

厚生労働科学研究費補助金
(政策科学総合研究事業 (臨床研究等 ICT 基盤構築・人工知能実装研究事業))
分担研究報告書

課題名 : 新薬創出を加速する症例データベースの構築・拡充/創薬ターゲット推定アルゴリズムの開発
研究分担者名 : 小倉 高志
神奈川県立循環器呼吸器病センター 所長 兼 臨床研究所所長 兼 間質性肺炎センター室長

研究要旨**(1) 統合する医療情報統合データベース 構築 :**

指定難病の一つである特発性肺線維症を主体とした間質性肺炎患者に関する診療情報と患者サンプル(血液、肺組織)の収集を行う。さらに収集した患者サンプルのオミックス解析を実施する。それにより得られた臨床情報(診療情報および患者サンプルから得たオミックスデータ)を格納した疾患統合データベースを構築する。さらに得られた臨床情報を医薬基盤研にて解析することにより、創薬ターゲット分子の発見・同定を行う。

間質性肺炎に関しては、臨床経過が非常に重要であり、経過の臨床情報の取得と転帰を確認する。

(2) Proteome 解析 :

特発性肺線維症を主体とした間質性肺炎患者の疾患診断、進行、予後予測のバイオマーカーの探索は臨床的にも大変有用である。今回、IPFを含む間質性肺炎(疑いを含む)の診断となった患者の血清からのプロテオーム測定を医薬基盤研で実施し、エクソソームの中に含まれるタンパク質を DIA (Data-Independent Acquisition) 法により同定・定量している。そして、疾患診断、鑑別、予後などの関係性を評価する。

(3) 画像・病理定量解析 :

特発性肺線維症を主体とした間質性肺炎患者に対して施行された胸部画像検査や採取された肺生検の病理組織標本を用いて、定量解析を行う。

一つ目は IPF を含む間質性肺炎(疑いを含む)の診断となった患者から得られた、胸部 X 線画像、CT 画像に対して、アノテーションツールを用いて、線維化病巣の定量的解析を行う。

現在ザイオ株式会社と連携し、画像特徴量の抽出を委託している。ザイオ株式会社では、キャノン社製高精細 CT (Precision)にて撮影された画像で従来の再構成方法 (FBP ベース) に 512 マトリックスで再構成された CT 画像データの全肺野領域及び指定された領域を解析することになっている。現在、解析対象となる領域(多くは病変)を選び、その中から画像的特徴抽出が行われている。

Zaio(画像会社)の協力で、画像特徴量抽出に特化した PyRadiomics (無料公開しているソフトウェア)を用いて約 100 例以上の特徴量を抽出している。

PyRadiomics (<http://www.radiomics.io/pyradiomics.html>)

必要な画像特徴量を選択し、統計学的手法によって臨床情報と相関関係のあるものや、機械学習を活用して重要度の高いものを選択する。

二つ目は、令和 1 年に購入した病理組織の定量的組織インフォマティクスである HALO (indica labs) を用いた得られた検体からのスライド組織の迅速かつ定量的評価を行い、ILD の予後予測システムを構築する。

(4) 遺伝子発現 (GWAS とトランスクリプトーム) 解析 :

IPF を含む間質性肺炎に対して施行された肺生検から得られた組織検体が、Precision Medicine に寄与するかを評価した。

採取組織検体から DNA と RNA の同時採取ができるかを確認し、バイオマーカー探索に有用であるかを評価した。また、組織の病理組織標本を用いてトランスクリプトームも行った。

今回トランスクリプトーム解析のために用いた VISIUM は、組織のパラフィン切片を使用して、組織構造内の情報を維持した状態での全遺伝子発現を解明できる画期的な方法である。

この方法は、組織形態と遺伝子活動を重ね合わせて視覚化できる利点がある。主要なバイオマーカーの発現量と発現部位の同定を行う。①各検体の mRNA の標準化、mRNA 発現プロファイルに基づいたクラスタの二次元分布パターンとの相違に基づいた分類法の確立を目指す。②これまでの別の解析から間質性肺炎病巣で特異的に発現する候補分子を得ている、候補分子の発現示すスポットの空間的分布と発現プロフ

