

# 化学療法が対象となる間質性肺炎合併肺癌患者の死亡率の動的予

## 測 : Joint frailty-copula モデルの臨床データへの適用

北里大学大学院臨床統計学部門 高見澤 勇

Isamu Takamisawa

Department of Biostatistics

Kitasato University

### 1 背景

間質性肺炎合併肺癌患者における手術、抗癌剤、放射線などの治療は間質性肺炎の急性増悪の危険因子となり重篤な合併症を引き起こす可能性が示唆されている<sup>1)</sup>。そのため間質性肺炎を合併していない肺癌患者と比較して、間質性肺炎合併肺癌の病状は複雑で標準的治療がいまだ確立していないのが現状である。そこで現在、間質性肺炎合併肺癌患者の治療ガイドライン策定を目的とした調査研究が実施されている<sup>2)</sup>。

このような現状の中で患者の予後予測は治療方針を決定する際の重要な指針になりうる<sup>3)</sup>と期待されている。一般的に患者の予後は治療開始時点における予後因子（共変量）の情報に基づいて判断されてきた経緯がある<sup>3)</sup>。しかし近年は間質性肺炎における急性増悪のような治療途中で起こる中間イベント（治療開始時点では得られない）の情報（急性増悪あり群となし群）を用いた解析の手法も提案されてきた<sup>4)</sup>。なかでもランドマーク解析が臨床研究で広く使用されてきた手法である<sup>5)</sup>。中間イベントあり群に分類される患者は中間イベントを経験するまで生存する必要がある、この期間生存できない患者は中間イベントなし群に分類される。したがって中間イベントなし群に分類された患者は中間イベントあり群に分類された患者と比較して生存期間が短くなる傾向がある。この偏りを **guarantee time bias** という。ランドマーク解析は、ある任意の時点（ランドマーク）まで生存した患者を対象に中間イベントあり群となし群に群わけすることで、ランドマーク以前に死亡した患者を除き **guarantee time bias** を抑制する解析手法である。この手法は単純であり、解釈が容易であるが早期に死亡した患者の情報を完全に捨てていること、ランドマークの設定が恣意的であるといった問題がある。

近年、こうした中間イベントの生存関数と予後の生存関数の依存性を考慮した同時分布関数を構成し、中間イベントの有無と時点に依存して予後予測する **joint frailty-copula** モデルが提案された<sup>6)</sup>。さらに、この手法では施設間または試験間の異質性を考慮した同時分布の構成が可能のため、データの有効利用という観点からも有効な手法である。

本研究ではこの手法を多施設で行われた間質性肺炎合併肺癌患者の化学療法対象症例（ガイドラインを引用）に適用し、層別（例えば性別）に急性増悪の有無と時点で予後予測することとした。また臨床データへの適用例の少ない joint frailty-copula モデルの有用性や問題点を検討し、さらなる実用性の向上を目的とする案を議論する。

## 2 Joint Frailty-copula モデル

### Data structure

今回のモデルの構築には複数の独立の施設で行われた臨床試験を含むメタアナリシスデータを必要とする。各臨床試験には複数の患者が参加している必要がある。ここで  $G$  つの異なる施設の臨床試験の  $i$  番目の施設が  $N_i$  例の症例を有しているとしたとき、その中の任意の 1 症例  $j$  の試験開始時から中間イベントが発生した時点、死亡した時点とセンサーとなった時点をそれぞれ  $X_{ij}, D_{ij}, C_{ij}$  と表す。そのとき患者の中間イベントの有無  $\delta_{ij} = I(T_{ij} = X_{ij})$ 、及び中間イベントが起きた時点  $T_{ij} = \min(X_{ij}, D_{ij}, C_{ij})$ 、最終イベントである死亡の有無  $\delta_{ij}^* = I(T_{ij}^* = D_{ij})$  及び観測開始から死亡またはセンサーまでの時点  $T_{ij}^* = \min(D_{ij}, C_{ij})$  と表す。ただし  $I(\cdot)$  は指示関数である。

### Joint frailty model

複数の臨床試験データを統合して解析する際、その研究間の差(異質性)を考慮する必要がある。研究間の異質性を考慮する変数を  $u_i$  とする。  $u_i$  は  $\eta$  を分散とするガンマ分布

$$f(u_i) = \frac{1}{\Gamma(1/\eta)\eta^{1/\eta}} u_i^{\frac{1}{\eta}-1} \exp\left(-\frac{u_i}{\eta}\right)$$

