

## X-Learner を用いた抗がん剤治療による間質性肺疾患誘引リスクの評価 Evaluation of the risk of interstitial lung disease by anti-cancer drug treatment with X-Learner

水谷 圭佑<sup>1,2</sup>      塚田 啓介<sup>1</sup>      徳山 健斗<sup>1</sup>  
Keisuke Mizutani      Keisuke Tsukada      Kento Tokuyama

### 1. はじめに

間質性肺疾患 (ILD) は、肺の間質に炎症や損傷が起こり、線維化が進むことで肺が硬くなり、拡がりづらいために呼吸が困難になる病気である。ILD の原因は多岐にわたり、ヒトの遺伝的背景に加え、薬剤の副作用や関節リウマチなどの膠原病、生活上での粉塵の慢性的な吸引など様々であることが知られており、原因を一つに特定することは難しい。ILD は肺がん患者に対する抗がん剤治療の重篤な副作用の一つであり、近年抗がん剤治療が進歩している状況にも関わらず、治療が継続できない要因となりうる[1]。抗がん剤は、細胞障害性抗がん剤、免疫チェックポイント阻害剤 (がん免疫療法)、分子標的薬の 3 種に分けられるが、いずれの治療薬投与後においても ILD 発症例が報告されている。ただし、それぞれの症例における患者背景も異なるため、薬剤の投与が原因となったか、または他の要因が原因となったかを後向きに解明することは困難である。抗がん剤治療薬の投与による ILD 発症リスクを評価することが出来れば、患者ごとに副作用リスクの少ない治療方法を選択するなどの個別化医療の実現が期待される。

薬剤投与という介入の有無による結果 (Outcome) の差を調べるためには、介入時と非介入時の結果を評価しなければならない。しかし、同一の患者に対して介入と非介入による効果検証を行うことは反実仮想のため実現不可能である。また、薬剤の主効果でない、副作用リスクの評価を目的にしたランダム比較化試験を実施することも難しい。近年、観察的研究から機械学習モデルを用いて反実仮想データを推論し、異質介入効果を調べる Meta-Learner という計算フレームワークが考案された[2,3]。Meta-Learner の一つである X-Learner は、選択バイアスを考慮して介入効果を評価するアルゴリズムであり、医師による薬剤投与という選択バイアス下においても介入効果を適切に評価できる手法として期待される [4]。本研究では、肺がん患者約 23 万人のリアルワールドデータを用いて機械学習による ILD 発症予測モデルの開発と X-Learner による解析を行うことで、各薬剤のリスク評価と ILD 発症に寄与する要因の解析を試みた。

### 2. 手法

#### 2.1 使用データ

本研究では、リアルワールドデータ (RWD) 社の提供する the RWD-DB の肺がん患者データを用いた。本データベースは RWD 社の支援を受けた健康・医療・教育情報評価推進機構 (HCEI) によって管理・保持されている。

1 中外製薬株式会社デジタル戦略推進部 Chugai Pharmaceutical Co., Ltd.

2 東京工業大学工学院経営工学系

本データは病院電子カルテから抽出されており、肺がん患者約 23 万人・約 940 万レコードのデータが含まれ、患者の年齢・性別などの背景情報や、検査結果、処置等が記録されている。抗がん剤治療による ILD 発症の介入効果推定を行う為に、説明変数として患者ごとの背景情報・既往歴・投与薬物・臨床検査値を用いた。ILD の発症は、国際疾病分類であり、RWD 社データでも利用されている、ICD-10 コードに基づき定義し、これらの ICD-10 コードを持っている症例を ILD 発症例とした (表 1)。

表 1 間質性肺疾患に関する ICD-10 コードと症例数

主な傷病名名称	icd10	症例数	全症例に対する割合 (%)
間質性肺疾患	J849	76373	0.8127
急性間質性肺疾患	J841	11276	0.1200
急性薬物誘発性間質性肺疾患	J702	52	0.0006
慢性薬物誘発性間質性肺障害	J703	13	0.0001
薬剤性間質性肺炎	J704	1715	0.0182

