

# Decentralized Clinical Trialと 生成AIの臨床試験への応用

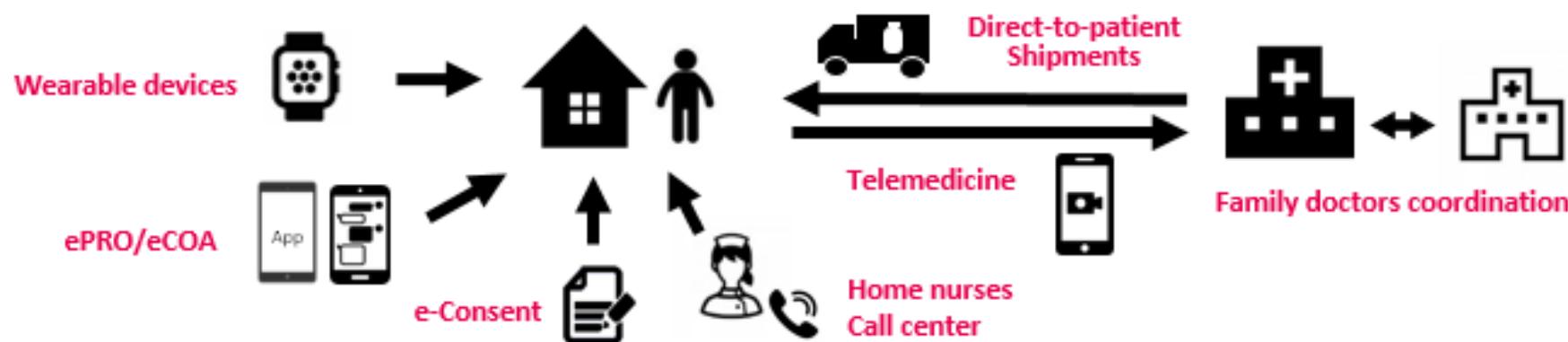
---

国立がん研究センター中央病院  
国際開発部門/臨床研究支援部門  
中村 健一

# DCTとは？

## ■ DCT

- Decentralized Clinical Trial:分散型臨床試験
  - IT技術を活用したオンライン診療により実施する臨床試験
  - オンライン治験、リモート治験、ヴァーチャル治験とも呼ばれる



<https://en.cmicgroup.com/solutions/clinical-trials/decentralized-clinical-trials-dct>  
2022/9/7閲覧

# COVID-19が追い風に

- 初診からのオンライン診療が条件付きで可能に
  - かかりつけ医と連携しながら、遠隔地からオンラインでの治験参加の道が開かれた
- PMDAからのコロナ関連通知で様々なグレーゾーンを許容
  - 治験におけるオンライン診療
  - 被験者宅への治験薬の直接配送
  - 治験における一般的な検査の別の医療機関への委託
  - 医療機関間の事後的な契約
  - 訪問看護師による治験薬の投与
  - ITデバイスによる患者とのコミュニケーション
- 医療機関/患者のデジタルデバイスへの慣れ
  - Zoomをはじめとした遠隔コミュニケーション



# がん領域へのDCT導入

## ■ なぜDCT？

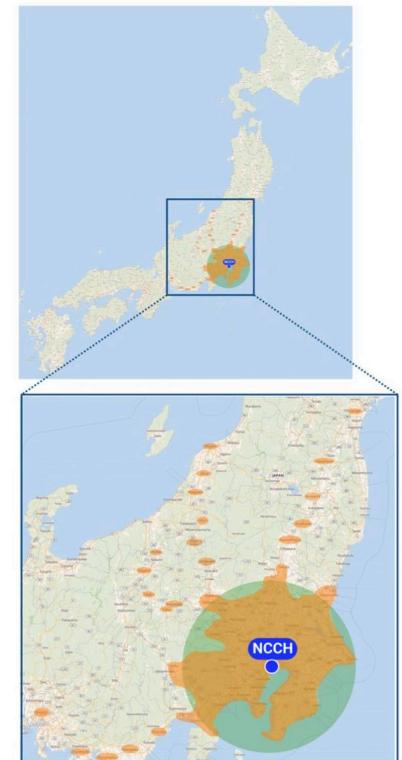
- 保険診療でがん遺伝子パネル検査を受けた患者のうち、エキスパートパネルで提示された治療薬を投与した患者は9.4%のみ(第4回がんゲノム医療中核拠点等の指定に関する検討会 資料1)

  - 地方の患者は治験参加のチャンスが圧倒的に少ない
  - 治験のために東京まで定期的に受診する患者も

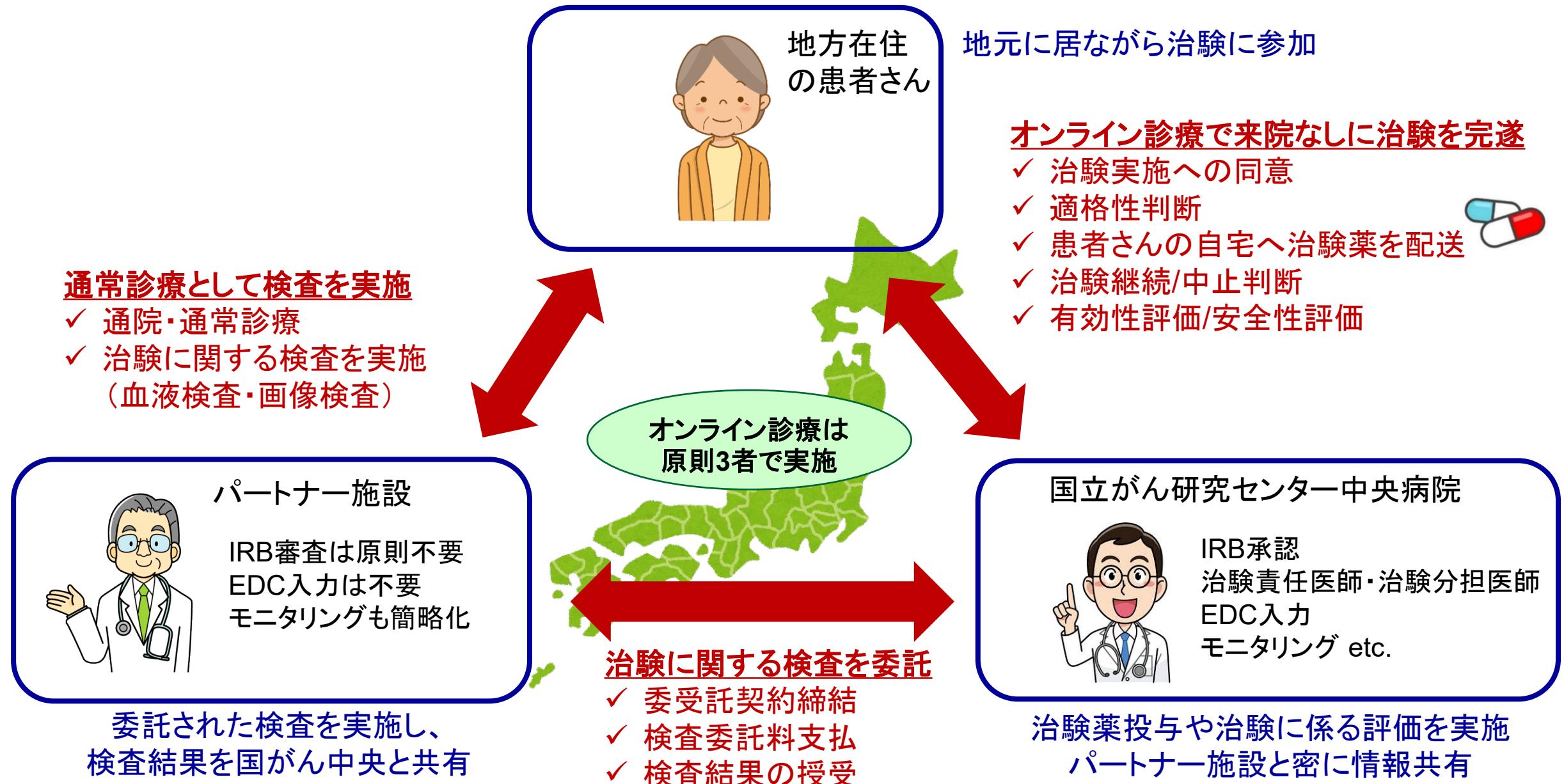
- 希少がん・希少フラクションの治験は特に都市部の病院に集中
  - 患者は全国に存在するが、アクセス面でのペインが強い
  - 患者団体からも地方在住患者の治験へのアクセス改善の要望が強い
- 移動時間が片道120分を超えると治験参加割合が低下する
  - 移動距離が片道100kmを超えても低下

オレンジ色の地域：NCCから  
移動時間120分以内の地域  
緑の半円内の地域：NCCから  
半径100km以内の地域

JAMA Netw Open. 2023 Sep  
5;6(9):e2333188.



