\tau}_{ab} & 0 & \hat{\tau}_b^2 \end{pmatrix} \tag{31}$$

この Σ_{DL} は半正定値行列である必要がある。そこで、Σ_{DL} をスペクトル分解により固有値 λ_k と固有ベクトル x_k とし、固有値の負の値は 0 とした上で再構成することにより、半

正定値行列である Σ_{DL+} を得る。

$$\Sigma_{DL+} = \sum_{k=1}^{K} max(0, \lambda_k) x_k x_k^t \tag{32}$$

$$\Sigma_{DL+} = \sum_{k=1}^{2} max(0, \lambda_k) x_k x_k^t \tag{33}$$

最終的に、本研究のアウトカムである感度 $\mathbf{p}_{(FEM)}$、$\mathbf{p}_{(UVRM)}$ 及び $\mathbf{p}_{(BVRM)}$ は、ロジット変換に基いた $\mathbf{m}_{(FEM)}$、$\mathbf{m}_{(UVRM)}$ 及び $\mathbf{m}_{(BVRM)}$ を逆変換することで求める。

$$\mathbf{p}_{()} = \frac{exp\left(\mathbf{m}_{()}\right)}{1 + exp\left(\mathbf{m}_{()}\right)} = \begin{pmatrix} \hat{p}_{()} \\ \hat{q}_{()} \end{pmatrix} \tag{34}$$

信頼区間の算出は、$\mathbf{m}_{(FEM)}$、$\mathbf{m}_{(UVRM)}$ 及び $\mathbf{m}_{(BVRM)}$ とその分散に基き、以下の式で計算した値を逆変換することで求める。

$$\mathrm{CI}_{\mathbf{m}_{()}} = \mathbf{m}_{()} \pm z_{\alpha/2}\sqrt{\mathrm{Var}(\mathbf{m}_{()})} \tag{35}$$

$$\mathrm{CI}_{\mathbf{p}_{()}} = \frac{exp\left(\mathrm{CI}_{\mathbf{m}_{()}}\right)}{1 + exp\left(\mathrm{CI}_{\mathbf{m}_{()}}\right)} \tag{36}$$

4 数値実験

本研究で誘導した $\hat{p}_{(FEM)}$、$\hat{p}_{(UVRM)}$、$\hat{p}_{(BVRM)}$ 及び $\hat{q}_{(FEM)}$、$\hat{q}_{(UVRM)}$、$\hat{q}_{(BVRM)}$ のバイアスと MSE、並びに 95% 信頼区間の被覆確率を検討するために、2 名の読影者の条件のもとでモンテカルロシミュレーションを実施した。シミュレーションデータは、以下の 2 段階の過程を経て作成した。

1. 任意に設定した感度 p と q をロジット変換した値と、σ_{a1}^2、σ_{a2}^2、σ_{a12}、σ_{b1}^2、σ_{b2}^2、σ_{b12}、τ_a^2、τ_b^2、τ_{ab} の値による 4 変量正規乱数を発生させ、その乱数を逆変換することにより、感度と特異度の擬似データを作成
2. 感度と特異度の擬似データに基き、ベルヌーイ乱数による擬似データを作成し、シミュレーションデータに利用

シミュレーションの結果を表 2〜13 に示す。なお、シミュレーションの繰り返し回数は 10,000 回とした。

バイアスは、いずれの条件においても BVRM による最尤推定量が最も小さく、UVRM に基く最尤推定量のバイアスは BVRM に比べてわずかに大きい傾向を示した。一方、FEM に基く最尤推定量のバイアスは他に比べて著しく大きくなる傾向を示した。MSE は、いずれの条件においても各最尤推定量は同様の傾向を示した。また、バラツキの条件を大きくすると、いずれの最尤定量もバイアスが大きくなる傾向が認められたが、MSE ではそのような傾向は認められなかった。

被覆確率については、BVRM と UVRM に基く信頼区間が名義の水準 95% に近く、FEM に基く信頼区間は名義の水準を下回った。また、バラツキの条件を大きくすると、いずれの信頼区間の被覆確率も名義の水準を下回る傾向が認められた。

