

# データマネジメントの革新的変化（フルバージョン）

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  
2021年4月

- はじめに
- 臨床試験環境の大きな変化
  - 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
- CDMに求められる新たなスキルセットとコンピテンシー
  - CDMの現状とCDMに求められる役割
  - CDSの役割とスキルセット
    - リスクに基づくDM業務のアプローチ
    - データレビューの進化
    - 高度なCDMのコンピテンシー
  - CDSへの進化がビジネスモデルに及ぼす影響
  - 役割の進化の影響
- さいごに

# はじめに

[目次に戻る](#)

## ➤ 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とコンピテンシー、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
  - Part 1 : Drivers (4 June 2019)<sup>1</sup>
  - Part 2 : The technology enablers (5 March 2020)<sup>2</sup>
  - Part 3 : The evolution of the CDM role (31 August 2020)<sup>3</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://scdm.org/wp-content/uploads/2020/12/SCDM-Reflection-Paper-Evolution-to-Clinical-Data-Science.pdf>
2. SCDM, March 2020, The Evolution of Clinical Data Management to Clinical Data Science (Part 2: The technology enablers) – A Reflection Paper on how technology will enable the evolution of Clinical Data Management to Clinical Data Science. Available at <https://scdm.org/wp-content/uploads/2020/03/SCDM-Reflection-Paper-Evolution-to-Clinical-to-Data-Science-Part-2.pdf>
3. SCDM, August 2020, The Evolution of Clinical Data Management to Clinical Data Science (Part 3: The evolution of the CDM role) – A Reflection Paper on the evolution of CDM skillsets and competencies. Available at <https://scdm.org/wp-content/uploads/2020/08/SCDM-Reflection-Paper-CDM-Role-Evolution-Part-3-.pdf>



- 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（Part 1）、2020年3月（Part 2）、2020年8月（Part 3）にリリースされた
- COVID-19の影響でDMの変革が待たなしの状況下で、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

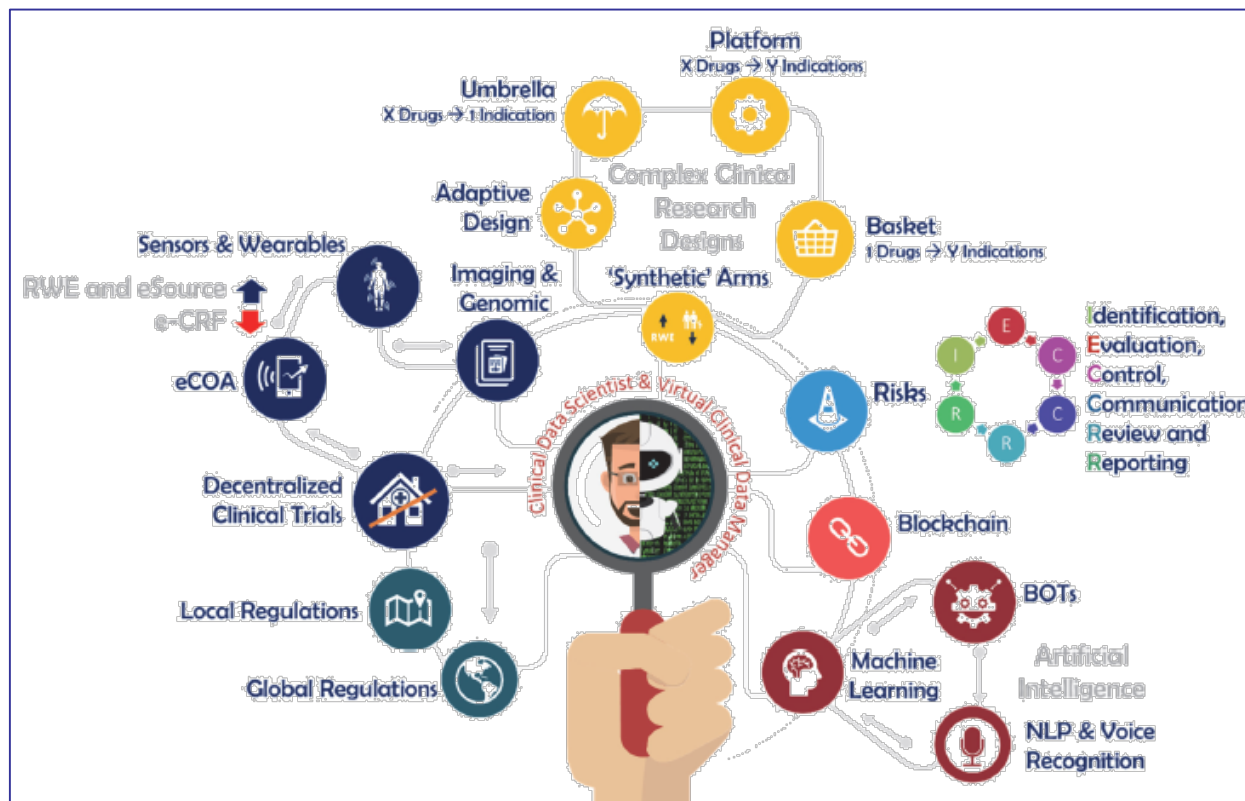
## ➤ なぜスライド形式？

- 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？

- 2021年度以降のDS部会TF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



Part 1, 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、RWD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）との境目が益々小さくなっており、大規模・多種多様で変化に富むデータの処理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

**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、基本的な適性は変わらない。**
  - 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じ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（データインテグリティ、データクオリティ、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**

# 臨床試験環境の大きな変化

[目次に戻る](#)

## 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

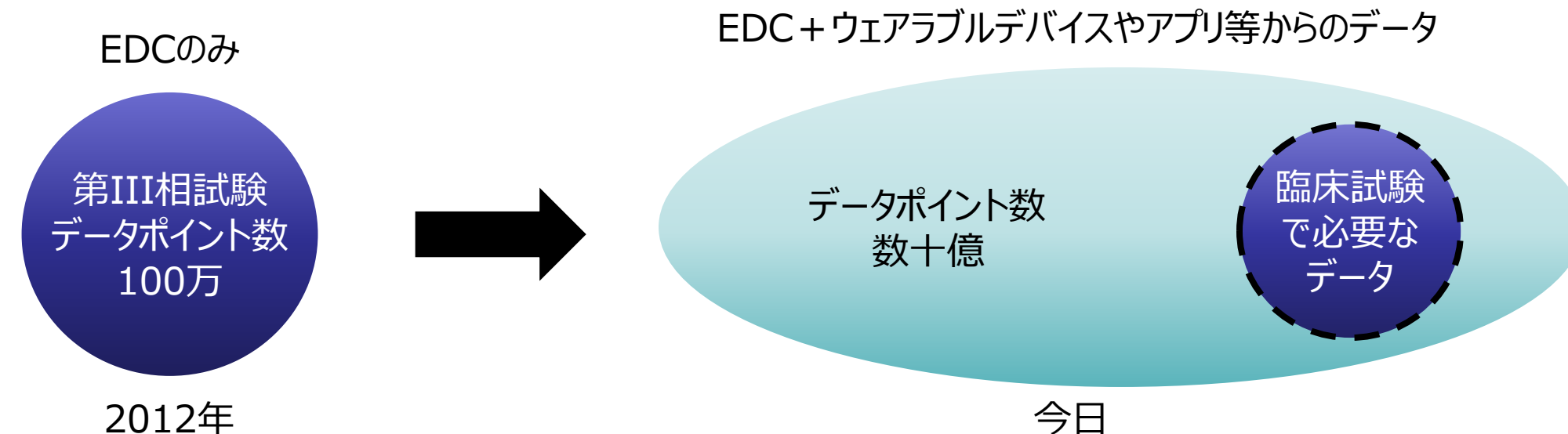
# 目的に応じた臨床データ戦略 (5つのV)

[目次に戻る](#)

# 目的に応じた臨床データ戦略 (5つのV)

## ①Volume: 量

SCDM Part 2 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<sup>※1</sup>：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

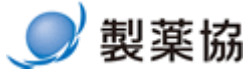
※1 クオリティ・バイ・デザイン(QbD)、品質リスクマネジメント(QRM)及び医薬品品質システム(PQS)に関連する用語集 <https://www.pmda.go.jp/files/000222953.pdf>



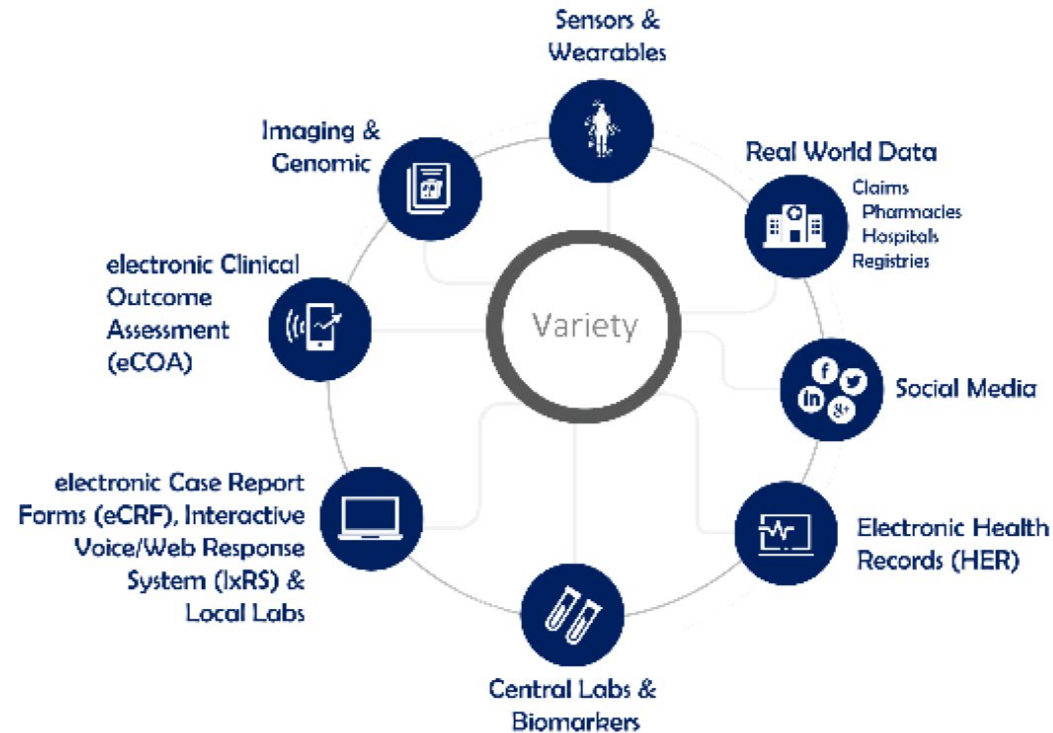
# 目的に応じた臨床データ戦略 (5つのV)

## ②Variety: 多様性

SCDM Part 2 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り、従来型のEDC中心の考え方を見つめ直すことになる。



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、データがどこでどのように生まれ利用されるのか、論理面と物理面の両面から考える必要がある。

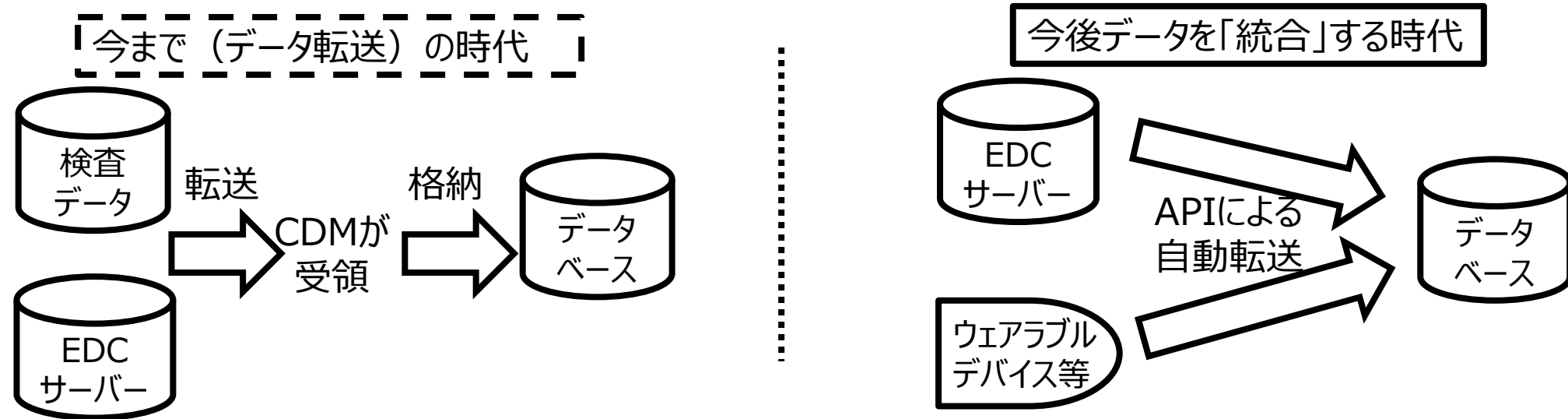
# 目的に応じた臨床データ戦略 (5つのV)

## ③Velocity: 速度

SCDM Part 2 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。

### ④ Veracity: 真実性

➤ 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ることで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。

■ e-SourceやRWDのような修正ができないデータにも、このアプローチが必要になる

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

➤ 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。

- CDMは、臨床試験データを本来の目的に加えて、新たな目的（2次利用）も加味し、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。
- 臨床データの2次利用の例
  - Synthetic Arms（合成対照群）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
- 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。
  - 本来の目的でのデータ利用以外に、どのような目的でデータを利活用でき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・コンセントで事前に患者の承認を求める。

- 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～

	現在のDM業務のフレームワーク	望まれるDM業務のフレームワーク
<b>量</b> 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>
<b>多様性</b> 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>
<b>速度</b> 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>
<b>真実性</b> 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>
<b>価値</b> 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

[目次に戻る](#)

# CDMS、EDCに求められる変化

SCDM Part 2 6. Intelligent Clinical Data Management Systems (CDMS)  
SCDM Part 2 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…??

