

一般社団法人医療福祉情報システム工業会(JAHIS)

標準化推進部会業務報告会

# リアルワールドデータの利活用と標準化 －現状と今後への展望－

2023年 3月 9日(木)

岡田美保子

一般社団法人医療データ活用基盤整備機構



# COI開示

発表者：岡田美保子

所 属：一般社団法人医療データ活用基盤整備機構

私が発表する演題について開示すべきCOIはありません

口演中に述べる意見は報告者個人のものであり、所属する組織・参加活動する組織の公式見解ではありません

# リアルワールドデータ (Real World Data: RWD)

# RWD：国際的に合意された、ただ一つの定義は存在しない

## 定義の例

ISPOR(International Society for Pharmacoeconomics and Outcomes Research)	意思決定に利用される、伝統的なRCTsにより収集されたものではないデータ
ABPI(英国製薬工業協会)	(このガイダンスでは) RWDは、通常の診療業務で起きていることを記述する、非介入の方法で得られるデータを指す
FDA(米国食品医薬品局)	様々なソースから日々収集される(routinely collected)患者の状態や診療の提供に関するデータ※

※ The Framework for FDA's Real-World Evidence Program (FDA, Dec 2018)

医薬品の安全性と、限定的範囲での有効性(ラベリング情報の適応追加・変更、用量や用法、投与経路の変更、等)の評価におけるRWEの活用が対象

※ RWDはスコープが広く議論のコンテキストにより範囲が異なる

様々な定義に共通的にみられる特徴

- RCTにより取得されたものでなく日常的な患者状態や提供される診療に関するデータ
- 非介入的な方法で、観察的に得られる
- データソースとしてEHR、管理的データ、医療費請求に係るデータの他セルフモニタリング機器から得られるデータや患者レジストリーも含まれる

## **ICH Reflection on “GCP Renovation”: Modernization of ICH E8 and Subsequent Renovation of ICH E6, January 2017**

ICH GCP (ICH E6ガイドライン)の改訂にむけてRWD関係の部分として  
現行GCPのフォーカスを維持し、より広い研究のタイプ・データソースを想定した議論

- 治験の研究データ分析は、伝統的な方法だけではない
- 収集・蓄積されているリアルワールドで得られる大量データの活用
  - 例) EHR(電子カルテ)、退院サマリ、診療報酬請求データ、レジストリ
- 日々の患者ケアに臨床研究を統合化する努力
  - 例) Case Reportの電子カルテ・インタフェースでデータ収集の重複を最小化

リアルワールドのソースに世界的に  
合意された標準はない

➢ 大規模医療記録を治験に活用する  
際の恐らく唯一にして最大の課題

## ICH E6の改訂経緯・状況

- 1996年 ICH-E6 「Guideline for Good Clinical Practice (GCP)」の合意
  - 治験実施の国際基準
  - 日米欧で規制当局に提出される治験データを収集する際に遵守を要する
- ICH E6 (R2) 2016年11月 Step 4 Final Guideline
- ICH E6 (R3) 改訂作業中

### E6(R3) RWDに関わる部分の議論

多様な臨床試験デザイン、データソースへの柔軟な対応(Flexibility)

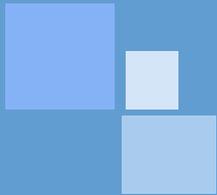
- 疾患レジストリ、医療Data BaseをはじめとするReal World Data(RWD)、Decentralized Clinical Trial(DCT)等の承認申請への活用
- 活用による開発期間の短縮, コスト低減
- 活用のためのデータ、試験の信頼性確保の基準(被験者保護を含む)の設定
  - 依頼者/責任医師/医療機関の信頼性確保の責任範囲の明確化 等
- 活用するための要件、指針の設定
  - 得られるエビデンスの位置づけ、Non-traditional 介入試験の考慮事項、効率的な運用に繋がる具体的な基準、被験者へのリスクに応じた柔軟な規制 等

# やや広い範囲のReal World Data※

	研究目的		非研究目的			
	試験データ	レジストリー	日々収集されるRWD 議論の中心はこの範囲			調査データ
一次目的	治験、その他の臨床試験	観察研究	電子カルテ 医療記録	DPC、レセプト等	セルフモニタリング、 PHR、ウェアラブルデバイス等	国民調査 病院調査 患者調査等
データ収集	CRFを用いた収集	患者レジストリ、臨床DB、ゲノムDB、etc.	記録の蓄積	所定の様式	ソースにより多様	調査様式による収集
二次目的	これまでは適用されないことが多かった	観察研究利用				

