



千年カルテ二次利用：最適な研究デザインを考える

株式会社NTTデータ

製造ITイノベーション事業本部 第4製造事業部 VC担当 西田陽介

1-1 千年カルテとは

千年カルテは次世代医療基盤法のもと医療分野における社会課題の解決の一助となることを目指しています。

認定匿名加工医療情報作成事業者
(一社) ライフデータイニシアティブ



豊富な研究実績を保有する有識者にて構成
アカデミア中心の運営による社会的信頼



認定医療情報等取扱受託事業者 **NTT Data**
(株) NTTデータ

社会インフラ構築で求められる高度なセキュリティ技術
多業界のシステム構築実績にて培った高いデジタル

千年
カルテPJ

産業
製品開発力低下

製品申請タイムラグ是正
医薬品効能・副作用に関わる
情報精度/スピード向上
医療技術の国際展開

解決すべき社会課題

行政
医療費増大

無駄な医療の低減
個別医療の実現

社会
超高齢社会突入

自身の健康管理拡大
QOL向上による予防医療
持続可能な保険医療提供

認定事業者

(一社) ライフデータイニシアティブ



豊富な研究実績を保有する有識者にて構成
アカデミア中心の運営による社会的信頼

受託事業者

(株) NTTデータ

社会インフラ構築で求められる高度なセキュリティ技術
多業界のシステム構築実績にて培った高いデジタル技術



(サービス名) **千年カルテ**

NTT DATA

* 厳格なセキュリティ管理、匿名加工医療情報作成技術を
基に、国内初の次世代医療基盤法認定を受けました。

安全・安心



