

一般口演 | 電子カルテ・EHR

一般口演1

EHR・PHR

2021年11月19日(金) 09:10 ~ 11:10 E会場 (2号館2階222+223)

[2-E-1-01] がん患者の臨床アウトカムにおける電子カルテデータベースを用いた評価方法の後ろ向き研究

*荒木 賢二¹、米本 直裕²、東郷 香苗²、大木 恵美子³、徐 凌華²、長谷川 義行⁴、井上 裕文⁴、山下 咲子⁴、松元 信弘⁵ (1. 宮崎大学医学部附属病院 病院IR部, 2. ファイザー株式会社 ヘルスアンドバリュー統括部, 3. ファイザー株式会社 オンコロジーメディカル・アフケアーズ部, 4. 株式会社NTTデータ 製造ITイノベーション事業本部, 5. 宮崎大学医学部附属病院 呼吸器内科)

*Kenji Araki¹, Yonemoto Naohiro², Kanae Togo², Emiko Ohki³, Ryoka Jo², Yoshiyuki Hasegawa⁴, Hirofumi Inoue⁴, Sakiko Yamashita⁴, Nobuhiro Matsumoto⁵ (1. 宮崎大学医学部附属病院 病院IR部, 2. ファイザー株式会社 ヘルスアンドバリュー統括部, 3. ファイザー株式会社 オンコロジーメディカル・アフケアーズ部, 4. 株式会社NTTデータ 製造ITイノベーション事業本部, 5. 宮崎大学医学部附属病院 呼吸器内科)

キーワード : Electronic Medical Records, Cancer, Outcome, Unstructured Data

【背景】電子カルテ（EMR）データの研究利用への期待は大きいですが、臨床アウトカムに関わる多くのデータは非構造化データであるため、解析が困難という課題がある。本研究ではEMRを用いて、がん患者を対象に薬物治療効果などのアウトカムを客観的に評価する手法を検討した。

【方法】宮崎大学医学部附属病院に2018年4月から2021年3月に通院または入院した、肺がん患者15例（グループA）及び肺がん及び乳がん患者100例（グループB）を対象として、後ろ向きにEMRデータを収集した。評価者の医師2名により、グループAの医師の経過記録、腫瘍に対する放射線レポート及び病理検査レポートの記載に基づいて、薬物治療の効果判定を行った。グループBに対しては、グループAにおける治療効果の評価手法を当てはめた。また、遺伝子検査結果がEMRから収集可能か検討した。

【結果】グループAの15例28治療ラインにおける治療効果（奏効、安定、進行）の評価者間の最良効果一致率は0.59であった。治療効果判定に用いられたデータは経過記録が61%、放射線レポート29%であった。判定の基となった用語（キーワード）は、奏効では「縮小」、「効果」、「著変」（否定表現）で、「縮小」は感度、特異度ともに高かった。これらのキーワードはグループBでも確認できた。また、グループBにおいて、遺伝子検査結果は延べ296件中77%で結果を抽出できた。

【考察】EMRの非構造化データからがん患者の治療効果判定に有用な情報を得られることが分かった。今後、複数施設の症例への適応及びキーワード間の関係について検討が必要である。

がん患者の臨床アウトカムにおける電子カルテデータベースを用いた評価方法の後ろ向き研究

荒木 賢二^{*1}、米本 直裕^{*2}、東郷 香苗^{*2}、大木 恵美子^{*3}、徐 凌華^{*2}、長谷川 義行^{*4}、井上 裕文^{*4}、山下 咲子^{*4}、松元 信弘^{*5}

*1 宮崎大学医学部附属病院 病院 IR 部、*2 ファイザー株式会社 ヘルスアンドバリュー統括部、*3 ファイザー株式会社 オンコロジーメディカル・アフケアーズ部、*4 株式会社 NTT データ 製造 IT イノベーション事業本部、*5 宮崎大学医学部附属病院 呼吸器内科

Clinical outcomes in cancer patients using Japanese electronic medical records: a retrospective study

Kenji Araki MD^{*1}, Naohiro Yonemoto PhD^{*2}, Kanae Togo PhD^{*2}, Emiko Ohki PhD^{*3}, Ryouka Jo^{*2}, Yoshiyuki Hasegawa^{*4}, Hirofumi Inoue^{*4}, Sakiko Yamashita^{*4}, Nobuhiro Matsumoto MD^{*5}

