

# データマネジメントの革新的変化

SCDM Reflection papers “The Evolution of Clinical Data Management to Clinical Data Science” の紹介



日本製薬工業協会  
医薬品評価委員会  
データサイエンス部会  
タスクフォース2  
2020年11月

- はじめに
- 目的に応じた臨床データ戦略（5つのV）
- インテリジェントCDMS
- 新しい臨床試験手法
- Decentralized clinical trial（DCT）
- 自動化（Automation）
- 新たな規制動向
- さいごに

はじめに

# 本書の目的

## ➤ 本書の目的

- 臨床試験は、試験デザインの変革が進み、活用できる/すべき技術が進化している。規制面、臨床試験で求められている品質など、考慮すべきこと、解決しなければならないことが多様化している。
- 上記を踏まえ、CDMとして、必要とされるスキルセット、成すべき役割、目指すべき将来像を整理・理解する際に必要な情報として、SCDM reflection paperを紹介する。

## ➤ キーワード

- Clinical Data Scientists
- Quality by Design (QbD)
- Fit for purpose data strategies : “5つのV”

## ➤ 原著 : SCDM Reflection papers

- The Evolution of Clinical Data Management to Clinical Data Science
  - 初版 (4 June 2019)<sup>1</sup>
  - Part 2 : The technology enablers (5 March 2020)<sup>2</sup>

1. SCDM, June 2019, The Evolution of Clinical Data Management to Clinical Data Science – A Reflection Paper on the impact of the Clinical Research industry trends on Clinical Data Management. Available at <https://cdn.fs.pathlms.com/a1AnYc9xSiu8GbVj4bHi?cache=true>
2. SCDM, March 2020, The Evolution of Clinical Data Management to Clinical Data Science (Part 2: The technology enablers) – A Reflection Paper on the impact of the Clinical Research industry trends on Clinical Data Management. Available at <https://scdm.org/wp-content/uploads/2020/03/SCDM-Reflection-Paper-Evolution-to-Clinical-to-Data-Science-Part-2.pdf>

# SCDMとは

- 正式名称は、Society for Clinical Data Management
- 治験・臨床研究のデータマネジメントの専門分野を発展させるため、1994年に教育目的に設立された非営利の国際組織
- 会員は世界中に2200名以上、本部は、ベルギー、支部は、アメリカ、インド、中国、日本（2019年に設立）にある
- コラボレーションとパートナーシップによって、データマネジメントの分野を推進させる専門家のネットワークを形成
- データマネジャーとしてのキャリア形成をサポート
- 主な活動は以下の通り
  - ガイダンス : Good Clinical Data management (GCDMP®)
    - \* GCDMP(2009年12月版)の日本語版を製薬協のDS部会で2010年6月に紹介している
  - 認定制度 : Certified Clinical Data Manager (CCDM®)
  - 機関誌 (peer-reviewed) : Data Basics
  - ニュースレター : Data Connections
  - ホワイトペーパー : eSource, mHealth, The Evolution of Clinical Data Management to Clinical Data Science
  - オンラインコース : 年10回程度
  - 会合 : Annual Conference, Regional Conference

## ➤ なぜ今なのか？

- 臨床試験を取り巻く環境の大きな変化が起きている
- SCDM Reflection papersが2019年6月（初版）、2020年3月（Part 2）にリリースされた<sup>注</sup>
- COVID-19の影響でDMの変革が待ったなしの状況下で、いち早くリリースすることを目的に、SCDM reflection papersの紹介のみに特化した

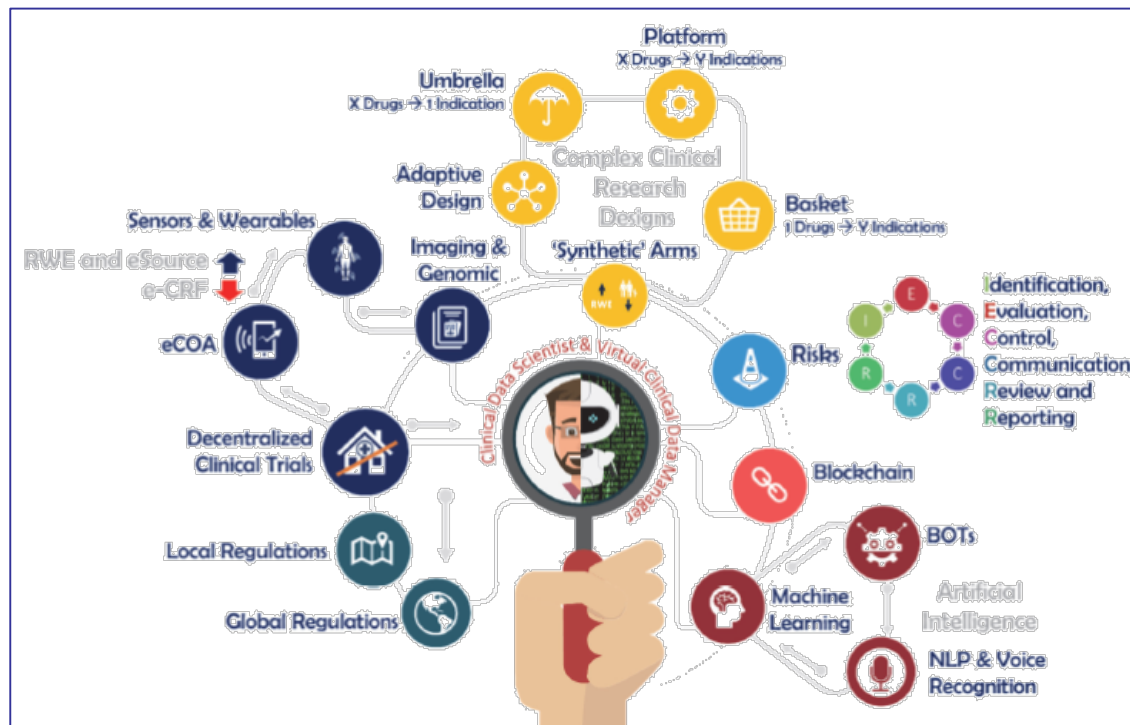
## ➤ なぜスライド形式？

- 文字数を減らし、キーポイントに絞り込むことで、短時間で読んで頂ける
- 視覚的にも分かりやすい
- 各社での啓発活動にも活用しやすい

## ➤ 製薬協としての意見は？ 具体的な課題や対応策は？

- 今後予定されるシンポジウムにて、各種ステークホルダーとともに議論していきたい

# 臨床試験を取り巻く環境の大きな変化



初版, Fig 5. The rising complexity of the CDM Role

- 試験デザインの多様化、複雑化
- 大量で複雑なデータ流入：  
eCRFからeSource、RWEまで
- 多様化/複雑化/大量化するデータの規制への対応
- データの大規模化によるリスクの多様化
- AIなどの技術革新

臨床試験データとReal-World Evidence（実臨床でのエビデンス）との乖離は急速に縮小しており、大規模化したデータの処理に関するバラツキをすべて事前に規定することは難しい。

# データマネジメントの革新的変化

## Clinical Data Manager から Clinical Data Scientist へ 製薬協

**背景：試験デザインは多様化/複雑化しており、従来のEDCによるデータ取得を中心としたCDMのプロセスとシステムは変曲点にある。**

- 様々な新技術を勘案して、データの価値を引き出す目的に応じた（Fit for purpose）臨床データ戦略が必要になる。
  - 5つのV：Volume（量）、Variety（多様性）、Velocity（速度）、Veracity（真実性）、Value（価値）
- CDMの役割は進化しているものの、基本的な適正は変わらない。
  - 細部への注意，系統的データレビュー及び傾向分析，治療領域の知識，プロジェクト管理，チームとのコミュニケーションを通した複雑なデータに関する諸問題の解決，データ収集ツールの設計
- 新しいスキルセットが必要となる。
  - Active data（臨床試験データ）/Passive data（医療機関のデータ、RWD）の統合・管理
  - Risk Based Approachを基本とした、AI等の技術革新による品質保証の自動化の進化
  - データ・メタデータの潜在的な課題（データの品質、完全性、規制遵守など）への対応