大規模な電子カルテ データベース



先進医療を担う特定機能病院等を中心に数多くの
医療機関から医療情報を収集
薬剤や治療の効果測定、副作用解析に必要な検査
値を含む電子カルテ由来の情報を保持

次世代医療基盤法に基づく厳格なセキュリティ管理
個人の特定期間リスクを低減する匿名加工技術

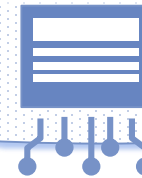
匿名加工された 医療情報



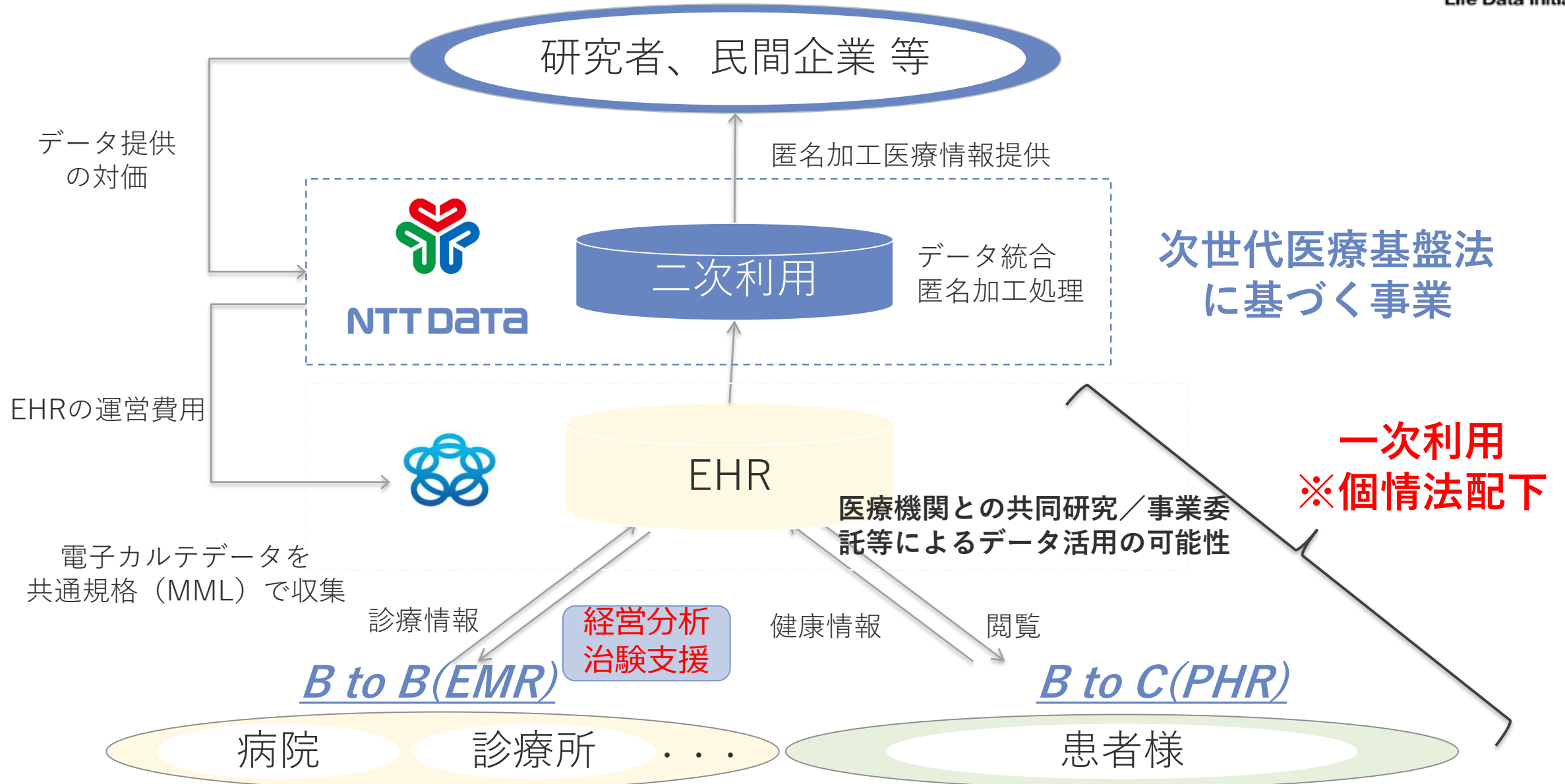
疾病予測や未病・予防解析、リスクファク
ター特定など新たな価値の創出に貢献

医療分野の研究開発促進

高度なIT技術

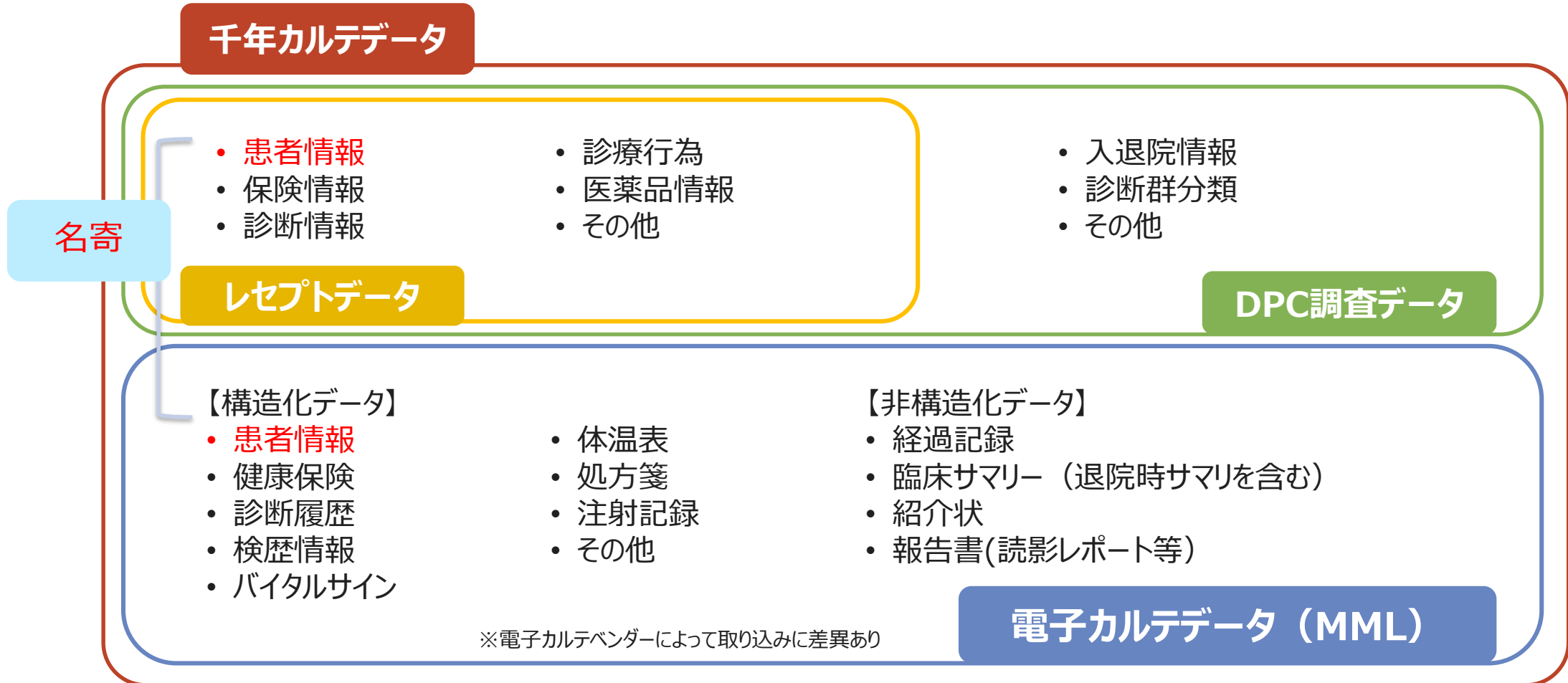


ビッグデータ化した診療情報を用い、必要なAI解析（機械
学習・深層学習・自然言語処理）などを支援



千年カルテ データセット

千年カルテデータは、電子カルテ、レセプト、DPC調査データを保有しており、実臨床における治療実態の把握やアウトカム評価に用いるなど、より詳細な分析をすることが可能です。



千年カルテ 利活用審査概要 (現在_全18件)



Life Data Initiative

2020年6月からのサービス開始より、研究実績※が増えています。(9件目まで掲載)

承認日	課題名	活用データ項目	活用者区分
2020年10月20日	乳癌のサブタイプ別、治療実態を探るための千年カルテデータのFeasibility	電子カルテデータ、DPC調査データ、レセプトデータ	アカデミア
2020年10月20日	がん患者の臨床アウトカムにおけるEHRデータベースを用いた評価方法の後ろ向き研究	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2021年1月15日	自己免疫疾患領域における寛解指標のフィージビリティ確認	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2021年3月5日	検査値等を用いたウイルス性肝炎患者研究のフィージビリティスタディ	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2021年5月26日	検査項目の多施設突合手法開発を目的とした研究	電子カルテデータ	アカデミア
2021年7月15日	非構造化データの評価方法確立を目的とした研究	電子カルテデータ	民間企業
2021年7月15日	希少疾病領域における症状把握を目的としたフィージビリティ検証	電子カルテデータ、DPC調査データ、レセプトデータ	アカデミア/ 民間企業
2021年7月15日	乳がんデータ項目に関するフィージビリティ調査	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2021年8月31日	匿名加工医療情報のAI研究への利活用可能性の検討	電子カルテデータ、DPC調査データ、レセプトデータ	アカデミア

癌、希少疾病 テキスト活用したアウトカムリサーチが主

※審査概要につきましてLDI HP上で公表の義務が生じます

千年カルテ 利活用審査概要 (現在_全18件)