ファイルから、間質性肺炎の病態の一端に迫る。

(5) テロメア長の解析：

IPF を含む間質性肺炎の予後とテロメア短縮との関連は言われている。

今回、IPF を含む間質性肺炎の各疾患に対して、血液を使用して、テロメア長の測定を施行した。

A. 研究目的

(1) 統合する医療情報統合データベース 構築

特発性肺線維症(IPF)及び間質性肺炎(疑いを含む)患者に関する診療情報および患者サンプルの収集を行い、臨床情報(診療情報および患者サンプルから得たオミックスデータ)を統合し格納した疾患統合データベースを完成させる。また経過と転帰を確認する。

(2) Proteome解析

特発性肺線維症を主体とした間質性肺炎患者の疾患診断、進行、予後予測のバイオマーカーの探索を行う。

(3) 画像・病理定量解析

特発性肺線維症を主体とした間質性肺炎患者の画像、病理定量からの予後予測システムを構築する。

(4) 遺伝子発現(GWASとトランスクリプトーム)解析

特発性肺線維症を主体とした間質性肺炎患者の組織検体からの全遺伝子発現(トランスクリプトーム)を行う。

(5) テロメア長の解析

特発性肺線維症を主体とした間質性肺炎患者のテロメア長との関係性を検討する。

B. 研究方法と結果

(1) 統合する医療情報統合データベース 構築

研究開始後の全実績：

同意取得 1045 例(うち同意撤回 7 例)

組織採取：血液 1032 例、クライオ生検 237 例、VATS 78 例、手術肺 15 例

令和 4 年 3 月 31 日をもって、予定症例数の登録が終了し、再 MDD も施行し、当センターにおける PRISM 症例の MDD 診断を一旦確定するも、経過と転帰を引き続き確認している。

MDD 診断結果によると、IPF が全体の 32%、330 例であった。その他、UCIP が 15%、161 例、HP が 17%、182 例、NSIP が 10%、106 例であった。典型的な IPF は組織採取の対象とならないことから、血液についても解析の対象とすることとしたが、上記のような分布となった。

(2) Proteome 解析

IPF330 例と HP182 例を対象に血清中のエクソソーム中に含まれるタンパク質の差を解析中である。

(3) 画像・病理定量解析

1) 画像定量研究

当院で肺生検を行い、MDD で診断された IPF 87 例、CHP 93 例において臨床・画像的特徴を後方視的に検討した。画像に関しては臨床的な画像評価に加え、Pyradiomics を用いて画像特徴量を抽出した。現在解析進行中である。

両者の違いが特徴づけられた画像特徴量としては、現在 4 つ認められているが、その画像的意味を検討中である。臨床情報に加えて Pyradiomics の情報を加えることで、診断力を向上させる可能性が示唆されている。

2) 病理定量研究

間質性肺炎切除肺生検(およびクライオ生検)の HE 染色組織切片のバーチャルスライド画像を、病型分類名を「解」としてディープラーニングで学習させた。得られた人口知能を用いて、客観性・再現性のある、予後予測システムを構築した。

現在、すでに診断目的に生検されている 370 症例の組織像を上葉・下葉に分けてそれぞれ学習させ、おなじ 370 症例について上葉・下葉それぞれの病型分類の組成を求め、その組成値（比）に基づいた階層分析（ワード法）から樹形図を描き、上葉・下葉それぞれで 4 個のクラスタを得た。

その組み合わせ（16 通り）の生存分析（カプランマイヤー・ログランクテスト）を行い、予後不良群（5 年生存率 28% ; $P < 0.001$ ）を選別の可能な分類システムの構築を達成した。