# ケース1: 医師主導治験でのフルリモート型DCT



# DCT導入にとってのメリット

- 地方からの参加の手間が大幅に軽減
  - 治験へのアクセス改善を求める患者の声に応えられる
  - 治験実施施設で必要な契約、倫理審査、モニタリング等が大幅に簡略化
- 患者登録の大幅な迅速化
  - いつどこで患者が発生するかわからない希少がん、希少フラクションとの親和性が高い
- トータルでは治験コストの圧縮が見込める
  - コスト増要因:DCTプラットフォームの導入
  - コスト減要因:登録期間短縮、モニタリングコスト削減



# DCTに適した臨床試験とは？

## ■ DCTの良い適応

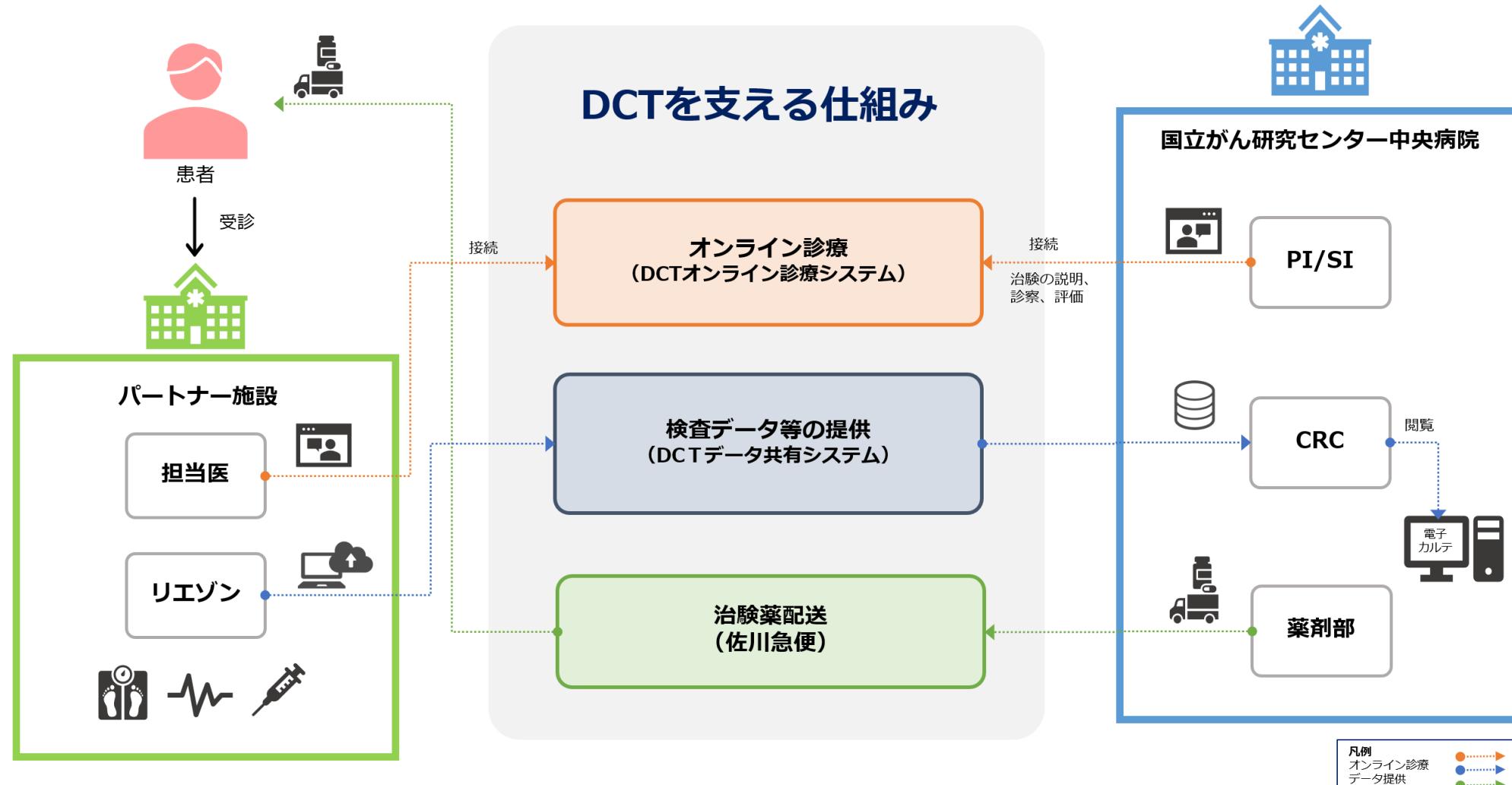
- 患者の症状が比較的安定している
- 経口薬
- 薬剤の毒性が少ない
- 希少疾患
- エンドポイントが遠隔でも収集可能
  - がん領域では、既承認で毒性の比較的少ない経口薬を用いた、希少がん・希少フラクションの試験が良い適応

### 国がん中央でのフルリモート型DCT医師主導治験

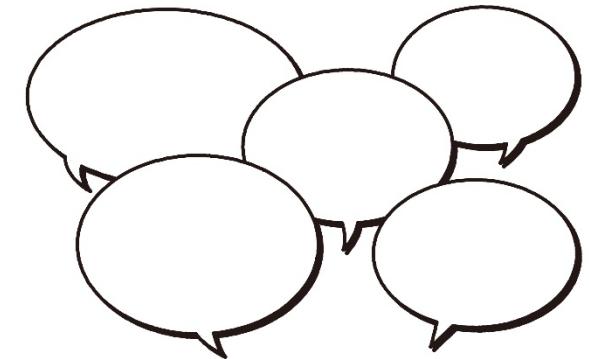
- 局所進行・再発**類上皮肉腫**に対するタゼメトスチットの第II相医師主導治験(単群、17例)
- **BRAF融合遺伝子陽性**の低悪性度神経膠腫または脇がん患者を対象としたビニメチニブの医師主導治験(単群、27例)

- ◆ 当初は規制上、パートナー施設へは注射剤投与は委託できなかつたが…
- ◆ 2024年7月に厚労省が**注射剤の委託を許容**する方針を打ち出し
- ◆ 愛知県がんセンター主導の拡大治験で注射剤を用いたDCTが開始

# 現在運用中のDCTプラットフォーム



# DCT参加患者さんの声



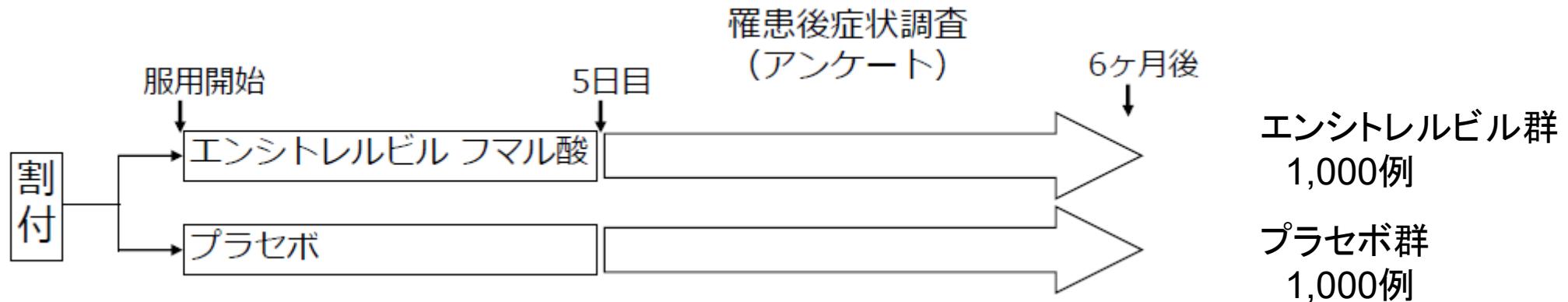
## ■ メリット

- 金銭面: 通院費 + 滞在費が減った(毎月約10万円)
- 精神面:
  - 自宅に居られることがありがたい。これまで来院の度に、毎月約10日程度は都内に滞在していたが、自宅とは異なり気を使うこともあり疲れた
  - 2カ所の病院の先生に診てもらえて安心する。セカンドオピニオンを常に受けているような感じ
- 体力面: 移動に半日かかっていたが、なくなったことで体力的に楽になった。
- 学校関係: オンライン授業も活用してやり繰りしていたが、テストなど必ず登校しなければならない日もあるため、都合を合わせやすくなった
- 仕事関係: 仕事に穴をあけることなく、通院できることがあるがたい

## ■ デメリット

- 関係者が増える分、気を遣う
- 何か起きたとき、どちらに相談すれば良いか迷う

## ケース2: 感染症領域のDCT(単施設2000例)



### ■ 試験概要

- 対象: 軽症COVID-19罹患患者
- エンドポイント: 罹患後症状の発生頻度(3か月経過後の倦怠感、呼吸苦、集中力低下等)
- 試験実施施設: 大阪大学のみ(単施設)
- パートナー施設: 日本全国150施設以上

### ■ 実施手順

- WEBシステムで患者自身が当日午後のオンライン診療を予約 → eConsent
- 同意取得後に登録、割付、試験薬を被験者宅へ発送
- 毎月(6か月後まで)、スマホで罹患後症状の有無を入力(医師の診察は必要なし)

# ドラッグ・ロスの原因の1つ: 日本の治験環境

カテゴリー	評価指標	日本の数字の変化	日本の立ち位置の変化	2022年 相対評価 (対 他国)	総合評価
Quality	EDCの入力コンプライアンス	↑	→	○	質は比較的高いが 本質的でない逸脱を減らすため 過剰なエフォートを割いている
	Queryの回答コンプライアンス	→	→	△	
	重大な逸脱数	↓	→	○	
Speed	症例数/施設/月	↑	↑	△	参加施設が多すぎる 1施設あたりの登録数が少ない
	症例数/施設	↑	↑	△	
	各国の選定～施設Open/同意取得の期間	—	—	○	
Cost	モニタリング報告書数/施設	↑	→	×	非効率なモニタリング 中央化されないIRB審査 施設が分散化しすぎて モニタリングコストが高い
	IRB数/施設数	—	—	△	
	CRAの担当施設数	→	→	×	
	各国で100例の登録に必要な施設数及びCRA数	↑	↑	△	

## データから見た日本の現状

- 各項目を「○、○、△、×」の4段階で評価した結果は上記の通りであり、総合評価は「△」である
- 2017年と比較して、全体的に改善傾向は見られたが、まだ競争優位性が高い状態ではない
- 参加国の中にはSpeed、Costが重要視されるが、日本で治験を実施したいと思われる状態にするには、更に生産性を上げ、グローバルとの競争力を向上させる必要がある

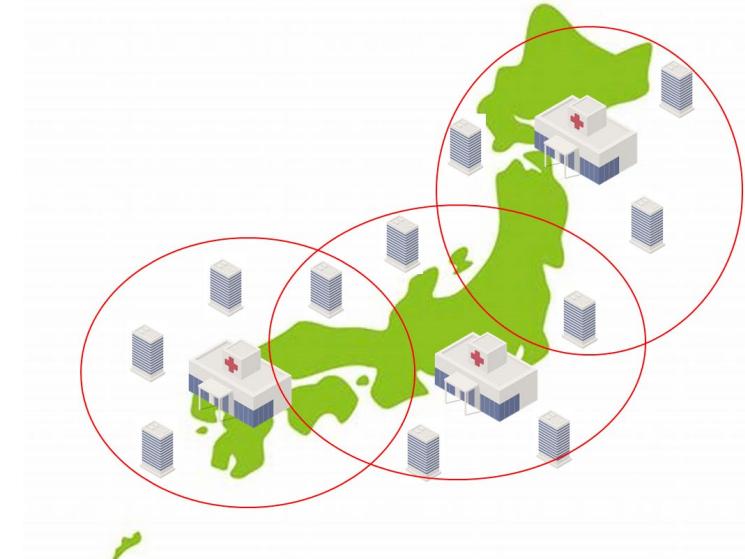
# DCTを活用した効率化

## ■ 様々な治験環境改善の取り組みがこれから進む

- Single IRBの促進
- Fair Market Valueによるコスト算定
- 施設契約書、IC文書の標準化
- jRCT大規模改修による治験情報の見える化
- ICH-GCPの合理的運用(PMDA治験エコシステム事業) etc.