Life Data Initiative

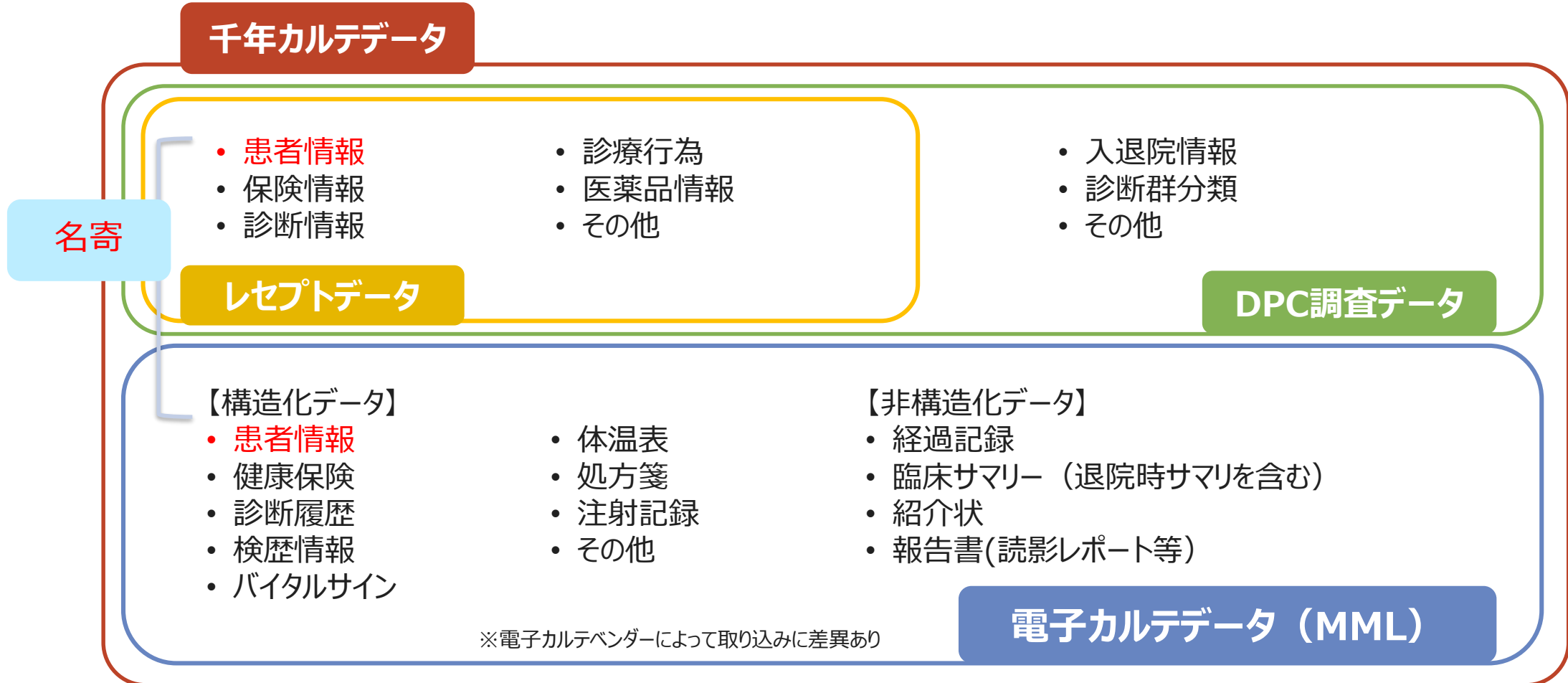
2020年6月からのサービス開始より、研究実績※が増えています。(10～18件目まで掲載)