5 応用事例

本章では、アルツハイマー病発症の原因と考えられている脳内 β アミロイド集積の画像化を目的として開発された診断薬である ^{18}F-Florbetapir の第 III 相臨床試験のデータに対する本提案法の応用事例を示す。本臨床試験は、アルツハイマー病の疑いで ^{18}F-Florbetapir の投与による PET 診断を受け、死後に脳の剖検が実施された 35 名の被験者を対象としたものであり、剖検における β アミロイドの免疫染色の結果からアルツハイマー病の被験者が 20 名、非アルツハイマー病の被験者が 15 名と判断された。^{18}F-Florbetapir による PET 診断で得られた 35 名の被験者の画像は、独立した 3 名の読影者により 0-4 の 5 段階で ^{18}F-Florbetapir の集積度が評価され、この 5 段階評価のうち 0 及び 1 を陰性、2, 3, 4 を陽性として感度と特異度が算出された。表 14 は米国の医薬品審査機関である FDA により公開されている承認資料から、β アミロイドの免疫染色の結果をスタンダードとした 3 名の読影者による ^{18}F-Florbetapir の感度と特異度の解析結果を引用したものである (FDA (2012))。

3 名の読影者による感度と特異度を要約するために本提案法を適用した結果、FEM、UVRM 及び BVRM により算出された感度、特異度及び各々の 95% 信頼区間は、$\hat{p}_{(FEM)}$ が 74.5% (65.5, 81.7)、$\hat{p}_{(UVRM)}$ が 78.4% (50.4, 92.9)、$\hat{p}_{(BVRM)}$ が 78.4% (50.3, 92.8)、$\hat{q}_{(FEM)}$ が 87.1% (71.6, 94.7)、$\hat{q}_{(UVRM)}$ が 90.9% (65.8, 98.1)、$\hat{q}_{(BVRM)}$ が 90.8% (66.2, 98.0) であった。

6 考察

本研究では、fix effects model、univariate random effects model 及び bivariate random effects model に基づき、メタアナリシスの手法を応用することで、複数の読影者により評価された診断方法の感度及び特異度の要約、並びに 95% 信頼区間の推定方法を提案した。

今回検討した三つのモデルのうち、複数の読影者が同一の被験者集団を評価することによる相関、被験者間の変動及び同一被験者による感度と特異度の相関の全てを考慮した bivariate random effects model による最尤推定量と信頼区間が、バイアス、MSE 及び被覆確率の観点から最も優れた性質を持っていることが、数値実験の結果から示唆された。しかしながら、感度及び特異度のバラツキが大きい状況では、本提案法による信頼区間の被覆確率が名目の水準を下回る傾向が認められたことから、この課題を解決するために更なる検討が必要であると考える。

参考文献

[1] Lehr, R. G. and Kashanian, F. K. (2009). Three persistent issues in analysis of clinical trials involving diagnostic contrast agents. *Drug Information Journal* **43,** 525–532.

[2] FDA. (2004). Guidance for industry. Developing medical imaging drugs and biological products. Part 3: design, analysis, and interpretation of clinical studies. *Food and Drug Administration (FDA), Guidance for Industry.*

[3] Obuchowski, N. A. and Lieber, M. L. (2008). Statistics and methodology. *Skeletal Radiology* **37,** 393–396.

[4] CHMP. (2009). Appendix 1 to the guideline on clinical evaluation of diagnostic agents (CPMP/EWP/1119/98 REV. 1) on imaging agents. *Committee for Medical Products for Human Use (CHMP) of the European Medicines Agency (EMA), EMEA/CHMP/EWP/321180/2008.*

[5] Saeki, H. and Tango, T. (2011). Non-inferiority test and confidence interval for the difference in correlated proportions in diagnostic procedures based on multiple raters. *Statistics in Medicine* **30,** 3313–3327.

[6] DerSimonian, R. and Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials* **7,** 177–188.

[7] Jackson, D., White, I. R., and Thompson, S. G. (2010). Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analysis. *Statistics in Medicine* **29,** 1282–1297.

[8] FDA. (2012). Drug Approval Package. Amyvid (Florbetapir F 18 Injection) Statistical Review(s). *Food and Drug Administration (FDA), Drug Approval Package* http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/202008Orig1s000StatR.pdf [25April2014].

