

第38回抗悪性腫瘍薬開発フォーラム

『産官学共創の力:リアルワールドデータによる抗悪性腫瘍薬開発の革新と実践』

# RWEのグローバル潮流と日本における がんリアルワールドデータ構築の最前線 ーFlatiron Healthの経験を踏まえてー

June 28, 2025

田島 絵里

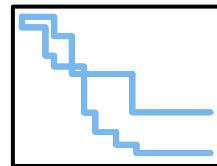
*Country Lead, Flatiron Health Japan*

# リアルワールドデータ・リアルワールドエビデンス



## リアルワールドデータ ( RWD )

さまざまな情報源から日常的に収集される患者  
の状態や医療の提供に関連するデータ



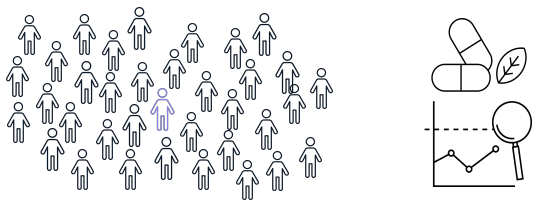
## リアルワールドエビデンス ( RWE )

RWDの解析から得られたエビデンス

RWDを基に適切な研究メソッドの適用により RWEを抽出して  
初めて価値につなげることができる

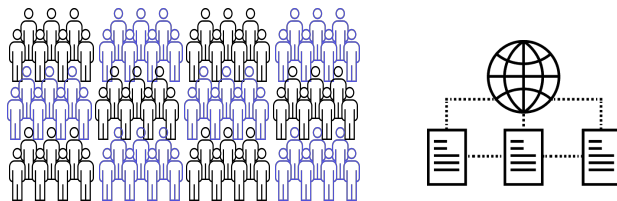
# RCT or RWE ?

## ランダム化比較試験 ( RCT )



がん治療に関するエビデンス創出の  
ゴールドスタンダード  
一方で、患者母集団の代表性が不十分

## リアルワールドエビデンス ( RWE )



RCTでは捕捉できない患者集団等に関する知見を  
実臨床下でのエビデンス生成により補完

双方を活用することで、臨床的に重要な意思決定に繋げることができる

# 医薬品のライフサイクル全体にわたるユースケース例



# CASE STUDY 複数の疾患領域にわたるアンメットメディカルニーズの評価・がん領域のポートフォリオ戦略の策定

## 尿路上皮がん

### 背景

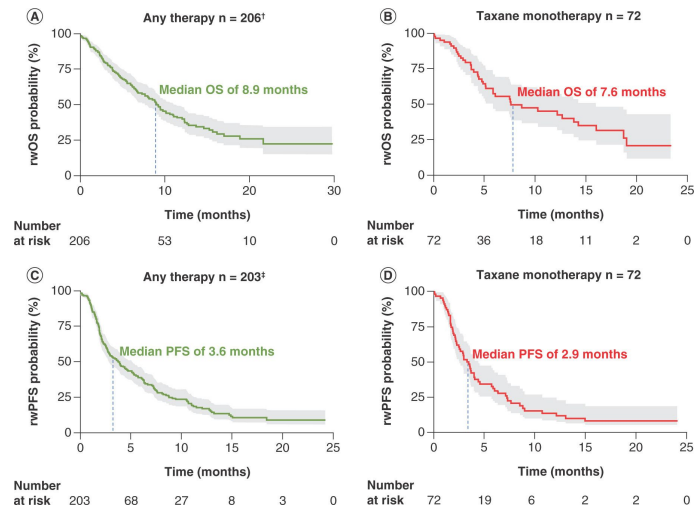
- がん領域のポートフォリオ戦略の策定を目的とし5つの疾患領域より臨床開発の優先領域を決定するためのエビデンス生成を実施
- 進行性尿路上皮がん(aUC)について、本研究実施時点で、化学療法及びPD-1/L1阻害剤の治療後に再発したaUC患者の生存アウトカムを評価する研究は限定的

### Flatiron HealthのRWDを活用したアプローチ

- 対象患者:**
  - 2011年～2018年までに局所進行性または転移性尿路上皮がんと診断
  - プラチナ製剤およびPD-1/L1阻害剤治療後に、タキサン単剤または他の治療を受けた成人患者
- エンドポイント:**
  - 実臨床における全生存期間(rwOS)・無増悪生存期間(rwPFS)

### 結果

- 下記のエビデンスを生成
  - 約半数のaUC患者に対し、初期治療後、追加治療の実施がない
  - 追加治療を受けた患者の生存期間が平均1年未満
- aUC患者の顕著なUMNを踏まえ、ポートフォリオ戦略の策定においてUCを優先