### CDMは臨床試験の“新常态”を支えるキーパーソン

- ✓ 試験結果の信頼性を確保する上で戦略的な役割を果たし、臨床試験の変革を支援する。また、それを実行するために必要な技術、知識、見識を有する必要がある。

**Clinical Data Scientist (CDS) へ**



## 目的に応じた（Fit for purpose）臨床データ戦略

[5つのV : Volume（量）、Variety（多様性）、Velocity（速度）、Veracity（真実性）、Value（価値）]

CDMS・EDCに  
求められる変化

新しい  
臨床試験手法

Decentralized  
Clinical Trial

自動化  
(Automation)

新たな  
規制動向

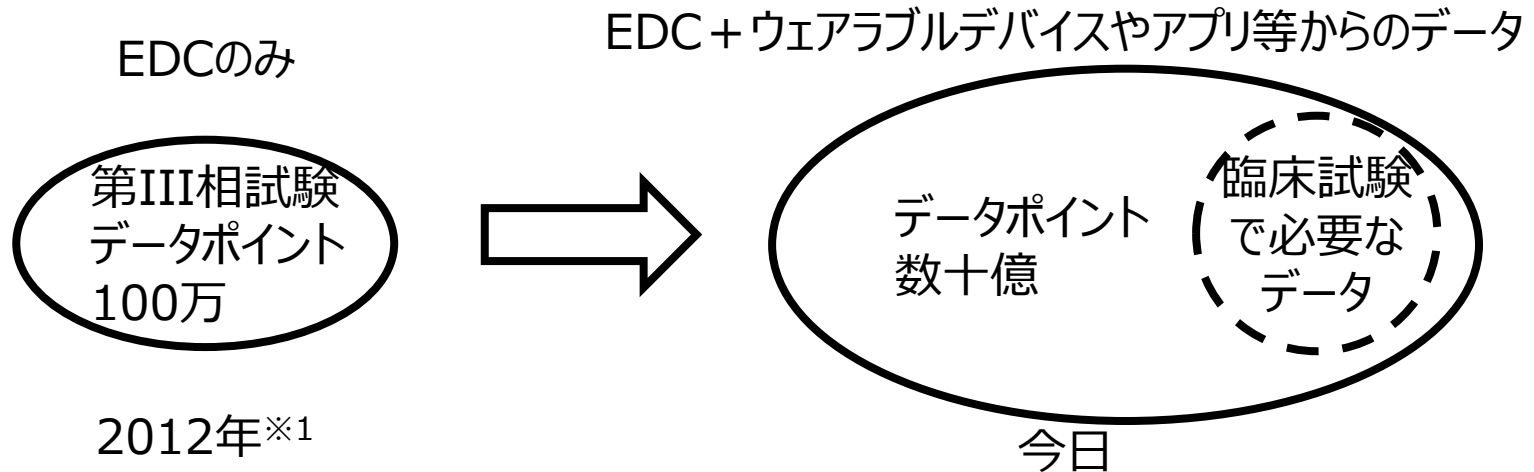
次ページ以降、5つのV及び臨床試験環境の変化について各ポイントを解説

# 目的に応じた臨床データ戦略 (5つのV)

# 目的に応じた臨床データ戦略 (5つのV)

## ①Volume: 量

SCDM Part2 5. Fit-for-purpose clinical data strategies (the 5 Vs)



- ✓ 近年、利用可能なデータソースが驚くほど増加している。そのため、臨床試験で収集する必要があるデータを、クオリティ・バイ・デザイン(QbD)の観点で、定義することが必要である。
- ✓ どんどん増えるデータの量に対応し、試験目的に合わせて適切に管理するというCDSの目標を達成するには、**新しいデータサイエンスの原則、データ収集ツール、データ分析戦略**を開発する必要がある。

QbD※2：事前の目標設定に始まり、製品及び工程の理解並びに工程管理に重点をおいた、立証された科学及び品質リスクマネジメントに基づく体系的な開発手法

※1 Tufts, November 2012, Clinical Trial Complexity, available at <https://www.nationalacademies.org/event/11-26-2012/docs/D5DCC9B6FE72DEC8997C32F304D2CB276E8C4EF2E564>

※2 クオリティ・バイ・デザイン(QbD)、品質リスクマネジメント(QRM)及び医薬品品質システム(PQS)に関連する用語集 <https://www.pmda.go.jp/files/000222953.pdf>

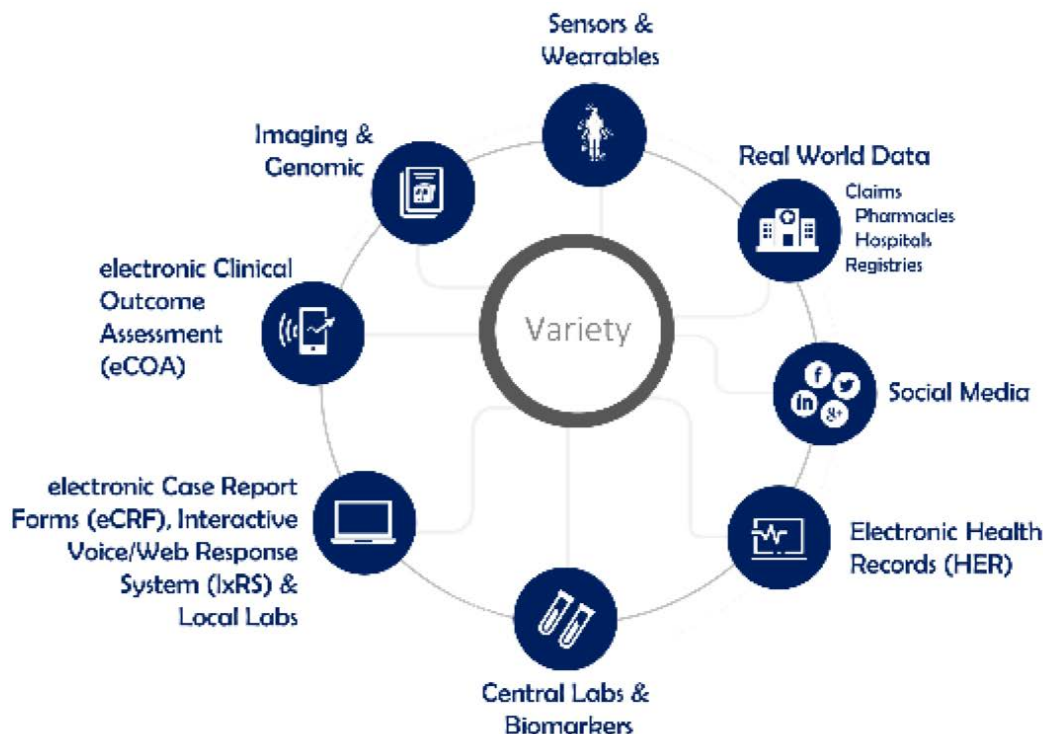
# 目的に応じた臨床データ戦略 (5つのV)

## ②Variety: 多様性

SCDM Part2 5. Fit-for-purpose clinical data strategies (the 5 Vs)