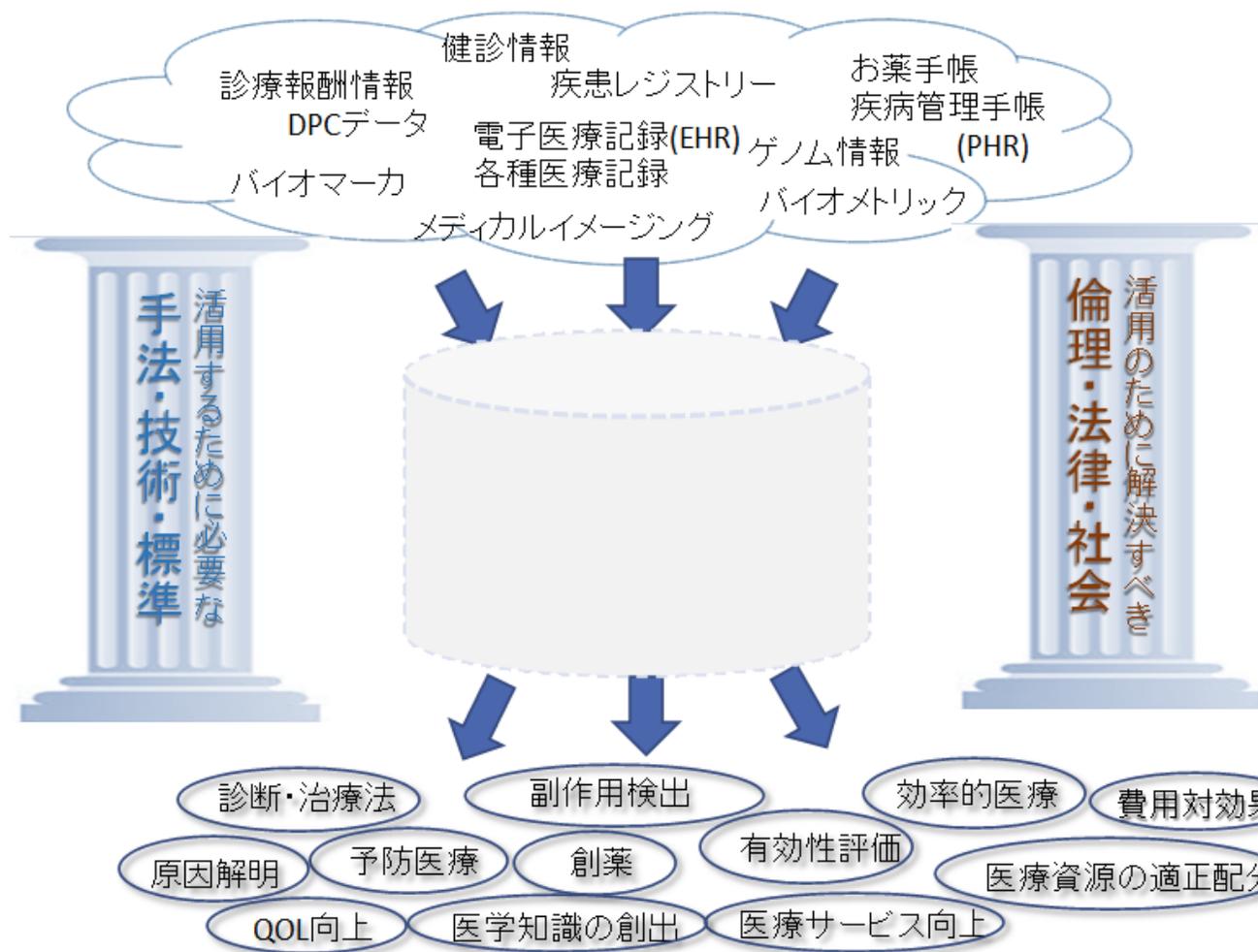
※いずれかの一つの公式定義によるものではない

# 国内リアルワールドデータの実際・課題



# 日本におけるデータ2次利用の課題

データソース



- 手法・技術・標準
- 倫理・法律・社会

以下、主として前者の観点から

効果

## リアルワールドデータ(RWD)

- 国際的に合意された、ただ1つの定義が存在するわけではない
  - およそEHR (Electronic Health Record)や医療費データ等、日々の医療提供で生成、収集されるデータが中心
  - セルフモニタリングデバイス等から収集されるデータや患者レジストリーデータも
- ※より広義に捉えることも少なくない