*1 Department of Hospital Institutional Research, University of Miyazaki Hospital, *2 Health & Value, Pfizer Japan Inc., *3 Oncology Medical Affairs, Pfizer Japan Inc., *4 Manufacturing IT Innovation Sector, NTT DATA Corporation, *5 Division of Respiriology, Department of Internal Medicine, University of Miyazaki Hospital

Researches utilizing real world data (RWD) have been increasing as a role to overcome or supplement limitations of clinical trials. However, most clinical outcomes are unavailable from structured data but available from unstructured data. This study was planned to examine a method of generating outcomes in cancer patients utilizing electronic medical record (EMR). The study included patients ≥ 18 years of age who visited or were admitted to the University of Miyazaki Hospital between April 2018 and September 2020; and had a diagnosis of lung cancer or breast cancer (Group A: 15 lung cancer patients, Group B: 100 lung cancer or breast cancer patients). Adjudication of the treatment response was performed by two evaluators (experienced oncologists) using clinician's progress notes, radiology reports and pathological reports. Best response to each treatment was identified, and the concordance correlation of adjudicating results of the best response from two evaluators was 0.59. For the treatment response of objective response, "reduction/shrink" was the most common key term (69%) used to describe tumour response with the highest sensitivity and specificity in the progress notes. It was found that unstructured data from EMRs can provide useful information for determining treatment effects in cancer patients. In the future, it will be necessary to examine the indications for cases at multiple institutions and the relationship between the keywords.

Keywords: Electronic Medical Records, Cancer, Outcome, Unstructured Data

1 緒論

近年、医療情報データベースなどのリアルワールドデータ (RWD) を活用した研究が増加しており、臨床試験でのリミテーションを補う役割としてその重要性が増している。本邦におけるデータベース研究は、現在、レセプトや DPC データなどの診療報酬請求データベースを用いていることが多い。米国では電子カルテなどの電子健康記録 (EHR) 由来データベース (EHR データベース) を民間企業が提供しているが、本邦においてそのような民間のデータベースは限られている。EHR データベースは、保険償還を目的として入力される診療報酬請求データよりも、より実際の診療に近く、また、より多くの項目が評価可能であり、本邦でも EHR データベースの活用がより促進されることが望まれる。

一方、EHR データベースを用いる際にも課題がある。臨床アウトカムに関わる多くのデータは非構造化データであり、データの構造化をしなければ解析が困難である。米国では、がん領域において EHR データを用いたリアルワールドのアウトカム (real-world outcome) を行う手法が研究されている。

Griffith¹⁾ は進行非小細胞肺癌 (NSCLC) 患者を対象にした EHR データベースを用いた後ろ向き研究を実施し、固形がんの治療効果判定のためのガイドライン (RECIST)²⁾ に基づく腫瘍増悪の評価はできないことを報告した。しかし、臨床医の経過記録や放射線レポートなどを基にしたリアルワールドの腫瘍増悪の評価方法は実用可能な水準であった。Khozin³⁾ は、免疫療法を受けた進行 NSCLC 患者においてリアルワールドの腫瘍評価と全生存期間の関連性を評価し、適度な関連性が認められた。

本邦では宮崎大学および京都大学を中心とした EHR システム「千年カルテプロジェクト」がある⁴⁾。本プロジェクトと連携したライフデータインシアティブが 2019 年 12 月に次世代医療基盤法認定事業者として初めて認定されたため⁵⁾、今後広く活用されることが期待される。しかし、本邦の EHR データベースを用いた、経過記録等の非構造化データを基に臨床アウトカムを生成する手法に関する研究は、著者の知る限り報告されていない。本研究では、千年カルテの中心である宮崎大学医学部附属病院の電子カルテを用いて、がん患者を対

象に薬物治療に対する効果などの臨床アウトカムを非構造化データより生成する手法を検討した。

2 目的

本研究では、がん患者を対象に電子カルテのデータベースを用いて、経過記録や放射線レポートなどの非構造化データを基に薬物治療効果などの臨床アウトカムを客観的に評価する手法を検討した。また、遺伝子検査結果などの診療報酬請求データベースでは得られ難い情報についても、電子カルテから収集可能かを検討した。