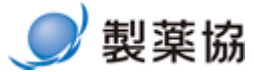
必要なデータを扱うための機能や手順  
とは…??  
いろいろなデータをただ組み合わせれば  
よいのか…??  
データ品質の説明が適切にできるだろ  
うか…??

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

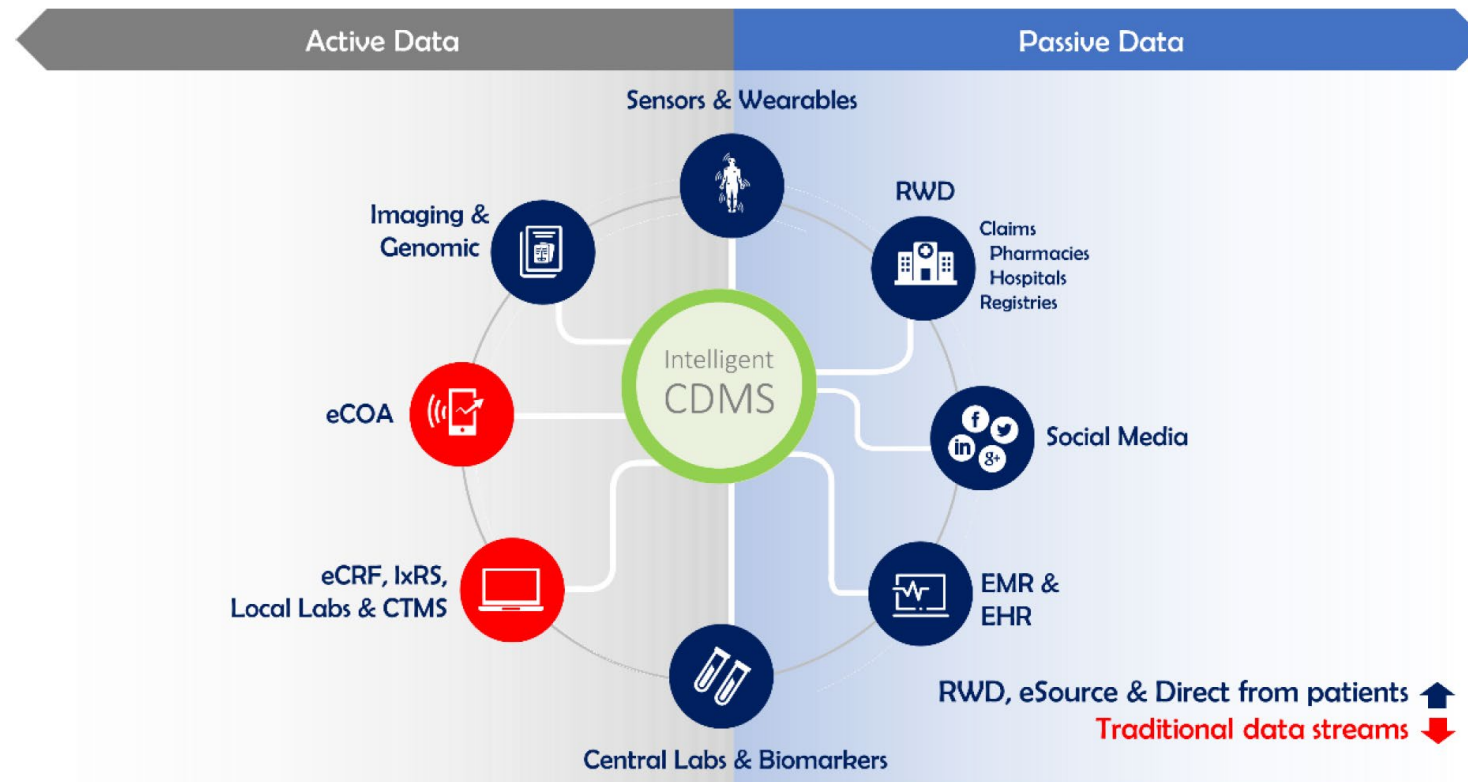


# インテリジェントCDMS：多様なデータを扱うために

SCDM Part 2 6. Intelligent Clinical Data Management Systems (CDMS)  
SCDM Part 2 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 Part 2 6. Intelligent Clinical Data Management Systems (CDMS)  
SCDM Part 2 7. Maximizing the value of traditional EDC



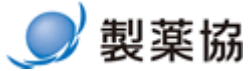
## ➤ Active DataとPassive dataの相違点

	Active data (≒Prospective)	Passive data (≒Retrospective)
データ収集の目的	臨床試験用（例：CRF、PRO）	臨床試験 <u>以外</u> （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（SDV、ロジカルチェックなどで疑義抽出し、修正が可能）	<u>困難</u> （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（扱いが容易）	<u>多い</u> （データ統合のためのマッピング等が必要。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 Part 2 6. Intelligent Clinical Data Management Systems (CDMS)  
SCDM Part 2 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）

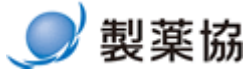
医療機関側システム（例：RWD、EHRなど）と各種デバイス（ePRO、モバイルアプリ、ウェアラブルなど）と連携し、多様なデータソースの利用が可能になる。

## 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(データインテグリティ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 Part 2 6. Intelligent Clinical Data Management Systems (CDMS)  
SCDM Part 2 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

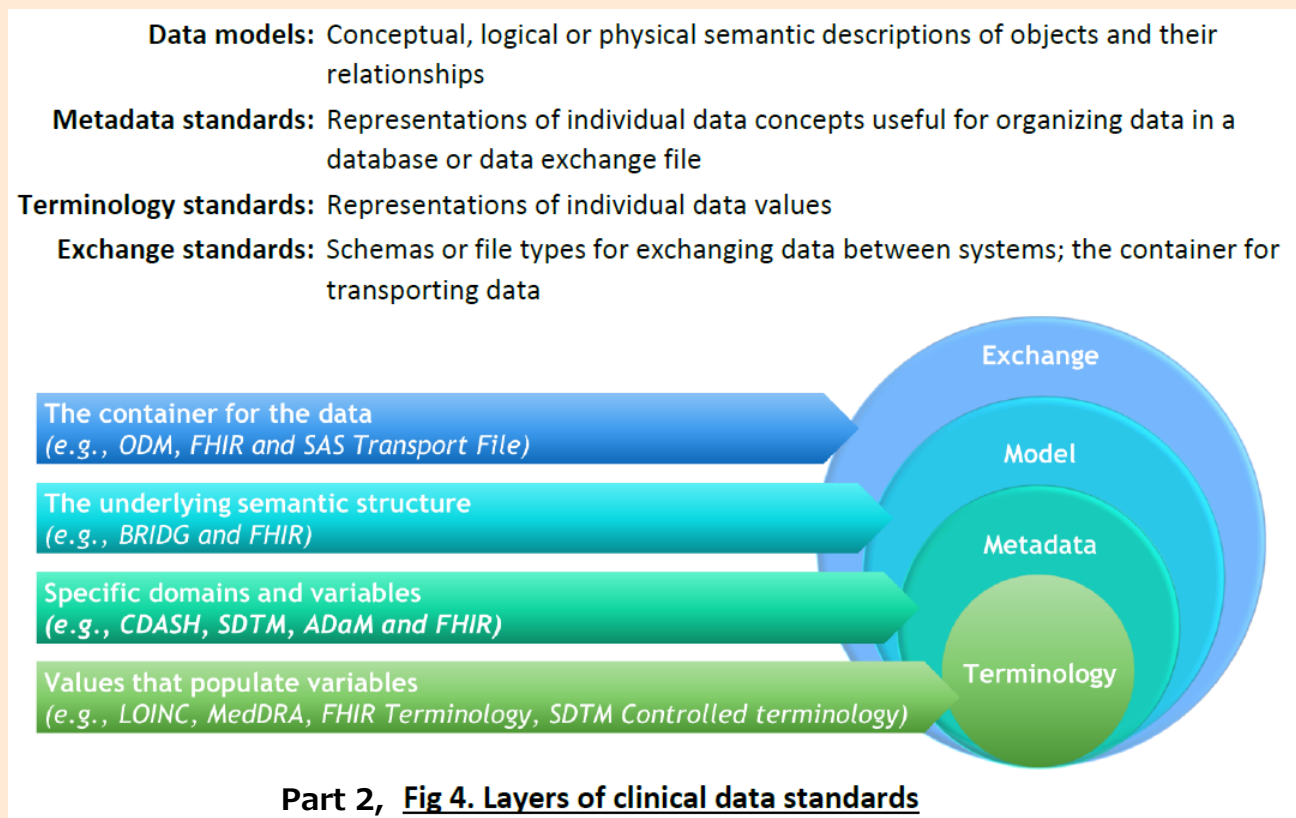
### 新たな可能性の模索

- 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應させることで価値の最大化を目指すことができる。

## ☆ 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。

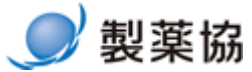


# 新しい臨床試験手法

[目次に戻る](#)

# 複雑化する臨床試験の手法

SCDM Part 2 8. Impact of new clinical research approaches on technology



- **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** : Adaptive design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 Part 2 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 Part 2 8. Impact of new clinical research approaches on technology



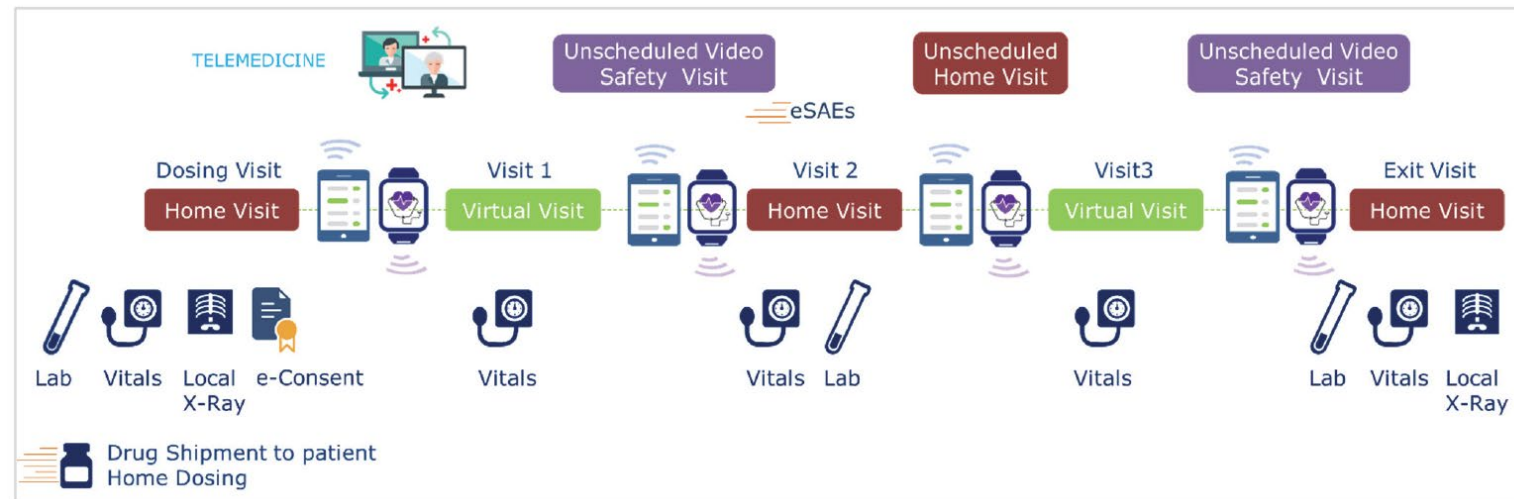
- Adaptive design等の複雑化した試験デザインでは、運用途中の変更に対応可能な柔軟性を備えたシステムとプロセスが必要である。
  - 例えば、Adaptive designを適用した用量反応試験では、随時得られる検査値から用量反応曲線を作成し、そこから次の患者に使われる用量を選択する。このとき、IxRS\*とEDC、安全性データベースを連動させることができれば、プロセスをスムーズに回すことができる。
    - \* : IxRS : 最初に入力された被験者ID、治験実施医療機関を特定し、無作為化、層別化、投与の割付けを実施する。

- 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)

[目次に戻る](#)

- 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）



Part 1, 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

## ■ 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 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～

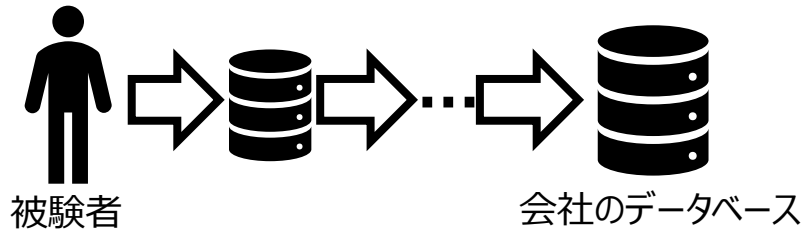
ウェアラブルデバイスやセンサーの利用は、多くの企業が興味を持っている。

## 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

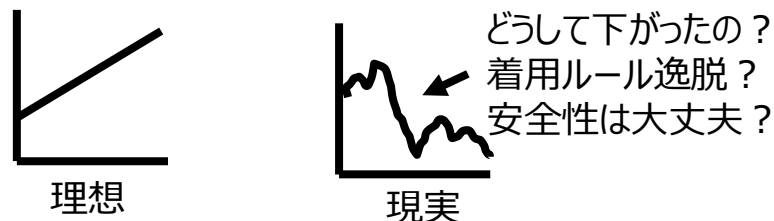
ただし、以下のような**注意点**がある。

### ①データフローの把握



**データインテグリティを担保するために**、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ります**。

### 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、不自然な変動を示すデータを収集することがある。**不自然な変動を示すデータは、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)

[目次に戻る](#)