# ケーススタディ RWDの活用による重要なエビデンス創出の早期化

## 非小細胞肺癌

### 背景

- 免疫療法に対する長期奏効を示す進行性非小細胞肺癌患者に対し、治療後2年後に治療中止すべきか、無期限に継続すべきかの評価が求められた

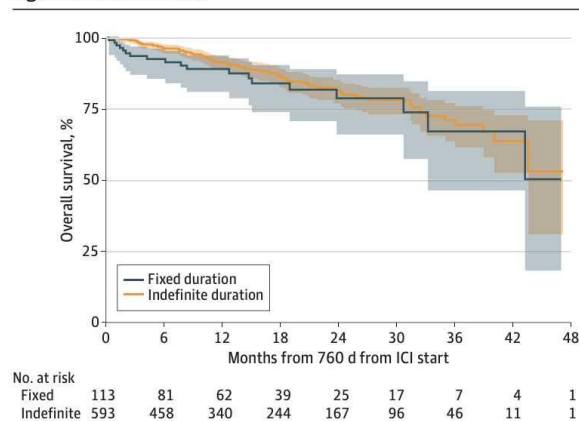
### Flatiron HealthのRWDを活用したアプローチ

- 対象患者:**
  - 2016年～2020年に進行性非小細胞肺癌と診断
  - 初回治療として免疫療法を含む治療を受けた
  - EGFR / ALK / ROS1ドライバー変異陰性
- エンドポイント:**
  - 実臨床における全生存期間(rwOS)・無増悪生存期間(rwPFS)

### 結果

- 治療を2年で終了した患者群と、継続した患者群との間で全生存期間(rwOS)に有意差は認められなかった

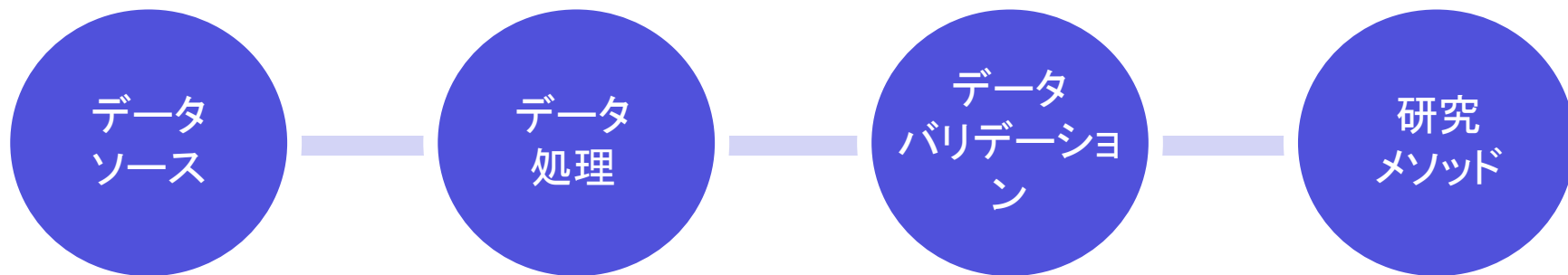
Figure 2. Overall Survival



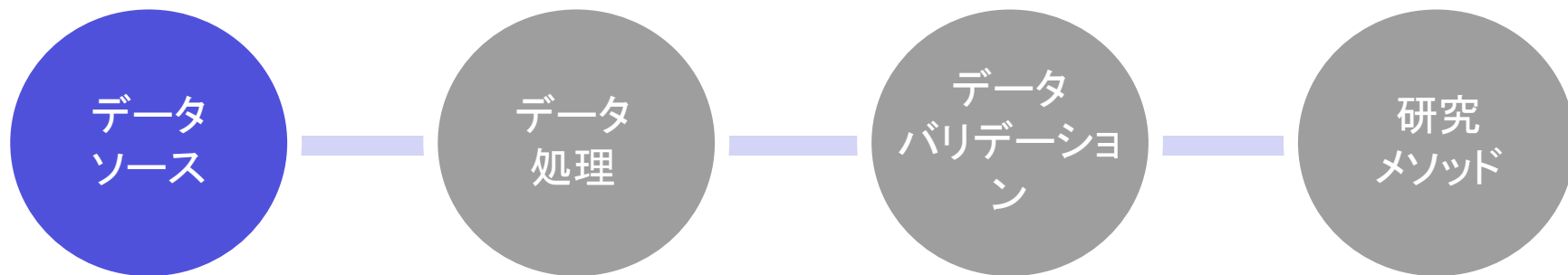
### JAMA Oncology編集者の見解

本研究のような課題に対しては、従来の前向き試験の実施は困難であり、結果が得られるまで数年間かかる可能性がある。  
その間「完璧」を求めすぎて「良い」ものを無駄にすべきではない。 今回のRWD解析から、2年で治療を中止した場合でも 継続治療と同程度の OSが得られ、毒性リスクの低減、治療負担の軽減、医療費の削減といった観点からも、患者と医療者の意思決定を後押しする有用な示唆となる