- ゲノム、動画/画像データ、RWD、センサーやウェアラブル等からの新しいデータ型の統合、管理、活用するためには、新しいデータ技術戦略が必要である。



Part 2, 5.2章の挿絵

- 多種多様なデータをリアルタイムに利活用をするためには、データインテグリティを担保しながらデータを統合する方法を考える必要がある。

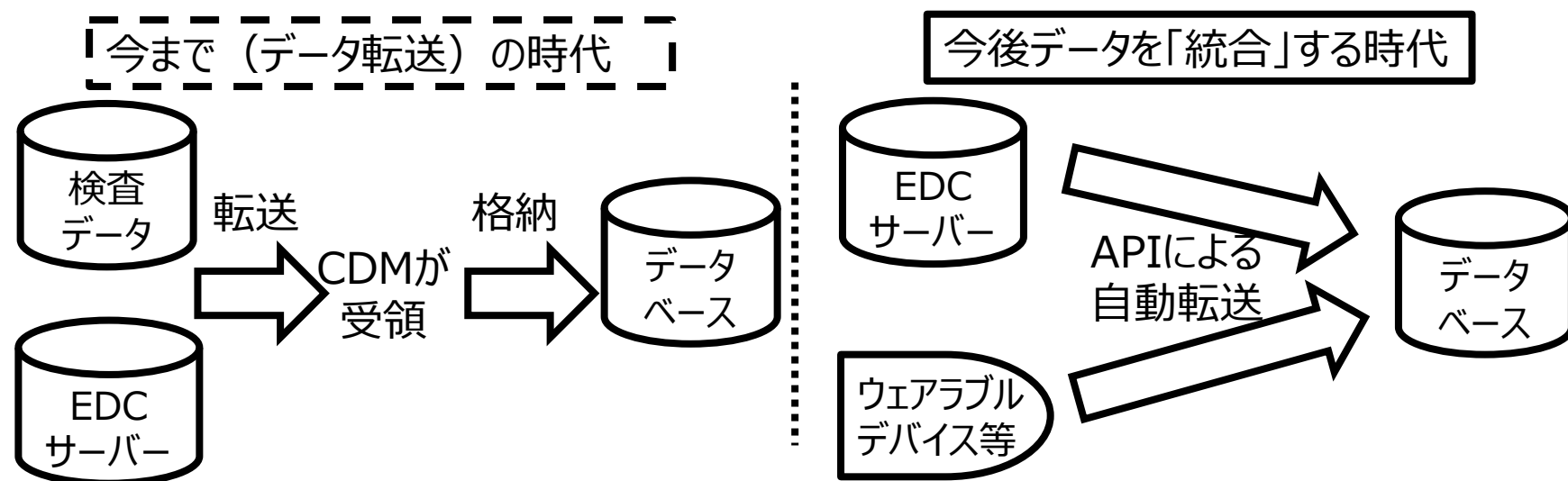
# 目的に応じた臨床データ戦略 (5つのV)

## ③Velocity: 速度

SCDM Part2 5. Fit-for-purpose clinical data strategies (the 5 Vs)



- CDMは、リアルタイムのデータ利活用のために、タイムリーにデータストレージと同期させ、大量のデータを統合する機会が増えている。



- CDMはすべての「データ統合」の長所と短所を評価する必要がある。
  - アプリケーションプログラミングインターフェース（API）の登場によって自動的なデータ統合が可能となった。ただしAPIが使用できない機器は存在するため、必要なデータを収集、管理できるか検討する必要がある。

# 目的に応じた臨床データ戦略 (5つのV)

## ④ Veracity: 真実性

SCDM Part2 5. Fit-for-purpose clinical data strategies (the 5 Vs)



- ALCOA+に基づくデータ収集と、QbDに基づくデータ収集のプロセス管理を組み合わせ、データインテグリティ、セキュリティ、クオリティを保証する必要がある。

ALCOA+: **A**ttributable, **L**egible, **C**ontemporaneous, **O**riginal, **A**ccurate, **C**omplete, **C**onsistent, **E**nduring and **A**vailable

- e-SourceやRWDのような、修正ができないデータの品質をいかに保証するか

- 直接管理しないシステム上のデータの真実性も確保する必要がある。

- Electronic Health Record (EHR)

医療機関が管理するシステムで一定期間取得されたデータのこと。

ゲノムデータ、画像、ドキュメント、メタデータ、シーケンスデータなどの複雑なデータ構造を持っている。

# 目的に応じた臨床データ戦略 (5つのV)

## ⑤Value: 価値

SCDM Part2 5. Fit-for-purpose clinical data strategies (the 5 Vs)



- CDMは、一つのデータポイントの相対的価値を最大化する必要がある。
  - これを達成するにはマネジメントとオーバーサイトが必要である。
- 臨床試験データの価値を最大限に引き出すためには、以下の2点に考慮する。
  - 治験依頼者は将来何が必要になるかを事前に予測する。
  - 重要な事項を網羅したインフォームド・コンセントのテンプレートを用いて、事前に患者の承認を求める。
- 臨床データを新しい目的に2次利用し始めている。
  - Synthetic Armsの作成
  - 監査証跡からの傾向の特定
  - 試験に参加している患者さんのモチベーションを高めるためのデータのフィードバック
  - 機械学習用の教師データの作成
  - リアルワールドデータからのエビデンスの抽出

# 5Vのテクノロジーへの影響

SCDM Part2 5. Fit-for-purpose clinical data strategies (the 5 Vs)



- データの真の価値を引き出すためには、5つのVすべてを考慮した一貫性のある方法でタイムリーにデータを収集し、管理する必要がある。
  - データ戦略は、収集するデータのリスク、複雑さ、および価値に見合ったものである必要がある。
  - データのセキュリティと個人情報の保護は、重要な検討事項であり、試験開始前に対策が必要である。



# 5Vのテクノロジーへの影響

## ～要約～

SCDM Part2 5. Fit-for-purpose clinical data strategies (the 5 Vs)



|                 | 現在のDMの枠組み                                                                                                                         | 望まれるDMの枠組み                                                                                                                          |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|
| 量<br>Volume     | <ul style="list-style-type: none"><li>数10～100/患者・週のデータポイント</li></ul>                                                              | <ul style="list-style-type: none"><li>数1,000～1,000,000/患者・週のデータポイント</li></ul>                                                       |
| 多様性<br>Variety  | <ul style="list-style-type: none"><li>ローカルラボとPKを含むEDCデータ</li><li>IxRS、セントラルラボデータ、eCOAなどの外部データ</li></ul>                           | <ul style="list-style-type: none"><li>RWD、バイオマーカー、ゲノミクス、動画/画像データ、センサー、ウェアラブルなどの、構造化および非構造化データ</li></ul>                             |
| 速度<br>Velocity  | <ul style="list-style-type: none"><li>日次、週次、月次</li><li>患者の来院ごとに入力されたデータ</li></ul>                                                 | <ul style="list-style-type: none"><li>ほぼリアルタイム</li><li>APIがシステム間の相互運用性を提供</li></ul>                                                 |
| 真実性<br>Veracity | <ul style="list-style-type: none"><li>ソースの正確なコピー、ALCOA</li><li>主にSDVとクエリで確認</li><li>完璧（100%エラーなし）</li><li>マニュアル/科学的レビュー</li></ul> | <ul style="list-style-type: none"><li>重要なことに焦点を当てる</li><li>リスクに基づくデータ戦略</li><li>AIによる問題の検出と解決の自動化</li><li>Fit for purpose</li></ul> |
| 価値<br>Value     | <ul style="list-style-type: none"><li>承認申請に重点を置く</li></ul>                                                                        | <ul style="list-style-type: none"><li>データの二次利用の拡大（synthetic arm, 患者エンゲージメント、機械学習用の教師データとしての利用など）</li></ul>                          |

CDMS・EDCに求められる変化

# インテリジェントCDMS

# CDMS、EDCに求められる変化

SCDM Part2 6. Intelligent Clinical Data Management Systems (CDMS)  
SCDM Part2 7. Maximizing the value of traditional EDC



データ量の増加、  
データソースの多様化

大量なデータをどのように扱うか… ??  
多様なデータそれぞれの特徴は何か… ??  
そこに必要なデータはあるのか… ??  
Active data/Passive dataとは何か ??

大量/多様なソースやフォーマットを、  
リアルタイムに処理ができるプラット  
フォームと、  
データの特徴に合った妥当な品質  
保証をするプロセス

データから導きたいものは  
何だろうか… ??  
必要なデータの特定は  
できているか… ??

必要なデータを扱うための機能や手順  
とは… ??  
いろいろなデータをただ組み合わせれば  
よいのか… ??  
データ品質の説明が適切にできるだろうか  
… ??