自動化技術が成熟し、DM業務での活用も可能に！

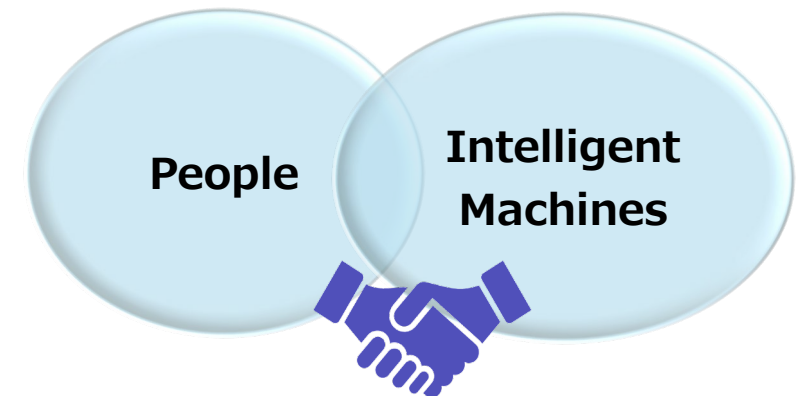
“人”と“Intelligent Machines”が協働する環境を考慮し、  
CDMからCDSへのロードマップを策定していく必要がある

<<Road map to CDS>>

Clinical Data Manager

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

Clinical Data Scientist





## ➤ 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 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

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

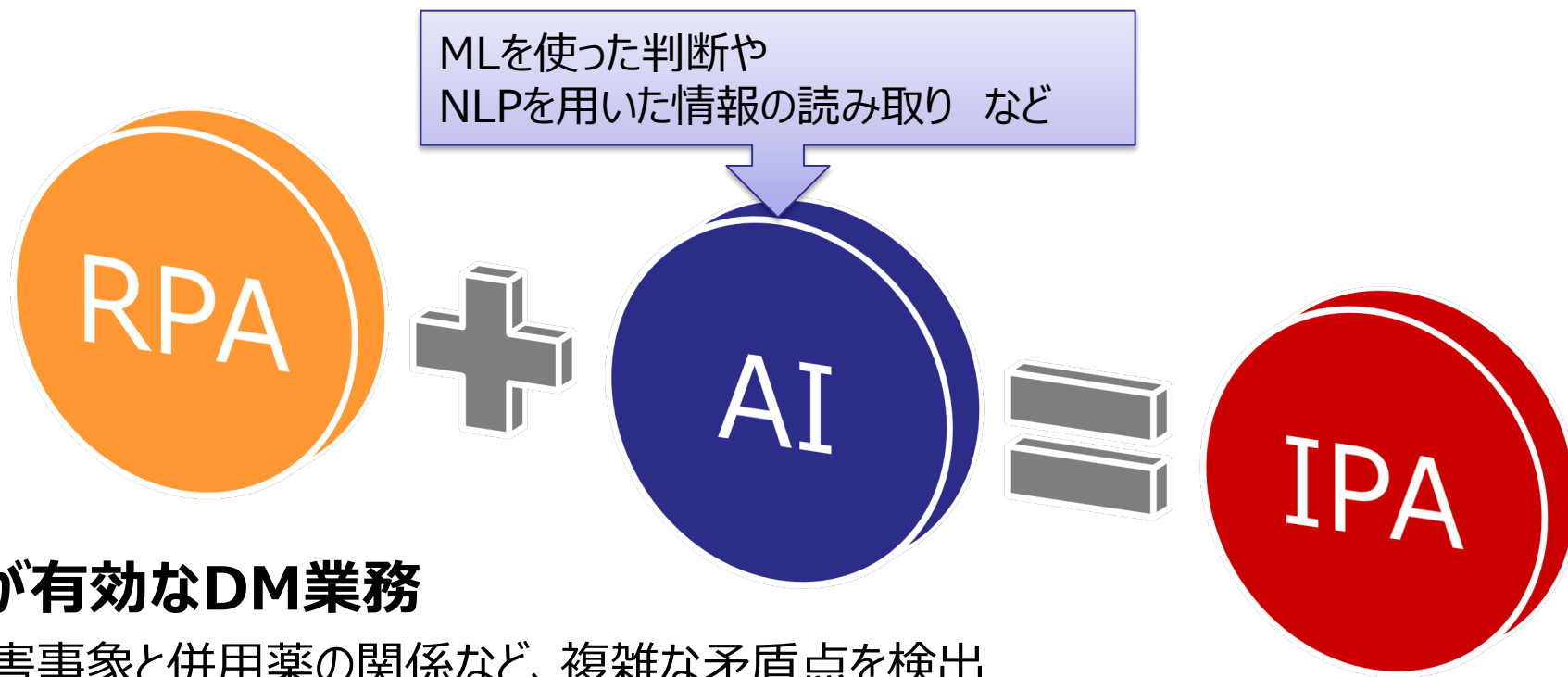
### ■ 既存システムの変更が不要

- 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

- レポーティ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

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

### 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

## 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

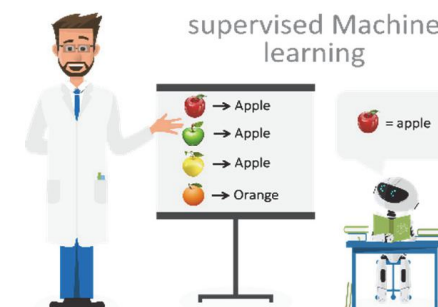
- 有効なDM業務

- ✓ 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

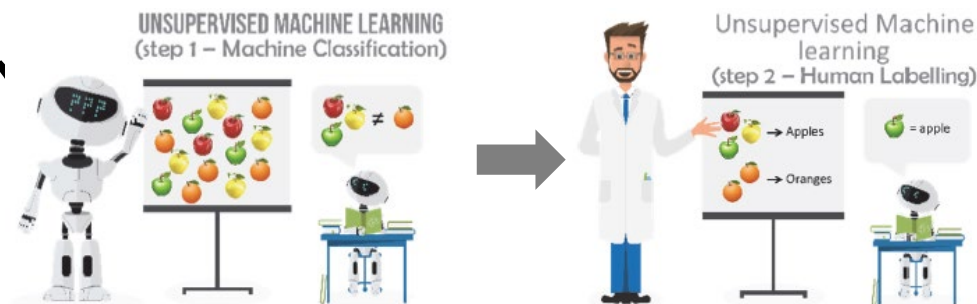
- 教師なし学習：正解のないデータを使用し、データの類似性を判断させる

- 有効なDM業務

- ✓ 画像診断などに活用



Part 1, 5.3.b)の挿絵

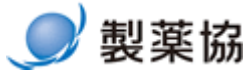


Part 1, 5.3.b)の挿絵

Part 1, 5.3.b)の挿絵

# Artificial Intelligence (AI)

SCDM Part 2 9. Automations / SCDM Part 1 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」



Part 1, 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

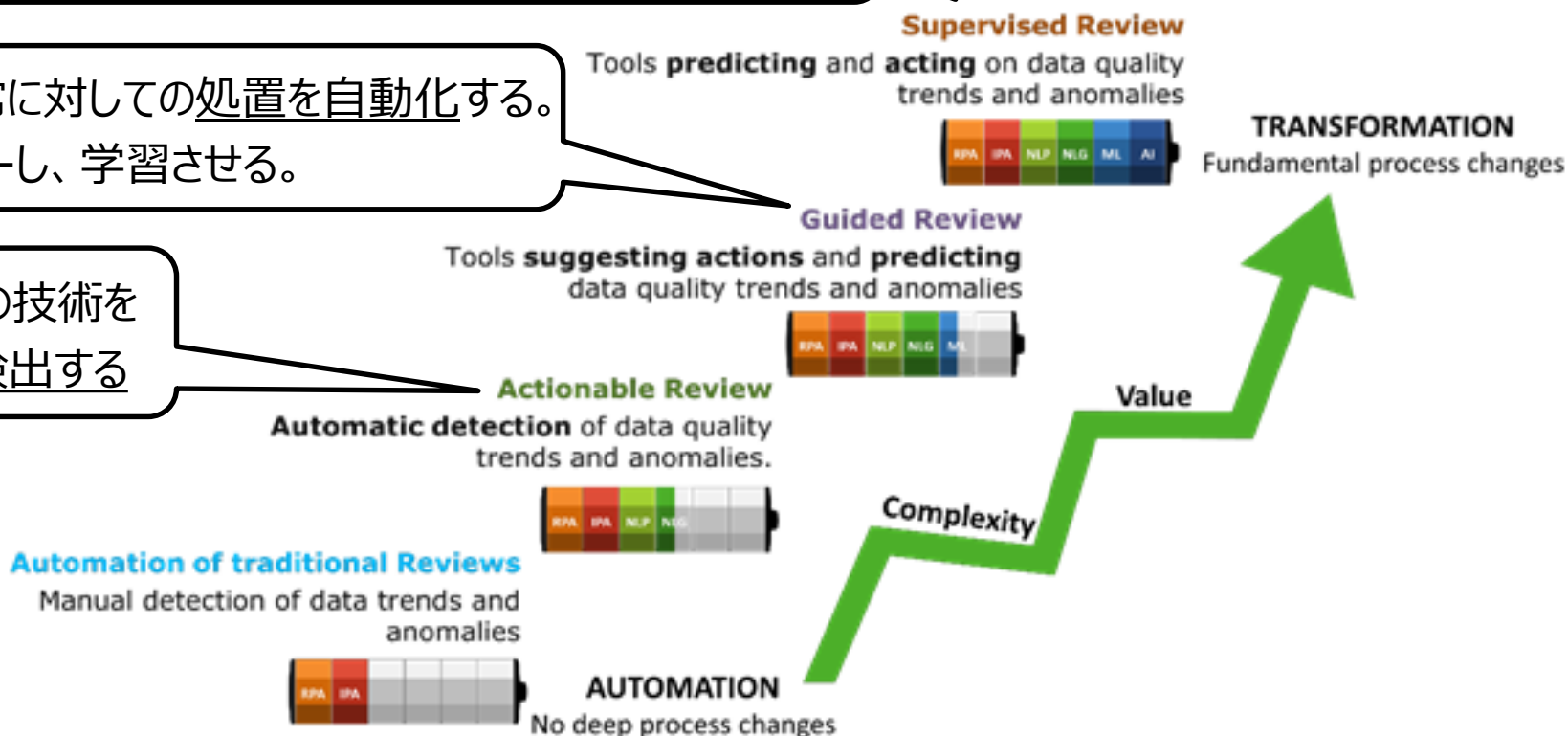
## データレビューの自動化は、以下の4段階を経て進化すると考えられる

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。

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



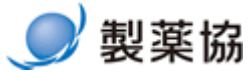
# 新たな規制動向

[目次に戻る](#)



# データクオリティとは？ - 規制による定義

SCDM Part 1 5.2 Impact of regulations on Data Quality and CDM Practices

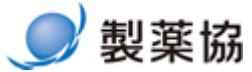


- データクオリティを保証するには、データインテグリティの主要な基準を満たすだけでは十分ではない
- 一般的なガイダンスの定義は
  - データインテグリティ：データが正しい（適切な）方法で管理されていること
  - データクオリティ：データが信用及び信頼できる（credible and reliable）こと
- データが信頼できる（reliable）とは？
  - 例えば、下記ポイントについてデータレビューを行うことでreliableかそうでないかが分かる
    - データの異常値、データのばらつきが不自然
    - 施設内および施設間でのデータ範囲、整合性、ばらつきなどのデータの傾向
    - 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インテグリティの問題

**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 Part 2 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

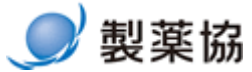
- RBA → “重要な”コンプライアンス違反の検出

- 日常的な監査証跡レビューにより、データクオリティやデータインテグリティ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 Part 2 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**



# 臨床試験環境の大きな変化 まとめ

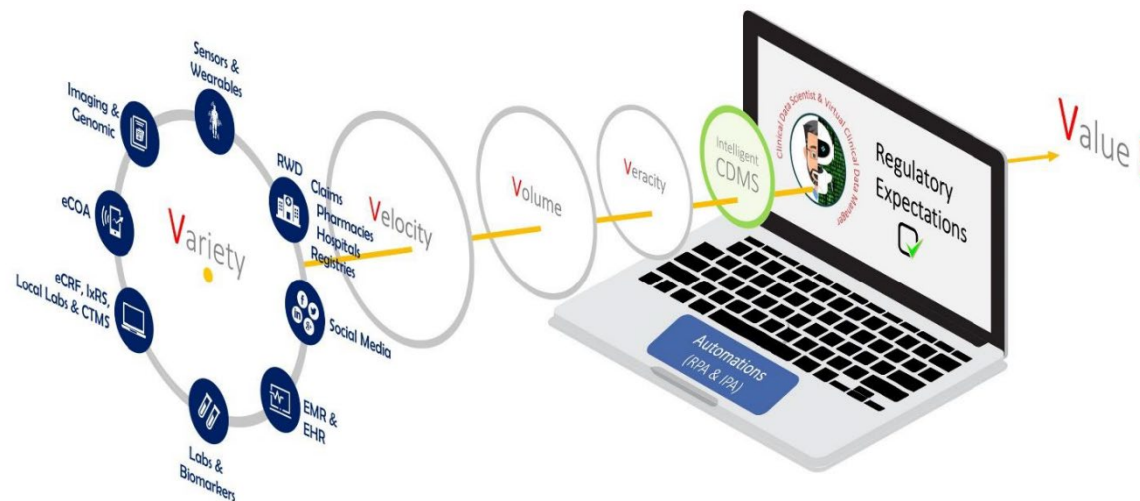
[目次に戻る](#)

# 臨床試験環境の大きな変化 まとめ

ここまでは、5つのVを中心に、臨床試験環境の変化に対応してデータの価値を最大化するために必要な臨床データ戦略が示されてきました。

ここからは、その臨床試験環境の変化と新しいデータ戦略のもと、「CDMに求められる新たなスキルセットとコンピテンシー」という“人”に焦点を当てた内容を示し、組織・キャリアのロードマップ作成をアシストするための情報をご紹介します。

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



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

# CDMに求められる 新たなスキルセットとコンピテンシー

[目次に戻る](#)

## スキルとコンピテンシーの進化

- 臨床試験がより一層複雑化しており、CDMが対応するためにはスキルとコンピテンシーの進化が必要
  - スキル：習得された能力
  - コンピテンシー：スキルを正しく適用できる能力