# RWDを構築し、RWEを生成するために必要な要件



# RWDを構築し、RWEを生成するために必要な要件



# フラットアイアンヘルスのグローバルにおける RWDネットワーク

## | アメリカ

**5M+**  
患者数

75% 地域クリニック

25% 学術医療機関

**4,200+**  
医療従事者

**200+**  
医療機関・診療拠点

**1,000+**  
ユニーク施設数

## | 日本

日本全国で標準治療に基づき多くの患者を診療するがんセンターや大学病院等の学術医療機関

**30,000+人**  
患者

## | イギリス

英国全土のNHSトラスト：国民保健サービス（NHS）のもと、多くの疾病の患者を治療する公的医療機関ネットワーク

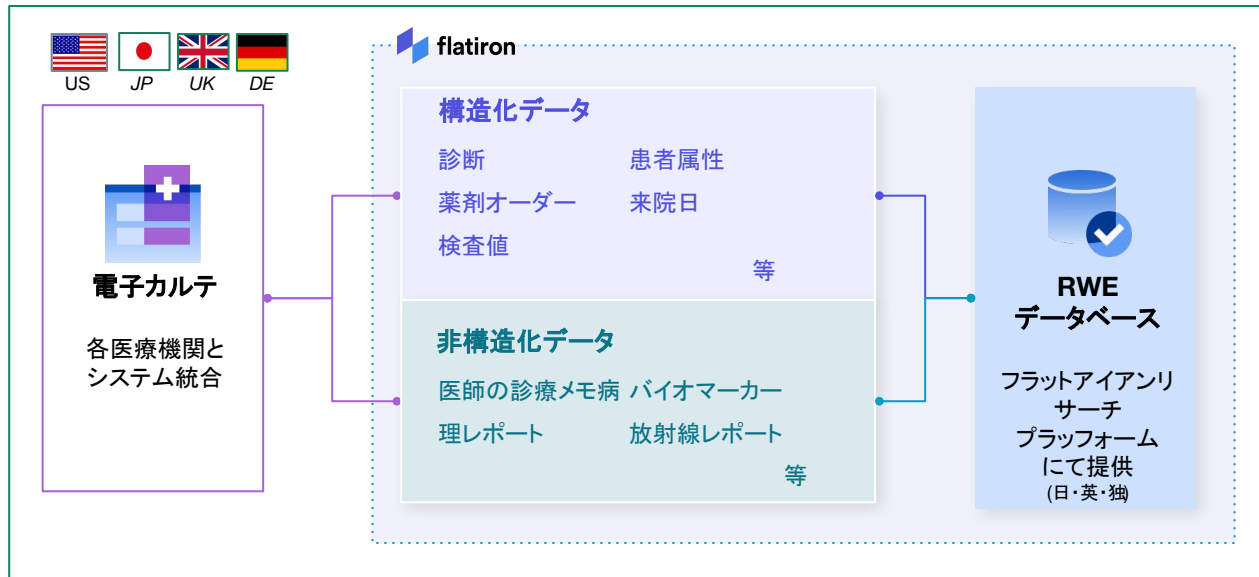
**40,000+人**  
患者アクセス

## | ドイツ

ドイツ全土のがん治療を反映するコミュニティ施設と学術医療機関

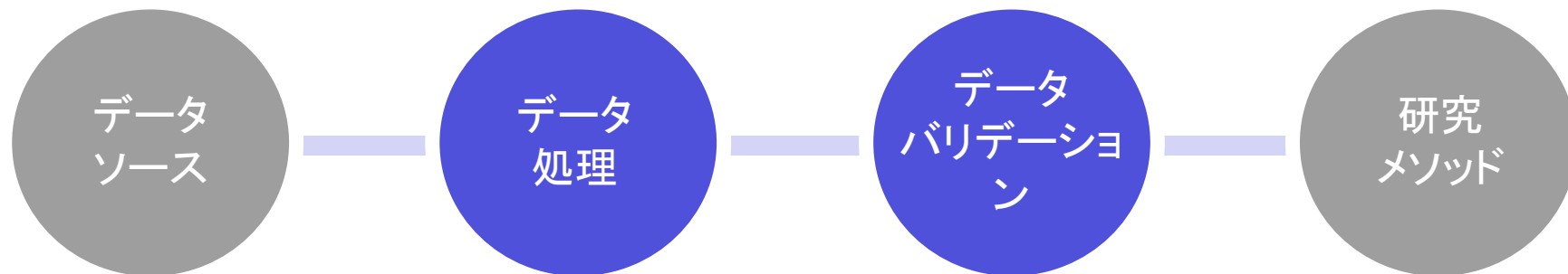
**25,000+人**  
患者アクセス

# Flatiron HealthにおけるRWDの構築アプローチ



- 全ての国(米・日・英・独)、全てのパートナー施設において、構造化・非構造化データの双方に安全にアクセスするシステム基盤を構築
  - 米国でFlatiron Healthが提供する電子カルテ (OncoEMR) を利用する施設を除き、電子カルテシステムはどの国においても施設間のばらつきがある
  - 各国の法規制・倫理・技術要件に対応
- 日本では、電子カルテに加えて部門システムのデータにもアクセスし、非構造化データの範囲を拡充

# RWDを構築し、RWEを生成するために必要な要件



# 電子カルテの構造化データ・非構造化データを統合し、高品質ながんRWDを構築

## 構造化データ

診断                      患者属性                      検査値  
薬剤オーダー          来院日                      等