に従うとする (Rondeau et al.)。

すると joint frailty model の中間イベントと死亡に関するハザードはそれぞれ

$$\begin{aligned} r_{ij}(t|u_i) &= u_i r_0(t) \exp(\beta_1' Z_{1ij}) \\ \lambda_{ij}(t|u_i) &= u_i^\alpha \lambda_0(t) \exp(\beta_2' Z_{2ij}) \end{aligned}$$

となる。  $\beta_1$  と  $\beta_2$  は回帰係数である。  $\alpha$  は死亡ハザードの研究間の異質性を示すパラメーターであり、通常、死亡ハザードが研究間で同じであれば  $\alpha = 0$ 、異質であれば  $\alpha = 1$  と置く。

本研究では 166 例と比較的少数人数であるため、単変量の層別解析を中心に解析を行った。

### Joint frailty-copula model

次に中間イベントに関する累積ハザードと死亡に関する累積ハザードから生存関数をそ

れぞれ導き、クレイトンコピュラで接合することで同時分布を構成する。

累積ハザードは $r_{ij}(t|u_i)$ と $\lambda_{ij}(t|u_i)$ をそれぞれ時間で積分し

$$R_{ij}(t|u_i) = \int_0^x r_{ij}(v|u_i) dv$$

$$\Lambda_{ij}(t|u_i) = \int_0^y \lambda_{ij}(v|u_i) dv$$

と表せ、生存関数は

$$S_{Xij}(t|u_i) = \exp \{-R_{ij}(t|u_i)\}$$

$$S_{Dij}(t|u_i) = \exp \{-\Lambda_{ij}(t|u_i)\}$$

と表せる。

急性増悪と死亡の同時分布は

$$\Pr(X_{ij} > x, D_{ij} > y|u_i) = C_\theta[S_{Xij}(t|u_i), S_{Dij}(t|u_i)]$$

となる。(Emura et al.)。ここで、 $C$ はコピュラと呼ばれる関数 (Nelsen) である。

使用するクレイトンコピュラの一般式は、

$$C_\theta(v, w) = (v^{-\theta} + w^{-\theta} - 1)^{-1/\theta}, \theta \geq 0$$

と表す。

コピュラパラメーター $\theta$ は互いの依存の強さを表しており、 $\theta = 0$ のとき $C_\theta(v, w) = vw$ 、独立コピュラとなる。

回帰係数、パラメーター各種推定

Likelihood under the joint frailty copula model

急性増悪と死亡の同時分布の尤度は以下のように表す。

$$\begin{aligned} & \ell(\alpha, \eta, \theta, \beta_1, \beta_2, r_0, \lambda_0) - \\ &= \sum_{i=1}^G \left[ \sum_{j=1}^{N_i} \{ \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij}^* \log \lambda_{ij}(T_{ij}^*) \} \right. \\ &+ \log \int_0^\infty \left\{ u_i^{m_i + \alpha m_i^*} \prod_{j=1}^{N_i} \psi_\theta [u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)]^{\delta_{ij}} \psi_\theta^* [u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)]^{\delta_{ij}^*} \right. \\ &\left. \left. \times \Theta_\theta [u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)]^{\delta_{ij} \delta_{ij}^*} D_\theta [u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \right\} f_\eta(u_i) du_i \right] \end{aligned}$$

補足

$$\begin{aligned}
m_i &= \sum_{j=1}^{N_i} \delta_{ij} , \quad m_i^* = \sum_{j=1}^{N_i} \delta_{ij}^* \\
D_\theta[s, t] &= C_\theta[\exp(-s), \exp(-t)] , \\
\psi_\theta &= D_\theta^{[1,0]}/D_\theta , \quad D_\theta^{[1,0]} = -\partial D_\theta/\partial s, \\
\psi_\theta^* &= D_\theta^{[0,1]}/D_\theta , \quad D_\theta^{[0,1]} = -\partial D_\theta/\partial t, \\
\Theta_\theta &= D_\theta^{[1,1]}/D_\theta/D_\theta^{[1,0]}D_\theta^{[0,1]} , \quad D_\theta^{[1,1]} = \partial^2 D_\theta/\partial s\partial t
\end{aligned}$$

ハザード関数をスプライン (Emura et al.) で近似し、滑らかな関数にするために、罰則付き尤度を考える。

$$\ell(\alpha, \eta, \theta, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, r_0, \lambda_0) - k_1 \int \ddot{r}_0(t)^2 dt - k_2 \int \ddot{\lambda}_0(t)^2 dt$$

$$\ddot{f}(t) = d^2 f(t)/dt^2$$

罰則付き尤度を最大化し、最尤推定量 $\hat{\alpha}, \hat{\eta}, \hat{\theta}, \hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \hat{r}_0, \hat{\lambda}_0$ を R パッケージ joint.Cox で算出した。