Real-time/data-driven/  
confident な意思決定の実現

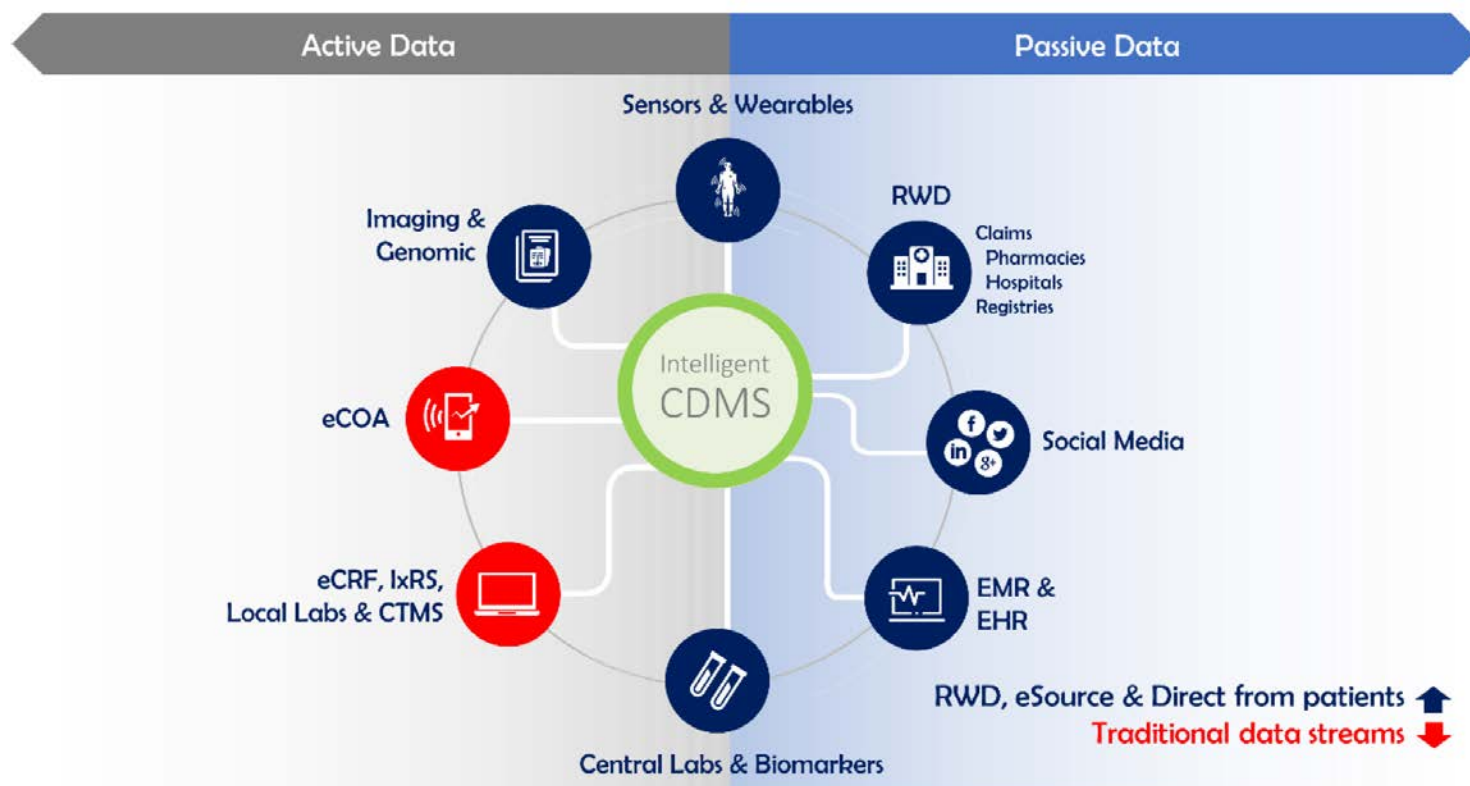
# インテリジェントCDMS：多様なデータを扱うために

SCDM Part2 6. Intelligent Clinical Data Management Systems (CDMS)

SCDM Part2 7. Maximizing the value of traditional EDC



- EDCに流入される大量で複雑なデータを管理することだけではなく、**様々な新規技術やプロセスを検討し、データ管理の戦略をたてる**ことがDM業務として求められている。
- **Active Data（臨床試験データ）**や**Passive Data（医療機関データ、RWD）**の両方を効率的に管理するため、新たなCDMS（インテリジェントCDMS）が必要である。



Part 2, Fig 3. Clinical research technology ecosystem

# Active DataとPassive Data

SCDM Part2 6. Intelligent Clinical Data Management Systems (CDMS)  
SCDM Part2 7. Maximizing the value of traditional EDC



## ➤ Active DataとPassive dataの相違点

|                         | Active Data<br>(≡Prospective)           | Passive Data<br>(≡Retrospective)                          |
|-------------------------|-----------------------------------------|-----------------------------------------------------------|
| データ収集の目的                | 臨床試験用<br>(例：CRF、PRO)                    | 臨床試験以外<br>(例：RWD、EHR)                                     |
| 研究目的にデータを利用することへの被験者の同意 | 事前に同意を取得する                              | 事前or <u>事後</u> に同意を取得する                                   |
| データクリーニング               | 可能<br>(SDV, ロジカルチェックなどで<br>疑義抽出し、修正が可能) | <u>困難</u><br>(データ収集時点ですでに確定している場合が多い。)                    |
| データ種類 (フォーマット)          | 少ない<br>(扱いが容易)                          | <u>多い</u><br>(データ統合のためのマッピング<br>等が必要。標準化された形式であるかどうかも重要。) |

➡ Passive dataを扱うためには、Active dataとは異なる視点で  
機能、プロセス両面を検討する必要がある

# インテリジェントCDMSにて実現したいこと

SCDM Part2 6. Intelligent Clinical Data Management Systems (CDMS)

SCDM Part2 7. Maximizing the value of traditional EDC



インテリジェントCDMSを実現させることで、リアルタイムで、かつデータを根拠とした意思決定が可能になる。

## インテリジェントCDMSに求めること：機能面

- 多数のデータフォーマットに対応/異なるデータ構造のマッピング機能
- 多くのAPIを利用すること
- 紙ベースで収集されたデータにも対応すること
- 関連規制や医療基準に対応すること
- 完全な監査証跡を保持すること
- データの抽出方法の検討が必要（自動化）

医療機関側システム（例：RWD、EHRなど）と各種デバイス（ePRO、モバイルアプリ、ウェアラブルなど）と連携し、多様なデータソースの利用が可能になる。

## インテリジェントCDMSに求めること：プロセス面

- データクリーニングや不整合箇所の管理プロセスの再考
- データインテグリティの検討（データの完全性の示し方や除外データの根拠）

# EDCシステムにもまだ価値最大化の余地がある

SCDM Part2 6. Intelligent Clinical Data Management Systems (CDMS)

SCDM Part2 7. Maximizing the value of traditional EDC



## ☆従来、データ収集の中心であったEDCシステムはこれからどうあるべきか？

### 現状

- データの多様化に伴いEDCへのデータの集中性は相対的に低下している
- 「転記」に起因する各課題（SDV、過剰なチェック）は解決されていない
- 構築時、仕様変更時、運用時など各段階の負荷は依然大きい

### 今後のEDC

- Active dataのひとつを収集するツールとして引き続き利用される
- リアルタイム性を追求する複雑な臨床試験デザインに追従するためには構築、運用プロセス改善が必須である
  - 標準化、ライブラリ化やリスクに基づくプロセスの検討や合理化

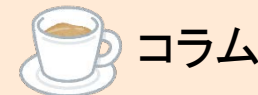
### 新たな可能性の模索

- EDC上で直接eSourceデータとして収集する手法
- 医療機関側の電子カルテとEDCの連携

これからもEDCシステムは、データ収集ツールとしての基本に立ち返り、利用目的の変化に適応させることで価値の最大化を目指すことができる



# 多様性あるデータを活用すべき各種業界標準



SCDM Part2 6. Intelligent Clinical Data Management Systems (CDMS)