本研究は宮崎大学倫理審査委員会にて2020年12月4日に承認された(受付番号:0-0845)。

3 方法

3.1 試験デザイン

本研究は宮崎大学医学部附属病院の電子カルテを用いた後ろ向き研究である。宮崎大学医学部附属病院に2018年4月から2020年9月に通院または入院した、肺がん患者15例(グループA)および肺がんまたは乳がん患者100例(グループB)を対象とした。

3.2 評価項目

主要評価項目として、評価者の医師2名により、肺がん患者15例(グループA)の医師の経過記録、腫瘍に対する放射線レポートおよび病理検査レポートの記載に基づいて、薬物治療の治療効果を評価した。治療効果の判定はRECISTを参考に、日常診療における薬物治療の治療効果として奏効、安定、進行の判定基準を設定した。具体的には、ベースラインを各薬物治療ライン開始日の1か月前後とし、ベースラインと比較して画像で何らかの腫瘍の縮小が認められる場合を奏効と定義した。ベースラインと比較して悪化または効果不十分もしくは不認容のため薬物治療を中止した場合を進行と定義した。奏効や進行に該当せず、薬物治療を継続している場合を安定と定義した。評価者間の結果の一致率は κ 係数を用いて算出した。

また、グループAの薬物治療効果判定に関連するキーワードを評価した。評価者が抽出した効果判定に重要な記載に対して、形態素解析を用いてキーワードを特定した。各キーワードについて、効果(奏効、安定、進行)に対する感度および特異度を推定した。同様に、グループBに対して、グループAにおける治療効果のキーワードを当てはめた。

副次的評価項目として、薬物治療ラインの情報や遺伝子検査結果が電子カルテから収集可能かを検討した。

4 結果

4.1 患者背景

グループAの肺がん患者15例の平均年齢±標準偏差(SD)は64.2歳±9.9で、約70%が男性であった。原疾患ステージは60%がステージ4であり、次いでステージ3(20%)であった。原疾患の再発なし、転移なしの症例はそれぞれ93%、80%であった。

グループBの肺がん患者70例の平均年齢±SDは67.9歳±10.4で女性が多く(61%)、乳がん患者30例の平均年齢±SDは66.1歳±10.8であった。グループBの肺がん患者の原疾患ステージは約半数がステージ4で、乳がん患者ではステージ1から3が各20%であった。原疾患の再発は肺が

んおよび乳がん患者のほとんどの症例でなく、転移のない症例がそれぞれ67%、77%であった。

4.2 肺がん患者の薬物治療

電子カルテの投与記録から薬物治療ラインをアルゴリズムにより判定した。ただし、他院での治療は考慮せず、本院での最初の肺がん薬物治療を一次治療とした。その結果を評価者の医師2名により確認したが、治療ラインに変更はなかった。グループAの全ての肺がん患者15例が一次治療を受けており、約50%が二次治療を受けていた。15例で合計28の薬物治療ラインを確認した。一次治療としてオシメルチニブメシル酸塩が最も使用され(33%)、二次治療においては、デュルバルマブが最も使用されていた(38%)。

4.3 治療効果判定

評価者の医師2名により、グループAの肺がん患者15例の経過記録2039文書、放射線レポート131文書および病理検査レポート60文書の記載に基づいて、治療効果の判定を行った。また、治療ライン毎に最良効果を判定した。最良効果の判定に用いられたデータは経過記録が61%、放射線レポート29%で、経過記録の割合が最も高かったものの、放射線レポートは文書数(131)が少ないにもかかわらず判定に利用された文書数が多かった(16/131)。15例28治療ラインにおける治療効果(奏効、安定、進行)の評価者間の一致率は0.59であった(表1)。

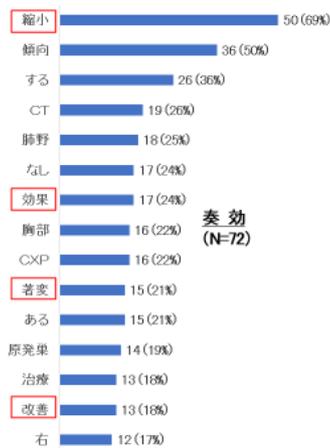
表1 評価者レビューに基づく評価者間の最良効果

肺がん		評価者B							計
		奏効	安定	進行	判定不可	対象外	評価不可	記載なし	
評価者A	奏効	15	0	0	0	0	0	0	15
	安定	2	2	0	0	0	0	0	4
	進行	1	0	3	0	0	0	0	4
	判定不可	0	0	0	0	0	0	0	0
	対象外	0	0	0	0	0	0	0	0
	評価不可	0	3	0	0	0	1	0	4
	記載なし	0	1	0	0	0	0	0	1
	計	18	6	3	0	0	1	0	28