## 国内における主なリアルワールドデータのソース



## 【参考】リアルワールドデータ ≠ ビッグデータ

両者は同義ではないが臨床研究目的ではあえて厳密に使い分けしていないことも多い  
ここではリアルワールドデータを対象とする



# 国内で活用されている大規模データベースの例

## ■ 国が主導する大規模データベース

- ① MID-NET: 厚生労働省/PMDA
- ② NDB (National Data Base): 厚生労働省

## ■ 医療機関の診療データに基づいたデータベース

- ③ CCT: コンバージェンス・シーティー・ジャパン株式会社Convergence CT Global Research Network (CGRN)
- ④ MDV: メディカル・データ・ビジョン株式会社EBM Provider

## ■ 健康保険組合からのレセプトに基づいたデータベース

- ⑤ JMDC: 株式会社日本医療データセンターMedical Data Bank
- ⑥ JammNet: ジャムネット株式会社

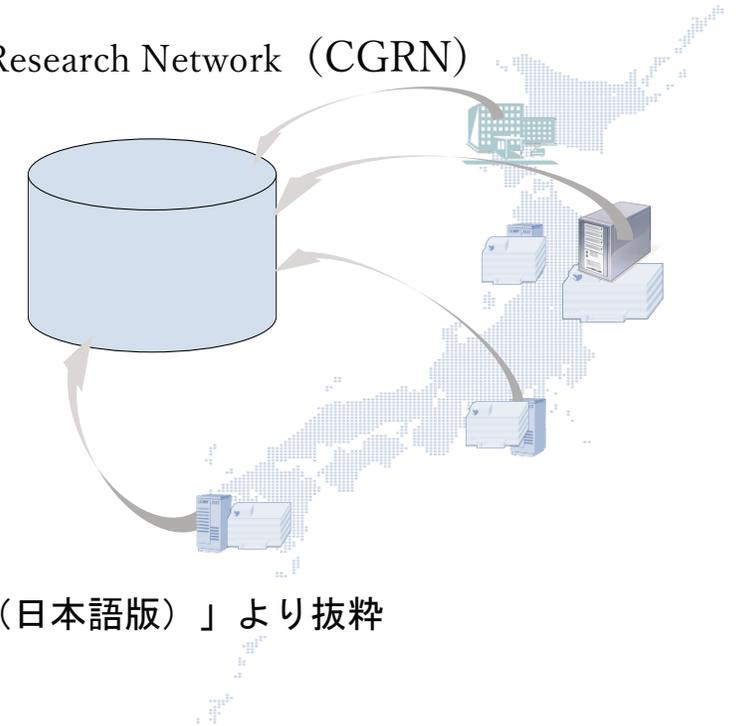
## ■ 調剤レセプトに基づいたデータベース

- ⑦ メディトレンド: 株式会社メディカルフロントメディトレンド
- ⑧ IMS NPA data: 株式会社アイ・エム・エス・ジャパン
- ⑨ JMIRI: 株式会社医療情報総合研究所処方情報データベース
- ⑩ 日本調剤株式会社処方箋データベース

①～⑩は日本薬剤疫学会「日本における臨床疫学・薬剤疫学に応用可能なデータベース調査結果（日本語版）」より抜粋  
[http://www.jspe.jp/mt-static/FileUpload/files/JSPE\\_DB\\_TF\\_J.pdf](http://www.jspe.jp/mt-static/FileUpload/files/JSPE_DB_TF_J.pdf)

## ■ 学会・国立病院・日本医師会などが主導するデータベースの例

- NCD (National Clinical Database): 一般社団法人National Clinical Database (日本外科学会等の外科系諸学会)
- NCDA (National Hospital Organization Clinical Data Archives): 国立病院機構診療情報集積基盤
- JROAD (The Japanese Registry Of All cardiac and vascular Diseases): 日本循環器学会 循環器疾患診療実態調査
- J-DREAMS: 国立国際医療研究センター・日本糖尿病学会
- J-CKD-DB (Japan Chronic Kidney Disease Database): 日本腎臓学会 包括的腎臓病データベース
- J-DOME (Japan medical association Database Of clinical Medicine): 日本医師会 かかりつけ医 診療データベース研究事業