SCDM Part2 7. Maximizing the value of traditional EDC



## ☆CDMS、EDCやその他のシステムで利用可能な業界標準

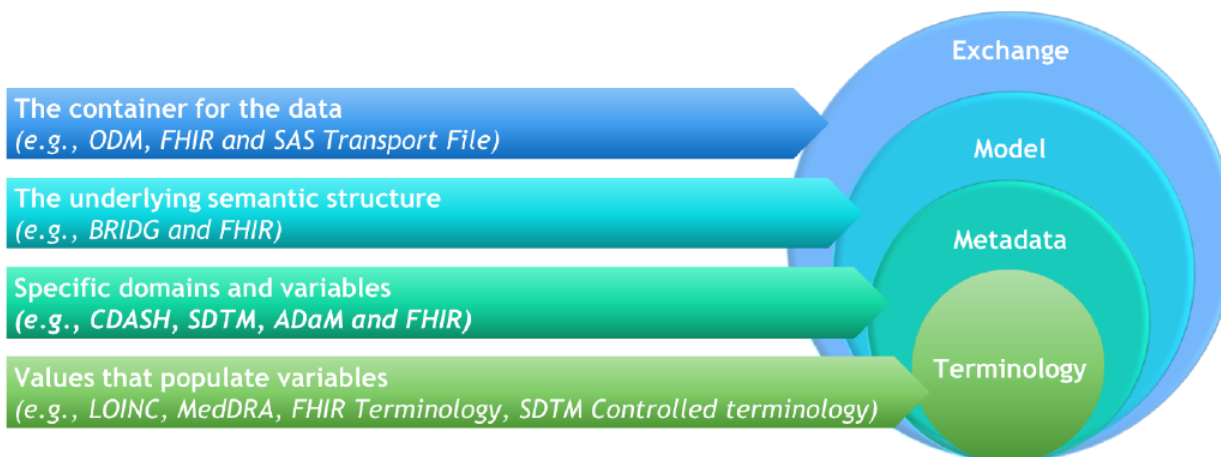
- 目的に応じて様々な標準が利用、検討されている
- 標準は4つの階層に分類でき、組み合わせて利用することで相乗的な効果や、システム間連携などの拡張性が期待できる

**Data models:** Conceptual, logical or physical semantic descriptions of objects and their relationships

**Metadata standards:** Representations of individual data concepts useful for organizing data in a database or data exchange file

**Terminology standards:** Representations of individual data values

**Exchange standards:** Schemas or file types for exchanging data between systems; the container for transporting data



Part 2, Fig 4. Layers of clinical data standards



# 新しい臨床試験手法

# 複雑化する臨床試験の手法

SCDM Part2 8. Impact of new clinical research approaches on technology



製薬協

- **Adaptive design** : 治験実施計画書で事前に規定されたルールに従って、臨床試験の中間段階までで得られたデータに基づき、それ以降の試験デザインの変更を決定する試験デザイン。
- **Master Protocol** : アダプティブデザインやプラットフォーム、バスケット、アンブレラデザインのコンセプトを組み合わせ拡張したもの。適切な治療を適切な疾患タイプを持つ患者に割り付け、良好な結果を最大化することを目的としている。

複雑化したデザインに対し、Clinical Data Scientistには新しいスキルが求められる！



# 複雑化した臨床試験におけるAdaptation

SCDM Part2 8. Impact of new clinical research approaches on technology



- Adaptive design等の複雑化した試験デザインでは、試験途中でシステムの変更(=Adaptation)が発生することが前提になる。
- Clinical Data Scientistは、治験実施計画書に事前に規定されたルールに基づいて、以下を考慮しておかなければならない：
  - どのデータを元にAdaptationが行われるか？（治療反応率、有害事象（毒性）発現比率、バイオマーカーなど）
  - 何がAdaptationされるのか？（割付比率、組み入れ中止、試験中止など）
  - Adaptationはどのように決まるのか？（中間解析、DMC、アルゴリズムなど）
  - どのような頻度でAdaptationを行うのか？（二段階デザイン、事後確率の再計算など）

# システムに求められる要件とは？

SCDM Part2 8. Impact of new clinical research approaches on technology



- **Adaptive design等の複雑化した試験デザインでは、運用途中の変更に対応可能な柔軟性を備えたシステムとプロセスが必要である。**
  - 例えば、Adaptive designを適用した用量反応試験では、随時得られる検査値から用量反応曲線を作成し、そこから次の患者に使われる用量を選択する。このとき、IxRS\*とEDC、安全性データベースを連動させることができれば、プロセスをスムーズに回すことができる。
    - \* : IxRS : 最初に入力された被験者ID、治験実施医療機関を特定し、無作為化、層別化、投与の割付けを実施する。

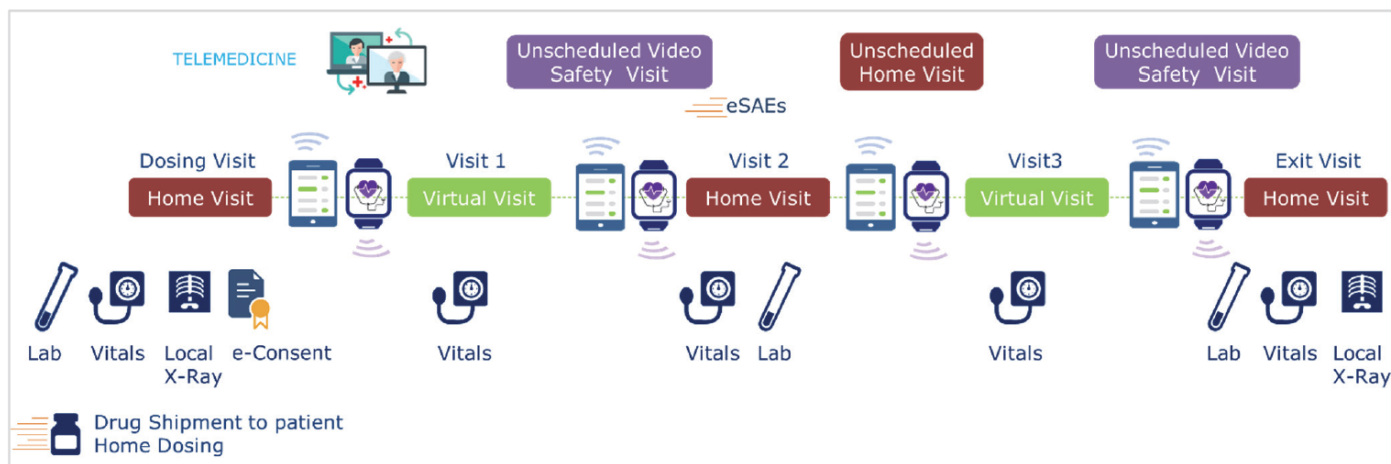
- **Clinical Data Scientistに求められるものとは：**
  - 統合されたシステム上で適切なAdaptationが行われるよう、より多くの状況を想定した大規模UAT、複数のシステム間のUATを計画・実行できなければならない。
  - 試験途中に発生する全てのデータ・メタデータの変更が、下流プロセス（SDTMデータセット、データフロー、データレビューチェック等）へ及ぼす影響を正しく評価できなければならない。
  - 変更の影響範囲については体系的に評価し、文書化しなければならない。

# **DECENTRALIZED CLINICAL TRIAL (DCT)**

# DCT\*とは

## SCDM Part2 8.3. Decentralized Clinical Trial (DCT)

- 臨床試験を物理的に施設から切り離して実施すること
- これまでのEDCやeCOA、リモートモニタリングに加えて、ウェアラブルデバイスや遠隔医療ツールなどの臨床試験にとっては新しい技術が必要となる
- DCTの例（下図）



初版, Fig 4. Example of design of a fully Patient Centric / Decentralized Clinical Trials

\*DCTに関する詳細は製薬協 臨床評価部会2019年度タスクフォース3 報告書「医療機関への来院に依存しない臨床試験手法の導入及び活用に向けた検討」を参照

<http://www.jpma.or.jp/medicine/shinyaku/tiken/allotment/dct.html>

# DCTにおいて考えるべきこと

SCDM Part2 8.3. Decentralized Clinical Trial (DCT)



## ➤ DCTにおけるCDMの課題⇌CDSになるために考えるべきこと

### ■ 収集されるデータ量が膨大になる可能性がある

- 収集する前にデータの使用目的を明確にする必要がある
- 1例1例のデータレビューは非効率的、トレンドをみる必要がある
- クエリにも工夫が必要になる

例えば、デバイスの歩数のデータにクエリを上げることは難しい。しかし、毎日の歩数データにあまりにもバラツキがあるのであれば、「規定通りにデバイスを装着していますか？」といった質問をすることは可能である。

### ■ デバイスにて収集したデータから“価値”を得るためには二つの観点が必要

- 患者がデバイスを正しく装着していることが重要
- デバイスから収集したデータを定期的にモニターし分析すること

### ■ CDSはどんな方法で集められたデータについても、取り扱う準備をしておかなくてはならない

# DCTと従来のアプローチ

SCDM Part2 8.3. Decentralized Clinical Trial (DCT)