(4) 遺伝子発現（GWAS とトランスクリプトーム）解析

気管支鏡下クライオ生検施行された 237 例を対象とした。検体半割し、片方を標本、一部で空間的遺伝子発現解析に使用した。一方を DNA、RNA 抽出のために使用した。

NucleoSpin®RNA 及び RNA/BufferSET を用いて、DNA 及び RNA 抽出。Bioanalyzer による RNA 分解度合い確認では、RIN 値 4 以上であった。これら RNA 検体を用いて次世代シーケンサーのためのライブラリ作成、また RNA シークエンスも問題なく可能であった。

(5) テロメア長の解析

テロメア長は間質性肺炎の種類で異なり、予後と関連するとされる。

IPF において胸部 CT 検査での Pleuroparenchymal fibroelastosis (PPFE) 様所見は予後不良因子と報告されているが、PPFE 様所見を有する IPF 患者のテロメア長の報告はない。

今回は、PPFE 様所見の有無で IPF 患者のテロメア長を比較検討した。

当院でテロメア長を測定した間質性肺炎 492 例のうち、IPF、PPFE と診断された患者を抽出し、さらに IPF 患者を PPFE 様所見の有無で分け、3 群で比較した。

PPFE 様所見のない IPF (IPF/UIP 群) 151 例、PPFE 様所見のある IPF (IPF/PPFE 群) 29 例、PPFE (PPFE 群) 31 例が登録された。観察期間内で呼吸機能低下、急性増悪を来した患者の割合は、IPF/PPFE 群 > IPF/UIP 群 > PPFE 群の順で有意に多かった。一方、テロメア長は IPF/PPFE 群 < IPF/UIP 群 < PPFE 群の順で有意に短かった ($p < 0.01$)。3 群間で年齢調整したテロメア長を比較した解析でも、テロメア長は IPF/PPFE 群 < IPF/UIP 群 < PPFE 群の順で有意に短かった ($p < 0.01$)

IPF 患者 180 例において 1 年間の呼吸機能の低下をアウトカムとした多変量解析では、年齢 (OR: 0.96、 $p = 0.035$) とテロメア長短縮 (OR=0.77; $p = 0.040$) が有意なリスク因子であった。

（倫理面への配慮）

課題名「新薬創出を加速する症例データベースの構築・拡充/創薬ターゲット推定アルゴリズムの開発」について、当院の倫理委員会の承認を得て研究開始。同意説明文書を用いて医師が患者へ説明を行い、同意を取得した患者に研究協力をして頂いている。また、同意撤回の機会を保障している。

C. 考察

(1) 統合する医療情報統合データベース 構築

コホート内の疾患分布、特に UCIP の分布が問題になることが多い。UCIP の疾患頻度は、報告により様々であるが、本邦における外科的肺生検を実施した検討では、中央判定の MDD によっても 36%であり 1) Guler によるメタアナリシスでは、11.9%と報告されており、2) 今回のコホートにおける UCIP は適切と考えられる。

また、本コホートには肺組織採取例が 300 症例入っており、大変貴重なコホートであると考えられる。

Reference

- 1) Fujisawa T, et al. Eur Respir J 2019;53:1802243
- 2) Guler SA, et al. Ann Am Thorac Soc 2018; 15: 854-863

(2) Proteome 解析

今回の研究におけるエクソソームからの抽出されたタンパク質は、類似して動くタンパク質が多いこと、また肺外（特に肝臓由来）と関係するタンパク質が多いために、解析には時間がかかっている。

IPF と HP の鑑別は、臨床的にも非常に難しく、診断に有用なタンパク質を同定することは非常に臨床的にも意義のあるため、引き続き解析を進めていく。

(3) 画像・病理定量解析

今回のコホートでは、CT 検査情報に対して、画像特徴量抽出に特化した PyRadiomics (無料公開している

ソフトウェア)を用いて約 100 項目以上の特徴量を抽出している。現在、研究対象である IPF と HP 症例を主に解析し、有意な画像特徴量を選択し、統計学的手法によって臨床情報と相関関係のあるものや、機械学習を活用して重要度の高いものを選択している。しかし実際の抽出された特徴量の評価が非常に難しく、慎重に検討する必要がある。