	(オーダリング) 処方情報	調剤	医薬品コード	処方日	処方量	処方日数	注射剤情報	ワクチン接種情報	Kコード	処置・手術名	処置日	処置時間	入院日	入院中の薬物治療	退院時サマリ	診断名	診断名コード	(オーダリング) 検査	検査コード	検査結果	画像診断結果	誕生年	誕生月	誕生日	身長	体重	血圧
①MD-NET	○	△	○	○	○	○	○	△	○	○	○	×	○	○	△	○	○	○	○	○	×	△	△	△	△	△	○
②NDB	○	○	○	△	○	○	○	×	○	○	○	×	○	○	×	○	○	○	○	×	×	○	△	×	△	△	△
③CCT	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	△	○	○	○	○	○	○
④MDV	○	×	○	○	○	○	○	×	○	○	○	×	○	○	○	○	○	○	○	△	△	△	△	△	△	△	×
⑤JMDC	○	○	○	○	○	○	○	△	○	○	○	△	○	○	×	○	○	○	○	×	×	○	○	×	△	△	△
⑥JammNet	○	○	○	○	○	○	○	×	○	○	○	×	○	○	×	○	○	○	○	×	×	○	○	×	×	×	×
⑦メディトレ	×	○	○	○	○	○	△	×	×	×	×	×	×	×	×	×	×	×	×	×	×	○	○	×	×	×	×
⑧IMS NPA	×	○	○	○	○	○	△	×	×	×	×	×	×	×	×	×	×	×	×	×	×	○	×	×	×	×	×
⑨JMIRI処方 情報DB	×	○	○	○	○	○	△	×	×	×	×	×	×	×	×	×	×	×	×	×	×	○	○	×	×	×	×
⑩日本医薬総 合研究所	○	○	○	○	○	○	△	×	×	×	×	×	×	△	×	△	×	×	×	×	×	○	○	○	△	△	△