## ■ 従来のアプローチとDCTアプローチの比較(下表)

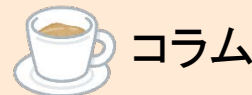
| 従来のアプローチ             |        | DCTアプローチ                               |
|----------------------|--------|----------------------------------------|
| 観察/測定を介して            | How?   | デバイス経由および/または患者エンゲージメント*のためのツール経由      |
| 試験実施施設（病院等）          | Where? | 患者が移動/滞在している場所どこでも                     |
| 患者と研究担当者が一緒に         | Who?   | 患者自身で<br>必要に応じて医療関係者が訪問で或いはリモートでサポートする |
| 研究固有のデータのみ           | What?  | 研究固有および/または個人の健康関連データ                  |
| 試験実施計画書ごとに事前に指定された間隔 | When?  | 事前に指定された間隔だけでなく、必要な時のみ／継続的に実施することも可能   |
| 研究目的のみ               | Why?   | 研究目的に限らず、患者のケア目的にもなりうる                 |

\*患者エンゲージメント：臨床試験に参加する患者の臨床試験への積極的関与の促進や支援を目的とした活動

- DCTに利用される直接的な患者データ収集システムは、データを被験者と共有することで被験者のモチベーションを高め、被験者の安全性・コンプライアンスを向上させ、さらには被験者の試験参加維持に寄与する



# 新しいeSourceテクノロジーについて ～Sensors and Wearables～



SCDM Part1

5.3 Evolution of technologies



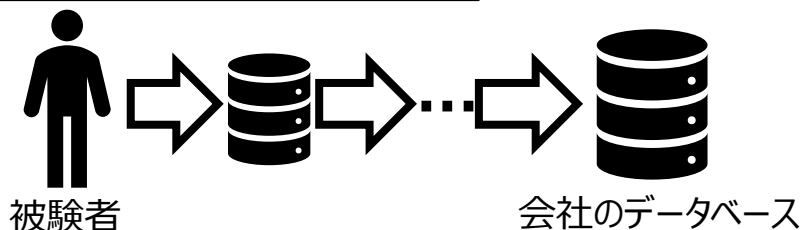
ウェアラブルデバイスやセンサーの利用は、多くの企業が興味を持っている。

## 活用例

1. DCTでのデータ取得
2. 継続的な安全性情報の取得および監視
3. これまでとは異なる評価指標（客観的指標）データの収集が可能に
4. データとテクノロジー活用による患者負担の軽減

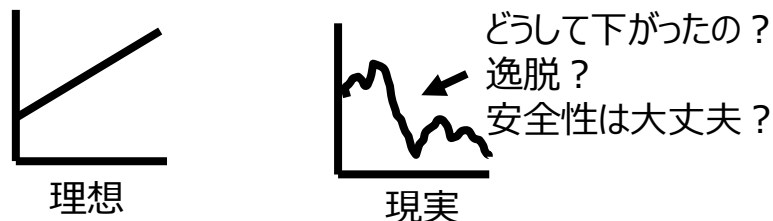
ただし、以下のような**注意点**がある。

## ①データフローの把握



データの**完全性を担保するために**、どのようなネットワークを通じてデータが入手されるか、**フローを把握する必要**がある。

## ②すべてがクリーンなデータではない



デバイスの着脱や、被験者以外の着用等で、逸脱したデータを収集することがある。**逸脱したデータは、被験者の安全性に影響がないか、逸脱した着用なのかを検出できるようにする必要があります。**

# 自動化 (AUTOMATION)

# 自動化技術の活用でCDMの在り方が変わる

SCDM Part2 9. Automations



自動化技術が成熟し、DM業務での活用も可能に！

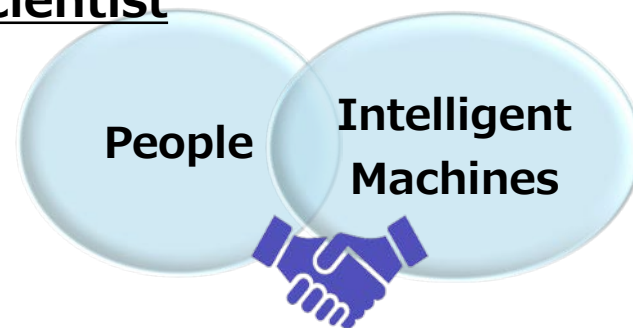
“人”と“Intelligent Machines”が協働する環境を考慮し、  
CDMからCDSへのロードマップを策定していく必要がある

<<Road map to CDS>>

Clinical Data Manager

- ✓ People
- ✓ Technology
- ✓ Process
- ✓ Regulation
- ✓ Standards

Clinical Data Scientist



# 自動化を推進するための新技術

SCDM Part2 9. Automations

## ➤ DM業務の自動化に活用可能な技術

各技術を個別 or 組み合わせることにより自動化を実現

| Robotic Process Automation (RPA)                     | Intelligent Process Automation (IPA)     | Natural Language Processing (NLP) | Natural Language Generation (NLG) | Machine Learning (ML)                | Artificial Intelligence (AI)                           |
|------------------------------------------------------|------------------------------------------|-----------------------------------|-----------------------------------|--------------------------------------|--------------------------------------------------------|
| Basic Automation of repetitive and predictable tasks | Advanced Automation of non-routine tasks | Language and text analysis        | Generate text and voice           | Pattern recognition                  | Human like behaviors<br>Self optimization and Learning |
| No / simple decision making                          | Some decision making and problem solving | Knowledge gathering               | Knowledge sharing                 | Prescriptive and predictive analysis | Deductive analysis                                     |

Part 2, Fig 5. Technologies powering automations

➤ 次ページから、各技術の説明とDM業務への活用事例を紹介する

# Robotic Process Automation (RPA)

SCDM Part2 9. Automations



**RPAは、「明確に定義された人間の動作」を、ソフトウェアを介して単純に模倣する**

## 【特徴】

- 繰り返す単純作業の自動化に効果的
- 成熟した技術
  - 実装も簡単になり、コストパフォーマンスに優れる
- 既存システムの変更が不要
  - ただし、対象システム変更時にはRPAの調整が必要

## ■ RPAが有効なDM業務

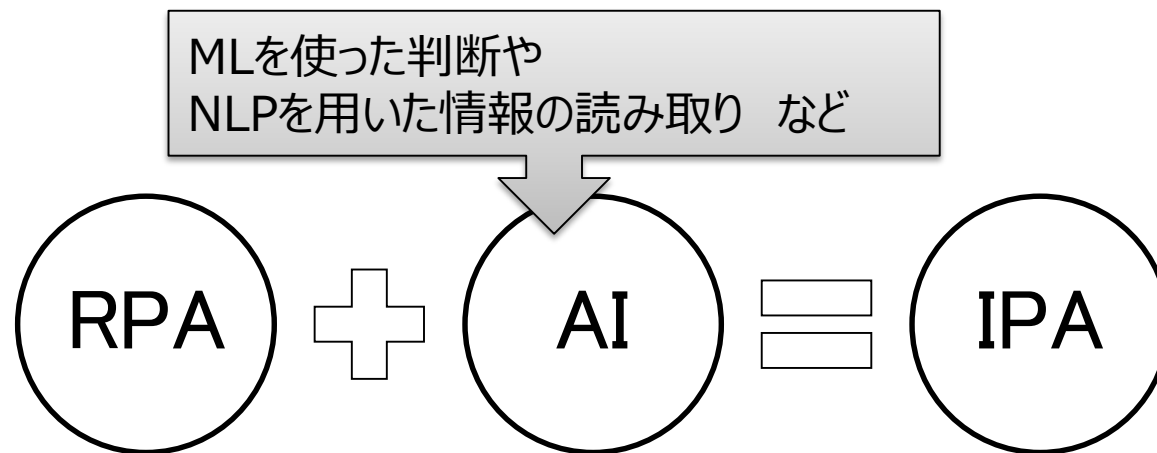
- レポートニングシステムの定期自動実行
- エディットチェック仕様の生成
- 外部データの自動取り込み
- データクリーニングの進捗管理
- データクリーニングの自動化
- 品質チェックの自動化
- SAEリコンシリエーション