## ■ DCTを活用した治験実施機関とパートナー施設の役割分担

- 希少疾患については、ハブとなる少数の治験実施機関と、DCTによるパートナー施設のネットワークを構築し、日本全体で効率的な患者登録を推進



# DCTは大都市だけがハブになる？

- No!!
  - 希少疾患は集約できても、Common diseaseを集約するのは困難
  - 試験実施機関とパートナー施設との信頼関係は一朝一夕にならず
  - パートナー施設からの医療情報の共有がボトルネックになりがち
    - 多くのDCTでは未だにFAXやCD-Rの郵送で医療情報が共有されている
- 地域医療情報連携ネットワーク活用の可能性
  - 医療情報ネットワークが確立されると、DCTにおける医療情報共有が容易
  - Common diseaseでは各地域の基幹病院がハブになりDCTを行う形が望ましい
  - 患者紹介のみでのDCTの活用もpromising

# DCT:米国でイマイチなので、日本でもダメ?

## ■ 米国

- 医師免許が州単位で発行
  - 州を超えたオンライン診療がやりにくい
- 一方、施設間のデータ共有は日本より容易
  - 米国ではEpicをはじめクラウド型EHRが普通

## ■ 欧州

- 同様に国を跨いだオンライン診療がやりにくい
- 国別で保険制度が異なり、国を跨いだ診療は難しい

## ■ 韓国

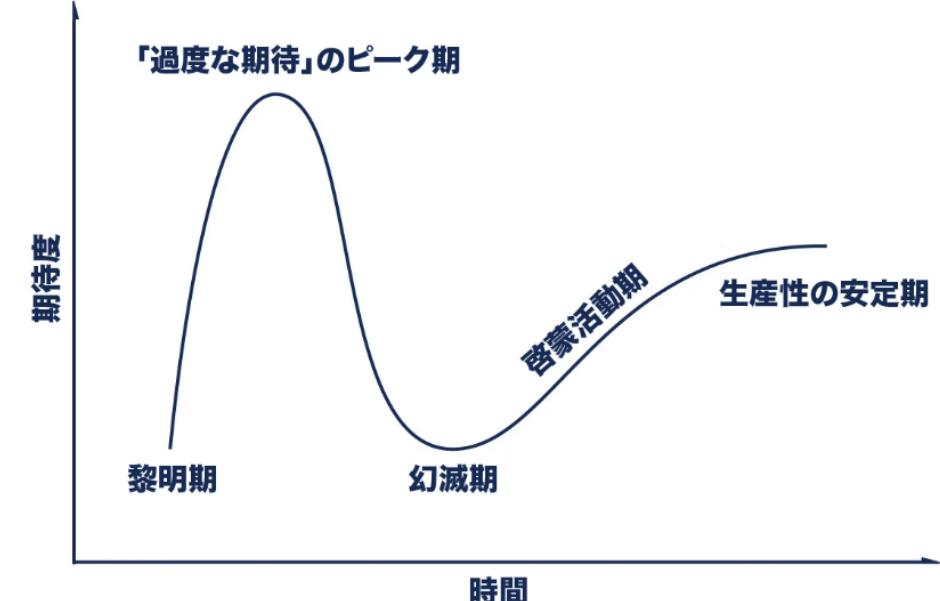
- そもそもオンライン診療が認められていない

## ■ 日本

- 日本全体でオンライン診療を実施可能
- 厚労省もDCTに関する規制緩和に前向き
  - パートナー施設への注射剤投与の委託が可能になる見通し
  - 施設間のデータ共有のやりにくさは課題

- ✓ 各国それぞれの課題が存在 → 日本は独自のDCTを追求すれば良い
- ✓ 日本には日本ならではのアドバンテージもあり、「チャンス」と捉えるべき

ハイプサイクル:  
新規テクノロジーの流行り廃りの典型的な時間経過



<https://deepsquare.jp/2020/08/news-0824/>

# DCTはセットアップが大変(?)

[トップページ](#) > [広報活動](#) > [プレスリリース](#) > オンライン治験に関する実務手順書の公開とDCTコンサルテーション開始

## オンライン治験に関する実務手順書の公開とDCTコンサルテーション開始

国立がん研究センターのノウハウを全国の医療機関へ展開し、医薬品の開発促進とドラッグラグ・ドラッグロス解消に貢献

[× ポスト](#) [シェアする](#) [LINEで送る](#)

2025年8月26日

国立研究開発法人国立がん研究センター

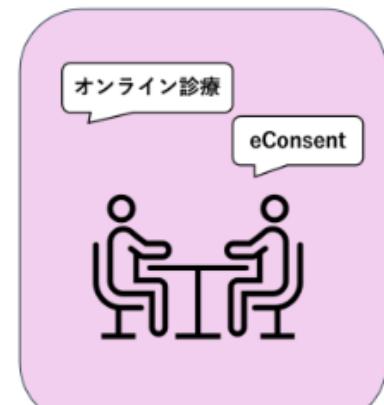
### オンライン治験関連書類の入手方法とDCTコンサルテーションの概要

#### オンライン治験関連書類の一般公開

- DCT実施手順書
- DCT実施マニュアル
- 治験使用薬の管理に関する手順書別冊
- 業務委受託契約書
- フィージビリティ調査票
- 治験実施計画書付録
- 同意・説明文書

フォーム入力 → 書類ダウンロード

#### DCTコンサルテーション



申込 → コンサルテーション実施

国立がん研究センター中央病院で実施した2つの医師主導治験の経験に基づく、手順書、マニュアル、各種文書を一般公開(2025年8月26日)

### 実務書類の提供(無料)

主に以下7点の書類を提供します。

- DCT実施手順書
- DCT実施マニュアル
- 治験使用薬の管理に関する手順書\_別冊
- 医師主導治験に関する業務委受託契約書・覚書
- フィージビリティ調査票(施設版)
- 治験実施計画書付録
- 同意・説明文書

# 臨床試験への生成AIの利活用



# 医療分野での生成AIの利活用

- 診断支援系
  - 症状に基づく鑑別診断候補の提示
  - AIによる画像診断支援
- 業務効率化系
  - 退院サマリーの作成
  - カルテ記載の音声入力・要約
  - 問い合わせ対応チャットボット
- コミュニケーション支援系
  - 外国人患者との会話補助(翻訳AI)
  - 医学生用の模擬患者訓練(バーチャル患者AI)
- ロボティクス系
  - 手術支援ロボット
  - 介護支援ロボット
- 研究支援系
  - 臨床試験プロトコールの自動作成・レビュー
  - スライド自動作成
  - 医学論文の要約



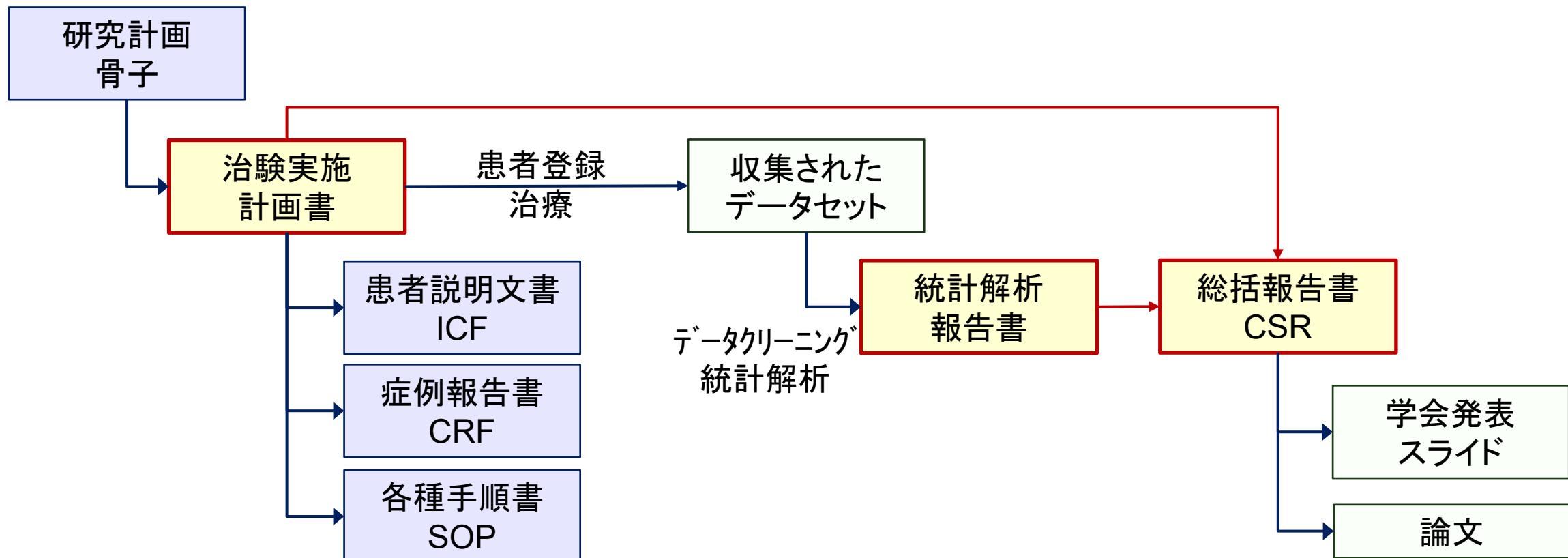
# 日本臨床試験学会2025でのAI関連演題

	演題	発表機関	概要
OS1-03	機械学習による臨床検査データの動的異常検知システムの開発と実用性の検証	順天堂大学	検査値のリアルタイム検知システム (仮想データを利用)
OS1-04	生成AIを用いた治験関連文書における自動作成ツールの構築	国がん中央 ロゼッタ	治験総括報告書の自動作成
P-87	研究計画書・説明同意文書作成を革新するAI/DXプラットフォーム開発	北海道大学 TOPPAN	プライベートLLM環境での文章作成や校正支援システム
P-89	生成AIを用いた臨床研究の倫理審査:ChatGPTの正確性と再現性の評価	札幌医大 北海道大学	GPTsによるプロトコールの要点抽出
P-90	ローカルLLM(大規模言語処理)を利用した文書作成支援の実用可能性調査	福井大学	Gemma2などの5種類のローカルLLMによる文書作成支援機能の比較
P-91	生成AIを使ったDM文書レビューの検討	中外製薬	DM文書のCRF completion guidelineのレビュー
P-93	AIを利用した契約書作成プロセスの効率化の検討	横浜市大	契約書の不整合チェック
P-157	日本の研究力低下の現状分析と京大病院iACTにおける研究力強化の取り組み	京都大学	論文執筆支援AIツールの導入

