

## 博士論文（要約）

電子化診療記録の計算機処理における特徴と課題に  
関する研究 —e-phenotyping 技術確立に向けて—

香川璃奈

# 論文の内容の要旨

論文題目：電子化診療記録の計算機処理における特徴と課題に関する研究 —  
e-phenotyping 技術確立に向けて—  
氏名：香川璃奈

## <背景>

本邦の医療現場において EHR システムの導入率が増加しており、電子化されたデータを大規模な研究材料として研究に直接利用する環境が整いつつある。当然のことながら必要な情報を人手で確認してはコストがかかる、しかし、患者の EHR データ全体の中の病名情報のみを用いても、診察して EHR に記録を行った医療者が考えた phenotype を正しく識別できないことが知られている。そこで、患者が特定の疾患、すなわち phenotype を持つか否かを計算機処理により自動で EHR データから直接正しく識別する技術が求められている。この技術は e-phenotyping と呼ばれる。しかし既存の e-phenotyping 研究は個々の手法の言及にとどまり、新しく e-phenotyping 手法を開発する際に有益な情報は少ない。その理由として、国あるいは施設の違いにより研究毎に入力データが異なること、計算機処理は技術革新が速く新しい手法が次々と適用される傾向にあること、phenotype は研究目的に応じて無限に定義可能であり疾患概念も時代とともに変わり続けることがあげられる。このことから、特定の入力データ、特定の計算機処理技術、特定の phenotype に依存しない包括的なモジュールを作成し、それらを再利用しながら疾患に応じて必要な部分だけを新規開発することで効率化を図るべきだと言われている。

本研究では、医学知識を踏まえた深い意味の理解が求められる EHR データをどのような方法で計算機処理すべきかを検討する際に、医師が日常診療において患者の EHR データから疾患に関する意味を総合的に判断していることに着目した。annotation において、annotator が phenotype の有無を想起可能とした情報が、annotation においてどのような役割を持つか、をその情報が持つ「意味」と定義する。特定の phenotype に依らない形でその意味を抽象化し、さらにデータ形式を問わない計算機処理技術の観点から整理した annotation criteria を作成すれば、その annotation criteria が特定の入力データや phenotype、計算機処理技術に依存しない e-phenotyping 技術のフレームワークになると着想を得た。このフレームワークに基づけば、研究毎に利用可能なデータ形式に応じて、e-phenotyping 技術の開発前に採用すべき計算機技術を検討することができる。さらに annotation において正例と負例だけではなくより細かな分類を行うことで、研究目的によって無限に定義されうる phenotype を包括的に管理でき、選択した phenotype に応じた技術選択が可能となる。このためには EHR データを医療者が解釈する際に用いる情報を、医師が理解する「意味」に基づいて分割する必要がある。本研究では、医師が疾患の有無に

ついて想起できた、という観点から EHR データを要素に分割し、EHR データ要素から想起できた「意味」を分類する。

#### <手法と目的>

対象とする phenotype は多様な特徴を持つ 10 疾患とした。2012 年 1 月 1 日–12 月 31 日の 1 年間に東京大学医学部附属病院(以下、東大病院)外来受診あるいは入院した患者のうち、2009 年–2014 年の間に健診以外で 2 回以上受診した患者 104,522 人の EHR データを実験対象として、3 つの研究を行った。

研究 1 では、医師が annotation に用いる EHR データ要素を意味の観点から特定の phenotype に依らずに類型化し、データ形式を問わない計算機処理技術でまとめ、annotation criteria を作成する。

研究 2 では、作成した annotation criteria を基にした annotation の精度検証を行う。さらに本研究の annotation 結果と、既存データベースである院内がん登録と、既に報告されている annotation criteria に基づく結果との比較を行うことで、本 annotation criteria の有用性を明らかにする。

研究 3 では、作成した annotation criteria に基づいた e-phenotyping フレームワークを提案する。これに基づいて、目標精度や利用する評価指標に応じて、用いるべき計算機処理技術の類型化を目指す。

#### <結果と考察>

研究 1 において、医師が annotation に用いる EHR データ要素を医師 1 名で整理を行った。その結果、医師が EHR データから疾患の有無を推測する際に用いる情報のうち、意味の深い理解が必要となる情報として、病名記載の意味を持たない病名記載、病名記載以外の情報、病名記載の変化、検体検査結果と医師の判断が異なるもの、の 4 種類が存在し、対象疾患全てにこれらの情報のうち少なくとも最初の 3 種類が存在することを示した。この結果に基づき計算機処理における難易度を考慮した annotation criteria を作成した。病名記載や診断基準といった疾患の有無に関する直接的な情報と、病名記載/診断基準以外ではあるが対象疾患の有無を想起できる間接的な情報の 2 軸に分けた annotation criteria とした。annotator はこの 2 軸のそれぞれにどの情報が存在するかを確認すれば、自動的に annotation が可能になる。特徴は、e-phenotyping のフレームワークに転用できること、言語や施設ごとの患者の特性に依存しないため汎用性が高いと期待されることにある。既存の annotation criteria では判断に迷うような症例を、両軸の正解付けの結果から annotation の根拠を共有しながら明示的に annotation できる利便性も示唆された。