表 2　読影者 2 名での $p=q=80\%$、$n=m=25$ におけるバイアスの評価 (Sens: sensitivity, Spec: specificity, F: fix effects model, U: univariate random effects model, B: bivariate random effects model)

					Bias		
$\sigma_{a_k}=\sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a=\tau_b$	τ_{ab}		F	U	B
0.05	-0.01	0.05	-0.01	Sens	-0.0100	-0.0039	-0.0027
				Spec	-0.0098	-0.0037	-0.0025
0.1	-0.01	0.05	-0.01	Sens	-0.0117	-0.0056	-0.0044
				Spec	-0.0118	-0.0057	-0.0045
0.05	-0.01	0.1	-0.05	Sens	-0.0118	-0.0056	-0.0044
				Spec	-0.0116	-0.0055	-0.0043
0.1	-0.01	0.1	-0.05	Sens	-0.0139	-0.0078	-0.0066
				Spec	-0.0136	-0.0077	-0.0065
0.05	-0.01	0.3	-0.05	Sens	-0.0207	-0.0148	-0.0137
				Spec	-0.0206	-0.0149	-0.0138
0.1	-0.01	0.3	-0.05	Sens	-0.0223	-0.0165	-0.0154
				Spec	-0.0218	-0.0163	-0.0152

表 3　読影者 2 名での $p = q = 80\%$、$n = m = 25$ における MSE の評価

$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		MSE		
					F	U	B
0.05	-0.01	0.05	-0.01	Sens	0.0032	0.0032	0.0032
				Spec	0.0032	0.0032	0.0032
0.1	-0.01	0.05	-0.01	Sens	0.0032	0.0032	0.0032
				Spec	0.0033	0.0032	0.0032
0.05	-0.01	0.1	-0.05	Sens	0.0032	0.0032	0.0032
				Spec	0.0033	0.0032	0.0032
0.1	-0.01	0.1	-0.05	Sens	0.0033	0.0033	0.0033
				Spec	0.0034	0.0039	0.0033
0.05	-0.01	0.3	-0.05	Sens	0.0036	0.0036	0.0035
				Spec	0.0037	0.0035	0.0035
0.1	-0.01	0.3	-0.05	Sens	0.0037	0.0036	0.0036
				Spec	0.0038	0.0036	0.0036

表 4　読影者 2 名での $p = q = 80\%$、$n = m = 100$ におけるバイアスの評価

$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		Bias		
					F	U	B
0.05	-0.01	0.05	-0.01	Sens	-0.0055	-0.0042	-0.0039
				Spec	-0.0051	-0.0039	-0.0036
0.1	-0.01	0.05	-0.01	Sens	-0.0075	-0.0063	-0.0060
				Spec	-0.0074	-0.0061	-0.0059
0.05	-0.01	0.1	-0.05	Sens	-0.0075	-0.0063	-0.0060
				Spec	-0.0076	-0.0063	-0.0061
0.1	-0.01	0.1	-0.05	Sens	-0.0097	-0.0085	-0.0082
				Spec	-0.0097	-0.0085	-0.0082
0.05	-0.01	0.3	-0.05	Sens	-0.0166	-0.0154	-0.0152
				Spec	-0.0160	-0.0148	-0.0146
0.1	-0.01	0.3	-0.05	Sens	-0.0180	-0.0169	-0.0166
				Spec	-0.0182	-0.0170	-0.0167

表 5　読影者 2 名での $p = q = 80\%$、$n = m = 100$ における MSE の評価

					MSE		
$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		F	U	B
0.05	-0.01	0.05	-0.01	Sens	0.0008	0.0008	0.0008
				Spec	0.0008	0.0008	0.0008
0.1	-0.01	0.05	-0.01	Sens	0.0008	0.0008	0.0008
				Spec	0.0008	0.0008	0.0008
0.05	-0.01	0.1	-0.05	Sens	0.0009	0.0008	0.0008
				Spec	0.0008	0.0008	0.0008
0.1	-0.01	0.1	-0.05	Sens	0.0009	0.0009	0.0009
				Spec	0.0009	0.0009	0.0009
0.05	-0.01	0.3	-0.05	Sens	0.0011	0.0011	0.0011
				Spec	0.0011	0.0011	0.0011
0.1	-0.01	0.3	-0.05	Sens	0.0012	0.0011	0.0011
				Spec	0.0012	0.0011	0.0011