承認日	課題名	活用データ項目	活用者区分
2021年9月28日	心不全データベース研究のためのFeasibility調査	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2021年10月26日	感染症に対するTreatment flow 及び 関連医療費の推計	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2021年10月26日	がん患者の臨床アウトカムにおけるEHRデータベースを用いた評価方法の後ろ向き研究－自然言語解析－	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2021年11月30日	肺がん・乳がん患者の治療実態把握及び病気の進展に関する因果探索	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2021年11月30日	電子カルテのテキストを活用したRECIST評価の辞書作成	電子カルテデータ、DPC調査データ、レセプトデータ	アカデミア
2022年3月8日	電子カルテ情報を活用した、臨床試験の新規手法論開発	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2022年3月8日	電子カルテのテキスト情報を用いた癌患者の治療実態に関する検討	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業
2022年4月13日	希少疾病の罹患リスク予測モデル構築	電子カルテデータ、DPC調査データ、レセプトデータ	民間企業

※審査概要につきましてLDI HP上で公表の義務が生じます

千年カルテ データセット

千年カルテデータは、電子カルテ、レセプト、DPC調査データを保有しており、実臨床における治療実態の把握やアウトカム評価に用いるなど、より詳細な分析をすることが可能です。



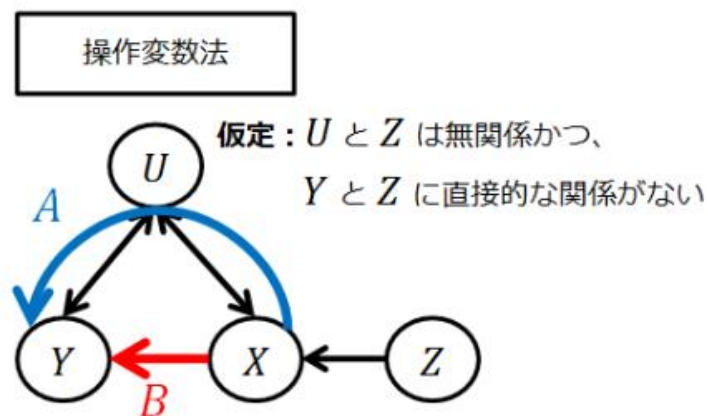
精緻なアウトカム定義、Criteria定義が期待される。

→アウトカムバリデーションのPPV向上

→感度分析のバリエーション期待

→傾向スコア、共変量の多様性

→操作変数法における操作変数Zの選択肢



Bの効果が存在するならば、Zが変動するとXを通じた間接的な効果のみによってYも変動する。この時、Zの変動でUは変動しないので、Aの効果は生じない。これを利用し、Bの効果のみを推定する方法が操作変数法。

電子カルテの臨床サマリには、こういった記載があるのか、サンプル情報を提示します。行為しか把握ができないレセプトに対し、電子カルテと行き来することで、検査結果や患者背景、治療内容を把握することが可能となります。

臨床サマリ情報のサンプル（一部）

※ダミーデータ

<mmlSm : history>

4/01 左乳房に疼痛自覚し、〇〇病院に受診された。妊孕性温存を希望され、当科へ紹介。

4/18 母が乳癌に罹患していたため、遺伝性乳がん卵巣がん症候群を疑い、遺伝カウンセリングを希望された。遺伝子検査の結果、**BRCA陰性**が判明した。

⋮

<mmlSm : clinicalRecord>

#右C領域 閉経前浸潤性乳管癌(cT1cN0M0,cStageI,luminal-B type)

⋮

6/22 右乳房部分切除術・センチネルリンパ節生検施行。
術後補助化学療法(TC4コース)の方針。

6/29 退院。

<mmlSm : testResults>

⋮

Immunohistology: **ER: Positive(100%)** **PgR: Negative(0%)**
HER2(IHC): Negative(score 0) **Ki-67:21%(106/500)**

→表記揺れ

例、骨粗鬆症

- DXA ? DEXA ?
- TRACP-5b ? TRAP – 5 b ?
- BRONJ → ONJ

→Feasibility時に表現系の辞書を詰めていく必要あり。



千年カルテの研究アイデアはいっぱいあるので、その一例を御紹介します。



NTT DATA

Trusted Global Innovator