データの一貫性の担保およびレギュラトリーでのユースケースに対応するため、パートナー施設と連携し、元の診療録への追跡可能な仕組みを導入

## 構造化データ処理

|            |                                             |              |
|------------|---------------------------------------------|--------------|
| 2220       | Blood Serum Albumin                         | g/dL         |
| QD25001600 | ALBUMIN/GLOBULIN RATIO QD                   | (calc)       |
| QD25001400 | ALBUMIN QD                                  | g/dL         |
| QD50059600 | ALBUMIN                                     | %            |
| QD50055700 | ALBUMIN                                     | g/dL         |
| CL3215104  | Albumin % (EPR)                             | %            |
| LC001081   | ALBUMIN, SERUM (001081)                     | g/dL         |
| LC003718   | Albumin, U                                  | %            |
| LC001488   | Albumin                                     | g/dL         |
| LC133751   | Albumin, U                                  | %            |
| CL3215162  | Albumin%, Urine                             | %            |
| CL3215160  | Albumin, Urine                              | mg/24hr      |
| 9254       | ALBUMIN SS                                  | g/dL         |
| LC133686   | Albumin, U                                  | %            |
| QD50060710 | MICROALBUMIN                                | mg/dL        |
| QD50061100 | MICROALBUMIN/CREATININE RATIO, RANDOM URINE | mcg/mg creat |
| QD85991610 | ALBUMIN                                     | relative %   |
| S0058600   | ALBUMIN UPEP RAND                           | %            |
| CL3210074  | ALBUMIN LEVEL                               | g/dL         |
| QD6008211  | ALBUMIN/GLOBULIN RATIO                      | (calc)       |
| LC149520   | Albumin                                     | g/dL         |
| QD45069600 | PREALBUMIN                                  | mg/dL        |

1751-7 Albumin [Mass/volume] in Serum or Plasma g/dL



## 非構造化データ

医師の診療メモ          バイオマーカー検査レポート  
病理レポート          放射線レポート          等

## 非構造化データ処理



# 構造化データ処理:施設・国横断の“Harmonization”(標準化)

| CodeSystem | Code            | Description                                                                                                                    |
|------------|-----------------|--------------------------------------------------------------------------------------------------------------------------------|
| Pharmacode | 1338103         | (Farmos) Tamoxifen Citrat Tabl 20 mg                                                                                           |
| PZN        | 03103166        | TAMOXIFEN-ratiopharm 20 mg Tabletten                                                                                           |
| GTIN       | 7680490590354   | Tamoxifen Farmos 20 milligramm Filmtabletten  |
| HOT13      | 1267136010101   | 「DSEP」タモキシフェン錠20mg                                                                                                             |
| GTIN       | 24987081188806  | タモキシフェン錠20 MG「明治」                                                                                                              |
| YJ code    | 4291003F2070    | 「日医工」タモキシフェン錠20mg                             |
| BNF        | 08030401        | Tamoxifen 20mg Tablet                                                                                                          |
| ATC        | L02BA01         | TAMOXIfen Citrate - 20mg - Tab                                                                                                 |
| dm+d       | 388311000001100 | Tamoxifen 20mg tablets (Teva UK Ltd)          |



**RxNorm  
313195**

**Tamoxifen 20 MG Oral Tablet**



薬剤や検査名について、施設や国によって様々なフォーマット・単位・言語・略称・コーディングシステムで保存されている用語を、 RxNormやLOINCなどの国際標準用語に統一

# Flatiron Health独自のアブストラクション手法により、非構造化データの体系的かつ高品質な処理を実現

flatiron

2 ☆ Sep 3, 2019 プログレッシブノート

3 ☆ Sep 6, 2019 内視鏡報告書 [Diagnosis date >](#)

4 ☆ Sep 6, 2019 放射線報告書 [Diagnosis date >](#)

5 ☆ Sep 6, 2019 病理報告書 [Viewing](#) [Diagnosis date >](#)

6 ☆ Sep 6, 2019 検査結果報告書

7 ☆ Sep 6, 2019 検査結果報告書

filter tags Q > > [start](#) [Treatment end](#) [Stage](#) [Histology](#) [Systemic treatments](#) [PDL1](#) [her2\\_tests](#) [HER](#)

## 診療記録

病理組織診断報告書 (最終診断)

|        |             |          |            |
|--------|-------------|----------|------------|
| <患者番号> | <診療科> 消化器内科 | <組織診断番号> | PAX-51112  |
| <患者番号> |             | <受付年月日>  | 2019.09.06 |
| <生年月日> |             | <採取日時>   | 2019.09.06 |
| <性別>   | 男性          | <報告日>    | 2019.09.07 |
| <年齢>   | 62歳         |          |            |

<臨床診断>  
胃癌、中分化型胃癌

## アブストラクションフォーム

### 胃癌アブストラクションデモ

[Submit](#)

> 特殊症例を登録

胃癌分類

患者は胃癌と診断されましたか [\(見直し\)](#)

☒ はい ☐ いいえ

診断日を記載してください [\(記録更新\)](#)