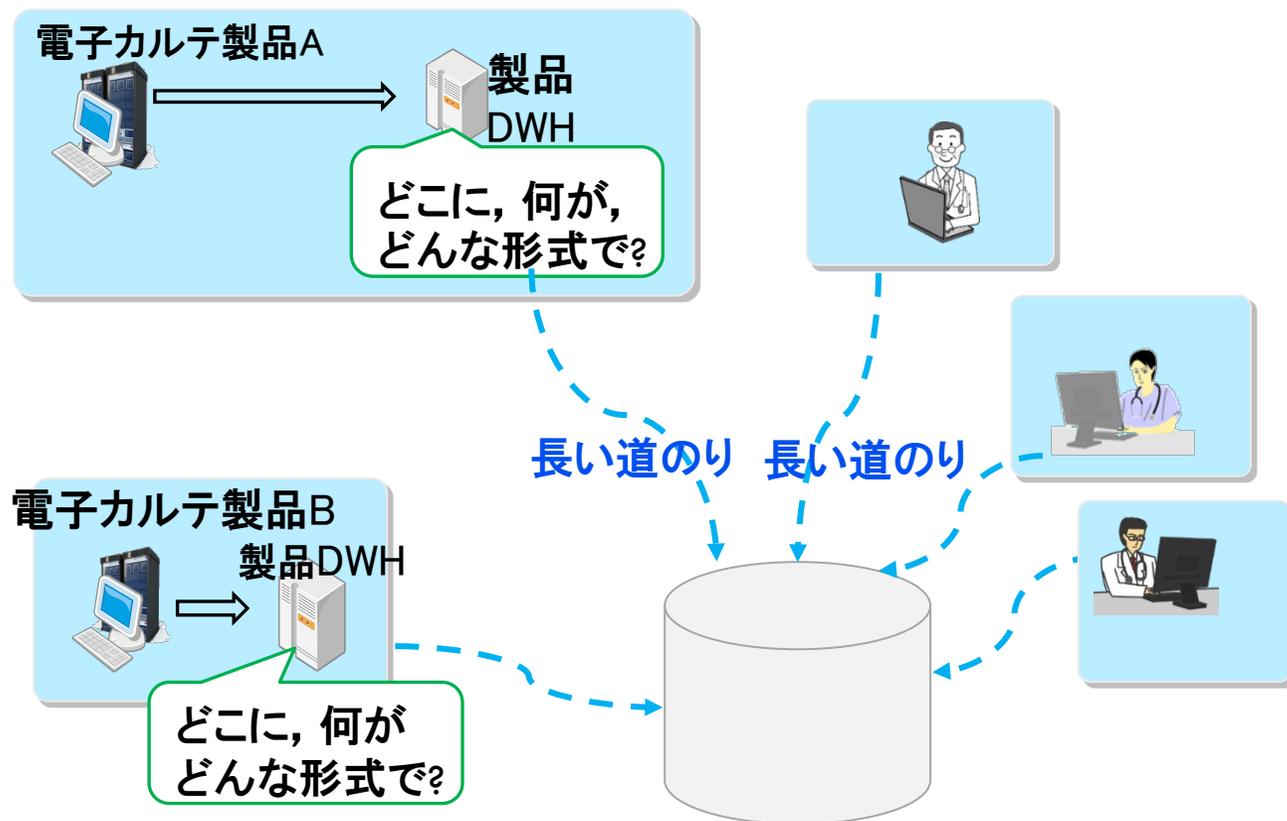
【引用】以下の文献より一部抜粋して引用

日本薬剤疫学会：日本における臨床疫学・薬剤疫学に応用可能なデータベース調査結果（日本語版）

[http://www.jspe.jp/mt-static/FileUpload/files/JSPE\\_DB\\_TF\\_J.pdf](http://www.jspe.jp/mt-static/FileUpload/files/JSPE_DB_TF_J.pdf)

※抜粋はあくまでも参考までである。引用される場合は、元資料を引用されたい

# 多施設からの診療データの収集



電子カルテ製品により何がどこに、どのような形式であるかバラバラ

- ▶ 手入力に頼らざるを得ない
- ▶ 臨床家の入力負担は増す一方



その都度データ収集  
整備に膨大な時間・費用・労力が

診療情報活用が進むとともに問題が顕在化 ➡ 課題解決に向けて様々な取り組み

J-CKD-DB構築事業（代表:日本腎臓学会理事長 川崎医科大学 柏原直樹教授）

- 有効な予防・治療戦略を立案、実施するためエビデンスに基づく治療指針（ガイドライン）の策定、改訂と普及、医療の質向上と均霑化の推進が必要
- 平成26年度補正予算厚生労働省事業として慢性腎臓病臨床効果データベース（J-CKD-DB）として平成27年10月1日からDB構築開始

### 自動的収集:

- SS-MIX2標準化ストレージを用い共通プログラムで多施設から自動収集
- 登録前に参加施設において検体検査項目と医薬品に標準コードを割当

### 一年分を纏めて一回で収集:

- 1年間分のデータを纏めて抽出し、1回でデータベースに登録

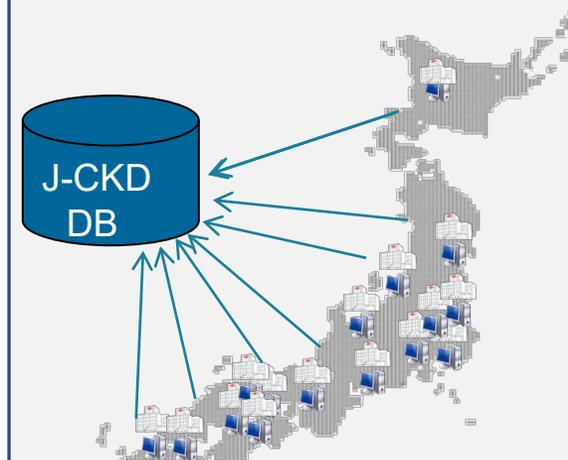
### 匿名加工:

- SS-MIX2からの抽出時に匿名加工
- オプアウト方式

### データクレンジング:

- 川崎医科大学に設置された事務局でデータベースを運営・維持
- 事務局にてデータクレンジングを実施

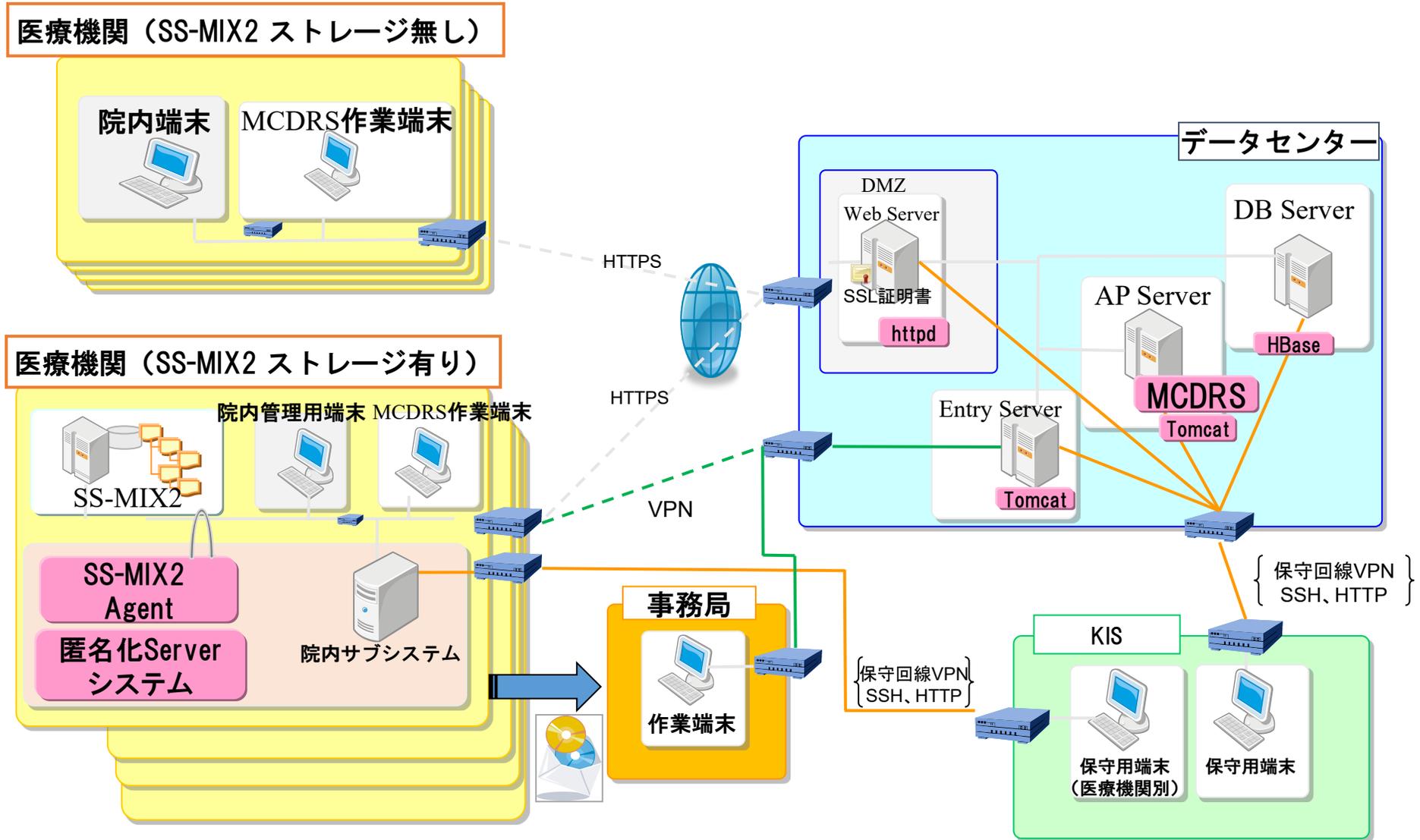
### J-CKD-DB



登録件数:  
15施設、約140,000件  
(2019年10月現在)