#### 2.2 X-Learner

##### 2.2.1 理論

X-Learner は機械学習を用いて条件付き平均処置効果 (CATE) を求める手法である。まず、対照群と介入群それぞれで介入 (Treatment) 以外の共変量 (Covariate, X) から結果 (Outcome, Y) を推定する機械学習モデル (M1, M2) を作成する。

$$\hat{\mu}_0(x) = M_1(Y^0 \sim X^0) \quad (1)$$

$$\hat{\mu}_1(x) = M_2(Y^1 \sim X^1) \quad (2)$$

このモデルを用いて、それぞれの群で介入の有無が逆だった場合を推定する。

次に、逆の介入だった場合の推定値と本来の介入の結果の差を処置効果 (Treatment Effect,  $\hat{D}_i^0, \hat{D}_i^1$ ) に相当するものとみなし、これを推定する機械学習モデル (M3, M4) を作成する。

$$\hat{D}^0 = \mu_1(X^0) - Y^0 \quad (3)$$

$$\hat{D}^1 = Y^1 - \mu_0(X^1) \quad (4)$$

$$\hat{t}_0(x) = M_3(\hat{D}^0 \sim X^0) \quad (5)$$

$$\hat{t}_1(x) = M_4(\hat{D}^1 \sim X^1) \quad (6)$$

最後に、患者個人が対照群、介入群のどちらに振り分けられやすいか、すなわち介入確率 $g(x)$ をその共変量から推定するモデルを作成する。患者個人ごとに、対照群だった場合と、介入群だった場合の処置効果を介入確率で重みづけすることで、最終的に患者個人の条件付き処置効果の推定値が得られる。

$$\hat{t}(x) = g(x)\hat{t}_0(x) + (1 - g(x))\hat{t}_1(x) \quad (7)$$

今回のデータでは各患者の属性、既往歴、薬剤投与履歴、検査値データを共変量とし、抗がん剤投与の有無を介入、ILDの発症の有無を結果とすることで検証を行った。

### 2.2.2 ILD 発症予測モデルの開発

X-Learner での解析を行うにあたり、結果 (Outcome) の予測については、最初の抗がん剤投与から 3 ヶ月以内の ILD 発症の有無を予測する二値分類モデルを構築し、利用した。特徴量として、年齢・性別・既往歴などの患者背景、薬剤投与履歴、直近の検査値記録を用い、これらのデータが取得できる患者を解析対象とした。既往歴と薬剤投与履歴については、最初の抗がん剤投与前の全期間データを遡り、履歴のある疾病名および薬剤名について One-hot エンコーディングを行うことで特徴量を生成した。特徴量加工後のデータは 9177 特徴量×80377 例となった。作成したデータセットを用いて LightGBM による ILD 発症予測モデルを開発した。

## 3. 結果

### 3.1 抗がん剤治療と ILD 発症に関する基礎集計

肺癌患者電子カルテデータにおける抗がん剤治療と ILD 発症に関するデータについて基礎集計を行った。抗がん剤治療法毎の ILD 発症割合は 3.20 ~ 4.03% であり、プラチナ製剤を上乗せした細胞障害性抗がん剤（以降プラチナ製剤と略記する）が最も高く、がん免疫療法が最も低かった（表 2）。同日に肺癌と ILD が記録されたデータは除いた上で、最初の抗がん剤投与から ILD 発症までの日数を調べたところ、1 か月後をピークとして、ILD 発症記録の 6 割以上が 3 か月後までに記録されていた（図 1）

表 2 抗がん剤治療方法毎の ILD 発症者数

	ILD 発症者数	ILD 未発症者数	全患者数	ILD 発症割合 (%)
細胞障害性抗がん剤	271	7193	7464	3.63
プラチナ製剤	264	6289	6553	4.03
がん免疫療法	24	725	749	3.20
分子標的薬	102	3003	3105	3.29
合計	661	17210	17871	3.70