研究 2 では、作成した annotation criteria を用いて医師が annotation を行った。対象疾患のうち 9 疾患で  $\kappa$  係数  $> 0.81$  と、本 annotation criteria の妥当性が示唆された。また、間接的な情報を用いなければ正しく annotation できない症例が肺塞栓以外のすべての対象疾患で認められ(平均 13.9%)、間接的な情報を annotation に用いる必要性、そして間接的な情報の利用を明示した本 annotation criteria の有用性が示唆された。さらに、annotation

criteria が異なると正例集合の性質が有意に異なり、開発された e-phenotyping 手法の精度の比較も十分に行えないことが示唆された。このことから、特定の正例集合の識別や特定の言語に依らない形で annotation の根拠を共有可能である本研究の annotation criteria の有用性が示唆された。

研究 3 では、作成した annotation criteria に基づいて縦軸と横軸それぞれを正しく識別する計算機処理を行うことで、目的とする正例を識別するものを e-phenotyping のフレームワークとした。このフレームワークに基づいて、病名記載と検体検査結果のみを用いた e-phenotyping 手法を用いた理論値の計算結果から、e-phenotyping 研究において、目標精度が感度または特異度 95%以下、または陽性的中率 80%以下の場合には病名記載と検体検査値の診断基準を正しく識別するための技術開発を行うべきことが示された。これはすでに実用レベルでの研究が進んでいるタスクであり、実現可能性は高いと考える。その一方で、100%に近い感度や特異度、あるいは先行研究で採用された 95%以上の陽性的中率を達成するためには間接的な情報も用いた識別が可能な技術開発が必要なことが示唆された。annotation 結果から理論値を計算でき、用いるべき計算機処理技術の検討ができる点で、作成した annotation criteria の利便性が示唆された。

本研究で作成した e-phenotyping フレームワークの汎用性について、医師が annotation に用いる情報を分類した結果が、複数の医師の知見を統合したものでもないこと、大学病院 1 施設のみでのデータを用いたことが限界である。そのため、今後施設を変え、より客観的なデータを収集する研究を行う必要がある。しかし本研究のフレームワーク作成においては、大学病院でないと扱わないような疾患や検体検査など、大学病院に特化した情報の分類は行わなかった。そのため、他施設のデータで実験を行った場合に、各 EHR データ要素の存在割合や患者の基本性質などの数値は変わりうるが、フレームワークは変わらないと考える。ただし、医師の専門性や患者の特徴などの人的資源の特徴以外にも、東大病院が採用している EHR のシステムに依存する結果がある可能性は否定しきれない。

研究 1 より、計算機処理が困難になる 1 つの要因である、医師でも解釈できない記載やあきらかな間違いに基づく記載が、疑い記載に対して 10 疾患の平均で 4.1%と最も多く認められた。これは臨床現場において利用される利点もないため、より多くの疾患を対象にして体系化を検討し、計算機的な入力支援や医療者への教育的指導を検討する必要がある。また、病名記載がなく病名記載以外の情報から病名の有無を推定しなければいけない症例が 10 疾患の平均で 29.2%存在することがあきらかになった。他職種が EHR データを介して患者の情報を理解する医療現場において、医療安全の観点からもさらなる調査が必要である。

#### <結論>

医師が疾患の有無を EHR データ要素から判断する際に用いる情報のうち、意味の深い理解が必要となる情報を、疾患や言語に依存しないかたちで 4 種類に類型化し、これを用いた annotation criteria を作成した。これを用いて 2 名の医師が annotation を行ったところ、

対象疾患のうち9疾患で  $\kappa$  係数  $> 0.81$  と、本 **annotation criteria** の妥当性が示唆された。異なる目的で作成されたデータセットとは得られた正例集合の性質が有意に異なり、**annotation criteria** が異なるだけで既存の **e-phenotyping** 手法の精度も再現できないことが示唆された。このことから、特定の言語や疾患、研究グループに依らない形で根拠が共有された **annotation** が **e-phenotyping** 研究に重要であり、本研究で作成した **annotation criteria** が有用だと示唆された。この **annotation criteria** を計算機処理技術で再現すると想定したところ、直接的な情報のみを用いた識別を正しく行う計算機処理技術という実現可能性の高い識別の開発のみを行った場合でも、感度と特異度は対象疾患を問わず 95%以上を、陽性的中率も 80%以上を目標精度として設定できることが示唆された。

本研究はパイロット研究であり、今後は、専門医も含めた複数の医師への調査や、他施設のデータを用いた研究を要する。