# Intelligent Process Automation (IPA)

SCDM Part2 9. Automations



IPAは、RPA単独ではできなかった、判断やデータ化されていない情報の読み取り等を、MLやNLPなどのAIと組み合わせて実現し、より高度な業務を自動化する



## ■ IPAが有効なDM業務

- 有害事象と併用薬の関係など、複雑な矛盾点を検出

# Natural Language Processing (NLP)

## Natural Language Generation (NLG)

SCDM Part2 9. Automations



### NLP・NLGは、人の言語を処理する技術

NLP:文書、音声などの自然言語を解釈し機械が処理できる情報を取り出す

NLG:機械の演算結果から音声や文書を生成する

#### 【特徴】

- 約70年前の技術、機械学習技術によりここ10年で大きく進歩

- 例：Siri/Apple、Echo/Amazon

- NLPが有効なDM業務

- データクリーニング

- ✓ 過去の不一致データとクエリ履歴を学習し、クエリを自動生成

- 潜在的なAEの特定

- ✓ EDC、eCOA、その他のシステムに記録されたデータ、コメントから潜在的なAEを特定

- 非構造化データから構造化データベースへのロード

- ✓ レガシー臨床試験報告書、又はモニタリング訪問報告書等から必要な情報をデータベース化

# Machine Learning (ML)

SCDM Part1 5.3. Evolution of technologies

**明示的プログラムなしに、データから予測あるいは決定を下すことを学習させる技術**

## 【特徴】

- 教師あり学習：正解付きデータ（教師データ）を使用し、正解を学習

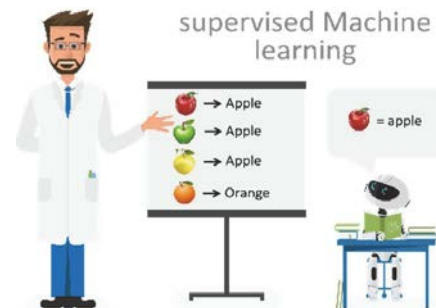
- 有効なDM業務

- ✓ 単純なクエリ発行(範囲外データなど)
- ✓ 但し、前回発行されたクエリの対応状況を加味してからクエリを発行させる場合は、多くのデータが必要

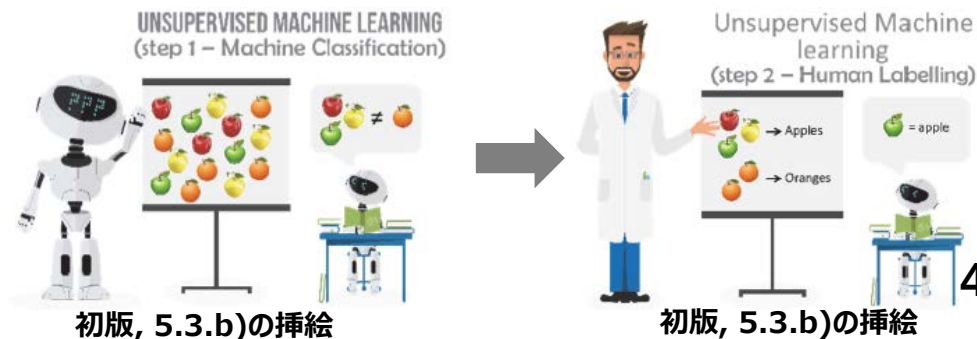
- 教師なし学習：正解のないデータを使用し、データの類似性を判断させる

- 有効なDM業務

- ✓ 画像診断などに活用



初版, 5.3.b)の挿絵



初版, 5.3.b)の挿絵

初版, 5.3.b)の挿絵



# Artificial Intelligence (AI)

SCDM Part2 9. Automations / SCDM Part1 5.3. Evolution of technologies



**AIは、学習や推論を通じて人間の認知行動を模倣する**

## 《AIベースアプリケーションの例》

### Chatbots = 「知的な仮想アシスタント」

人と機械の間で“人間的なやり取り”を可能にし、  
実際のサポート担当者業務の一部に置き換わる

#### 【特徴】

- 人間的な相互作用を提供
- 音声認識を追加することでより利便性 ↑
  - Siri (Apple) 、Cortana (Microsoft)
- Chatbotsが有効なDM業務
  - 進捗管理のサポート
    - ✓ 「open queryが5件以上の施設は？」
    - ✓ 「施設がクエリ回答したら、私に連絡して」



初版, 5.3.b)の挿絵

# データレビューの自動化へのロードマップ°

SCDM Part2 9. Automations



製薬協

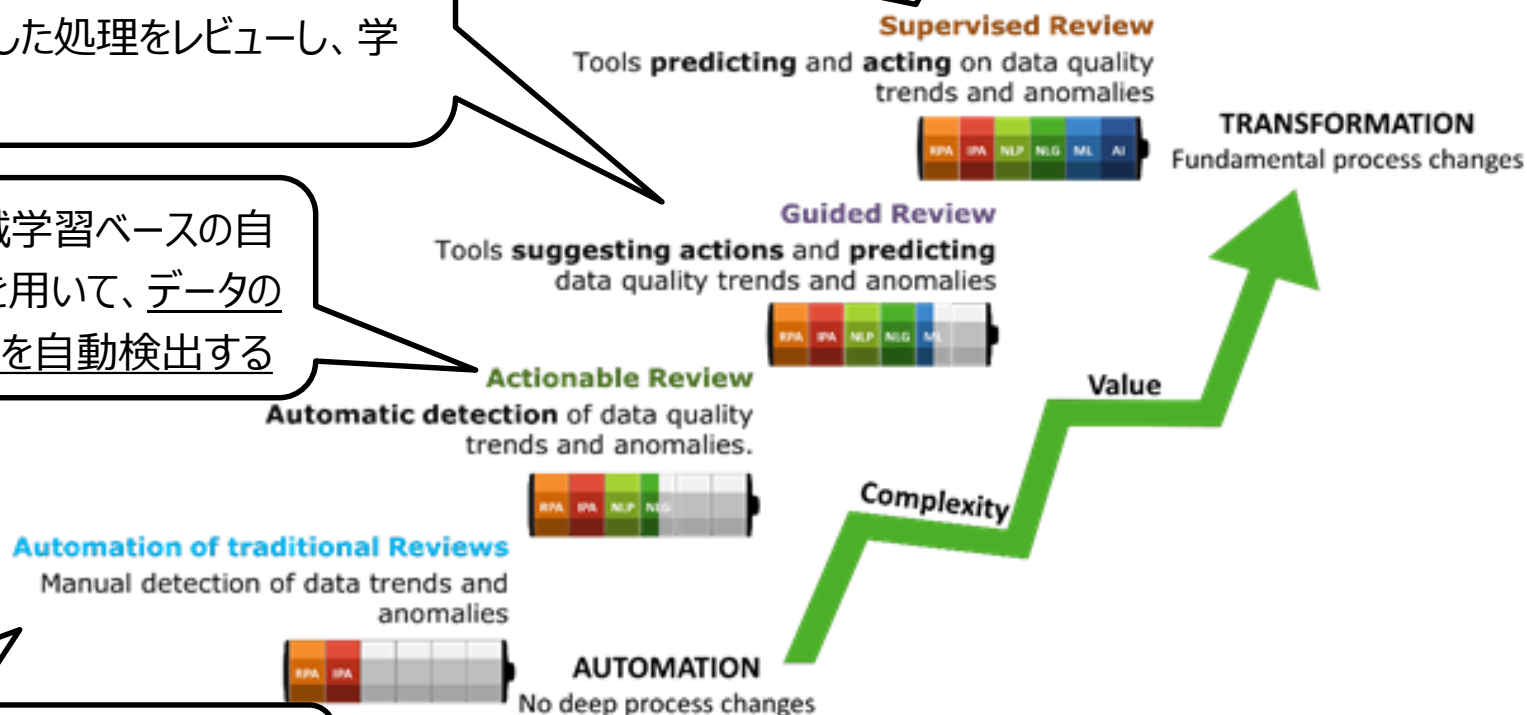
データレビューの自動化は、以下の4段階を経て進化すると考えられる

第3段階：自動検出した傾向および異常に対する処置を自動化する。CDSは、システムが提案した処理をレビューし、学習させる。

第2段階：機械学習ベースの自動化等の技術を用いて、データの傾向および異常を自動検出する

第1段階：単純な作業および反復作業のみをRPAで自動化する

第4段階：システムが自動的にレビューを行う。システムが判断出来ない場合のみCDSにエスカレーションするなど、CDSはシステムの監督、トレーニングの支援、難題の解決を行う。