効果判定の基となった臨床上重要なキーワードは、奏効では「縮小」(69%)、「効果」(24%)、「著変」(21%)、「改善」(18%)などであった(図1a)。一方、安定では「著変」(44%)、「浸潤影」(14%)、「増悪」(13%)が比較的良好に使用されており、進行では「効果」(38%)、「増大」(29%)などが良く使用される用語であった(図1b, c)。また、奏効における「縮小」は感度、特異度ともに高い結果であった(図2)。

これらのキーワードの態度表現(肯定・否定)が確認できたもののうち、奏効において「縮小」、「効果」、「改善」はすべて肯定表現として、「著変」はすべて否定表現として使用されていた(表2)。進行において「増大」はすべて否定表現として使用されていたが、「効果」は肯定と否定表現の両方で使用されていた。

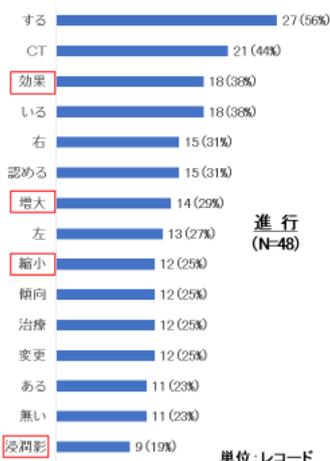
グループBの肺がん患者70例では、奏効の記録以外も含まれているものの、グループAで確認されたキーワードがグループBでも高頻度で認められた。



a. 奏効



b. 安定



c. 進行

図1 治療効果判定関連キーワード(経過記録)

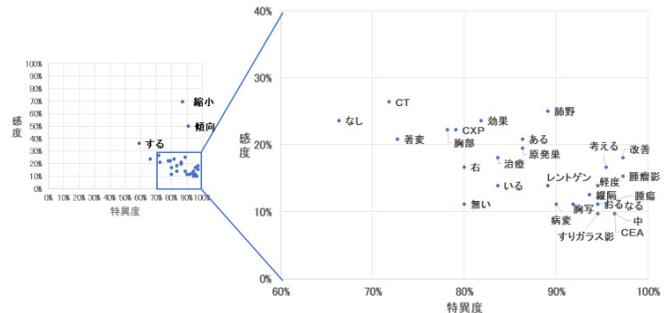


図2 奏効における治療効果判定関連キーワードの感度と特異度

表2 治療効果判定関連キーワードの態度表現

単語	態度表現	登録した表現の種類数	N (%)		
			奏効 (N=72)	安定 (N=71)	進行 (N=48)
縮小	肯定	23	49 (68%)	4 (6%)	8 (17%)
	否定	6	0 (0%)	1 (1%)	5 (10%)
効果	肯定	13	12 (17%)	1 (1%)	3 (6%)
	否定	6	0 (0%)	0 (0%)	2 (4%)
著変	肯定	0	0 (0%)	0 (0%)	0 (0%)
	否定	7	13 (18%)	27 (38%)	1 (2%)
改善	肯定	12	13 (18%)	2 (3%)	2 (4%)
	否定	0	0 (0%)	0 (0%)	0 (0%)
病変	肯定	9	1 (1%)	0 (0%)	5 (10%)
	否定	15	4 (6%)	4 (8%)	3 (6%)
浸潤影	肯定	7	1 (1%)	2 (3%)	6 (13%)
	否定	7	3 (4%)	5 (7%)	1 (2%)
転移	肯定	8	0 (0%)	0 (0%)	1 (2%)
	否定	4	0 (0%)	0 (0%)	0 (0%)
胸水	肯定	10	0 (0%)	0 (0%)	5 (10%)
	否定	5	4 (6%)	1 (1%)	1 (2%)
変化	肯定	17	0 (0%)	1 (1%)	1 (2%)
	否定	9	1 (1%)	4 (8%)	1 (2%)
増大	肯定	9	1 (1%)	1 (1%)	14 (29%)
	否定	2	1 (1%)	0 (0%)	0 (0%)
不変	肯定	1	1 (1%)	7 (10%)	0 (0%)
	否定	0	0 (0%)	0 (0%)	0 (0%)
増悪	肯定	7	0 (0%)	0 (0%)	2 (4%)
	否定	7	0 (0%)	7 (10%)	0 (0%)
サイズ	肯定	0	0 (0%)	0 (0%)	0 (0%)
	否定	16	0 (0%)	3 (4%)	2 (4%)