表 6　読影者 2 名での $p = q = 90\%$、$n = m = 25$ におけるバイアスの評価

					Bias		
$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		F	U	B
0.05	-0.01	0.05	-0.01	Sens	-0.0135	-0.0067	-0.0045
				Spec	-0.0141	-0.0072	-0.0051
0.1	-0.01	0.05	-0.01	Sens	-0.0154	-0.0083	-0.0062
				Spec	-0.0150	-0.0083	-0.0061
0.05	-0.01	0.1	-0.05	Sens	-0.0148	-0.0079	-0.0057
				Spec	-0.0151	-0.0083	-0.0061
0.1	-0.01	0.1	-0.05	Sens	-0.0165	-0.0095	-0.0073
				Spec	-0.0166	-0.0098	-0.0076
0.05	-0.01	0.3	-0.05	Sens	-0.0218	-0.0146	-0.0125
				Spec	-0.0226	-0.0155	-0.0135
0.1	-0.01	0.3	-0.05	Sens	-0.0235	-0.0162	-0.0141
				Spec	-0.0234	-0.0165	-0.0144

表 7　読影者 2 名での $p = q = 90\%$、$n = m = 25$ における MSE の評価

$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		MSE F	U	B
0.05	-0.01	0.05	-0.01	Sens	0.0020	0.0016	0.0016
				Spec	0.0020	0.0016	0.0016
0.1	-0.01	0.05	-0.01	Sens	0.0021	0.0017	0.0017
				Spec	0.0021	0.0017	0.0017
0.05	-0.01	0.1	-0.05	Sens	0.0020	0.0016	0.0016
				Spec	0.0020	0.0016	0.0016
0.1	-0.01	0.1	-0.05	Sens	0.0021	0.0017	0.0017
				Spec	0.0021	0.0017	0.0017
0.05	-0.01	0.3	-0.05	Sens	0.0024	0.0019	0.0019
				Spec	0.0025	0.0019	0.0019
0.1	-0.01	0.3	-0.05	Sens	0.0025	0.0020	0.0020
				Spec	0.0025	0.0020	0.0020

表 8　読影者 2 名での $p = q = 90\%$、$n = m = 100$ におけるバイアスの評価

$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		Bias F	U	B
0.05	-0.01	0.05	-0.01	Sens	-0.0052	-0.0032	-0.0027
				Spec	-0.0049	-0.0029	-0.0025
0.1	-0.01	0.05	-0.01	Sens	-0.0070	-0.0050	-0.0046
				Spec	-0.0068	-0.0048	-0.0043
0.05	-0.01	0.1	-0.05	Sens	-0.0070	-0.0050	-0.0046
				Spec	-0.0068	-0.0049	-0.0044
0.1	-0.01	0.1	-0.05	Sens	-0.0087	-0.0067	-0.0062
				Spec	-0.0086	-0.0067	-0.0062
0.05	-0.01	0.3	-0.05	Sens	-0.0140	-0.0121	-0.0117
				Spec	-0.0137	-0.0118	-0.0114
0.1	-0.01	0.3	-0.05	Sens	-0.0155	-0.0137	-0.0132
				Spec	-0.0156	-0.0137	-0.0133

表 9　読影者 2 名での $p = q = 90\%$、$n = m = 100$ における MSE の評価

					MSE		
$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		F	U	B
0.05	-0.01	0.05	-0.01	Sens	0.0005	0.0005	0.0005
				Spec	0.0005	0.0005	0.0005
0.1	-0.01	0.05	-0.01	Sens	0.0005	0.0005	0.0005
				Spec	0.0005	0.0005	0.0005
0.05	-0.01	0.1	-0.05	Sens	0.0005	0.0005	0.0005
				Spec	0.0005	0.0005	0.0005
0.1	-0.01	0.1	-0.05	Sens	0.0005	0.0005	0.0005
				Spec	0.0005	0.0005	0.0005
0.05	-0.01	0.3	-0.05	Sens	0.0007	0.0006	0.0006
				Spec	0.0007	0.0007	0.0007
0.1	-0.01	0.3	-0.05	Sens	0.0007	0.0007	0.0007
				Spec	0.0007	0.0007	0.0007