## 進化がもたらす組織やビジネスモデルへの影響

- 過去30年以上データインテグリティを最重要視してきたが、臨床試験環境の変化に伴い、その他のニーズの高まりを認識することが重要
- テクノロジーを応用してプロセスの改善を推進する最適な方法を理解するために、DM業務の核となる知識の活用が必須
  - 業務に対する深い専門知識と技術的なリテラシーが真の改善を生み出す

# CDMの現状とCDSに求められる役割



## ➤ 現在、CDMの主な業務範囲と考えられているもの

- データ収集ツールとデータ標準の管理
- データ収集からアーカイブまでのデータフロー管理
- 臨床試験データのレビューやリコンシリエーション
- ベンダーとベンダーから受領するデータの管理
- GCDMP<sup>©1</sup>のベストプラクティスの適用
- EDCの活用
- CDMに関連する業務のプロジェクト管理と文書管理

## ➤ CDMの業務の遂行に必要な4つの領域

① ベストプラクティス

④ ソフトスキル

② コンピテンシー

③ 基礎知識

# ① ベストプラクティス

## ➤ GCDMP<sup>©</sup>

- 紙CRFやEDCにおけるDM業務プロセスの参考資料
- 28章で構成され、各章でベストプラクティスが規定されている

GCDMP Chapters	
Data Privacy	External Data Transfers
Data Management Plan	Patient-Reported Outcomes
Project Management for the Clinical Data Manager	CDM Presentation at Investigator Meetings
Vendor Selection and Management	Training
Data Management Standards in Clinical Research	Metrics in Clinical Data Management
Design and Development of DCIs	Assuring Data Quality
Edit Check Design Principles	Measuring Data Quality
EDC - Concept and Study Start-up	Data Storage
EDC - Conduct	Data Entry Processes
EDC - Study Closeout	Coding Dictionary Management & Maintenance
CRF Completion Guidelines	Safety Data Management and Reporting
CRF Printing and Vendor Selection	Serious Adverse Event Data Reconciliation
Database Validation, Programming & Standards	Database Closure
Laboratory Data Handling	Clinical Data Archiving

Part 3, Fig 1. List of GCDMP Chapters

- 現在、GCDMP<sup>©</sup>の大幅な改訂が行われており、いくつかの章は改訂版がリリースされている

## ② コンピテンシー

試験デザイン	<ul style="list-style-type: none"><li>• EDCやeCOAといったデータ収集手段の特定とセットアップ</li><li>• データ標準を活用したデータ処理とレポート作成</li><li>• DM計画書やCRF入力の手引きといった文書作成</li></ul>
プログラミング	<ul style="list-style-type: none"><li>• eCRFやエディットチェック</li><li>• レポート</li><li>• アドホッククエリ</li><li>• データのインポート</li><li>• データ変換</li><li>• データの抽出</li></ul>
データ処理	<ul style="list-style-type: none"><li>• データの収集</li><li>• データの転送</li><li>• インポート</li><li>• クリーニング</li><li>• コーディング</li><li>• リコンシリエーション</li><li>• 品質評価</li></ul>
テスト	<ul style="list-style-type: none"><li>• 必要なツールのテスト計画の立案とその実施</li></ul>
トレーニング	<ul style="list-style-type: none"><li>• 組織全体としてのDM業務手順の理解</li></ul>
人員管理	<ul style="list-style-type: none"><li>• CDMスタッフのオーバーサイト</li></ul>
プロジェクト管理	<ul style="list-style-type: none"><li>• ベンダー管理を含めたDM業務のオーバーサイト</li></ul>
レビュー	<ul style="list-style-type: none"><li>• 試験やDM業務の成果物レビュー</li></ul>

### ③ 基礎知識

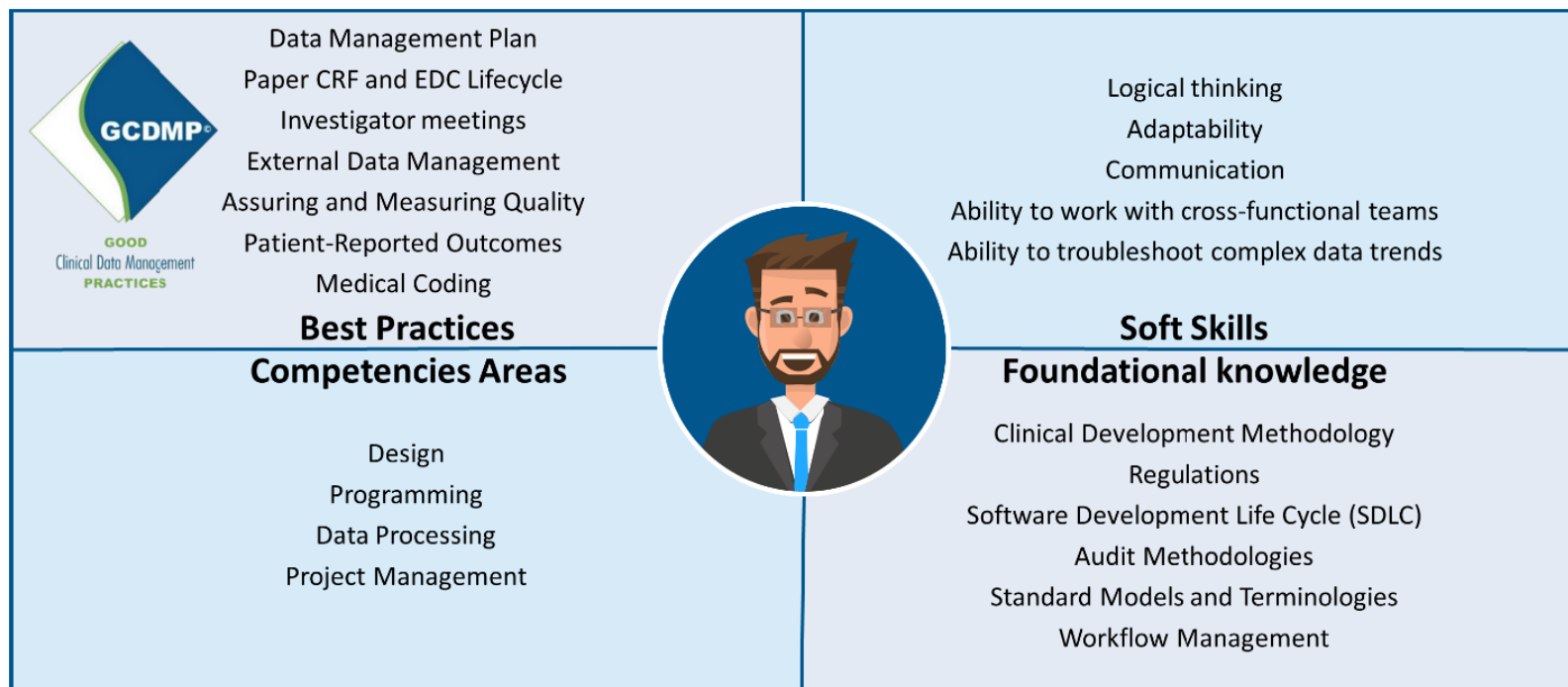
- 医薬品開発と臨床試験の基礎
- 科学的方法
- GCP及びその他のガイダンス
- ソフトウェア開発ライフサイクル（SDLC）の概念
- 監査方法
- プロジェクト管理の基礎
- 基本的な統計学的概念
- データやメタデータのモデル、標準及び用語（医学用語を含む）
- ワークフローの設計、解析、及び管理の基本

## ④ ソフトスキル

- 細部への注意力
- ロジカルシンキング
- 順応性
- 複雑なコンセプトを治験チームに明確に説明する能力
- 複雑なデータの傾向を調査し、問題に対処する能力
- 部門を横断して協力する能力

# CDMのフレームワーク

## SCDM Part3 4. Current state of the Clinical Data Management role



Part 3, Fig 4. CDM role framework

CDMの4領域は今後も重要な要素だが、CDSはさらに3つの要素が求められる。

① ベストプラクティス

④ ソフトスキル

② コンピテンシー

③ 基礎知識



リスクに基づくアプローチ

新たなデータレビュー手法

先進的な臨床試験手法  
と技術の知識

## ➤ 今後起こりうる業務の変化への対応

### ■ データインテグリティに加えてデータクオリティの確保が必要となる

□ データインテグリティ：データが正しい（適切な）方法で管理されていること

□ データクオリティ：試験結果が信用及び信頼できる (credible and reliable) こと

### ■ 全施設・全試験画一的な手法からの脱却

リスクに基づく アプローチ	新たな概念や規制要件への対応	
	<ul style="list-style-type: none"><li>QbD</li><li>Critical to Quality (CtQ) 因子</li></ul>	<ul style="list-style-type: none"><li>重要なデータやプロセス</li><li>リスクの評価や根本原因分析など</li></ul>
新たなデータレビュー 手法	データクオリティを確保する新たな手法の確立	
	<ul style="list-style-type: none"><li>データの多様化とEDC中心性の低下</li><li>DCT</li><li>データの信頼性</li></ul>	<ul style="list-style-type: none"><li>データやメタデータの量、多様性、速度</li><li>試験特有のデータの監視</li></ul>
先進的な臨床試験 手法と技術の知識	臨床試験を支える新たな技術の理解	
	<ul style="list-style-type: none"><li>Adaptive design / Master protocol</li><li>RWDの利用の増加</li></ul>	<ul style="list-style-type: none"><li>m-Healthの信頼性向上と価格の低下</li><li>AIの活用</li></ul>

## ➤ データクオリティ

### ■ MHRA（イギリスの医薬品・医療製品規制庁）の定義：

*“the assurance that data produced is exactly what was intended to be produced and fit for its intended purpose” \**

データが意図した通りに生成され、意図した目的に適合していることの保証

## ➤ 画一的な手法からの脱却

- 各施設的能力を最大限に活用し、患者中心のデータフローを構築することが求められる
- 国や施設、試験ごとにデータフローが形成されるため、試験の最適化に向けてチームをリードする必要がある

\* MHRA, March 2018, 'GXP' Data Integrity Guidance and Definition, Available at [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/687246/MHRA\\_GxP\\_data\\_integrity\\_guide\\_March\\_edited\\_Final.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/687246/MHRA_GxP_data_integrity_guide_March_edited_Final.pdf)



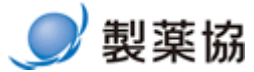
# CDSの役割とスキルセット

## リスクに基づくDM業務のアプローチ

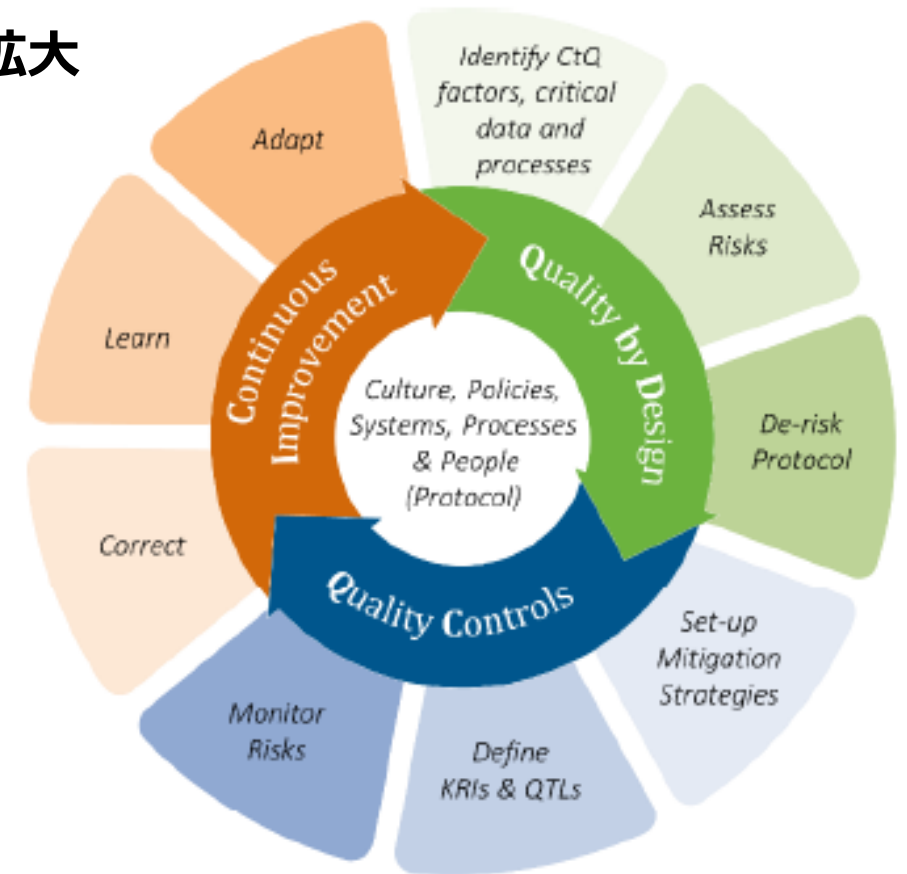
[目次に戻る](#)

## 5.1) リスクに基づくDM業務のアプローチ

SCDM Part3 5.1) Risk-based CDM approaches



- リスク管理のためにデータマネジメントプロセスと役割の再定義が必要  
…Part 1で明示されている
  - リスクに基づく品質管理（RBQM）の原則の適用範囲拡大
  - 被験者の保護と試験結果の信頼性を保証するため
- 伝統的なサイトモニタリングから  
リスクに基づくモニタリングへの進化と同様に、  
DM業務もリスクに基づいて考える必要がある
  - 可能な限りリスクを未然に防ぐために  
設計段階で品質を作り込む新しいパラダイムに移行



Part 3, Fig 4. CDS RBQM framework

## a) 試験デザインによる品質

- ICH E6(R2)で新たに設けられたリスクライフサイクルマネジメントは、リスクの軽減策やリスクのモニタリングよりも、QbDを使って最初からリスクを避けることを意味している
  - QbDは
    - 品質は前もって計画されるもので、QCによって回顧的に保たれるものではない
    - 適切な基盤に依存（企業文化、ポリシー、システム、プロセス、人材などを含む）
    - プロトコル設計の最適化、試験に大きな影響を与える重要なステップに依存
  