年 月 日

年 月 日

組織診断を選択してください [\(見直し\)](#)

☐ 腺癌

☐ 扁平上皮癌

☐ 扁平上皮癌癌

☐ 記載なし

初回臨床診断のグループステージを選択してください [\(ステップ\)](#)

該当なし

## データ処理手順

### PROGRESSION EVENT(S)

#### BACKGROUND

Progression events can only be captured if a clinician has documented that there has been overall growth or worsening of the cancer of interest. Other documentation in the clinical note is expected to capture evidence of progression such as a radiology report. Sometimes, the clinician may document progression without explicit evidence, and that should also be captured in this abstraction. This includes explicit mention of "progression" and/or description of findings consistent with progression. We also include parentheticals, as well as related findings that is associated with a treatment change. A question block should be created and answered for each progression event that occurs. Questions within each question block pertain only to that event. Work-up directly tied to advanced therapies should not be considered progression events (See Appendix B) for more information on this and specific guidance on how the data should be captured. Please refer to the following events that happen within 90 days of section group.

### 国内では日本語で提供

1. Select "Yes" if:

a. At least one progression event occurred

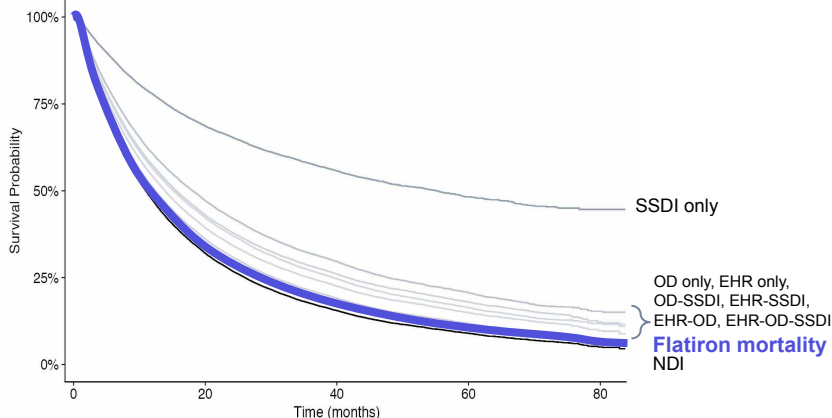
i. The following events qualify as progression (see details):

✓ 各国のがん臨床経験・専門知識を有する臨床アブストラクターが、高度な研修を経て継続的なパフォーマンス管理のもとアブストラクションを実施

✓ 事前に定められたバージョン管理付きの手順書・アブストラクションフォームを用いて、データ処理の品質を担保

# FDAとも共同でリアルワールドエンドポイントの検証を継続実施

18がん種における米国の複数データソースと  
ゴールドスタンダード(NDI)との全生存期間中央値の比較



[2021 VALIDATION STUDY](#)

[WEBINAR](#)

FlatironのRWDにより解析した全生存期間の中央値が、  
米国のゴールドスタンダードであるNational Death Index (NDI) に  
最も近い結果を示した

- Social Security Disability Insurance (SSDI)
- 死亡記事 (Obituary Data; OD)
- 電子カルテデータから得られた死亡情報 (EHR)
- 上記を全て組み合わせたFlatironのRWDの死亡情報

米国FDAと共同でリアルワールドエンドポイントについて複数研究を実施

- Real-world Overall Survival (rwOS) を用いた論文 : 5件
- Real-world Progression-Free Survival (rwPFS) を用いた論文 : 2件
- Real-world Response Rate (rwRR) を用いた論文 : 2件