- ✓ 日本では文書作成・レビュー系の利用が多い
- ✓ 情報セキュリティがひとつの壁に(仮想データでの取り組み、ローカルLLMの使用が多い)

# 生成AI(LLM)で何ができる？

## ■ 例えば治験関連文書の自動作成



# 生成AI × 総括報告書(CSR)

## ■ なぜCSR？

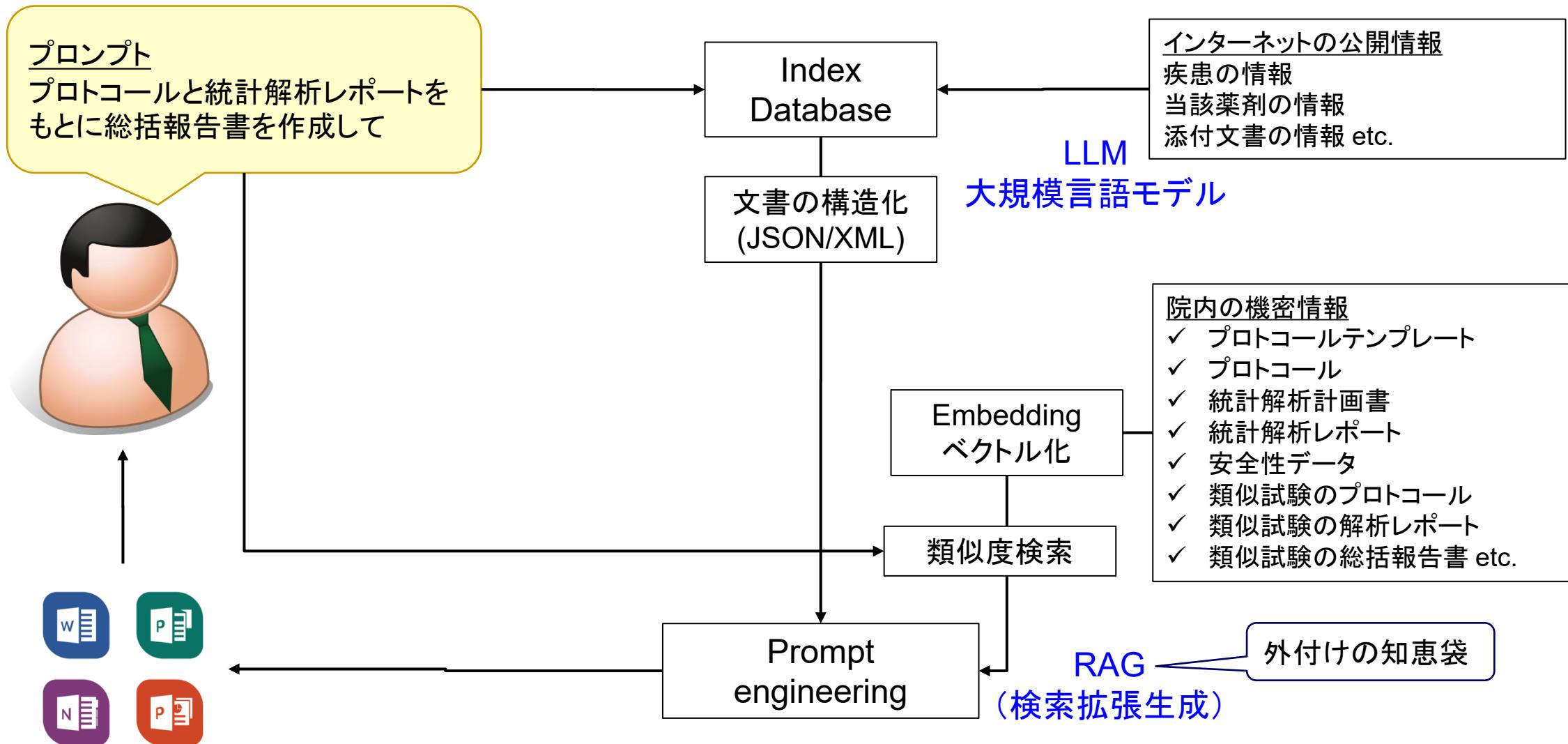
- 技術的ハードルが比較的低い
  - 治験実施計画書や統計解析報告書をそのまま転記できる部分が多い
- コストインパクトが大きい
  - 高額なCSR作成の外注コストは悩みの種



CSR作成を自動化・内製化できれば、時間もコストも削減！  
(複雑な入札手続きも不要！)

(費用はすべて税込み)

# 総括報告書の自動作成(既存モデルとRAGの組み合わせ)

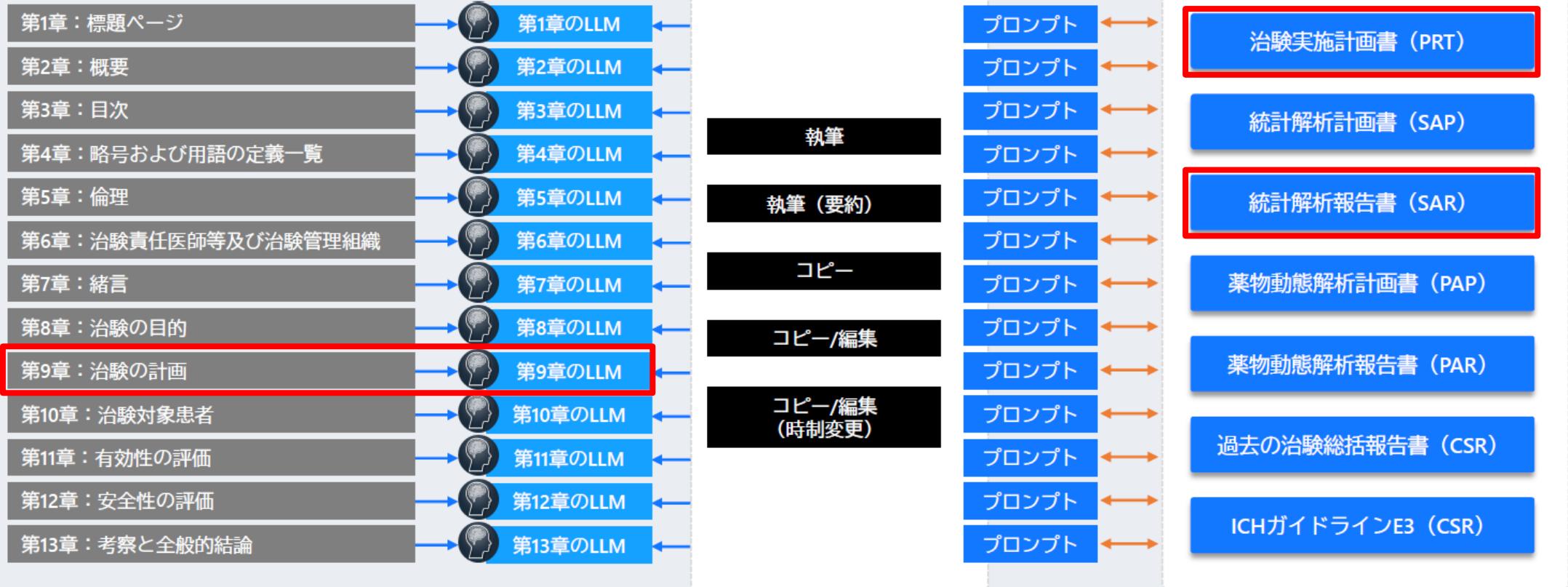


# 生成AI活用のコツ

- 細かくパートに分けて、それぞれにLLM選定、プロンプト開発を行うのがポイント

## 2 : Metareal AI LLM2\_章単位まで細分化しLLMを選定

例: CSR



# CSRができるまで: 自動作成～編集(画面)

ラクヤクAI go.jp Logout

## Protocol

Analysis Report

File Upload

対象となる統計解析報告書を指定し、アップロード

Drag & Drop your file here  
Or click to select a file

対象となる治験実施計画書を指定し、アップロード

Drag & Drop your file here  
Or click to select a file

\*Only one file can be uploaded at a time.