- QbDの原則に沿ってプロトコルの最終化前に、リスクとその軽減策を特定（可能であればリスクを回避するためにプロトコルを修正）
  - リスク軽減プロセスの一環として重要なことは
    - 試験のすべての側面が実施可能であること
    - 不必要な複雑さ、手順、データ収集を回避すること

## a) 試験デザインによる品質

➤ CDSの役割は、6つのCtQカテゴリに含まれる因子について、積極的にリスク管理すること

- 管理するリスク

- 重要なデータ及びプロセスに関連するCtQ因子など

CtQ Categories	CtQ factors
Protocol Design	Eligibility Criteria Randomization Masking Types of Controls Data Quantity (CTTI) – Data Quality (ICH E8) Endpoints Procedures Supporting Study Endpoints and Data Integrity Investigational Product (IP) Handling and Administration
Feasibility	Study and Site Feasibility Accrual (i.e., Enrollment Strategy)
Patient Safety	Informed Consent Withdrawal Criteria and Trial Participant Retention Signal Detection Safety Reporting Data Monitoring Committee (DMC) / Stopping Rules (if applicable)
Study Conduct	Training Data Recording and Reporting Data Monitoring and Management Statistical Analysis
Study Reporting	Dissemination of Study Results
Third-party Engagement	Delegation of Sponsor Responsibilities and Collaborations

Part 3, Fig 5. CTTI and ICH E8 CtQ categories and factors

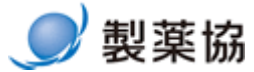
➤ CDSは、既知のリスクから生じる傾向を積極的に特定するため、頑健なリスクモニタリング戦略を計画する必要がある

- リスクモニタリング戦略は、統合品質管理計画書 (IQMP) またはDM計画書に記録する

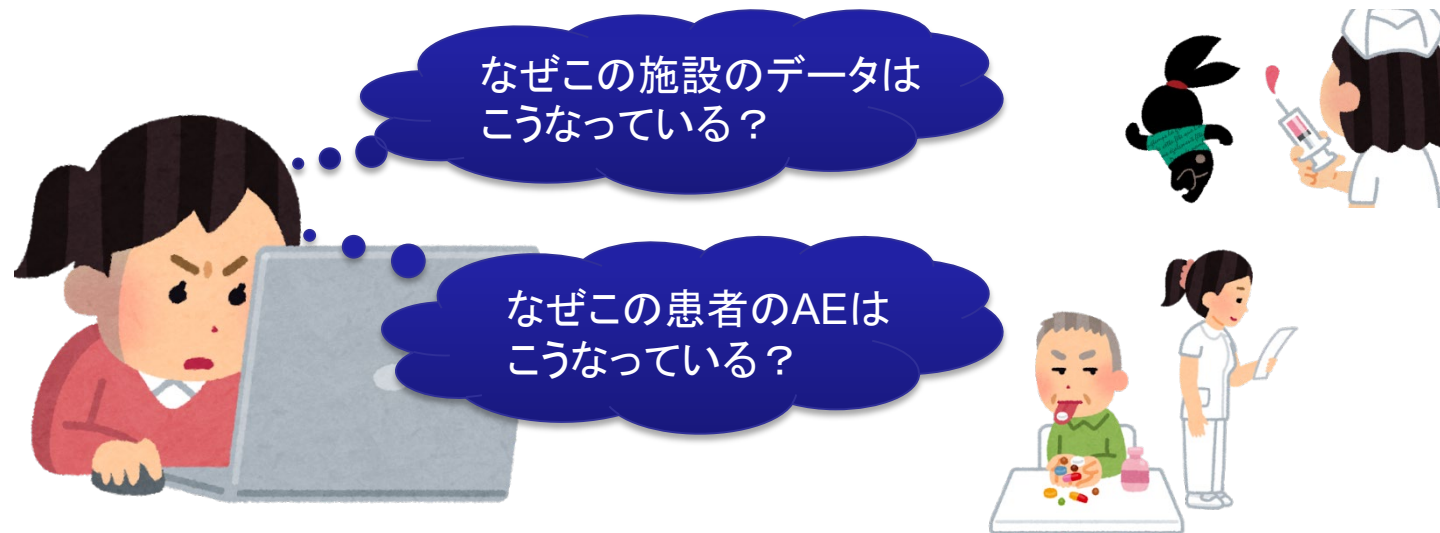
- 用いる分析ツール：主要なリスク指標 (KRI)、品質許容限界 (QTL)

## b) リスクに基づく試験の遂行（品質管理）

SCDM Part3 5.1) Risk-based CDM approaches



- CDSは、試験の進行中にデータに関連するリスクを継続的に監視することが期待される。
- 「シグナル」を理解し、データを深く掘り下げ、根本原因を分析、根本的な問題に対処し、フォローアップする。
- 単なるクエリ発行ではなく、プロセス上の問題や施設の特性に対処することになる。



値の入力ミスを疑うだけではなく、測定値の取り方や記録方法などの系統的な問題を見つけ、根本原因の分析並びに対処しなければならない。

## c) CDSプロセスの継続的な改善について

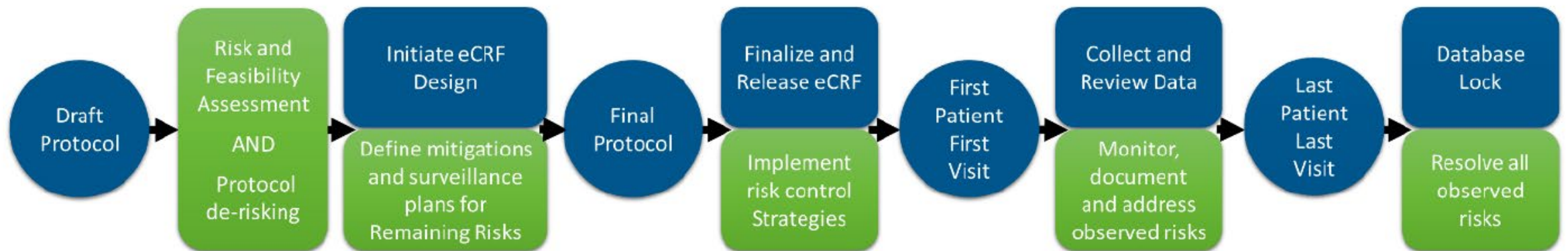
- 試験途中で得られた知見を積極的にプロセス改善に役立てることが期待される。
  - 系統的错误についてはCAPA（是正措置・予防措置）に基づく改善計画が提案されることもある。
  
- CDSはCAPAについても精通すべきである。
  - CDSはプロセス改善に際してリスクの性質を理解し、現実的かつ頑健な予防措置を提案するべきである。

# d) リスクに基づく手法がCDSにもたらす影響

- QbDとリスクに基づいた試験を支援するCDSに求められるものとは？
- CDMからCDSへの進化に伴い追加される重要な視点とは？

従来のDM業務の範囲	リスクに基づくCDS業務の範囲
アウトプットを重視したロジカルシンキング	アウトカムを重視したクリティカルシンキング
確定したプロトコルに基づく試験セットアップ	QbD：試験デザイン時から品質を視野に入れる
全試験画一的な手法	各試験に合わせたリスクに基づくプロセス
データインテグリティに重点を置く	データクオリティに重点を置く (data reliability)
データ収集後のレビュー	リスクに基づくデータモニタリング

- 従来のDM業務プロセス（青）にCDS業務プロセス（緑）が追加された例



Part 3, Fig 7. Example of a risk-based CDS process flow



# リスクに基づくアプローチを導入するためにCDSに求められる要件とは

SCDM Part3 5.1) Risk-based CDM approaches



ベストプラクティス	ソフトスキル
<ul style="list-style-type: none"><li>• リスクに基づく試験の実施</li><li>• KRIとQTLのライフサイクル</li></ul>	<ul style="list-style-type: none"><li>• クリティカルシンキングと根本原因究明</li><li>• 適応力</li><li>• 実用主義（实际的・実務的スキル）</li><li>• 影響力を発揮するリーダーシップ</li></ul>
コンピテンシー	基礎知識
<ul style="list-style-type: none"><li>• リスクライフサイクルマネジメント</li><li>• 高度な分析力</li><li>• プロセスマネジメント</li></ul>	<ul style="list-style-type: none"><li>• Adaptive design、Master Protocol等の新しい手法に関する知識</li><li>• DCTの手法とその技術について</li><li>• リスクに基づく手法と規制</li><li>• 多面的な問題を調査するための強力なデータフローとシステムの活用力</li></ul>



# CDSの役割とスキルセット

## データレビューの進化

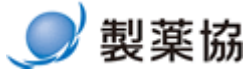
[目次に戻る](#)

- データレビューは試験全体のモニタリング戦略の一環である
- ICH E6には以下の内容が記載されている
  - 治験依頼者は、治験のモニタリングの適切な範囲と内容を決め、実施に当たって、優先順位を考慮し、リスクに基づく体系的な取組を策定すべきである
  - 治験依頼者は、「オンサイトモニタリング」「オンサイトモニタリングと中央モニタリングの組合せ」又は、正当な場合には「中央モニタリング」を選択することができる
  - 集積するデータの時宜にかなった遠隔的な評価は、適切な資格を有し、教育を受けた者（データマネージャー等）によって補助される
- 規制当局は、「すべてのデータに対し一貫した品質レベルを要求すること」から「重要なデータに焦点をあてて“データが意図したとおりに生成され、意図した目的に適合していること”を確かめること」へと、考えを大きくシフトさせた

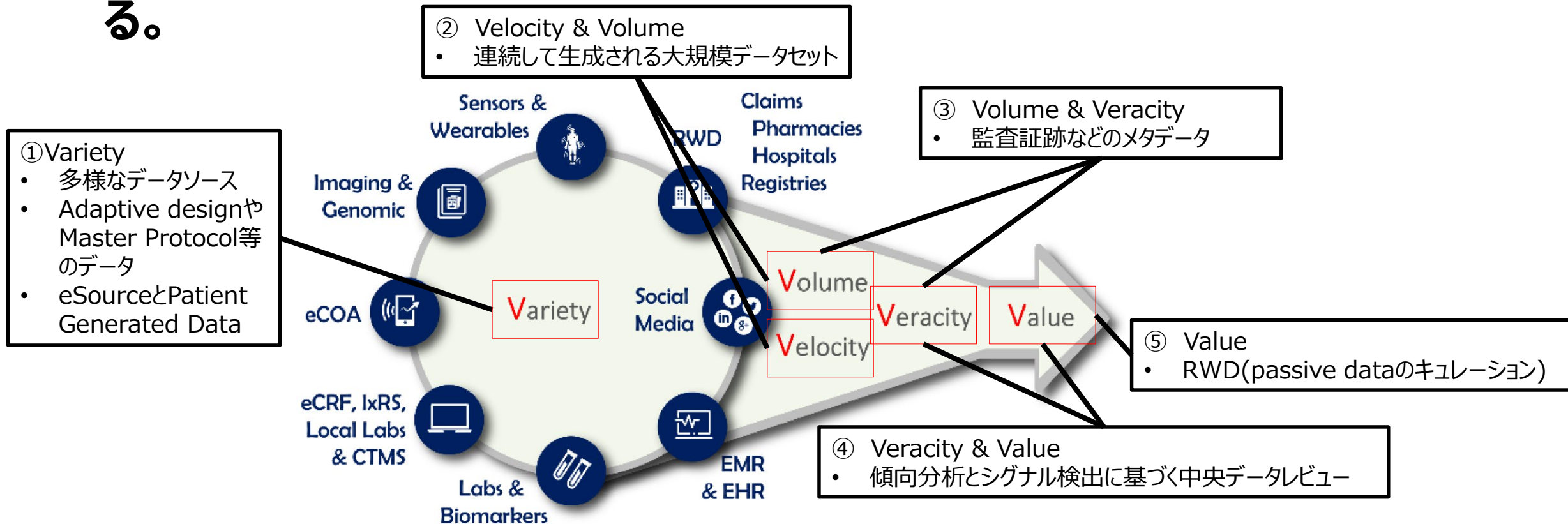
- データレビューの範囲は、患者データにとどまらず、監査証跡データの活用が含まれる
    - 監査証跡データには以下の貴重な情報が含まれる
      - ✓ 治験実施計画書がどのように運用されているか
      - ✓ データがどのように収集されているか
- ↓
- データの品質管理活動に利用
  - データインテグリティ確保
- 
- 適切なツールを活用する
    - インテリジェントCDMS
    - AI
- (スライド「インテリジェントCDMS」「自動化 (Automation) 」参照)

# データレビューの変化とCDMの役割への影響

SCDM Part3 5.2) The evolution of data reviews



- リスクに基づくDM業務フレームワークを前提に、大きく影響を与える5Vの観点から、データレビューの変化とCDMの役割への影響について解説する。



# データレビューの変化とCDMの役割への影響：各論①

SCDM Part3 5.2) The evolution of data reviews



## ① Variety

### 多様なデータソースのレビュー

- EDCや従来型のCDMSではDCTなどデータソースが多く複雑になると対応しきれない。加えて、それらのデータソースはMedDRAやCDISCといった臨床試験で慣れ親しんだ標準形式とは異なる形式である
- 従って、EDC中心のエディットチェックによるデータレビューでは不十分であり、CDSは以下のような対応が必要
  - ✓ シーケンスデータや非構造化データなど異なる構造のデータを統合
  - ✓ FHIRの様な医療情報標準、インテリジェントCDMS等の新しい技術、ICD、SNOMEDの様な各種用語等を理解