ただ、得られた研究成果からは、IPF と HP において、臨床情報に加えて Pyradiomics の情報を加えることで、診断力を向上させる可能性が示唆されている。更なる検討が必要である。

病理定量解析に関しては、すでに診断目的に生検されている 370 症例の組織像を上葉・下葉に分けてそれぞれ学習させ、おなじ 370 症例について上葉・下葉それぞれの病型分類の組成を求め、その組成値 (比) に基づいた階層分析 (ワード法) から樹形図を描き、上葉・下葉それぞれで 4 個のクラスターを得ている。

但し、まだ validation を行なっておらず、別コホートでの対応が必要である。

(4) 遺伝子発現 (GWAS とトランスクリプトーム) 解析

クライオ生検・切除肺生検材料を用いてビジウム解析を行った。バイオインフォマティクス解析で上記の項目をさらに完遂していく。現在は、切除肺生検もしくはクライオ生検パラフィン切片検体からのビジウム解析が完了している。ビジウムは、パラフィン切片検体から解析することができ、1 細胞解析とは少し異なるが、組織内の遺伝子の発現を見る上では大変重要な手法と考える。

(5) テロメア長の解析

ILD type、家族歴、CT 所見とテロメア長との関連が示された。既報に反して CVD-IP にてテロメア長の短縮を認めており、今後テロメア関連遺伝子の検討を行い、呼吸機能低下との関連について検討していく。

D. 結論

(1) 統合する医療情報統合データベース 構築

今回の登録患者様の中では、肺組織採取症例が 300 例であった。医薬基盤研との協力によりこれらの症例の臨床情報と検査結果を格納した新薬創出に資する世界に類をみない疾患統合データベースの構築を引き続き推進している。

(2) Proteome 解析

特発性肺線維症を主体とした間質性肺炎患者の疾患診断、進行、予後予測のバイオマーカーは、実地臨床では有用なものはない。バイオマーカーの同定は、臨床的にも非常に有用と考えられる。バイオマーカーが同定された場合、それにより治療が変わる可能性や、診断キット開発へとつながるものと考えられる。

(3) 画像・病理定量解析

IPF と HP を臨床的に鑑別することは難しいが、Pyradiomics からの画像抽出情報の追加が、2 群の鑑別に有用ならば、大変臨床的に有用であると考えられる。今回は、侵襲度が比較的低い CT 検査の結果を用いており、診断のための患者侵襲も減ることになる。

病理からの情報は、診断や治療効果の予測に使われることがあっても、予後との関連は評価できない。今回予後予測システムの構築は、病理組織を採取する意義が診断だけでなく、予後予測になる可能性がある点で非常に重要である。

今回の分類システムの有用性を確認するため、現在のコホートを用いた同様の解析を行う。

今コホートで肺組織にある症例は 300 例であり、その症例を目標とする。また、HE 染色の不安定性 (劣化、施設間の色調の違い) などが再現性に影響する可能性があり、多施設の染色、染色時期による解析結果への影響について検討したい。

(4) 遺伝子発現 (GWAS とトランスクリプトーム) 解析

引き続き解析を継続していく。

(5) テロメア長の解析

IPF 患者における PPF 様所見はテロメア長短縮と関連し、呼吸機能低下と予後との関連が示唆された。今後、各間質性肺炎症例に対しての年齢調整テロメア長と予後や進行に関して評価を考えている。また、テロメア長とテロメア関連遺伝子 (TERT など) との関連性を確認していきたい。

テロメア長が、間質性肺炎のバイオマーカーになる可能性が示されているので、更なる検討を計画している。

E. 研究発表

1. 論文発表
なし

2. 学会発表

- 1) 第62回 呼吸器学会総会 ポスター発表 発表者 室橋光太 小倉高志

F. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得
なし

2. 実用新案登録
なし

3. その他
なし