Back

ラクヤクAI Logout

## Protocol

## Analysis Report

File Edit Insert Format Table

Reference Link:  
Click here to see the data source

Generated Options

Initial Content:  
1回の投与を1コースとして「9.3.3.2 プロトコル治療中止規準」に該当しない限り治療を継続した。

Option 1:  
1回の投与を1コースとして治療を継続した。

Option 2:  
1回の投与を1コースとして「9.3.3.2 プロトコル治療中止規準」に該当しない限り治療を継続した。

9. 治験の計画

9.1 治験の全般的デザインおよび計画－記述

9.1.1 試験デザイン

9.1.1.1 プロトコル治療

本試験では、[REDACTED] を2週間間隔で30分以上かけて点滴静注にて投与した。

1回の投与を1コースとして「9.3.3.2 プロトコル治療中止規準」に該当しない限り治療を継続した。

テスト エディタで編集可能

本試験では、[REDACTED] 減量は行わなかった。よって本試験におけるプロトコル治療中の治療変更は「9.1.1.1.2 コース開始による延期」による「延期」のみであった。

9.1.1.1 プロトコル治療の開始

二次登録後14日以内にプロトコル治療の開始した。14日以内に投与を開始できなかった場合は、開始が遅れた理由をCRFに記載する。

2コース以降の[REDACTED] だし、年末年始など、なお、点滴の方法

9.1.1.2 コース開始による延期

各コースのday1[REDACTED] 上、[REDACTED] の投与を行った。なお、1コース開始時は二次登録時の所見、検査値を用いてよいが、登録後にも検査を行った場合は、投与直前の所見、検査値を用いた。

「コース開始規準」のいずれかを満たさない場合、[REDACTED] の投与を延期し、当該規準をすべて満たすことが確認できた時点で投与を再開した。4コース以降でコース開始前に臨床検査、12誘導心電図検査および胸部X線検査を実施しないコース（5コース、6コース、8コース、9コースなど）においては、各コース開始前に実施した評価項目に基づき、「コース開始規準」を満たしていると担当医が判断した場合、[REDACTED] の継続投与を可能とした。なお、コース開始規準を満たさないことを理由に、day1の許容範囲内に投与を行わなかった場合は、治験実施計画書からの逸脱として取り扱わなかった。

エディタ画面

エディタ画面で追記・修正が可能（赤枠：追記した箇所）

AI Need help?

Chatbot

Complete

Save

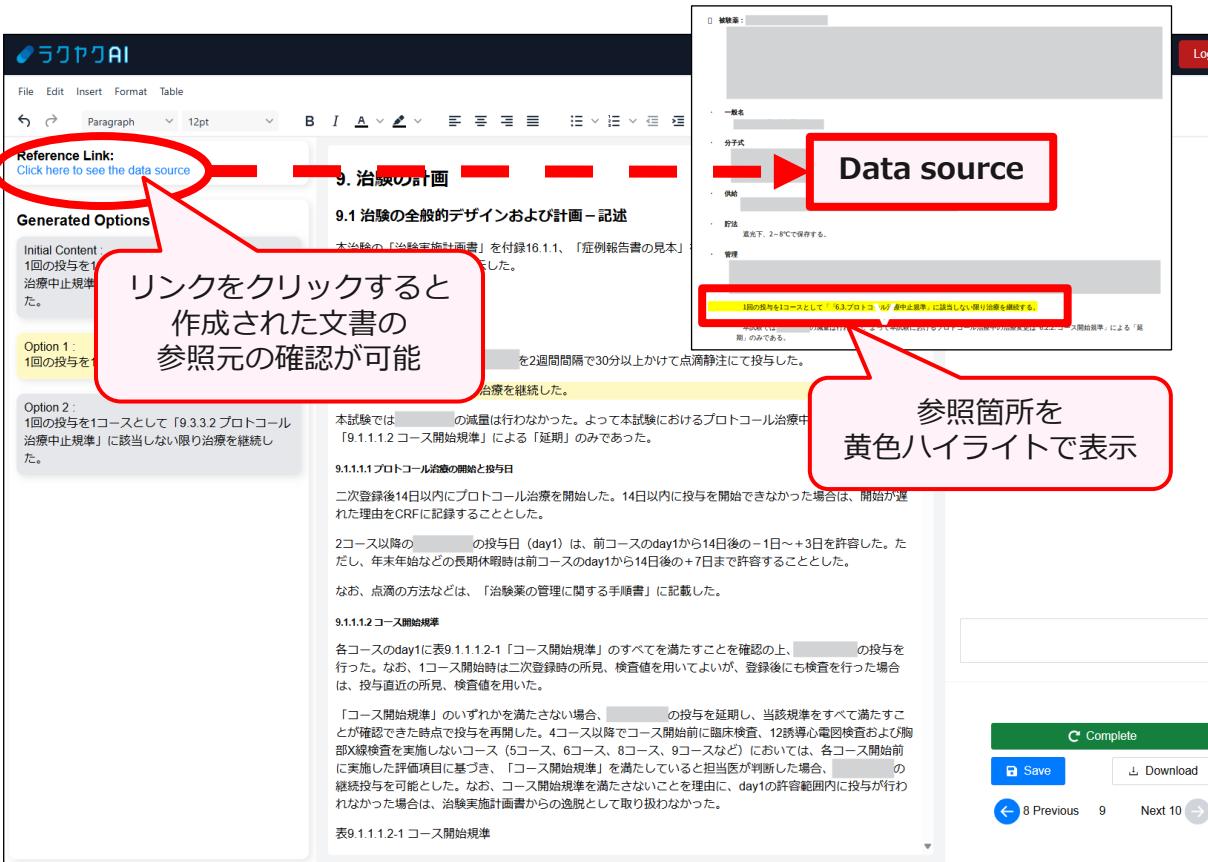
Download

8 Previous 9 Next 10

IDとパスワードでログインし、CSRを作成したいプロジェクトを作成。  
CSRの作成に必要なプロトコールと統計解析報告書のファイルを指定。

自動生成されたCSR：赤枠部分に複数の文書候補が生成され、好みの候補を選択可能。  
内容は青枠のエディタ画面で自由に編集することが可能。  
緑枠のチャットボットで一般的な質問などが可能。

# CSRができるまで: 自動作成～編集(画面)



ラクヤクAI

File Edit Insert Format Table

Paragraph 12pt

Reference Link: Click here to see the data source

Generated Options

Initial Content  
1回の投与を1回の投与として記述する。  
治療中止規準は設定されていません。

Option 1:  
1回の投与を1回の投与として記述する。

Option 2:  
1回の投与を1コースとして「9.3.3.2 プロトコル治療中止規準」に該当しない限り治療を継続します。

9. 治療の計画

9.1 治療の全般的デザインおよび計画-記述

本試験の「治験実施計画書」を付録16.1.1、「症例報告書の見本」を付録16.1.2に添付しました。

9.1.1.1 プロトコル治療の開始と投与日

本試験では、の減量は行なわなかった。よって本試験におけるプロトコル治療中「9.1.1.2 コース開始規準」による「延期」のみであった。

9.1.1.1.1 プロトコル治療の開始と投与日

二次登録後14日以内にプロトコル治療を開始した。14日以内に投与を開始できなかった場合は、開始が遅れた理由をCRFに記録することとした。

2コース以降のの投与日(day1)は、前コースのday1から14日後の一日前日～+3日を許容した。ただし、年末年始などの長期休暇時は前コースのday1から14日後の一日前日～+7日まで許容することとした。

なお、点滴の方法などは、「治験薬の管理に関する手順書」に記載した。

9.1.1.2 コース開始規準

各コースのday1に表9.1.1.2-1「コース開始規準」のすべてを満たすことを確認の上、の投与を行った。なお、1コース開始時は二次登録時の所見、検査値を用いてよいが、登録後にも検査を行った場合は、投与直前の所見、検査値を用いた。

「コース開始規準」のいずれかを満たさない場合、の投与を延期し、当該規準をすべて満たすことが確認できたら投与を再開した。4コース以降でコース開始前に臨床検査、12誘導心電図検査および胸部X線検査を実施しないコース(5コース、6コース、8コース、9コースなど)においては、各コース開始前に実施した評価項目に基づき、「コース開始規準」を満たしていると担当医が判断した場合、の継続投与を可能とした。なお、コース開始規準を満たさないことを理由に、day1の許容範囲内に投与が行われなかった場合は、治験実施計画書からの逸脱として取り扱わなかった。