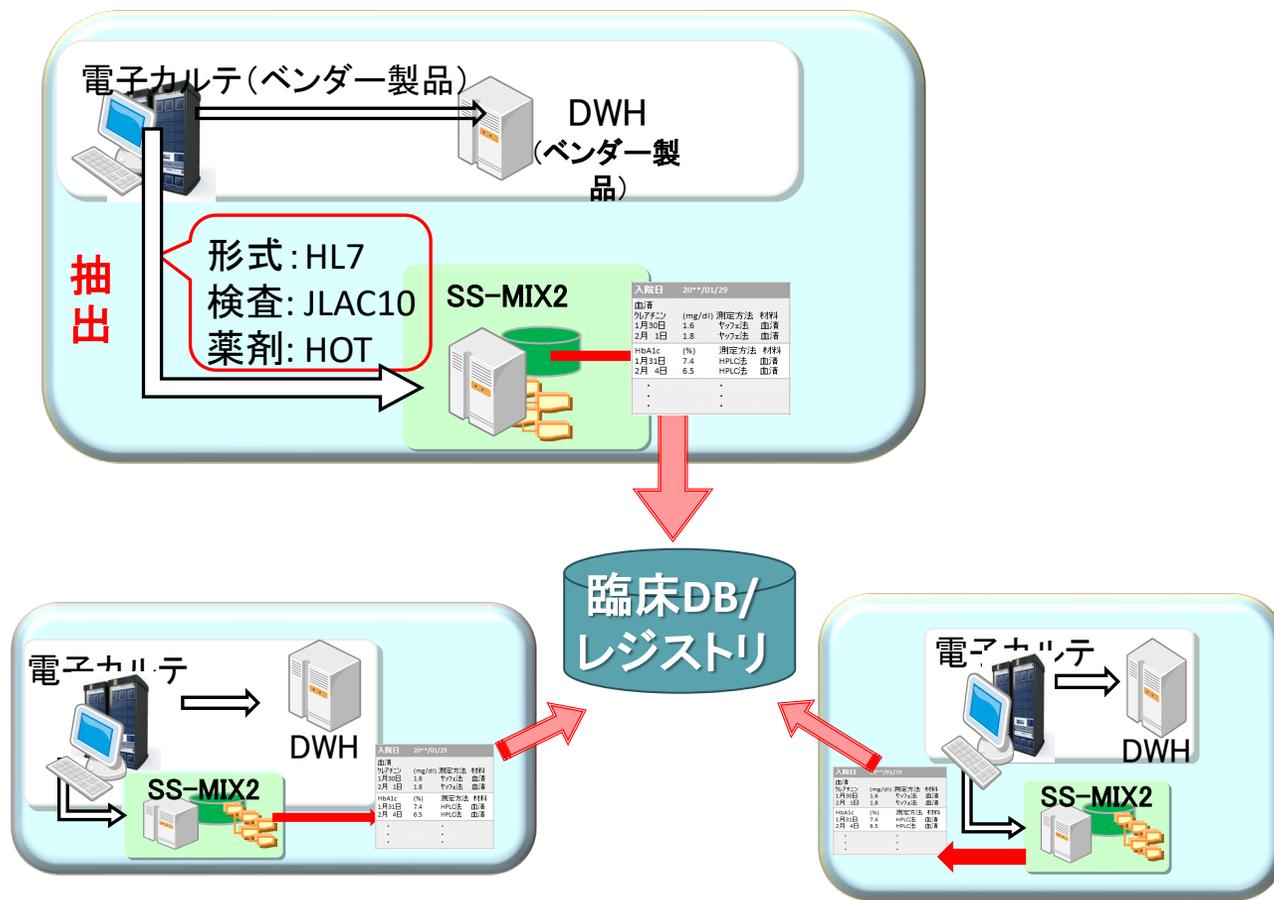
旭川医科大学、新潟大学、金沢大学、筑波大学、東京大学、横浜市立大学、名古屋大学、和歌山県立医科大学、京都大学、島根大学、岡山大学、川崎医科大学、香川大学、高知大学、九州大学