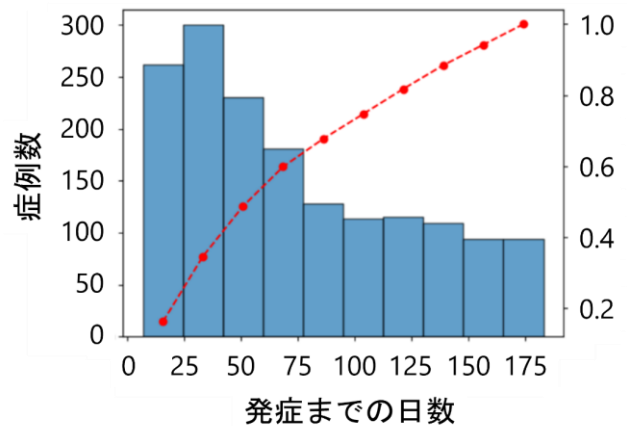


図 1 最初の抗がん剤投与から ILD 発症までの日数

### 3.2 間質性肺疾患発症予測モデルの開発

ILD の発症がある程度予測可能か確かめるために、介入を考慮しない ILD 発症予測モデルを開発した。最初の抗がん剤投与から 3 か月以内の ILD 発症の有無を予測対象とし、ラベルデータを作成した。機械学習のアルゴリズムは LightGBM を用いた。表 2 に示したように、ILD の発症割合は全データの 4% 以下であった。データの不均衡に対応するため、負例は 20% までダウンサンプリングし、学習を行った。テスト集合に対して ROC 曲線を用いて予測モデルの精度を評価したところ、AUC 値は 0.78 を示し、一定の精度で ILD 発症を予測できることが示唆された（図 2）。予測に重要であった特徴量を調べるため、shapley 値[5]を算出したところ、年齢や白血球数などの ILD との関連が報告されている因子が重要な特徴量として挙げられてくることが確認できた。（図 3）。