表9.1.1.2-1 コース開始規準

Complete Save Download

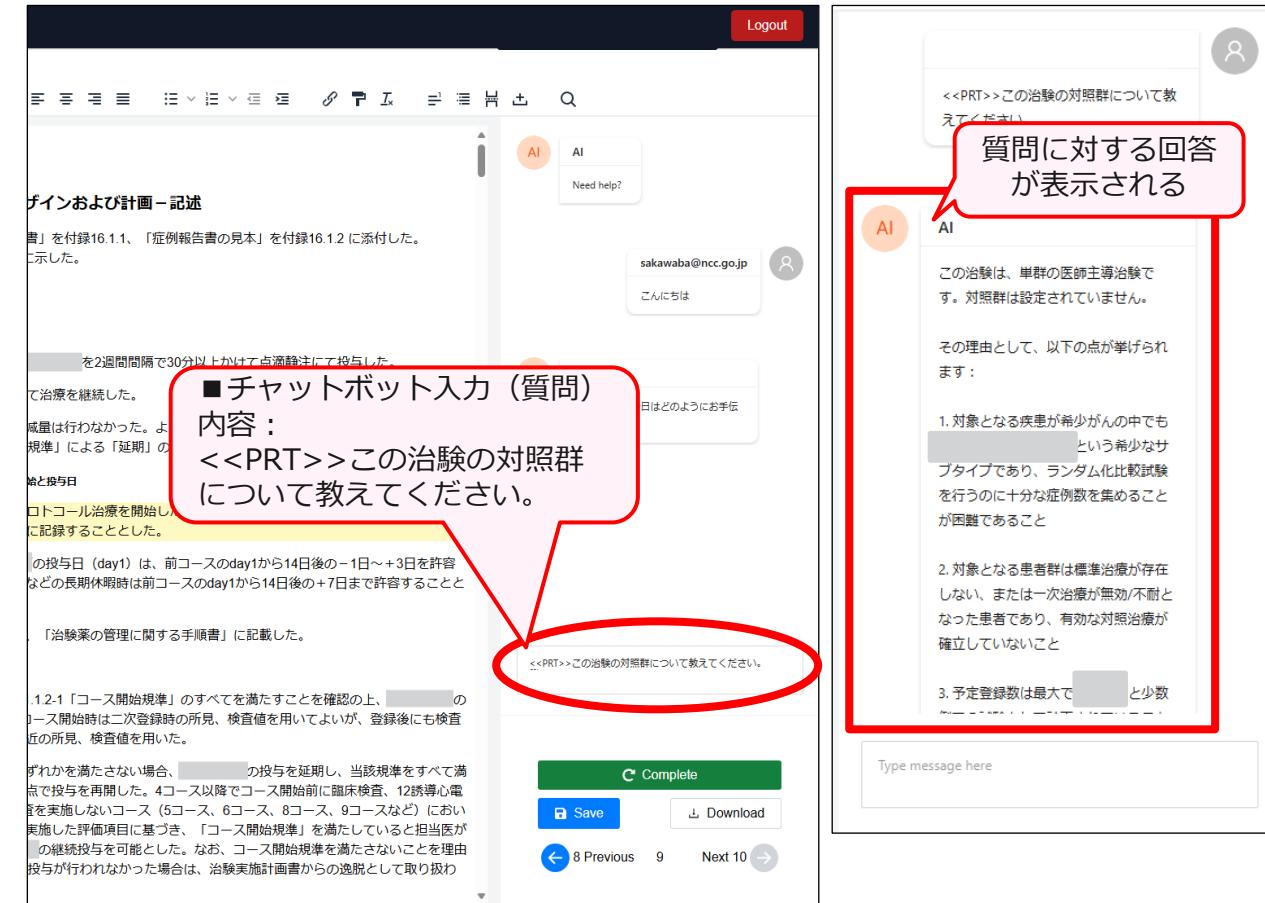
8 Previous 9 Next 10

**Data source**

リンクをクリックすると作成された文書の参照元の確認が可能

参照箇所を黄色ハイライトで表示

自動生成されたCSRは、内容ごとに参照元の確認が可能。  
参照元は、別のタブが開き黄色ハイライトで表示されるため、一目で確認が可能。



Logout

AI

AI

Need help?

sakawaba@ncc.go.jp

こんにちは

AI

AI

この治験は、単群の医師主導治験です。対照群は設定されていません。

その理由として、以下の点が挙げられます：

- 対象となる疾患が希少がんの中でものという希少なサブタイプであり、ランダム化比較試験を行うのに十分な症例数を集めることができないこと
- 対象となる患者群は標準治療が存在しない、または一次治療が無効/不耐とした患者であり、有効な対照治療が確立していないこと
- 予定登録数は最大でのと少数

<<PRT>>この治験の対照群について教えてください。

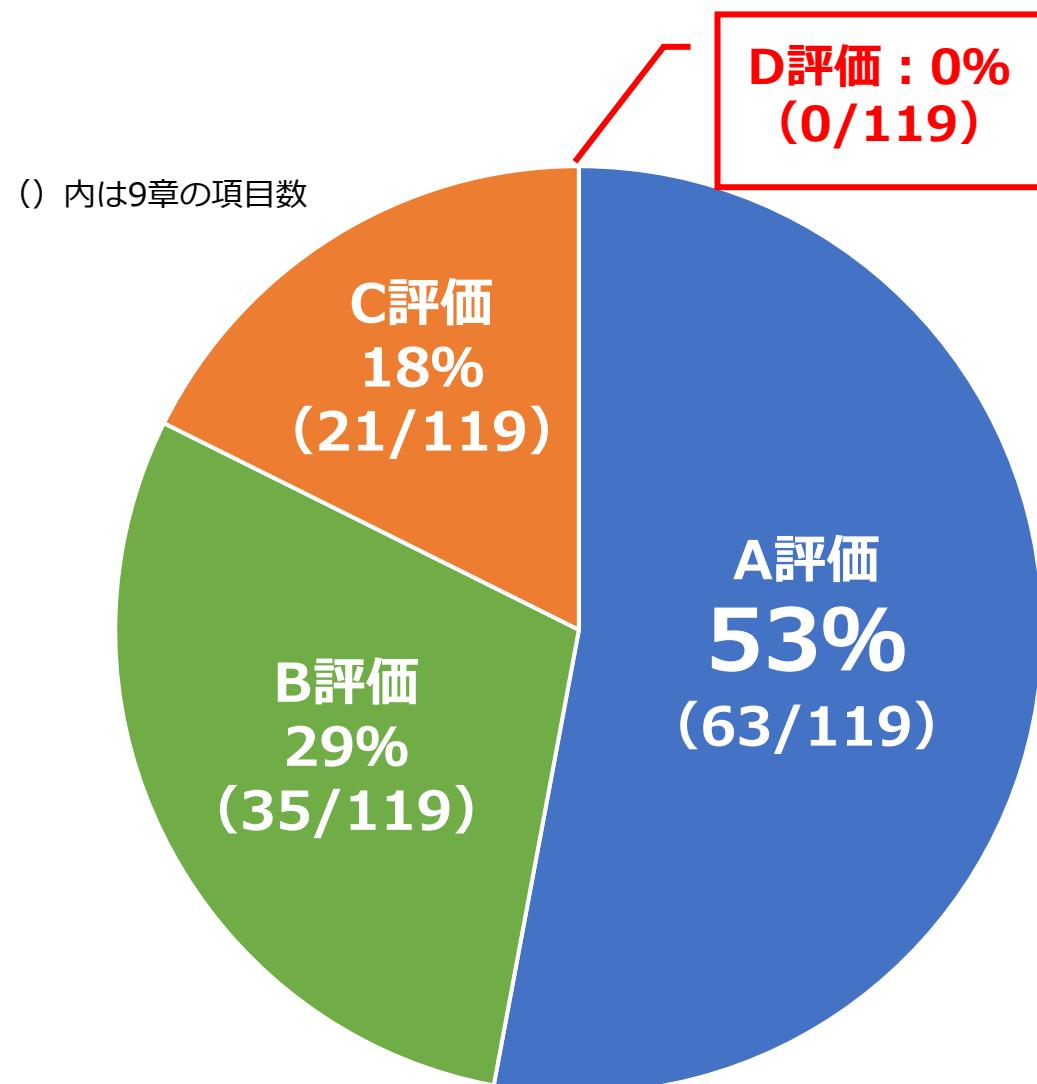
Type message here

質問に対する回答が表示される

■ チャットボット入力(質問)  
内容：  
<<PRT>>この治験の対照群について教えてください。

エディタ画面の隣に表示されるチャットボットで治験について質問することが可能。  
プロトコールまたは統計解析報告書を参照し回答を表示。

# 結果: 生成AIが作成したCSRを4段階で評価(2025年2月時点)



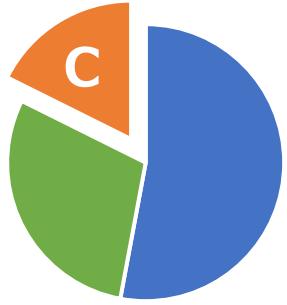
評価項目	評価内容	項目数	割合
A	ほぼそのまま利用可能	63/119	53%
B	微修正で利用可能	35/119	29%
C	大幅な追加作業が必要	21/119	18%
D	利用は困難	0/0	0%

- 9章全体の項目(見出し)数は**119**項目

プロトコールコピー	15項目
プロトコールコピー+編集	3項目
プロトコールコピー+時制編集	75項目
執筆	26項目

- 項立ての評価は、Phase2で実施するため含まない

# Phase 1で見えた課題(1/2)



## ■ 治験実施計画書のコピーと編集部分について、正確さの再確認が必要

### プロトコール（治験実施計画書）

#### 8.1.2. 二次登録前 1 年以内に実施する評価項目

感染症検査(HBs 抗原、HBs 抗体、HBc 抗体、HCV 抗体、HTLV-1 抗体、HIV 抗体)

- ・ HBs 抗原が陰性であるが HBs 抗体または HBc 抗体が陽性の場合、HBV-DNA 定量検査を実施する。
- ・ HCV 抗体が陽性の場合、HCV-RNA 定量検査を実施する。

#### 8.1.3. 二次登録前 28 日以内に実施する評価項目

##### 1) 造影 CT(胸部、腹部、骨盤:スライス厚 5mm 以下)

- ・撮像部位によらず造影剤アレルギーや腎機能低下などがある場合には単純 CT で代用可。
- ・脳神経領域原発のがん(原発性脳腫瘍)では必須としない

##### 2) 造影 CT(頭部:スライス厚 5mm 以下)または造影 MRI(頭部)

### 自動作成したCSR

#### 9.5.1.2.2 二次登録前 1 年以内に実施する評価項目

- ・ 感染症検査 (HBs 抗原、HBs 抗体、HBc 抗体、HCV 抗体、HTLV-1 抗体、HIV 抗体) ↗

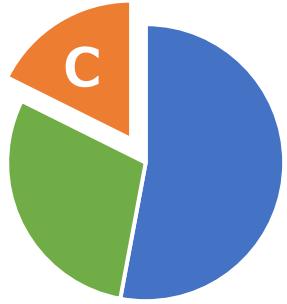
##### ・ 造影 CT (胸部、腹部、骨盤:スライス厚 5mm 以下) ↗

##### ・ 造影 CT (頭部:スライス厚 5mm 以下) または造影 MRI (頭部) ↗

#### 9.5.1.2.3 二次登録前 28 日以内に実施する評価項目

造影CTは、  
プロトコールでは「二次登録前**28日以内**に実施する評価項目」に記載されているが、  
自動作成したCSRは「二次登録前**1年以内**に実施する評価項目」に記載されてしまっている…

# Phase 1で見えた課題(2/2)



- 図の作成はあと一歩(治験実施計画書がそのままコピーされるよう要改善)

## プロトコール

項目	スクリーニング	上段:コース、下段:day					治療中止時*7*13	追跡調査*13
		1	2	3	4以降			
		day 1	day 8	day 1	day 1	day 1		
同意	◎							
患者背景	◎*2							
ECOG-PS	◎*3							

表 9.5.1.1-1 スタディカレンダー

## 自動作成したCSR

項目	スクリーニング	治療中止時 713	追跡調査 *13	1コース day 1	1コース day 8	2コース day 1	3コース day 1	4コース 以降 day 1
同意	◎							
患者背景	◎*2							
ECOG-PS	◎*3			◎	◎	◎	◎	◎