#### Joint frailty-copula model による動的予測

本モデルでは 2 通り場合の動的情報 (時点  $t$  での中間イベントの有無) を想定する。それぞれ、

$X > t$ ; 中間イベント  $x$  が時点  $t$  で観測されていない場合

$X \leq t, X = t$ ; 中間イベント  $x$  が時点  $t$  以前で起こった場合

と表せる。

$\mathbf{Z}$  は  $t=0$  で得られる患者の情報として、患者の死亡率は以下の 2 通りで分けられる。

中間イベント  $x$  が時点  $t$  で観測されていない場合

$$\begin{aligned}
F_\theta^{Joint}(t, t+w|X > t, \mathbf{Z}) &= Pr(D \leq t+w | D > t, X > t, \mathbf{Z}) \\
&= \frac{\int_0^\infty (C_\theta[S_X(t|u, \mathbf{Z}), S_D(t|u, \mathbf{Z})] - C_\theta[S_X(t|u, \mathbf{Z}), S_D(t+w|u, \mathbf{Z})]) f_\eta(u) du}{\int_0^\infty C_\theta[S_X(t|u, \mathbf{Z}), S_D(t|u, \mathbf{Z})] f_\eta(u) du}
\end{aligned}$$

中間イベント  $x$  が時点  $t$  以前で起こった場合

$$F_{\theta}^{Joint}(t, t+w|X=t, \mathbf{Z}) = Pr(D \leq t+w | D > t, X=x, \mathbf{Z})$$

$$= \frac{\int_0^{\infty} (C_{\theta}^{[1,0]}[S_X(x|u, \mathbf{Z}), S_D(t|u, \mathbf{Z})] - C_{\theta}^{[1,0]}[S_X(t|u, \mathbf{Z}), S_D(t+w|u, \mathbf{Z})]) u S_X(x|u, \mathbf{Z}) f_{\eta}(u) du}{\int_0^{\infty} C_{\theta}^{[1,0]}[S_X(x|u, \mathbf{Z}), S_D(t|u, \mathbf{Z})] u S_X(x|u, \mathbf{Z}) f_{\eta}(u) du}$$

補足

$$C_{\theta}^{[1,0]}(v, w) = \partial C_{\theta} / \partial v,$$

以上の2式で同じ患者背景  $\mathbf{Z}$  を持つ仮定の患者が時点  $t$  まで中間イベントを観測されていない場合と  $X=x$  で中間イベントを観測した場合の時点  $t$  から任意の期間  $w$  連続的に死亡率を予測し比較した。

### 3 データの Joint Frailty-Copula モデルへの適用

#### データ

本研究で用いるデータは間質性肺炎合併肺癌に対する治療ガイドライン策定に寄与する情報を得るために全国アンケート調査で収集された他施設合同の化学療法と緩和療法の患者 1033 例のデータである。2012 年 1 月から 2013 年末までに原発性肺癌 stageIV または術後再発と診断された患者とし、間質性肺炎の診断時期は肺癌診断と同時にまたはそれ以前としている。

データの内容は患者の年齢や性別などの基礎情報と組織型、PS(performance status)、肺気腫の有無など多数の病理的情報、さらに診断日、抗がん剤投与開始日、急性増悪の有無、急性増悪の発生日、患者の死亡の有無、また死亡した場合死亡観測日、生存の場合打ち切り日が登録されている。

本研究に用いたデータは後ろ向き研究であり、急性増悪発症との関連が疑われる患者背景因子を医師によって選定、収集しているが、一部の変数では欠損が多く、共変量として採用できなかった。本研究では、連続変数として、年齢と喫煙量、カテゴリー変数として年齢(70 歳以上、70 歳未満)、性別(男、女)、PS(0,1 を良好、2,3,4 を不良)、組織型(小細胞、非小細胞)、肺気腫(あり、なし)、HRCT(画像診断による通常型間質性肺炎、通常型間質性肺炎の可能性あり、通常型間質性肺炎でない)を層別予測の共変量として選定した。

化学療法を実施した施設は 105 施設であったが、患者数が少ない、イベント(急性増悪と死亡)の頻度が低い施設はモデル構築に適さない。本研究では以上の事項を考慮した結果、モデル構築に 9 施設を採用し、166 名のデータをデータセットとすることとなった。