表 10　読影者 2 名での $p = q = 80\%$、$n = m = 25$ における 95% 信頼区間の被覆確率の評価

					Coverage Prob.		
$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		F	U	B
0.05	-0.01	0.05	-0.01	Sens	91.8	94.7	94.7
				Spec	91.7	94.6	94.6
0.1	-0.01	0.05	-0.01	Sens	92.8	95.3	95.3
				Spec	92.3	95.6	95.6
0.05	-0.01	0.1	-0.05	Sens	90.9	94.2	94.2
				Spec	91.2	94.4	94.4
0.1	-0.01	0.1	-0.05	Sens	90.7	93.6	93.6
				Spec	91.8	94.9	95.0
0.05	-0.01	0.3	-0.05	Sens	89.5	93.8	94.0
				Spec	89.9	93.4	93.5
0.1	-0.01	0.3	-0.05	Sens	88.4	93.4	93.5
				Spec	86.5	91.8	91.8

表 11　読影者 2 名での $p = q = 80\%$、$n = m = 100$ における 95% 信頼区間の被覆確率の評価

					Coverage Prob.		
$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		F	U	B
0.05	-0.01	0.05	-0.01	Sens	91.3	94.0	94.4
				Spec	89.9	93.1	94.1
0.1	-0.01	0.05	-0.01	Sens	91.3	93.3	93.7
				Spec	91.8	94.3	95.1
0.05	-0.01	0.1	-0.05	Sens	92.1	94.9	95.0
				Spec	90.1	93.6	94.2
0.1	-0.01	0.1	-0.05	Sens	87.8	91.2	92.2
				Spec	87.4	91.6	92.0
0.05	-0.01	0.3	-0.05	Sens	85.6	90.5	91.3
				Spec	82.8	89.1	89.8
0.1	-0.01	0.3	-0.05	Sens	82.4	87.6	88.5
				Spec	82.7	87.8	88.9

表 12　読影者 2 名での $p = q = 90\%$、$n = m = 25$ における 95% 信頼区間の被覆確率の評価

					Coverage Prob.		
$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		F	U	B
0.05	-0.01	0.05	-0.01	Sens	93.1	96.4	96.9
				Spec	93.6	96.2	96.8
0.1	-0.01	0.05	-0.01	Sens	92.5	95.7	96.7
				Spec	91.6	94.9	95.3
0.05	-0.01	0.1	-0.05	Sens	92.7	95.1	95.7
				Spec	92.8	95.9	96.8
0.1	-0.01	0.1	-0.05	Sens	92.6	94.9	95.5
				Spec	92.2	95.3	96.1
0.05	-0.01	0.3	-0.05	Sens	88.8	93.1	93.7
				Spec	88.8	93.0	93.9
0.1	-0.01	0.3	-0.05	Sens	90.3	93.9	94.5
				Spec	88.3	93.0	94.1

表 13　読影者 2 名での $p = q = 90\%$、$n = m = 100$ における 95% 信頼区間の被覆確率の評価

					Coverage Prob.		
$\sigma_{a_k} = \sigma_{b_k}$	$\sigma_{ab_{kl}}$	$\tau_a = \tau_b$	τ_{ab}		F	U	B
0.05	-0.01	0.05	-0.01	Sens	93.5	94.8	95.3
				Spec	93.2	94.5	95.2
0.1	-0.01	0.05	-0.01	Sens	92.7	94.7	95.4
				Spec	89.6	92.8	94.3
0.05	-0.01	0.1	-0.05	Sens	91.7	94.4	95.4
				Spec	92.9	93.7	93.7
0.1	-0.01	0.1	-0.05	Sens	90.1	92.7	93.9
				Spec	90.0	92.4	94.0
0.05	-0.01	0.3	-0.05	Sens	84.4	89.4	90.8
				Spec	87.1	89.7	91.6
0.1	-0.01	0.3	-0.05	Sens	86.1	90.3	91.9
				Spec	84.9	88.9	90.4

表 14　免疫染色の結果をスタンダードとした、3 名の読影者による ^{18}F-Florbetapir の β アミロイド集積の診断における感度と特異度

Reader	Sensitivity (%)(CI)	Specificity (%)(CI)
1	90 (69, 97)	100 (82, 100)
2	55 (28, 79)	100 (82, 100)
3	85 (64, 95)	80 (55, 93)