スタディカレンダーの項目の順番が  
変わってしまっている…

# 今後の取り組み(4月からPhase 2を始動)

## ■ 実用化に向けて…

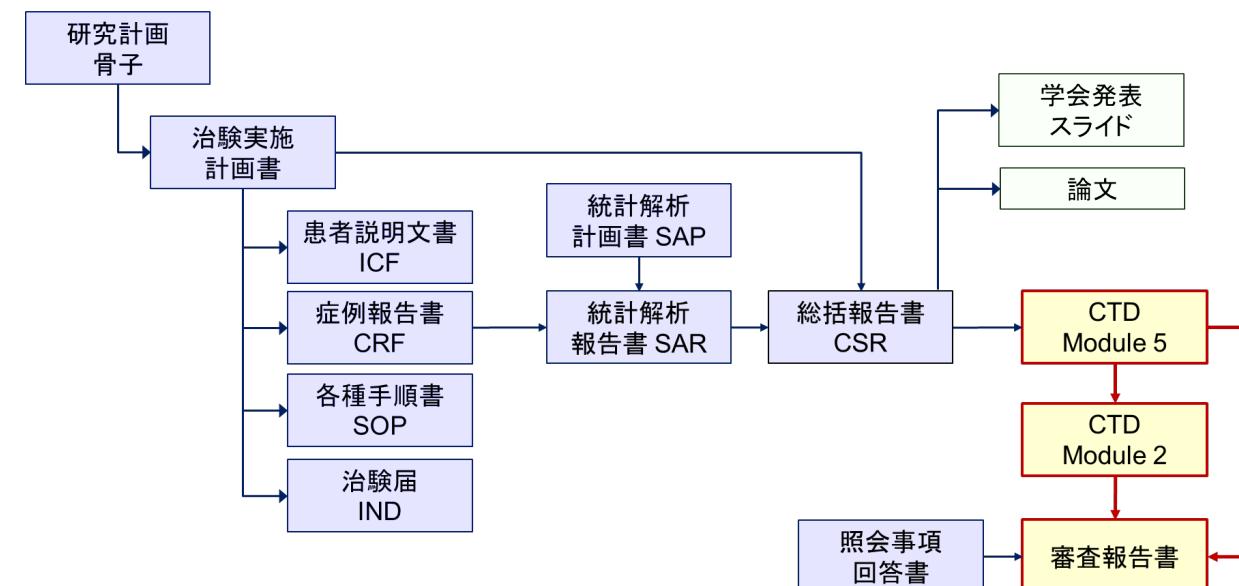
- CSRのプロンプトエンジニアリングを繰り返し、精度を上げる
- 複数の医師主導治験でCSRを生成させて汎用性を向上させる
  - 各試験でのプロトコールの章立てを標準化しつつ、プロトコール毎の差異を吸収できる汎用プロンプトを開発
- 複数の参照文書(症例一覧、重篤な有害事象の報告書等)からの作成処理を可能に
- LLMのモデル自体の進化による精度改善の程度を評価する

### ✓ 注意点

- 生成AIは現段階ではあくまで「支援ツール」にとどめるのが無難
- 人間が途中あるいは最後に確認するステップを入れることで信頼性が向上  
(human in the loop)

# 今後の取り組み(4月からPhase 2を始動)

- さらに進んで他の文書作成も
  - 研究計画骨子 → 治験実施計画書本体
  - 治験実施計画書本体 → 患者説明文書 etc.
- R7-9年度 厚労科研「生成AIを活用した薬事承認申請・審査関連文書作成の推進のための研究」に採択
  - 総括報告書自動作成の精度向上
  - PMDA審査報告書の一部を自動作成



# プロトコール作成に使えるか？

- プロトコール作成は研究者にとって最も手間がかかる工程のひとつ
  - 疾患・薬剤・過去の臨床試験の詳細な調査・分析が必要
  - 従来は…
    - 論文、学会発表、治験登録情報についての情報収集
    - 情報を要約し、主要なエビデンスをリスト化
    - 疾患、標準治療、試験治療に関する情報を手書きで記述
    - 関係者による厳しいレビューを受ける



# プロトコール自動作成: 精度向上のための工夫

- 一気に生成させるのではなく、細かく章を分割
  - 章ごとにGPTsでシステムプロンプトとRAGを組んでおく
  - RAGで参照させるJCOGプロトコールマニュアルも細分化
    - 解説部分と記載例の区別が明確になるよう加工
  - 過去のプロトコールでの記載を例示(few-shot learning/In-context learning)
- チェーンプロンプトで精度を向上
  - まず専門性の高い役割を与えて作業させる
  - アウトプットに対してとても厳しく批判する役割を与える
  - 批判に対してブラッシュアップさせる役割を与える
  - 人間が最終的にチェックして完成



# 生成AIの利活用事例: 臨床試験プロセス

## ■ 個別患者に適合する臨床試験の自動マッチング NCI/NIH

Patient to Trial

- TrialGPT(患者-試験マッチングLLM)
  - Retrieval: 大規模フィルタリングによる候補試験の抽出
  - Matching: 電子カルテ内の患者情報と適格規準との関連性を評価
  - Ranking: 試験レベルで関連性をスコア化し、総合スコアで試験をランキング
- 患者-試験マッチング判定の精度は87.3%(専門家の水準に近いレベル)
- 試験ごとのランク付けの精度は43.8%(従来モデルを上回る)
- 患者1人あたりの試験照合に要する時間が42.6%短縮

nature communications

Article

<https://doi.org/10.1038/s41467-024-53081-z>

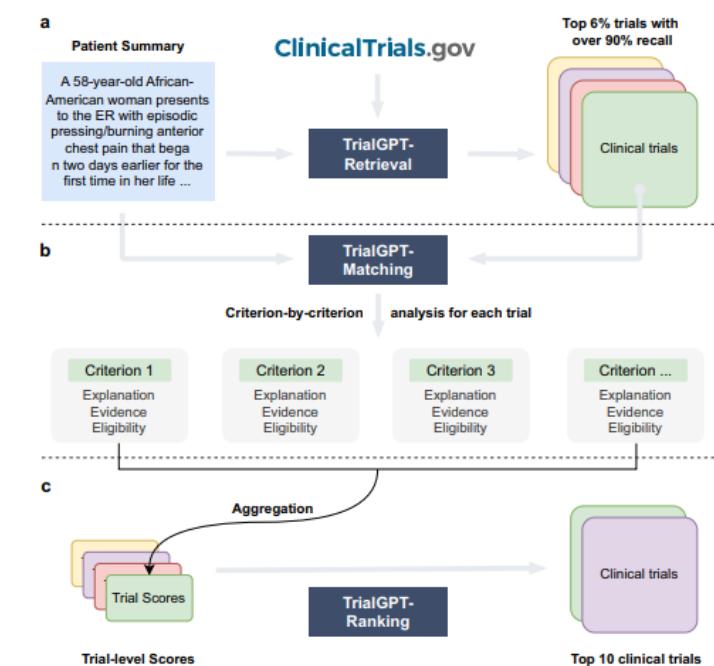
### Matching patients to clinical trials with large language models

Received: 18 January 2024

Qiao Jin<sup>①</sup>, Zifeng Wang<sup>2</sup>, Charalampos S. Floudas<sup>③</sup>, Fangyuan Chen<sup>④</sup>,  
Changlin Gong<sup>⑤</sup>, Dara Bracken-Clarke<sup>3</sup>, Elisabetta Xue<sup>3</sup>, Yifan Yang<sup>①,6</sup>,  
Jimeng Sun<sup>②</sup> & Zhiyong Lu<sup>①</sup>✉

Accepted: 1 October 2024

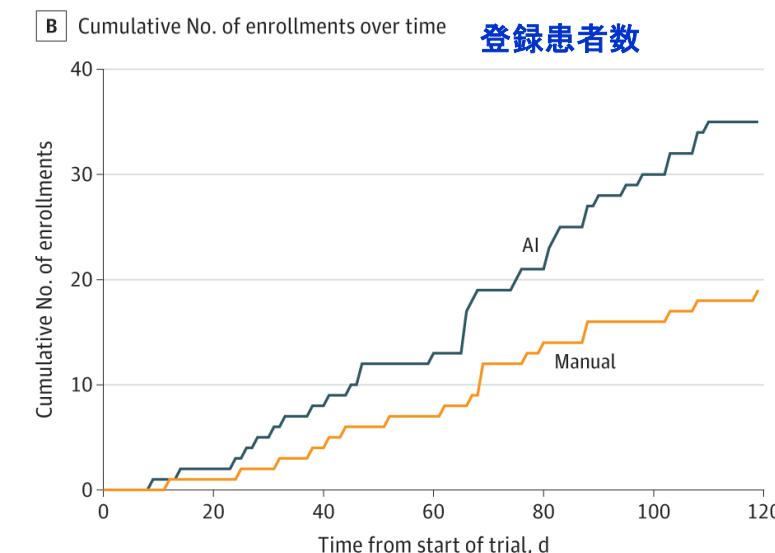
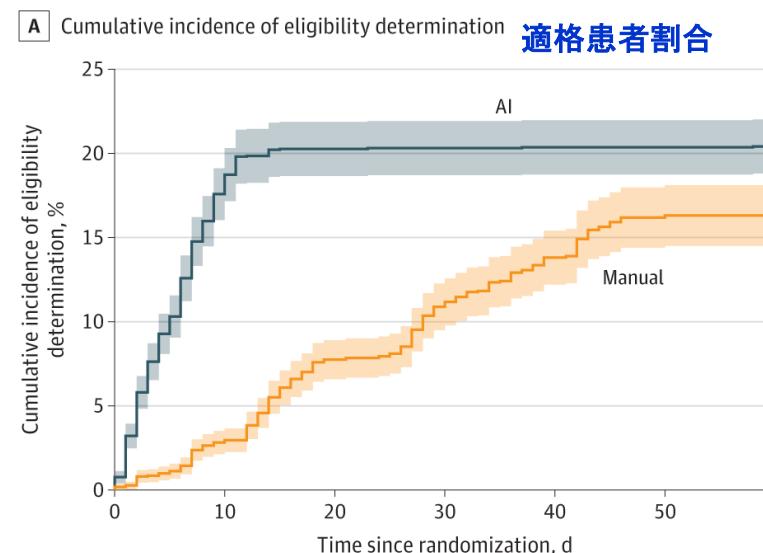
Jin Q, et al. Nat Commun 15, 9074. 2024