#### R パッケージ”joint.Cox”の適用

R パッケージ 'joint.Cox' の同時分布のパラメーター(共変量に関する回帰係数を含む)の推定には罰則付き尤度推定を行う関数 'jointCox.reg' を使用した(各種パラメーターの推定については論文<sup>6,7)</sup> 参照)。推定した最尤推定量 $\hat{\alpha}, \hat{\eta}, \hat{\theta}, \hat{\beta}_1, \hat{\beta}_2, \hat{f}_0, \hat{\lambda}_0$ を R パッケージ joint.Cox で算出した。

## 結果・考察

急性増悪を発症した場合、急性増悪を発症しなかった場合と比べて、常に死亡率が高く予測される結果となった。この傾向は層別予後予測、すべての変数に共通した。

本研究の結果から、患者の投与開始時点での背景因子及び病態にかかわらず間質性肺炎の急性増悪の有無が最も患者の予後に関与が大きいことが示唆された。この結果は「肺がん患者の化学療法による治療中に急性増悪を発症したとき(しなかったとき)予後はどうなるのか。」という臨床上の疑問に対する答えの一つとなった。

現在間質性肺炎合併肺癌患者の治療ガイドライン策定を目的とした調査研究が実施されている。そのような状況で本研究による化学療法治療中に急性増悪が発症した場合、急性増悪が発症しなかった場合と比較して予後が悪いという予測結果は、化学療法を受けることを検討するにあたって事前に知っておくべき内容であり、治療ガイドライン策定の際の参考ともなる。

今までの研究による生存予測モデルでは治療開始日の情報に基づいた生存予測であるため、治療途中の中間イベント(急性増悪)は考慮されない。本研究のような動的予測では患者が治療途中で急性増悪を発症し自分の予後がどう変化するのか知りたい場合、治療途中でありながら、その患者の生存予測について答えることが可能となった。本研究結果は患者の治療方針等の意思決定に大きく寄与すると考えられる。

また、すべての予測において急性増悪の発生が死亡率に大きな影響を与えているという結果は予後予測に中間イベントを取り入れて比較することの意義を考えさせる結果でもある。アウトカムと独立とはみなせない中間イベントは広範囲に及ぶ。例えば、臓器移植において移植を受けた群と受けていない群の生存率を予測する場合、臓器移植自体が中間イベントとなり得る。また、脳卒中患者の予後の解析において脳卒中の再発も中間イベントとして扱える。これらに本モデルを適用することで中間イベントの有無による予測が期待できる。

本研究では後ろ向き研究のデータを用いたため、欠損が多く、また、患者数が少なく患者背景である男女比に偏り(162 : 4)があったためパラメーターがうまく推定できず、利用できる変数に制約がかかった。後ろ向き研究ではデータの欠損が発生しやすいことや、患者背景をそろえることは困難であることは仕方ないことである。今後は前向き研究のデータに適用することを検討したい。

## 1. 参考文献

1. 日本呼吸器学会 びまん性肺疾患 診断・治療ガイドライン作成委員会 特発性間質性肺炎 診断と治療の手引き 改訂版 2 版 南江堂
2. びまん性肺疾患に関する調査研究班 特発性肺線維症合併肺がん部会 日本呼吸器学会 腫瘍学術部会 特発性間質性肺炎合併肺癌患者の内科治療に関する後ろ向き調査
3. Erika Graf<sup>\*</sup>, Claudia Schmoor, Willi Sauerbrei and Martin Schumacher : Assessment and comparison of prognostic classification schemes for survival data : Stat Med 1999 Sep 15-30;18(17-18):2529-45
4. Anita Giobbie-Hurder, Richard D. Gelber, and Meredith M. Regan : Challenges of Guarantee-Time Bias : Journal of Clinical Oncology volume 31 number 23 August 10 2013:2963-69
5. Urania Dafni, ScD : Landmark Analysis at the 25-Year Landmark Point. : Circ Cardiovasc Qual Outcomes 2011 May;4(3):363-71
6. Emura T, Nakatochi M, Murotani K, Rondeau V. : A joint frailty-copula model between tumour progression and death for meta-analysis. : Statistics Methods in Medical Research 2017, Vol. 26(6) 2649-66
7. Takeshi Emura, Masahiro Nakatochi, Shigeyuki Matsui, Hirofumi Michimae, Virginie Rondeau : Personalized dynamic prediction of death according to tumour progression and high-dimensional genetic factors: meta-analysis with a joint model : Statistics Methods in Medical Research, in press