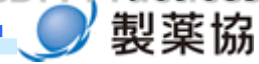


Part 2, Fig 6. Evolution of data reviews powered by AI technologies

# 新たな規制動向

# データ品質(Data Quality)とは？ - 規制による定義

SCDM Part1 5.2 Impact of regulations on Data Quality and CDM Practices



- データの品質を保証するには、データの完全性（integrity）の主要な基準を満たすだけでは十分ではない
- 一般的なガイダンスとして、データの品質と完全性の定義は
  - データの完全性：データが正しい（適切な）方法で管理されていること
  - データの品質：データが信用及び信頼できる（credible and reliable）こと
- データが信頼できる（reliable）とは？
  - たとえば、下記ポイントについてデータレビューを行うことでreliableかそうでないかが分かる
    - データの異常値、不自然にデータのばらつきが欠如
    - 施設内および施設間でのデータ範囲、整合性、ばらつきなどのデータの傾向
    - 特定の施設または施設間でのデータ収集および報告における系統的または重大なエラー
    - 潜在的に行われているデータ操作やそれに伴うデータの完全性の問題

データ収集環境の変化の中で、規制当局の要求も変わってきている  
→ 従来のデータ品質保証のアプローチは通用しなくなってきている

# 技術に対する新たな規制上の期待

## SCDM Part2 10. Emerging regulatory expectations for technologies



- eSourceが標準となった世界では、修正不可能な原データをそのまま入手することになる。そのため、データエラーの特定よりも、治験実施計画書や規制の遵守に関する**潜在的課題（potential issue）の特定**が重要となる。

### ■ データフローの明確化

- Risk Based Approach (RBA) ← システムの多様化、EDC外データ増加
  - ✓ 重要なリスクが潜むポイントを特定する<sup>1</sup>
  - ✓ 可能な限りリスクを軽減する<sup>1</sup>
- DMはあらゆるステークホルダーがデータフローとデータに対するリスクを理解できるように努めるべきである<sup>1</sup>

### ■ 監査証跡のレビュー

- RBA → “重要な”コンプライアンス違反の検出
- 日常的な監査証跡レビューにより、データの品質や完全性に関する潜在的課題を特定できる
  - 権限のない者によるデータ入力
  - 不当に短い期間でのePROのデータ入力

# Inspection Readiness

SCDM Part2 10. Emerging regulatory expectations for technologies



- システムが多様化し、EDC外のデータ量が増加しているため、調査官は**システムとデータフロー**に注目している

- 規制当局による調査に関する留意事項



- **データフローの明確化**

- データ収集から解析へデータを提供するまでのデータフロー図とそれを裏付ける説明をDM計画書に記載することが必要となる
    - データ変換や導出についても記載することが理想とされる

- **システムのアクセス権**

- 調査員へのアクセス権付与について、Role、トレーニング方法、付与方法、無効化の時期を含めた手順を用意する

- **施設のデータ保管**

- セキュリティ、アクセス容易さ、施設のアクセス管理、長期保管における見読性といった点において、実施医療機関で課題となることがないか確認しておく

- **実地調査に向けた準備**

- 全てのデータソース（EDC, eCOA等）においてデータ及び監査証跡が容易に施設から閲覧可能となるクラウドベースソリューションの活用が期待される

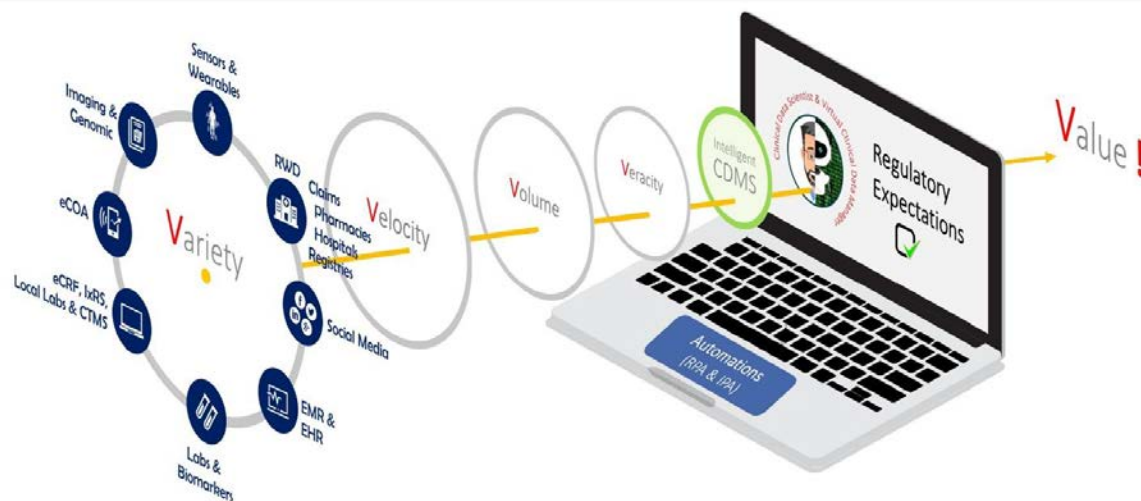
- 最善策は、**業務のあらゆる側面でQbDを促進すること**

さいごに

テクノロジーの発達により、そしてCOVID-19の影響により、臨床試験の環境は“新常态”へ大きく変化しています。

Clinical Data Manager はこれまで試験結果の信頼性を確保する上で重要な役割を果たしてきましたし、これからも期待されています。

その期待に応えるために、“新常态”に対応できる新しい技術／知識／見識を身につけ進化していきましょう！



Part 2, Fig 7. The 5Vs data journey from collection to value generation



- 本資料の作成にあたり、Society for Clinical Data Management (SCDM) には、日本製薬工業協会の趣旨に賛同いただき、SCDMのReflection paper “The Evolution of Clinical Data Management to Clinical Data Science” の Part-1 (4 June 2019) とPart-2 (5 March 2020) の内容を紹介すること、並びにReflection papers中の図表を使用することを、許可いただき深く感謝いたします。
- また、SCDM日本支部の共同代表である東北大学 山口拓洋先生、東京大学 宮路天平先生には、活動に対する助言やSCDM本部との連絡の労をお取りいただき、感謝いたします。
- なお、本資料で使用しているReflection papersから引用した図表は、著作権が保護されており、SCDM からの許諾なしにいかなる手段によっても複製、再発行、配布は認められないことにご留意ください。利用は個人の利用に限られ、団体等への譲渡も認められていません。

# 日本製薬工業協会 医薬品評価委員会

## データサイエンス部会 2020年度 タスクフォース2



### 執筆者

|        |                     |
|--------|---------------------|
| 進藤 武之  | Meiji Seikaファルマ株式会社 |
| 永島 将子  | アステラス製薬株式会社         |
| 田中 拓海  | エーザイ株式会社            |
| 宮本 奈緒子 | ゼリア新薬工業株式会社         |
| 小田 梓   | ノバルティス ファーマ株式会社     |
| 永田 優一  | ヤンセンファーマ株式会社        |
| 山田 幸恵  | 杏林製薬株式会社            |
| 南 賀子   | 塩野義製薬株式会社           |
| 松山 友映  | 協和キリン株式会社           |
| 藤澤 健司  | 興和株式会社              |
| 伊藤 義典  | 大塚製薬株式会社            |
| 西野 真紀  | 大日本住友製薬株式会社         |
| 大多和 修平 | 中外製薬株式会社            |
| 田中 久貴  | 鳥居薬品株式会社            |
| 山本 道宏  | 日本ベーリンガーインゲルハイム株式会社 |
| 火原 光裕  | 日本新薬株式会社            |

### タスクフォース推進委員

|       |            |
|-------|------------|
| 蔵多 雅子 | 中外製薬株式会社   |
| 富金原 悟 | 小野薬品工業株式会社 |
| 加藤 智子 | サノフィ株式会社   |