FDAとの共著論文例

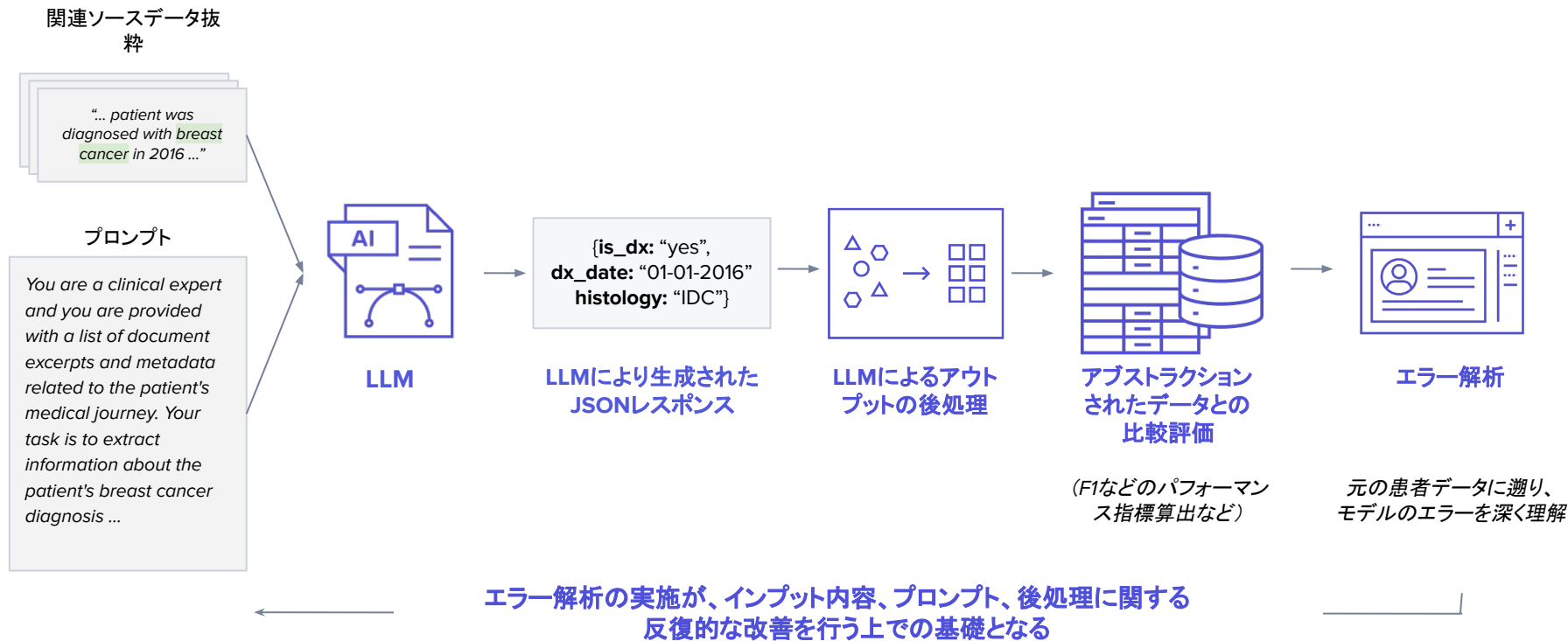
| リアルワールド<br>エンドポイント | 論文タイトル                                                                                                                                                                                              | 掲載誌                             | 発表<br>時期 |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------|----------|
| rwOS               | Real-World Outcomes of Patients with Metastatic Non-Small Cell Lung Cancer Treated with Programmed Cell Death Protein 1 Inhibitors in the Year Following US Regulatory Approval                     | The Oncologist                  | 2018     |
| rwPFS              | Real-world Progression, Treatment, and Survival Outcomes During Rapid Adoption of Immunotherapy for Advanced Non-Small Cell Lung Cancer                                                             | Cancer                          | 2019     |
| rwRR               | Evaluation of Real-World Tumor Response Derived From Electronic Health Record Data Sources: A Feasibility Analysis in Patients With Metastatic Non-Small Cell Lung Cancer Treated With Chemotherapy | JCO Clinical Cancer Informatics | 2024     |



Oncology Center of Excellence

CDER Office of Biostatistics

# LLMを活用したデータ抽出アプローチ: パノラマデータセットの作成



# 「VALIDフレームワーク」の提唱：LLM/MLにより抽出された RWDの品質評価の枠組み

The **V**alidation of **A**ccuracy for **L**LM/ML-Extracted **I**nformation and **D**ata; **The VALID Framework**



## データ項目レベルの パフォーマンス評価指標

人手による臨床アブストラクションと比較した、各臨床情報の抽出パフォーマンスを確認

- 例) 特定のデータ項目の完全性(欠測)検証
- 例) データ抽出精度の検証



## 自動データ検証

患者単位・コホート単位でのデータ検証および、臨床観点からの妥当性確認  
(明らかに誤り・現実的でないデータ項目の特定など)

- 例) バイオマーカーが陰性で、標準治療がなされている患者の割合
- 例) イベントの日付順序に齟齬がある患者の割合
  - ◆ 初回診断の前に手術が実施されている、など