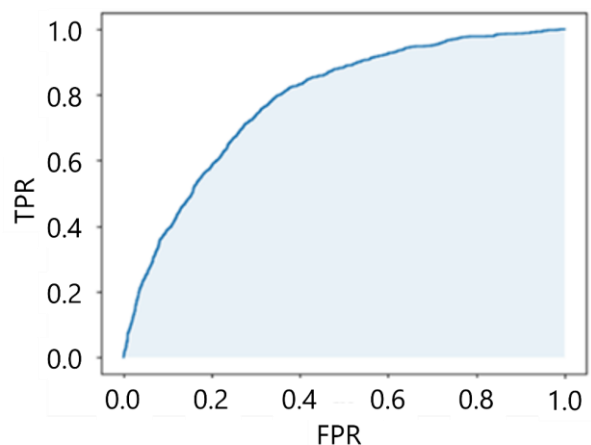


図 2 ILD 発症予測モデルの ROC 曲線

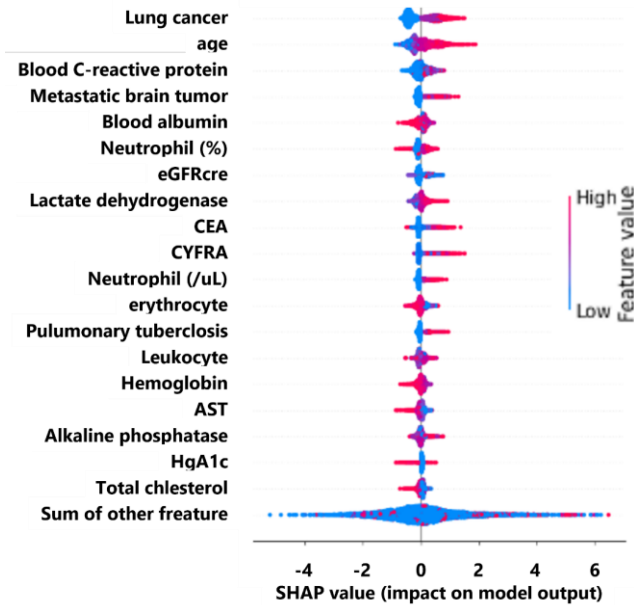


図 3 ILD 発症予測モデルにおける shapley 値の分布

### 3.3 X-Learner による薬剤投与リスク評価

X-Learner を用いた抗がん剤治療による ILD 発症リスクの評価を行った。X-Learner の各ステップで用いる機械学習モデルは、3.2 項と同じく LightGBM で学習を行った。薬物療法を受けているという条件を揃えるため、介入群は特定の種類の抗がん剤が投与された患者、対照群はその抗がん剤以外の抗がん剤が投与された患者とした。対象群の ILD 発症予測モデルと介入群の ILD 発症予測モデル、対象の抗がん剤の投与予測モデルを、異なる 4 種の治療ごとに開発した。モデル開発時には 4 分割の交差検証により、AUC 値を指標としたハイパーパラメータチューニングを行った。結果、精度にばらつきがあるものの、治療種別ごとの対象群および介入群の ILD 発症予測モデルが開発できた。また、抗がん剤の投与予測モデルは比較的高い精度を示しており、患者の背景情報や検査値記録は、どの種類の抗がん剤を投与するかと関連していると言える。

表 3 X-Learner の各二値分類モデルの AUC 値

	対象群の ILD 発症予 測モデル	介入群の ILD 発症予 測モデル	抗がん剤 投与予測 モデル
細胞障害性 抗がん剤	0.642	<b>0.781</b>	<b>0.781</b>
プラチナ製剤	0.626	0.651	<b>0.816</b>
がん免疫療法	0.628	<b>0.732</b>	<b>0.916</b>
分子標的薬	0.660	0.600	<b>0.877</b>

続いて、開発した各 ILD 発症予測モデルを用いて、抗がん剤の種類ごとの条件付き平均処置効果 (CATE) を求めた。その結果、プラチナ製剤の CATE が最も高く、ILD の

誘引リスクが高い可能性が示唆された。一方で、分子標的薬は CATE が最も低く、ILD 誘引リスクが低い可能性が示唆された (表 4)。

表 4 抗がん剤種別の条件付き平均処置効果

	条件付き平均処置効果
細胞障害性抗がん剤	0.0052
プラチナ製剤	0.0121
がん免疫療法	0.0065
分子標的薬	0.0005

X-learner では患者ごとに条件付き処置効果が算出されるため、発症に対して介入による影響が大きいと判断された患者において、どのような特徴量が高くなっているのかを調べるために、相関分析を行った。

条件付き処置効果と相関の強い特徴量を調べたところ、プラチナ製剤では、非小細胞肺癌 (non-small cell lung cancer) と性別が正の相関を、分子標的薬では負の相関を示した (表 5)。本データの性別は男性=1、女性=2 と変換されており、プラチナ製剤においては女性の CATE が強い傾向があるということになる。分子標的薬においては、肺癌で投与例の多い EGFR 阻害剤というサブグループにおいてみられる副作用 (asteatosis, diarrhea, acneiform eruption) が、処置効果と逆相関を示した。

表 5-1 処置効果と相関の強い特徴量 (プラチナ製剤)

特徴量名	相関係数
non-small cell lung cancer	0.2719
sex	0.1902
Metastatic liver tumor	0.1708
Paired creatinine clearance L/day	-0.1490
urticaria	0.1449
blood hemoglobin A1c	0.1405
blood sodium	0.1363

表 5-2 処置効果と相関の強い特徴量 (分子標的薬)

特徴量名	相関係数
non-small cell lung cancer	-0.2134
sex	-0.1953
blood total cholesterol mg/dL	-0.1330
asteatosis	-0.1021
diarrhea	-0.0971
acneiform eruption	-0.0971

#### 4. 考察

本研究でははじめに、大規模データでの ILD 発症予測可能性を示すために、ILD 発症予測モデルを学習した。学習されたモデルの shapley 値を解析したところ、年齢や血液検査値などが ILD 発症に寄与することが示唆された。個別の特徴量については既に ILD 発症に関連があるとされているもの[1]が多く見られたが、大規模データから機械学習を用いて疾患と関連がある特徴量を抽出し観察することは、疾患と現時点での治療実態を客観的にとらえることが出来る有用な手段の一つである。