### Adaptive designとMaster Protocolのデータのレビュー

- 患者背景やサンプルサイズ、治療群などがAdaptationされたり、試験内で複数の治験薬が複数の適応症に使用されたりする
- 従来型のデータレビュー、リコンシリエーションのスキームは働かなくなり、目的に合わせたデータレビュー戦略が必要
  - ✓ 患者やVisitによって収集するデータが異なり欠測データの検出が難化
  - ✓ 患者集団によって安全性・有効性レビューの焦点が異なる
- CDSは以下の対応が必要
  - ✓ データレビューテクノロジーの適応可能性を判断するために、プロトコルの変化が下流に与える影響を理解しなければならない
  - ✓ 多くの場合はデータが最新になってから試験デザインがAdaptationされるため、データの同時性にも注意を払わなければならない
  - ✓ Adaptationされるタイミングでデータベースロックが求められるかもしれない

### eSourceとpatient generated dataのレビュー

- 一度生成されたデータの変更はほとんど不可能。データクォリティ及びインテグリティに対するフィードバックはデータ生成時に提供することが必要で、後からCDMがデータ修正のクエリを出すことでの対応はほぼ不可能
- MHRAが導入した「データ除外」の概念へ対応するため除外データにタグ付けし、すべてのデータは（除外されている場合でも）元のデータセットと共にレビューできるように保持

# データレビューの変化とCDMの役割への影響：各論②

SCDM Part3 5.2) The evolution of data reviews



## ② Volume & Velocity

### 連続して生成される大規模データセットのレビュー

- センサーやウェアラブルを含むm-Healthの利用によりデータの量と速度は爆発的に増加
- マニュアルプロセスでのデータレビューはもはや不可能
  - ✓ データフィルタリングやトレンド把握だけでなく、ストーリー性のあるデータの可視化、統計モデル、MLモデル、さらにはIPAの活用等、様々な戦略の実装が必要
  - ✓ non-SQLのようなテクノロジーの専門知識が必要になる可能性

## ③ Veracity & Volume

### 監査証跡等のメタデータのレビュー

- 大量のデータがeSource等で収集され、従来のSDV、エディットチェック、目視による一覧レビュー等では対応できないため、監査証跡等のメタデータを活用する新しいデータレビュー戦略によりデータの妥当性を確保することが必要
- 監査証跡の形式はテクノロジー間で標準化されていない
  - ✓ カスタムでのデータ統合とレビュー戦略が必要
  - ✓ 監査証跡の量がデータ統合とレビュー戦略に影響
- 監査証跡のレビューについてe-Clinical ForumとSCDMがポジションペーパー<sup>1</sup>を2020年に公表

1. eClinical Forum and the Society for Clinical Data Management, October 2020, AUDIT TRAIL REVIEW: A KEY TOOL TO ENSURE DATA INTEGRITY - An Industry Position Paper. Available at [https://scdm.org/wp-content/uploads/2020/10/eCF\\_SCDM-ATR-Industry-Position-paper-Final-Editorial-Edits.pdf](https://scdm.org/wp-content/uploads/2020/10/eCF_SCDM-ATR-Industry-Position-paper-Final-Editorial-Edits.pdf)

# データレビューの変化とCDMの役割への影響：各論③

SCDM Part3 5.2) The evolution of data reviews



## ④ Veracity & Value

### 高度な傾向分析とシグナル検出に基づく中央データレビュー

- これまでのCDMのレビューはデータの欠測、一貫性、外れ値の特定に限定して焦点を当ててきたがICH E6(R2) においてデータレビューのスコープが拡張し、中央モニタリングとして集積されたデータのレビューを通じて試験全体の監督も担当
- 複雑なデータの傾向とシグナルを生成する統計・ML手法に基づく高度な分析ソリューションが必要
  - ✓ 例1：恣意的に改ざんされたデータを検出する仕組み
  - ✓ 例2：予防的にリスクを軽減する予測アルゴリズム
- 患者、施設、国ごとの異常を示すシグナルを調べるために、エンドツーエンドのデータフローに関するより深い知識が必要
  - ✓ シグナルが体系的・系統的なプロセスエラー、ずさんさ、意図的なバイアスを示している可能性がある一方で、誤検知である可能性もある。そのため、CDSは施設や患者などのステークホルダーに関連するものを含む治験のプロセスとシステムを包括的に理解する必要がある

## ⑤ Value

### RWDのレビュー（passive dataのキュレーション）

- passive dataは合成コントロール群、プロトコルの最適化、ベンチマークなどで利用可能
- データソースではデータは匿名化されておらず、臨床試験標準と一致せず、複数のシステムに散在
  - ✓ キュレート（匿名化、統合、整理、評価など）が必要
- 臨床試験で利用できる適切な二次利用データとしての価値とRWEを生成するために、インテグリティとクオリティを確認する客観的な方法が必要



実践的な観点から、CDSコンピテンシーは、急激なテクノロジーの変化に沿う必要がある。

- インテリジェントCDMSの管理
- non-SQL等の新しい技術の使用
- 統計手法を活用した分析ツールの使用
- ロボットによるインテリジェントなプロセス自動化の実装(RPA と IPA)
- MLなどのAI手法を活用したインテリジェントソリューションの実装

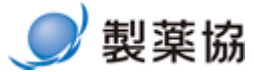
上記を考慮すると、CDMとCDSのデータレビュー範囲を以下のように比較することができる。

CDMのデータレビュー範囲	CDSのデータレビュー範囲
EDCに焦点	DCTテクノロジーに焦点
少量のデータとソース	大量のデータとソース
単純なデータフロー	複雑なデータフロー
アウトプットを重視したロジカルシンキング	アウトカムを重視したクリティカルシンキング
全試験画一的な手法	各試験に合わせたリスクに基づくプロセス
データインテグリティに重点を置く	データクォリティに重点を置く (data reliability)
データのクリーニング	データのレビュー、タグ付け、除外、キュレーション
臨床試験データ	臨床試験とヘルスケアデータ
従来のプログラミング (SQL、C#、SASなど)	ML (Python、Rなど) 、non-SQL



# 進化したデータレビューアプローチをサポートするための要件

SCDM Part3 5.2) The evolution of data reviews



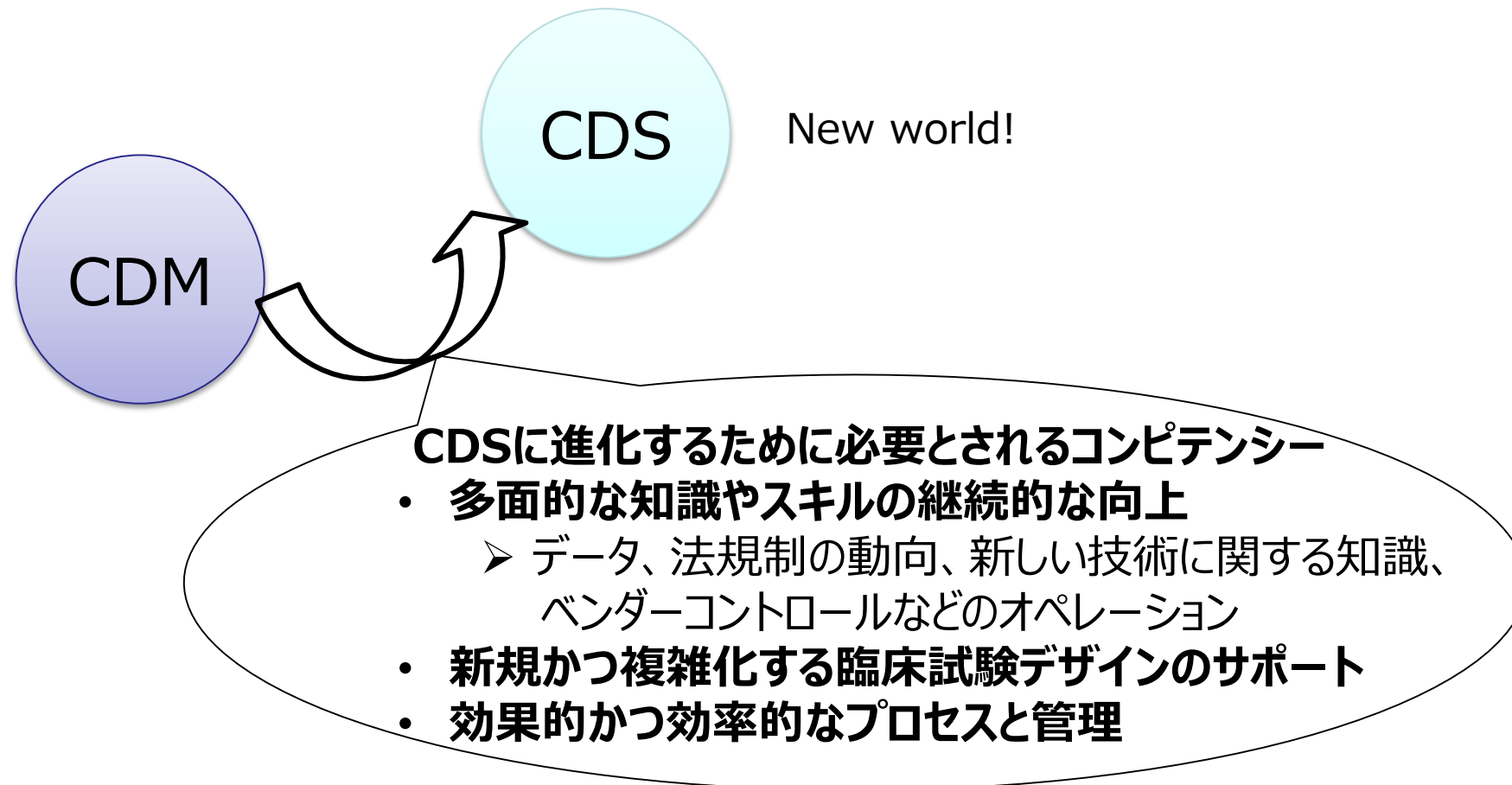
ベストプラクティス	ソフトスキル
<ul style="list-style-type: none"><li>• リスクに基づく試験の実施</li><li>• KRIとQTLのライフサイクル</li><li>• ストーリーテリングの視覚化</li><li>• 監査証跡のレビュー</li><li>• データのタグ付け、除外及びキュレーション</li></ul>	<ul style="list-style-type: none"><li>• クリティカルシンキング</li><li>• 複雑なデータフローを理解する能力</li></ul>
コンピテンシー	基礎知識
<ul style="list-style-type: none"><li>• 高度な分析力</li><li>• 高度なデータ照会手法 (non-SQL等)</li><li>• MLの手法</li></ul>	<ul style="list-style-type: none"><li>• Adaptive design、Master Protocol等の新しい手法に関する知識</li><li>• DCTの手法とその技術について</li><li>• リスクに基づく手法と規制</li><li>• シーケンスデータ、非構造化データ、データマイニング、MLなどの新しい概念の理解</li></ul>

# CDSの役割とスキルセット

## 高度なCDMのコンピテンシー

[目次に戻る](#)

- CDSは、期待される試験結果を得るために、クオリティとインテグリティが担保されたデータを生成するための高度な能力が求められる。



Adaptive designやMaster Protocolの臨床試験が近年増えてきており、2010年と比較して、2019年では9倍も実施されている。

Adaptive designの要素を持つ3つのタイプのMaster Protocol（アンブレラ型、バスケット型、プラットフォーム型）は新薬の全体的な開発を加速することができる利点もあるが課題もある。

## 全ての試験デザイン共通の課題

EDC、eCOA、およびIRTで複雑で動的な分岐ロジックを持つ柔軟なData Collection Instrumentが必要

- CDSは、プロトコルの科学的・運用的なコンセプトに精通し、技術的にデータシステムのセットアップを推進
- Adaptive design、Master Protocol、DCTは同時に実現することもできるため、多面的なシナリオプランニングが必要

## CDSの課題

デザイン、コントロール、データレビュー、および中間解析の頻度の増加への対応

- すべて適切に実行するには慎重な計画が必要
  - ✓ クリティカルシンキングが必要（もはやロジカルシンキングでは対応出来ない）であり、必要な基礎知識が大幅に増加

## CDSにとって重要なこと

### 1. データの価値を正しく理解する

全てのデータが同じように生成され、同じ価値を持たないことを理解することが重要である。データがプロトコルの主要目的にどのように影響するかに加えて、二次利用におけるデータの価値を理解した上で、データの管理方法を決定する必要がある。

### 2. 新たなデータの標準を理解する

臨床試験で収集される多様なデータをどこで統合するか、価値の創出に繋げるためにどのように変換するか、データがどのようなアクションに影響するかを評価することが重要である。

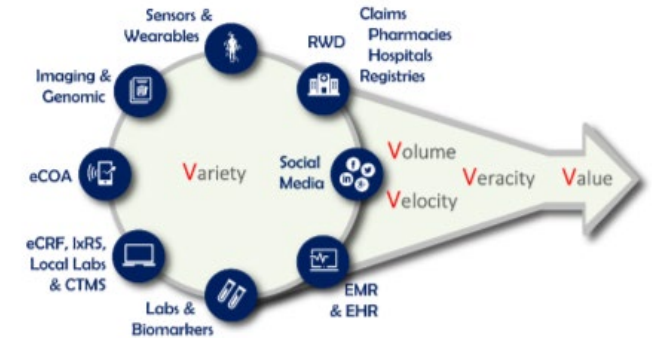
EMR/EHRデータのインテリジェントCDMSへの融合や、Direct Data CaptureやeSourceの取り組みにおいては、既存のデータ標準（CDISC）に加えて、他の標準（EHRの標準：HL7 FHIR）の習熟も必要である。

### 3. データの最適な管理手法を考える

大量のデータのクオリティやインテグリティ評価には自動化が必要である。即時性の高いデータソースに対しては、瞬時に異常シグナルを検出するような新たな仕組みを考える必要がある。

### 4. データの信頼性やインテグリティの評価方法を考える

CROの技術を利用したり、医療機関のデータを活用するなど、データの信頼性やインテグリティを直接管理出来ない状況においても、データの信頼性やインテグリティを評価する方法や、追加の管理施策を考える必要がある。



Part 3, 5.3) 挿絵

- 自動化の技術を迅速かつ低コストで導入することが可能になり、人の手を介さずにデータの処理を随時行うことが可能になった。
- 高度な自動化の技術を用いるために、その技術を用いたデータ処理のテスト、リリース、管理に対するアプローチを強化する必要がある。