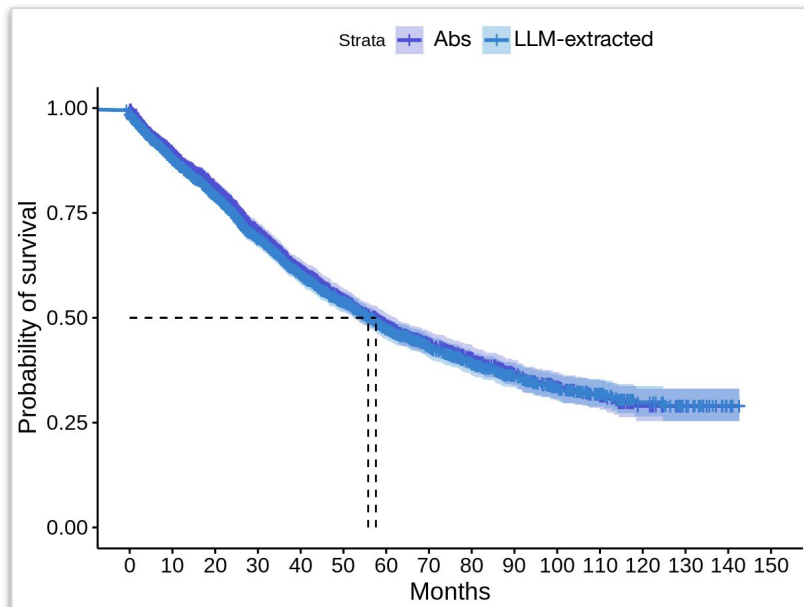
## 再現性評価・ ベンチマーク解析

同じ試験について、異なるデータ抽出方法のデータセットを用いて直接結論を比較

- 例) TNBC患者における術前・術後の治療パターンとアウトカムをアブストラクション・LLMによる抽出の双方のRWDで比較

# パノラマデータセットとアブストラクションされた RWDのrwOS分布

HER2陽性乳がんの 1L THP\* 患者のrwOS: **アブストラクション** vs. **パノラマ(LLM)**



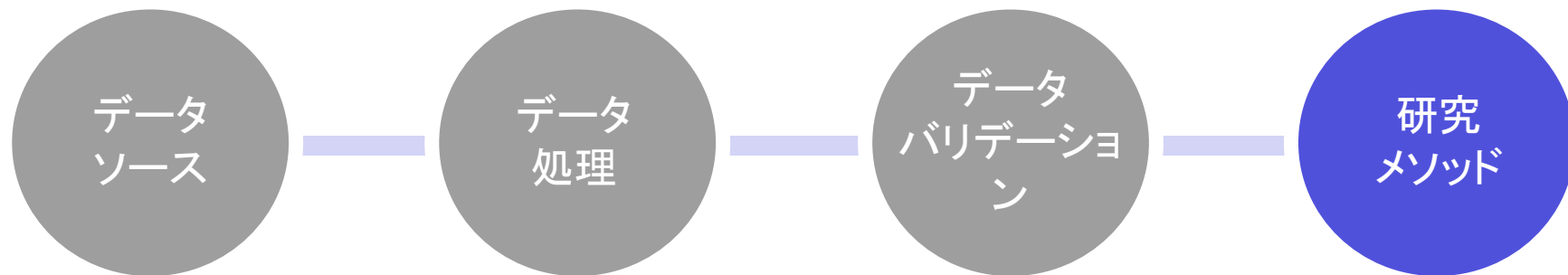
対象コホートで生存曲線が重複しており、2種類のデータ抽出アプローチ間で一貫した洞察を示している

- 1LのHER2+コホートのOS曲線は、アブストラクションによるデータとLLMによる抽出データで重複
- LLMによる抽出データには、以下のような複数のモデルを併用：
  - BC診断
  - 転移診断
  - HER2の状態
- 少数の患者サブセットにおけるモデル誤差はみられるが、結果はデータ抽出アプローチに関係なく一貫し、CLEOPATRA **試験データ**に基づく予想と一致（生存期間中央値 56.5カ月）

|                   | 生存期間中央値 | 下位95% CI | 上限95% CI |
|-------------------|---------|----------|----------|
| <b>アブストラクション</b>  | 57.6カ月  | 52.6カ月   | 62.1カ月   |
| <b>パノラマ (LLM)</b> | 55.8カ月  | 52.2カ月   | 59.8カ月   |

\*THP=ペルツズマブ、トラスツズマブ、ドセタキセル

# RWDを構築し、RWEを生成するために必要な要件



# Flatiron Health RWDを活用したスタディ実施の一例

## 【研究テーマ】

男性のHR+/HER-転移性乳がんにおけるアンメットニーズに関するエビデンス生成

### ✓ 深い臨床データを用いた対象集団の正確な選定

構造化データ

2011年1月1日以降にFlatiron Healthの乳がんデータベースに2回以上の受診あり

構造化データ

乳がんの診断コードの記録あり  
(ICD-9 174.x or ICD-10 C50.x)

非構造化データ

病的に確認された乳がんの診断あり

非構造化データ

ステージIVまたは再発転移性乳がんの診断あり(転移診断日が2011年1月1日以降)

構造化データ

男性(患者基本情報の構造化データをもとに、非構造化データの確認を経ての抽出)