X-Learner は介入確率を推定することで処置効果の重みを調整しているため、抗がん剤のような介入の起こりやすさが患者ごとに異なる疾病の場合、その真の処置効果を推定するのに有効であると言える。X-Learner による解析により、抗がん剤投与における選択バイアスを考慮した上で、条件付き平均処置効果 (CATE) 算出したところ、抗がん剤の種類ごとに CATE が異なる様子が見られた。また、CATE と患者の特徴量の相関分析を行うことで、抗がん剤治療薬の種類に応じて ILD 発症のリスク因子が異なる可能性も示唆された。プラチナ製剤では肝転移や他臓器状態を示す特徴量がヒットした一方で、分子標的薬では、ILD 以外の副作用として知られている症状が逆相関する特徴量として挙げられている。これらの症状は ILD の発症抑制に働いている可能性はあるが、何かしらの副作用が発生している際は減薬、休薬を行うことが一般的であるため[6]、他の副作用への配慮により、ILD の発症が結果として抑えられているという可能性も考えられる。

また、今回のデータでは共変量に含められていないが、ILD の強いリスク因子として喫煙が知られている。処置効果と関連のある特徴量として、非小細胞肺癌 (non-small cell lung cancer) と性別があげられたが、非小細胞肺癌でない肺癌である小細胞肺癌は喫煙との関連が強いこと [7]、男性で喫煙率が高いことから、これらが相互に関連していることには注意が必要である。X-learner は因果推論の考え方を元にした手法ではあるが、実臨床では副作用を考慮した治療や、事前にハイリスク群への投与が避けられることを鑑みても、本研究で見いだされた特徴量が ILD の原因となるかは、さらなる研究を必要とする。

本研究で利用した X-Learner の各二値分類モデル精度は改善の余地があるが、新たなリスク因子の追加などで精度が改善されれば、より正しく介入効果を推定できるようになることが期待される。また、患者属性や疾病歴、検査値に基づいて患者個別の ILD 発症リスク評価が行える可能性がある。

#### 5. おわりに

本研究では、肺癌患者のリアルワールドデータを用いて、間質性肺疾患 (ILD) 発症の予測、及び抗がん剤ごとに間質性肺疾患への介入効果推定を行った。ILD 患者の予測から、肺癌患者における ILD のリスクをある程度予測できる可能性があり、血液検査値データがリスク評価に有用であることが示唆された。

また、抗がん剤の間質性肺疾患に対する介入効果推定では、X-Learner を用いた効果検証によって、現状のリアルワ

ールドデータにおいて各種の抗がん剤が ILD 発症に及ぼしているリスクを可視化することが出来た。

本研究で用いた予測モデルの精度にはまだ改善の余地があり、データ数やデータ項目の充実による精度向上など、より正確に介入効果を推定できるような方法をデータ収集方法とアルゴリズム理論の両面から探索していくことが求められる。

#### 謝辞

本研究は HCEI にて開発、提供されている DB を用いて実施されました。この場を借りて御礼申し上げます。

#### 倫理

本研究で使用したヒト由来データは匿名加工情報として受領したものであり、組織内審査の上、人を対象とする生命科学・医学系研究に関する倫理指針の対象外とした。

#### 参考文献

- [1] 厚生労働省, “重篤副作用疾患別対応マニュアル 間質性肺炎” (令和元年 9 月改訂版)
- [2] Hill JL “Bayesian nonparametric modeling for causal inference.” *J Comput Graphical Stat* 20:217–240 (2011)
- [3] Athey S, Imbens G Recursive partitioning for heterogeneous causal effects. *Proc Natl Acad Sci USA* 113:7353–7360 (2016)
- [4] Künzel, S. R., Sekhon, J. S., Bickel, P. J., & Yu, B., “Metalearners for estimating heterogeneous treatment effects using machine learning”, *Proceedings of the national academy of sciences.*, 116.10 (2019)
- [5] Winter, E. “The shapley value. Handbook of game theory with economic applications”, 3, 2025-2054 (2002).
- [6] 里内 美弥子 “分子標的治療薬の使い方”, *Japanese Journal of Lung Cancer—Vol 61, Supplement, Nov 26, (2021)*
- [7] Sobue T, Yamamoto S, Hara M, Sasazuki S, Sasaki S, Tsugane S; JPHC Study Group. Japanese Public Health Center. “Cigarette smoking and subsequent risk of lung cancer by histologic type in middle-aged Japanese men and women: the JPHC study.” *Int J Cancer*.99(2):245-51. (2002)