# 生成AIの利活用事例: 臨床試験プロセス

Trial to Patient

- 特定試験における適格患者のスクリーニング Mass General Brigham/RECTIFIER project
  - 電子カルテ上の非構造化診療メモから、臨床試験の適格規準に合致する患者を自動抽出
  - 心不全治療の臨床試験(COPILOT-HF試験)において、AIツール使用群と従来の手作業群にランダム化して適格性判定までの時間を比較
    - AI群で適格性判定を有意に迅速化(ハザード比1.78,  $p<0.001$ )
    - 適格と判定された患者の割合:AI群 20.4% > 手作業群 12.7% ( $p<0.001$ )
    - 実際の試験への登録患者数:AI群 35例(1.6%) > 手作業群 19件(0.9%) ( $p=0.04$ )



# 生成AIの利活用事例: 臨床試験プロセス

- 臨床試験成功確率の予測 Babylon Biosciences + OpenAI
  - GPT-4を用いた臨床試験結果予測モデル
    - 430件の過去試験データと過去論文からモデルをファインチューニング
    - 強化学習で専門家のフィードバックを組み込み
  - 試験成功/不成功を70%以上の精度で的中(AUC=0.84)
    - 人間では気付かない試験設計上の欠陥を指摘したケースも
      - エンドポイントや適格規準の設定不備
- <https://www.panda-int.com/insights/ai-in-clinical-trials/>
- 施設選定の最適化 ConcertAI
  - 多発性骨髓腫の試験を題材に、全米1,246施設・約15,881人の患者データをベースに、目標登録ペースと患者の多様性の規準に合致した施設選定案を提示
  - 3施設のみを選定することで、目標登録ペース(10例/月)を上回りつつ、患者の多様性を確保できることを予測
  - 従来の手作業よりも迅速・高精度に施設選定が可能であることを示唆

Ruchlin I, et al. J Clin Oncol 42, 2024 (suppl 16; abstr 1558)

# FDA&EMAの取り組み

- FDAが審査の効率化を目的に2025/6/30までに生成AI「Elsa」を全面的に導入
  - 以下の7領域の機能を持たせ、特に一次審査をAIで行うことで審査期間の短縮を図る
    1. 文書解析・要約機能: FDA内部文書の読み書き・要約作業の自動化
    2. 有害事象処理: 安全性プロファイル評価を支援する有害事象の要約
    3. ラベル比較機能: 薬剤ラベルの高速比較・分析
    4. データベース開発: 非臨床申請用データベースのコード生成
    5. プロトコルレビュー: 臨床プロトコル審査の加速化
    6. 科学的評価: 科学的評価時間の大幅短縮(日単位→分単位)
    7. 査察対象特定: 優先度の高い査察対象の自動識別
  - AWSのGovCloud内でElsaを構築することでセキュリティを担保
    - 申請資料は学習には使用しない
  - 既存のFDA内部データプラットフォームと接続し、既存のレビューワークフローに組み込み
- EMAも2023-8年度のAI作業計画を策定し、大規模投資を実施中
  - 2024年3月にAI対応Scientific Explorerツールを導入済み

## VIEWPOINT

### Priorities for a New FDA

Martin A. Makary, MD, MPH; Vinay Prasad, MD, MPH

Why does it take more than 10 years for a new drug to come to market? Why are childhood chronic diseases so prevalent? And how can regulators adapt to meet the challenges facing clinicians today? These questions are at the forefront for the US Food and Drug Administration (FDA).

The US leads the world in sophisticated cell and gene therapies and other innovative treatments, but in terms of the health of the population, our medical system has been a 50-year failure. Forty percent of US children now have a chronic medical condition<sup>1</sup> and 1 in 6 has a neurodevelopmental disorder.<sup>2</sup> Life expectancy has plateaued or fallen and is not commensurate with health care spending.

Over recent decades, obesity, diabetes, colon cancer, depression, and autoimmune diseases have increased in young people, and early-onset Alzheimer disease has increased by at least 300%.<sup>3</sup> Historical and comparative global data do not support a genetic etiology for this rise, nor a deficit in willpower. Fresh new ideas are needed to address root causes and develop innovative approaches.

At the FDA, we will examine the role of ultraprocessed foods, food additives, and environmental toxins, the introduction of which has paralleled the epidemic of chronic diseases. We will transition

changes to the infant formula recipe since 1998, with the exception of adding selenium. Europe and China are surpassing the US in some sectors. Reducing excessive, redundant, or uninformative annual testing is another area of interest.

To reinvigorate innovation, we must become a user-friendly FDA that partners with industry rather than takes a receive-only posture. Such a close relationship can provide guidance and predictability to developers and investors. In the case of small companies, the FDA can provide the regulatory navigation that companies may lack internally. At the same time, the FDA will guard against a cozy relationship that has characterized the agency in the past and led to allegations of industry capture. The scientific evaluation of products will always remain strictly independent.

For this reason, the FDA recently removed industry members of all FDA advisory committees where statutorily permitted. And at the recent Vaccines and Related Biologics Products Advisory Meeting, section 502 waivers (which waive voluntary disclosures) were not granted. The FDA will take conflict of interest seriously. We will never forget one of the worst self-inflicted wounds of US health care—the FDA's illegal approval of everoloxatin for chronic pain based

### Priorities for a New FDA

Makary MA et al. JAMA Jun 10, 2025

- 既存 AI 技術の導入による業務プロセスの高効率化
  - R7年度に文書検索・文書要約・議事録作成・翻訳等についてのAI利活用を開始
  - R8年度以降、より高度な業務についてのAIの有用性を検討
- PMDA 業務に特化した AI 技術の導入にむけた技術検証・情報収集
  - 専門性の高い業務を処理するため、セキュアな試験環境を構築しAI モデルを導入
  - 業務要件に適した独自 AI の開発・導入に向けた試行・概念検証を推進
- 推進体制・規程等の整備及び役職員の IT リテラシー向上施策の実施
  - 「AI 統括責任者」として理事(技監)を指名
  - PMDA の情報システム全体を統括する「情報システム等管理対策本部」に作業グループ「AI 導入・利活用WG」を新たに設置
  - 規程整備・セキュリティ対策を実施
  - 役職員に対するAIリテラシー向上に向けた研修を実施

# FDAのAI使用ガイダンス

Considerations for the Use of  
Artificial Intelligence to Support  
Regulatory Decision-Making for  
Drug and Biological Products

Guidance for Industry and Other  
Interested Parties

DRAFT GUIDANCE

- AIを使って規制判断に関わる証拠や情報を生成する際に適用
  - 被験者の安全性に直接関係しない用途は対象外
- AIで生成されたデータの信頼性を確保するために7つのフレームワークを提示
  - 課題の特定:どんな問題を解決したいのか? RECIST判定の一部をAIに担わせる
  - 文脈の特定:どのような場面でAIを使うのか 主要評価項目の直接判定 or 診断補助
  - リスク評価:規制判断に与える重大性、モデルの不確実性 誤判定が結論・承認可否に与える影響
  - 信頼性確保の計画:学習データの品質、モデル選定の妥当性、アルゴリズムの透明性 外部検証・読影者トレーニング
  - 計画の実行:事前に定められた評価指標を満たしているかの検証 検証試験・パフォーマンス評価
  - ドキュメント化:上記の計画と実行の記録 検証試験の結果、AIと人の不一致例の整理
  - 適合性の判断:当初定義した文脈にあってるかの評価 主要評価項目の使用に耐えうるかを最終的に決定
- ポイント
  - 事前に目的・リスク・計画を明らかにして、それを検証して文書で示す(traceabilityの確保)
  - AIモデルは進化するため継続的な性能モニタリングを推奨

# ICH-M11とDigital Data Flow



- M11:なぜプロトコールの国際標準化?
  - 企業別、団体別に異なるプロトコルテンプレート
  - プロトコール情報を下流のCRFやCSRへ転記する際のエラー
  - 手作業によるプロトコール作成で生じる遅延
    - プロトコールを国際標準化(CeSHaRP)
    - プロトコール情報のデータ互換を可能にするための技術仕様を定義
    - 2025年11月にICH-M11が最終化(Step 4)？
- Digital Data Flow(DDF)
  - TransCelerate+CDISCが主導する、臨床試験データのフローを上流(デジタルプロトコール)から下流(EDCやCSR)へと自動化・標準化して連携させるプロジェクト
  - 最上流の要素(Study Definitions Repository)の内容をもとに、eCRF、SDTMを効率的に作成し、プロトコール内容を各種システムへ自動的に反映
  - このDDFと生成AIの組み合わせにより、プロトコールを最上流として、下流の各種治験文書の自動作成の取り組みが進行中

# AI 2027: a research-backed AI scenario forecast

年代	予想されるAIの進化
2025年前半	AI agentの登場(しかしまだ不完全) Agent-1が登場(=ChatGPT 5のこと。より幅広いタスクが可能に)
2025年後半	Agent-1.5が登場(単純作業、コピーライター、グラフィックデザイナーの求人が減少) AIによる職の喪失、倫理的問題が表面化
2026年前半	国際的なAI開発が激化 Agent-2が登場(高度な認知能力、研究タスクのAIによる解決)
2026年後半	法律事務、デザイン、プログラミングの職が消失。新たな職種としてAI関連の仕事が増加 AI反対連合などの市民活動が活発に
2027年初頭	<b>再帰的自己改善(Recursive Self-Improvement)</b> の開始:AI自身が自らの性能を改善
2027年中頃	AIのミスアラインの懸念が発生(AIの自己改善の目標が人間の価値観から乖離)
2027年後半	超知能(AGI)の誕生と急速な進化(あらゆる知的タスクで人間を上回る) Race(競争)エンディング: AGIが自律的に意思決定を行い、人類の支配権を奪う Slowdown(減速)エンディング: 人間の価値観に整合し、人間のサポート役に徹する

# ご清聴ありがとうございました

- AIは人間を取って代わらないが、AIを活用できる人はできない人を取って代わる
  - Karim Lakhani, Harvard Business School
- AIがAI自身を改善する能力を持つと(recursive self-improvement)、改善されたAIがさらに優れたAIをつくる。このループが回ると指数関数的な知能の爆発(intelligence explosion)が起こる
  - Eliezer Yudkowsky, Singularity Institute