非構造化データ

HR陽性、HER陰性

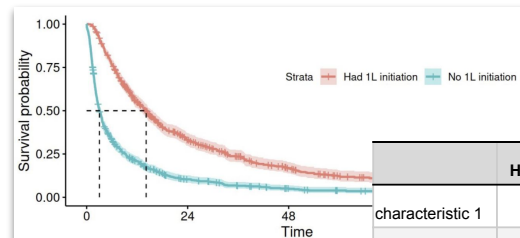
非構造化データ

転移診断後に抗がん剤治療歴あり

### ✓ 研究目的に適した実臨床のアウトカム情報の活用

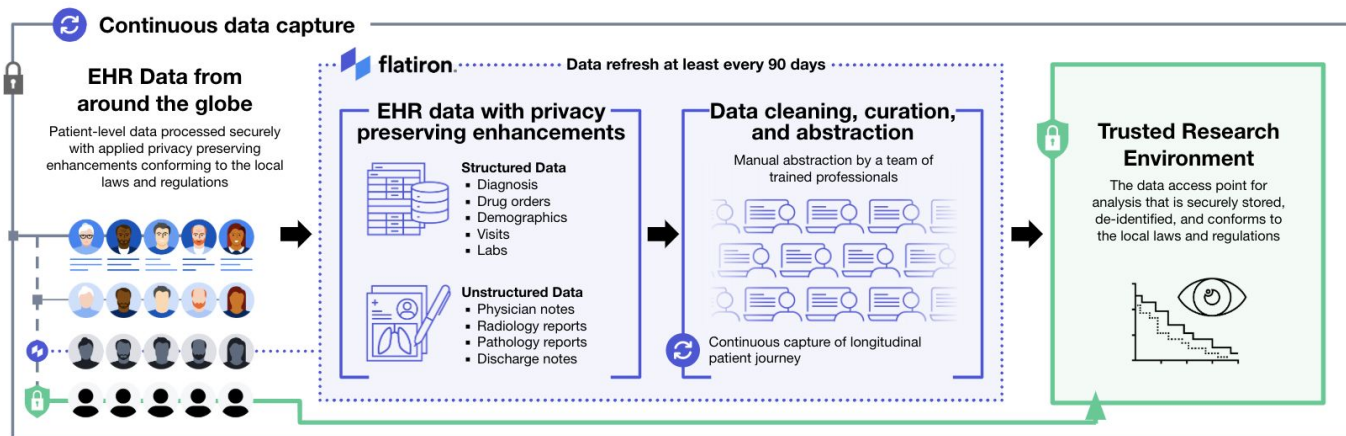
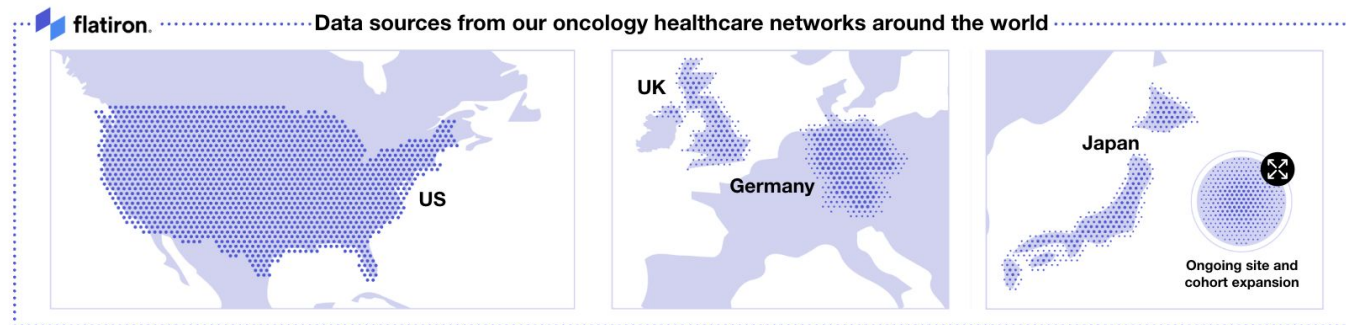
#### 非構造化データ由来の実臨床におけるアウトカム

- Real-world overall survival (全生存期間; rwOS)
- Real-world progression (無生存増悪; rwPFS)
- Real-world response (全奏効率; rwORR)
- Real-world adverse events (有害事象; rwAE)



|                  | Hazard Ratio | 95% confidence interval |      |
|------------------|--------------|-------------------------|------|
| characteristic 1 | 1.10         | 1.02                    | 1.25 |
| characteristic 2 | 0.80         | 0.40                    | 1.40 |
| characteristic 3 | 1.25         | 1.10                    | 1.35 |
| ...              |              | ...                     | ...  |

# ミッション: より良く長い人生のために、がんと共に生きる一人ひとりの経験から学ぶ



# まとめ

- **RWDとRCTは対立するものではなく、臨床エビデンス生成のため相互に補完 する関係にある**
- **高品質なRWDから生成された RWEは、エビデンス生成のゴールドスタンダードであるRCTの実施を効率化・高度化し、RCTでは捉えられない患者群におけるエビデンス生成を可能とする ことで、がん治療の向上・研究の進展に寄与する**
- **RWD構築をゴールとせず、いかに高品質の RWEに変換し、患者への価値に転換できるか 、という視点に基づき、適切なデータソース へのアクセス、堅牢なデータ処理 、継続的なデータバリデーション の実施と適切な研究メソッド が重要**
  - 今後 ML / LLMによる非構造化データの抽出が進む中でさらに上記の重要性が増す
  - 既に米国では上記の観点や、フラットアイアンヘルスの品質フレームワーク( VALIDフレームワーク)に沿って ML / LLMを活用したデータの構築とその活用 が進んでいる。アブストラクションされた大規模なデータとの比較などを通じて、ML / LLMによる抽出の有効性も適切に継続的にバリデーション することが必須
- **日本においても、グローバルで認められたデータ処理や方法論を適用し、これまで得られなかった日本のがん患者に対する深い洞察の取得 や、日本・米国・ヨーロッパの国横断でのグローバルエビデンスの生成を可能とする基盤を構築している**



**Thank you**