## RPA



RPAを用いることで、事前にスクリプトで定義された、反復的な処理を自動化することが可能である。概念としては、エディットチェックによく似ている。

RPAは人間と同じようにシステムにログイン・ログアウトしてタスク処理が可能である。

臨床試験で使用する場合は、RPAのタスク処理をシステムの監査証跡に残すとともに、RPAに付与するシステムIDと挙動をテストし、管理するための手順が必要である。

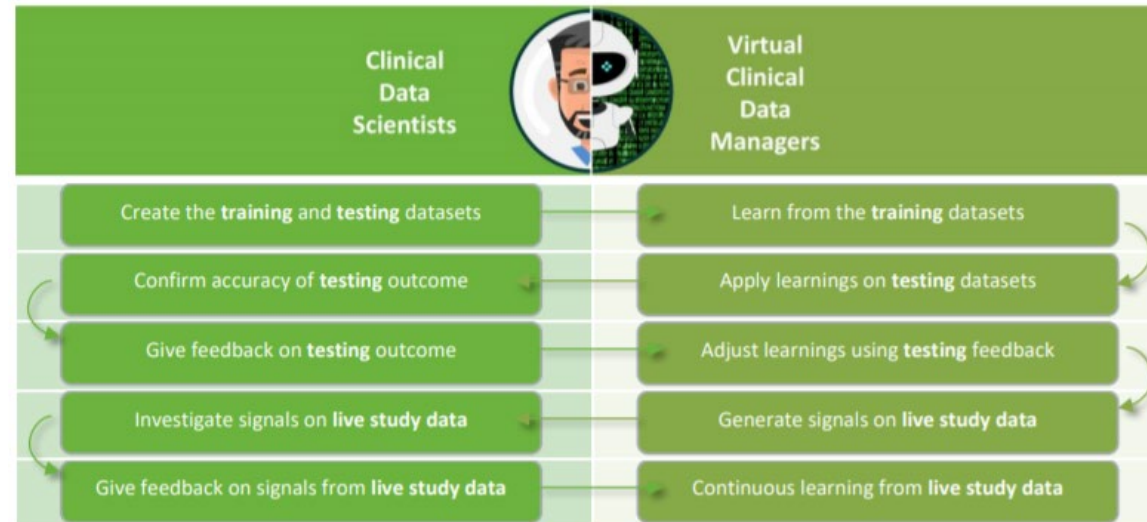
また、規制当局の査察に耐えうるように、CDSは、本番環境で使用されたRPAは、どのように開発され、テストされ、管理されたかを説明できるようにしておく必要がある。

RPAの活用が可能なユースケースを見出すことだけでなく、広い視点で時間や費用などのコストを考慮して、RPAの価値を評価することが求められる。

## AI/MLベースのソリューション：CDSを支援するバーチャルCDM DM業務において最も重要な変革が必要

CDSはバーチャルCDMを通じてデータレビューやその他のアクティビティを正確に実施する。

- CDSは、正しくAIに学習させる内容を定義するという重要な役割を担うことになる。
- 正しいデータセットを定義できないと、システムの動作に偏りが生じ、データレビューの精度に一貫性がなくなる可能性がある。



Part 3, Fig 8. Example of learning process for an ML-Based solution

学習モデルを常に最適化できるように、精度が保たれていることを監督し、定期的に再訓練を行う必要がある。



ベンダーオーバーサイトに関して、より多くの責任が治験依頼者に課されるようになってきている

**ステータスとタイムラインの管理から、パフォーマンスの監督に移行することが必要**

- 従来のベンダーからの報告を待つ受け身のアプローチからの脱却
  - ベンダーの運用データや試験データを集め、品質とパフォーマンスを継続的に分析できる能力が必要
- Clinical Operation、QA、Procurement など他部門との関係性や相互理解の強化も必要



## CDSは影響力のあるリーダーシップを発揮することが求められる

- CDSは試験データに対して責任を負う重要なステークホルダーとして、機能横断的なチームにおいて中心的な役割を担い、重要な意思決定に寄与することが期待されている
  - コアコンピテンシーやスキルに加えて、ビジネス感覚、技術力、変化への対応力を身につけることが必要
  - 以下のソフトスキルも重要
    - 幅広いステークホルダーの視点の理解
    - リスクや影響の評価と、最適な軽減策の決定
    - 実用的かつ革新的な代替案の提示
    - 複雑なテクノロジーや科学的なコンセプトの明確化
    - チーム決定の背景にある論理や問題の理解

## ➤ CDMと高度なコンピテンシーを備えたCDSの適応範囲の比較

CDMの適応範囲	CDSの適応範囲
one and two dimensional trials	three or more dimensional trials
プロジェクト管理	部門横断的なリーダーシップ
ランダム化比較試験	Adaptive、Master Protocol
ベンダー管理	ベンダーオーバーサイト
臨床試験標準	臨床試験及びヘルスケア標準
ロジカルシンキング	クリティカルシンキング

# 高度なCDSに求められる要件

## ベストプラクティス

- 二次利用データとしての価値の生成
- インテリジェントなシステム管理
- データやシステムの統合

## ソフトスキル

- 影響力を発揮するリーダーシップ
- 実用主義（实际的・実務的スキル）
- クリティカルシンキング
- 曖昧さや動的な環境を管理する能力

## コンピテンシー

- ベンダーオーバーサイト
- 患者中心のテクノロジー
- 高度な分析力

## 基礎知識

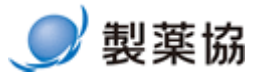
- Adaptive design、Master Protocol等の新しい手法に関する知識
- DCTの手法とその技術について
- シーケンスデータ、非構造化データ、データマイニング、MLなどの新しい概念の理解
- ヘルスケア標準と用語
- 自動化とAIの概念

# CDSへの進化がビジネスモデルに及ぼす影響

[目次に戻る](#)

# ビジネスモデルの背景(アウトソースモデルとFSPモデル)

SCDM Part3 6.1) Historical background on business models



**新しいテクノロジー、進化するプロセス、革新的な臨床試験戦略は、DM業務に関するもののビジネスモデルにも影響がある**

## アウトソースモデル

- サービスプロバイダー（CRO）が競争力を維持するためには、進化を続けるCDSに対する業界の期待に応えられるように、プロセス、テクノロジー、リソース能力を進化させる必要がある
- これらの期待は治験依頼者もCROも同じであり、この進化はすべての組織に必要となる

## FSP（Functional Service Providers）モデル

- サービスプロバイダー（CRO、BPO\*、テクノロジーサービスプロバイダー等）と治験依頼者間には、システム、プロセス、役割において依存関係があり、これらの変化に対応するためには両者でCDSに向けた進化について慎重に計画し、調整する必要がある
- 従来のFSPは、リソース能力を柔軟に増強するために確立された。また、内部の固定費を抑え、運用コストを削減する方法として確立された。アウトソースモデルと異なり、治験依頼者はFSPスタッフに自分たちのシステムやプロセスを使用させている

\*BPO: Business Process Outsourcing

## 小規模FSP

- スタッフの増強のためにFSPを使用している場合、進化するCDSの責任を考慮してスタッフを選択していく必要がある
- FSPスタッフも、治験依頼者の社員と同様のオンボーディングとトレーニングが必要になる可能性が高い

## 大規模FSP

- 人件費の低い国（インド、南アメリカ、メキシコ等）に大規模なオフショアセンターを設立しているような場合には、より大幅な調整が必要になる
- オフショアでは、新卒と経験豊富なリソースを組み合わせ採用しトレーニングすることで全体的なコストを抑えて、大量作業や反復作業（データ入力、ディスクレパンシーマネジメント等）を委託し、短期的に効果的なROIを実現していた
- 全体的な傾向は依然としてオフショアにシフトしているが、以下のような理由から、スキルの高い自社リソースへ業務を戻す場合もでてくる
  - FSPやCROに若手のスタッフが多く、経験不足が生じている
  - 近年多くの治験依頼者が、多くの主要な役割の自社リソース化につなげるよう、低コスト地域にある自社CDMセンターへ投資している

- **業界が発展し、新しい選択肢が出現するにつれて、様々な理由で低コスト・低スキルのビジネスモデルの実行可能性に課題が生じている**
  - 定着した低コスト拠点での賃金上昇
  - 臨床試験の複雑さが増大し必要なスキルレベルがあがったため、既存のFSP/CROのリソースのトレーニングではそのレベルに到達できない
  - データサイエンスや統計解析などの高度なスキルを持った人材の採用競争の激化
  - 単純で反復的な手作業を排除するRPAとIPAに基づく自動化ソリューションの信頼性と費用対効果
  - one-size-fit all及び予測可能なプロセスが例外となりつつある臨床試験デザインの多様化
  - 初期品質にフォーカスすることで、運用コストの削減から非品質コストの削減にシフト

**長期的に存続するためには、提供できるサービスを進化させ、ビジネスモデルを見直し、新しい認識に合わせる必要がある**

- サービスプロバイダーの中には、resource enabledモデル(中核ではない役割をサービスプロバイダーが担うビジネスモデル) に力を入れ続けるプロバイダーもいる一方で、将来のモデルをより根本的に再考するプロバイダーもいる。
- プロバイダーは、治験依頼者のプロセスやシステムの専門知識や深いノウハウを活用した、技術主導型のイノベーションを提供することで、ベネフィットを得ることができるとも思えない。
- technology enabled modelとプロセスへの投資が長期的には相殺されることが期待される。



# 従来型のresource enabledモデル

- ここ10年程の間で、ほとんどのDM業務が治験依頼者からオフショア FSPに委譲され、FSPもスタッフの技能を向上させたが、データベースロックやデータ分析のようなクロスコラボレーションを必要とするタスクの中には、まだFSPに移行されていないものもある。
- 変化のスピードは加速することが予想されるので、resource enabledサービスプロバイダーは、治験依頼者のCDSへの方向転換にリソースの需要を合わせていく中にチャンスがある。
- **この方向転換にはパラダイムシフトが必要。**
  - 新しい臨床試験アプローチや最先端技術に対してスタッフのスキルアップをするだけでは、コンピテンシー、基礎知識、ソフトスキルの変化に適応するのに十分ではない。
- **resource enabledサービスプロバイダーに今後求められること：**
  - 治験依頼者の技術やプロセスの変化の速度に合わせて、スタッフの役割の発展。
  - 治験依頼者の要求に確実に答えるため、引き続き社内のスタッフのトレーニングへの投資。
  - AI/MLなどの技術の専門家の採用。

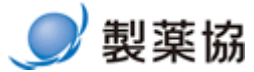
Tasks	Task Assignments to FSPs		
	2010		2020
	FSPs	In-house & Captive	FSPs
Database Development	Most	Most	All
Specification Writing & UAT	Most	All	All
Data Entry (CRF or Paper diaries)	All	Most	All
Study Planning	None	All	All
Discrepancy Management	All	All	All
Study Management	Few	All	All
Dataset Creation	Most	All	All
Coding	Most	All	All
Database Lock	Many	All	Most
Analytics and Reporting	None	All	Many
Quality Control	All	All	All
External Data Reconciliation	Many	All	All
Mainly Out of Scope or Future Planned		In scope or Limited Scope	

Part 3, Fig 9. The FSP Task assignments

治験依頼者は将来性のあるサービスプロバイダーを求めている。サービスプロバイダーは適切な採用活動および人材育成・確保への投資が必要になる。

# technology enabledビジネスモデルへの移行

SCDM Part3 6.2) Business models of the future



	パターン1 (従来resource enabledモデルのサポートしていたプロバイダー)	パターン2 (technologyのみを提供していたプロバイダー)
プロバイダーの特徴	<p>自身のテクノロジーを介してより高い予測可能性と品質を備えたサービスを提供。 治験依頼者のシステムに追加して使用できるサービスとしてのソフトウェア (SaaS) ソリューションを活用</p> <ul style="list-style-type: none"> <li>▪ ML based SDTM Mapping solutions</li> <li>▪ SDTM compliance QC tools</li> <li>▪ Metadata based eCRF Design creation tools</li> <li>▪ EDC design QC tools</li> </ul>	<p>自身の技術の知識とデリバリーチームの専門知識を活用して、サービスを提供。</p>
メリット	<ul style="list-style-type: none"> <li>▪ 治験依頼者は、容易に革新的技術を活用することができる。</li> <li>▪ 一部のプロバイダーは、ソリューションのライセンスを含め、IRT、EDC、eCOAなどのより複雑な第三者ソリューションに関するエンドツーエンドのサービスも治験依頼者に提供している。</li> </ul>	<ul style="list-style-type: none"> <li>▪ 治験依頼者は、中核技術やデータを社内に保持し、業務上の試験固有な対応の全面的な管理をtechnology enabledプロバイダーに移すことができる。</li> <li>▪ 治験依頼者は、時間とコストがかかる実行プロジェクトに投資する前に、新しいソリューションをテストすることができる。</li> </ul>

- **技術とサービスの両方に高度に特化したプレーヤー＝イノベーターの存在**
  - RBM、DCT、eSource、監督型クリーニング、将来性のあるプラットフォーム、AI/MLを活用できる新しいモデルを開発しているような組織
- **プロバイダーは、技術的な支援だけでなく、サービス力を生かして、技術とサービスの両方に高度に特化した上述のプレーヤーと協力していく必要がある。**

全てのモデルは、社内あるいは協力者の関係を通じて、望ましい知識と能力を獲得することにより、発展型のresource enabledサービス又はtechnology enabledサービスのいずれかに転換する必要がある。  
モデルに関係なく、本Reflection paperで示したCDSの要件や役割は変わらない。

# 役割の進化の影響

[目次に戻る](#)