4.4 遺伝子検査結果

グループ B の肺がん患者 70 例および乳がん患者 30 例において、遺伝子検査は延べ 296 件の実施が DPC データより確認された。そのうち 77% で検査結果を電子カルテから抽出できた。肺がん患者では EGFR、ALK、PD-L1、ROS1 および BRAF の 5 種類の検査のうち EGFR の検査結果が最も高く抽出され (90%)、その他の検査でも検査結果の抽出割合は高かった (76~85%) (図 3)。乳がん患者では HER2 とホルモン受容体の検査がそれぞれ 40 件および 75 件実施され、そのうち検査結果は 68% および 75% で抽出された。

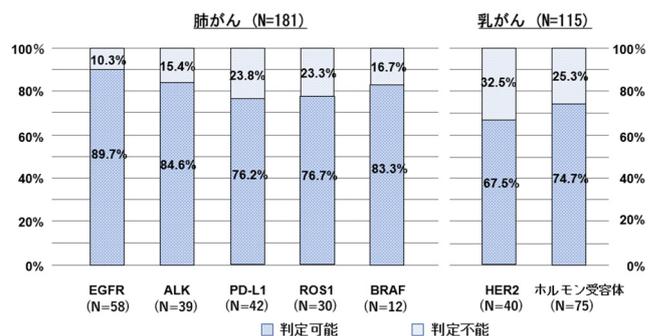


図3 グループ B における遺伝子検査結果

5 考察

本研究により電子カルテの非構造化データからがん患者の治療効果判定に有用な情報を得られることが分かった。経過記録や放射線レポートに基づく治療効果の判定で、頻度は低いものの、評価者間の結果が不一致となったケースがあった。判定の不一致は、経過記録に以前の診察時における記載の写しが多く含まれ、薬物治療効果に関連する記載を

見落としした場合や、薬物治療による効果かそれ以外の治療の効果に関する記載が分かり難い場合などで起こった。今後、以前の経過記録の記載の写しを除外したり、薬物治療情報などを効果的に参照できる仕組みを作ることで、多くの不一致は回避可能と考えられる。

薬物治療効果に関するキーワードについて、本研究の形態素解析で抽出された結果は臨床的にも意味があるものが多く含まれていた。キーワードが肯定的または否定的な文脈で用いられているかが効果の判定において重要である。そのためには、今回のキーワード毎の検討に加えて、今後は複数のキーワード間の関係についての検討や、文脈も考慮した AI による評価モデルを構築することも有用と考えられる。

本研究は宮崎大学医学部附属病院の電子カルテデータに限られていたため、他院からの紹介で来院した患者での他院の治療を考慮していない。紹介状のデータは存在するものの、データベース化することには課題がある。また、本院以外を含む複数施設の症例への適応も検討する必要がある。今後のさらなる研究では、複数施設の電子カルテデータを含む EHR データベースの活用が期待される。

参考文献

- 1) Griffith SD, Tucker M, Bowser B, et al. Generating real-world tumor burden endpoints from electronic health record data: comparison of RECIST, radiology-anchored, and clinician-anchored approaches for abstracting real-world progression in non-small cell lung cancer. *Adv Ther*, 2019;36:2122-2136.
- 2) 固形がんの治療効果判定のための新ガイドライン(RECIST ガイドライン)―改訂版 version 1.1―日本語訳 JCOG 版 ver.1.0
- 3) Khozin S, Miksad RA, Adami J, et al. Real-world progression, treatment, and survival outcomes during rapid adoption of immunotherapy for advanced non-small cell lung cancer. *Cancer* 2019;125:4019-4032.
- 4) 一般社団法人ライフデータイニシアティブ HP. [<https://www.lidi.or.jp/aboutus>]. (Cited 2021-Aug-01)]
- 5) 内閣府 HP. 一般社団法人ライフデータイニシアティブ及び株式会社エヌ・ティ・ティ・データ：次世代医療基盤法に基づく事業者の認定[について]. [<https://www8.cao.go.jp/iryounintei/nintei/jigyousha.html>]. (Cited 2021-Aug-01)]