# CDMからCDSへの役割の進化

CDMの焦点の概要	CDSの焦点の概要
データインテグリティ	データクオリティ
Quality controls	Quality by Design
アウトプットを重視したロジカルシンキング	アウトカムを重視したクリティカルシンキング
ランダム化比較試験	Adaptive and master protocols
医療機関で生成されたデータに焦点	DCT由来のeSourceデータに焦点
試験間の標準的なプロセス (one size-fits-all)	各試験に合わせたリスクに基づいたプロセス (focus on what matters)
少量のデータとソース	大量のデータとソース
単純なデータフロー	複雑なデータフロー
ベンダー管理	ベンダーオーバーサイト
データのクリーニング	データのレビュー、タグ付け、除外、キュレーション
プロジェクト管理	部門横断的なリーダーシップ
臨床試験標準	臨床試験及びヘルスケア標準
臨床試験データ	臨床試験及びヘルスケアデータ
従来のプログラミング (SQL、C#、SASなど)	ML (Python、Rなど) 、non-SQL
標準的なデータ照会 (SQLなど)	高度なデータ照会 (non-SQLなど)

- CDMに限らず臨床試験に関係するすべてのステークホルダーは、進化する規制、技術、及び臨床試験アプローチに適応し、役割を進化させていく必要がある

- 自分たちの向かうべきCDSの将来のために、最初にすべきこと

## ■ 自社の組織に合わせたCDSの将来像を明確に定義し、目標を定めること

- 現在会社が置かれている状況も考慮事項となる
  - 会社規模、地理的要因
  - CDMの役割、範囲、組織構造
  - 企業文化（デジタルリテラシー、誤りへの耐性、機敏性、イノベーター対フォロワーなど）
  - 合併・買収戦略
  - チーム構成 部門間の依存関係
  - 技術（メタデータレポジトリの利用可能性の有無）
  - 人材
  - 新たな機能と役割
    - 機能：スタートアップ、デザインセンター、デジタルイノベーション
    - 役割：クリニカルデータサイエンスヘッド、チーフデジタルオフィサー、デジタルインテグレーションスペシャリスト、Trial Innovation Lead

## ➤ 以下の側面が含まれた進化へのロードマップとチェンジマネジメント計画も必要

### ■ 人材戦略

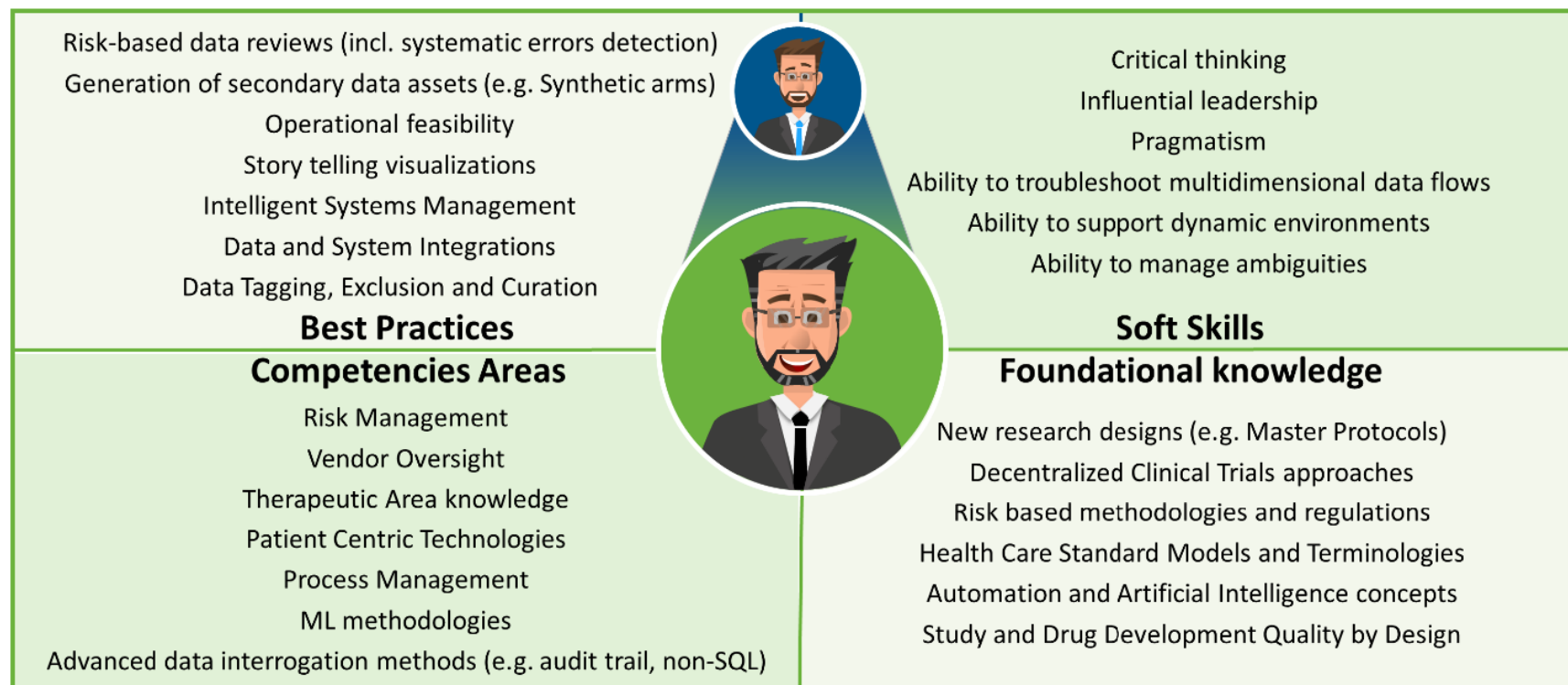
□ 職務分類、キャリアラダー、人材獲得、報酬、オンボーディング、研修、スキルアップ、メンタリング、評価

### ■ CDS運用モデル（例：インハウス、アウトソーシング、FSPモデル）

### ■ 内部及び外部のステークホルダーとの関係性構築

### ■ カルチャーを含めた組織改革

- CDMにとって、ソフトスキル、基礎知識、ベストプラクティス、コンピテンシーを進化させるチャンスであり、以下のような能力を期待されている



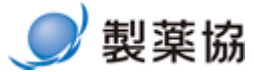
Part 3, Fig 10. CDS role framework

- このフレームワークは、CDMからCDSに向けた進化の出発点であり、今後さらなる技術や規制の進展に適応して発展させていく必要がある。

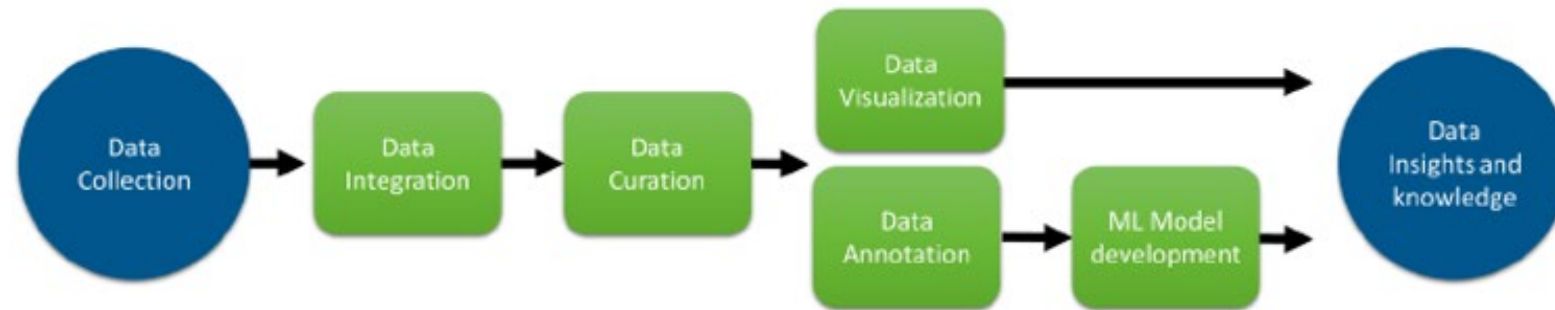


# 今後出現する可能性のある技術的・非技術的な役割の一例

SCDM Part3 7) Impact of the role evolution



## ➤ データからインサイト・ナレッジを創出する一連のCDSのフローを支える役割



Part 3, Fig 11. From data to knowledge data flow

### ① Data Integration Specialist

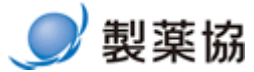
- ウェアラブルデバイスなど新しいデータ型が定期的に臨床試験に入ってくることに対応する
- 科学技術専門家と連携し、新しいデータ型を探求する
- 例：高速・大容量のセンサーやウェアラブルにより生成される時系列データをどのように統合するか？ EDCデータと同じ用には扱えない

### ② Data Mining and Profiling Specialist

- データマイニングとプロファイリングはデータ解析の最初のステップ
- 大規模なデータセットを探索して、初期のパターン、特性、及び注目すべきポイントを明らかにする
- データの重要な傾向や調査すべき重要なポイントの全体像をより詳細に描き出すことを目的とする
- データプロファイリングはまた、作業時間を短縮し、最初からより有用で実用的な洞察を見出すとともに、より良い分析を行うための明確な道筋を提示することによっても役立つ

# 今後出現する可能性のある技術的・非技術的な役割の一例

SCDM Part3 7) Impact of the role evolution



## ③ Data Curator

- 匿名化、統合、整理、探索が含まれる
- RWDからRWEなどの適切な二次利用データとしての価値を生成するために、そのインテグリティとクオリティを客観的に確認する

## ④ Data Annotator

- ラベリングのプロセス。テキスト、ビデオ、画像など、さまざまな形式で利用できるデータとする
- 教師ありMLの場合、機械が入力を容易かつ明確に理解できるように、ラベル付きデータセットが必要である

## ⑤ Data Visualization Expert (“Storyteller”)

- データを可視化し、情報やデータを視覚的に表現する

## ⑥ ML Model Builder

- 統計学的手法やプログラミング手法を活用してML modelを開発・教育する
- ML modelのトレーニングとテストのためにキュレーションされたデータセット・注釈付きデータセットを選択する

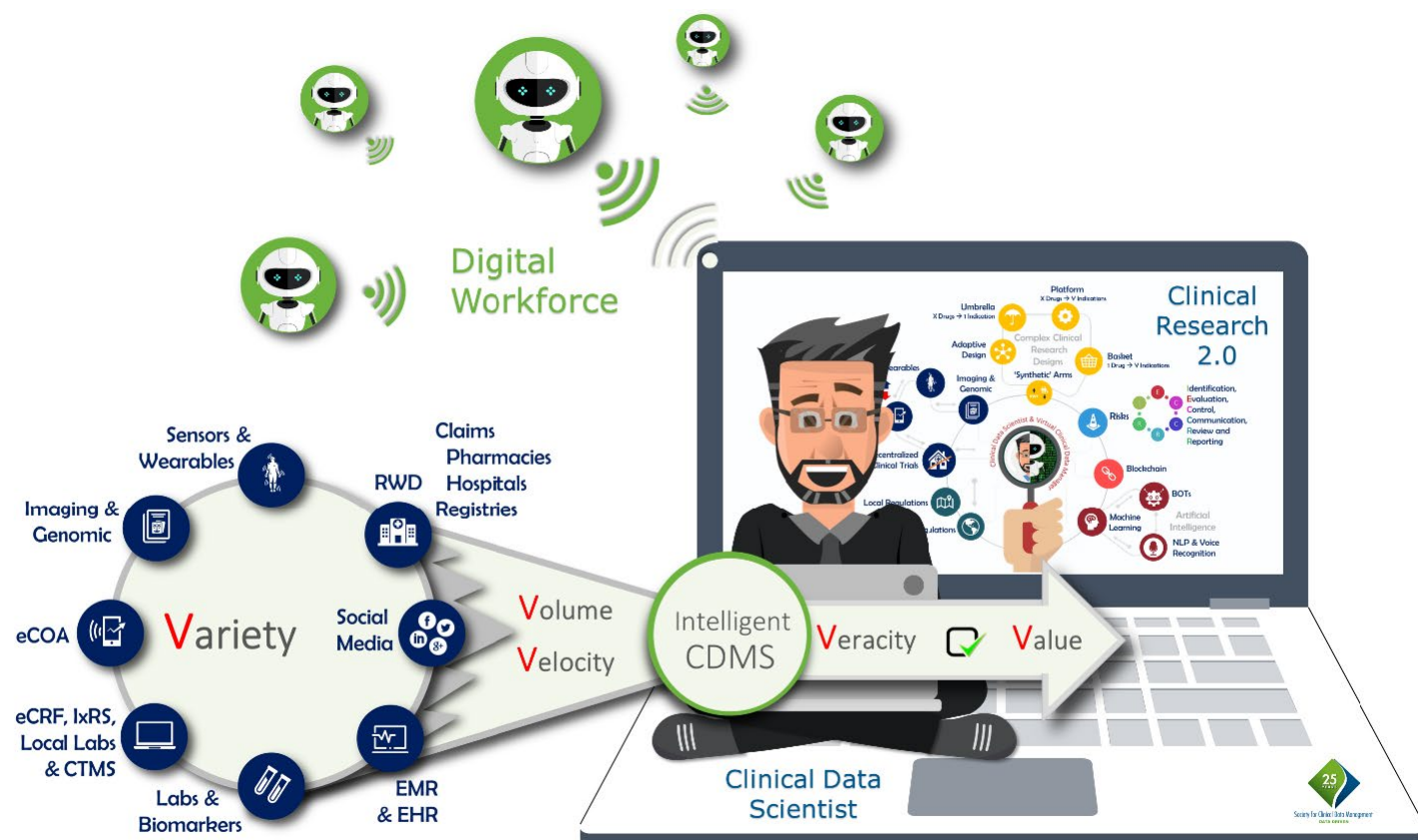
# さいごに

[目次に戻る](#)

本書は、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の紹介に特化しました。

CDMは、本書が述べているCDSの役割を担うべく進化することが期待されており、これまでのCDMの範囲を超えた新たなスキルセット、コンピテンシーを身につけることが求められています。

製薬協DS部会ではCDMの進化に対応しみなさまの行動につながるような活動を引き続き検討していきます。



- 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) 、Part 2 (5 March 2020) 、Part 3 (31 Aug 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 からの許諾なしに如何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（～2021年2月まで）
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