# J-CKD-DB構築システム



# 多施設から研究目的で診療データを収集する場合の課題

例) SS-MIX2を利用した収集



電子カルテ製品により何がどこに、どのような形式であるかバラバラ

- ➡ 手入力に頼らざるを得ない
- ➡ 臨床家の入力負担は増す一方



その都度データ収集

整備に膨大な時間・費用・労力が



可能な部分の共通化の努力

## 【例】SS-MIX2標準ストレージとテンプレート入力の併用(ハイブリッド型)

- SS-MIX2標準化ストレージにあるデータは自動抽出
- それ以外の必要情報は電子カルテテンプレートから入力
- 両者を統合してデータベース登録

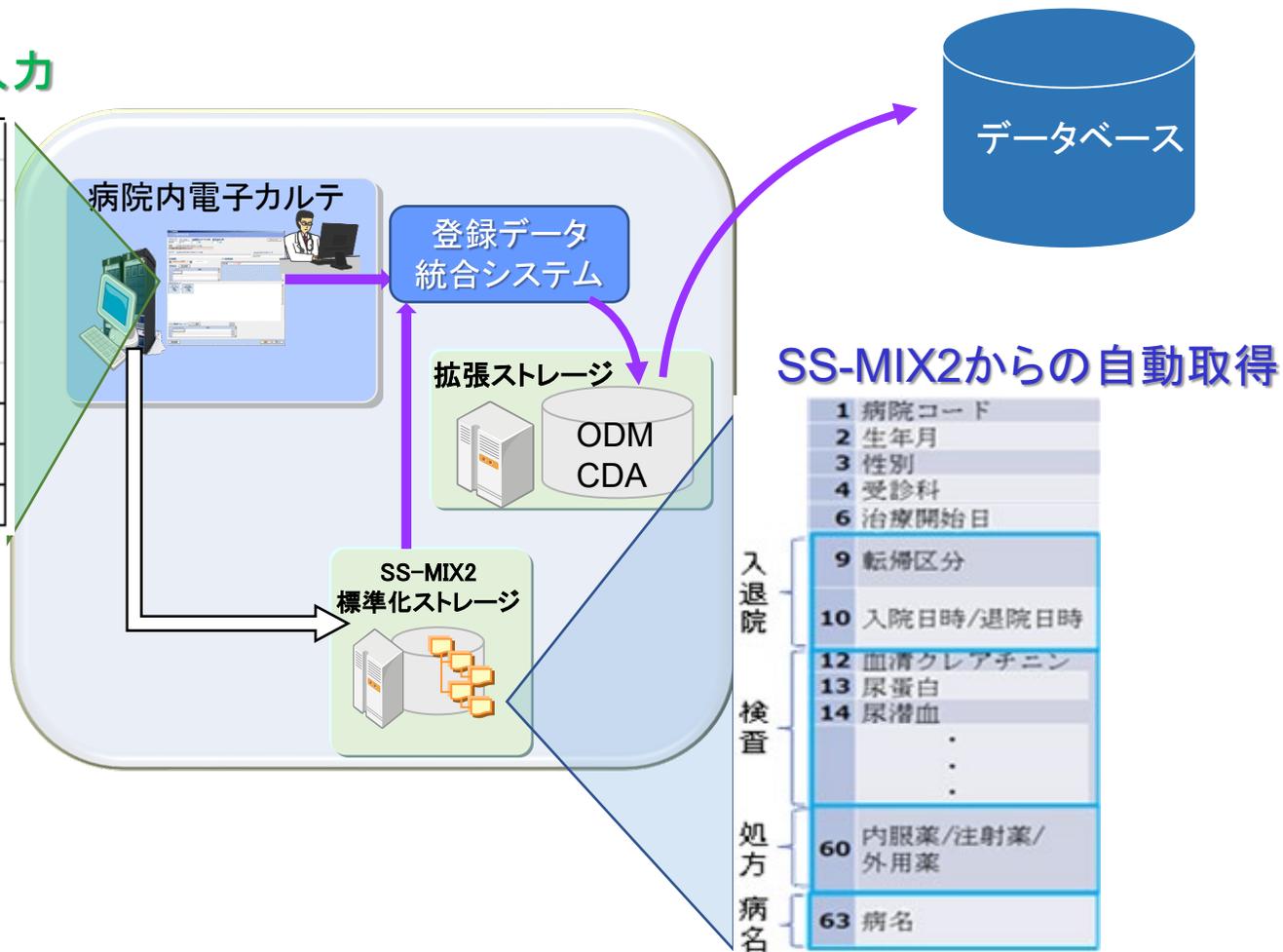
### 電カルテンプレートからの手入力

手入力できないと  
取得できない

身長
体重
収縮期血圧
拡張期血圧
原疾患
家族歴
喫煙
イベント情報
既往歴(患者由来情報のみ)
.....



- 電子カルテの診療情報の多くは標準化されていない、入力されていない、テキストで記載しているなどプログラムによる抽出は困難
- テンプレートの機能は電子カルテベンダーにより異なる





# 国内におけるリアルワールドデータの課題

## RWDにおける薬剤に関する情報の課題

- 多施設から収集されるRWDでは薬剤識別に苦勞
- 蓄積データでは投与量の把握が困難、一定期間の投与量がほとんど計算できない
- 調剤薬局で調剤された医薬品が反映されているわけではない
- 実際に服薬した量はわからない
- 臨床家が研究上で必要とする疾病・症候等に基づいた医薬品の選定が容易でない

## 薬剤識別の問題

- 病院では流通(GS1)、レセプト(レセ電算コード)、電子カルテ(YJ等)と目的に応じたコードが普及
- 処方には基本YJコードがあるが、ローカルコード(ハウスコード)も多く病院内固有の運用
- 厚労省標準はHOTコードであるが、業務にHOTが利用されているわけではなく普及が進まない
- 各施設には院内の独自ルールがあり標準コードか否かとは別の問題がある
- コードが特定できないとき名称を頼るが、マスター上の薬剤の名称は改変されがち、表記揺れが大きい

## 検体検査データに関する標準化の現状と課題

### ■ 単位の問題

例) WBCの単位に  $\times 1000/\mu\text{L}$ 、 $10^3/\mu\text{L}$ 、 $\times \text{千}/\mu\text{L}$ 、等古くから知られている問題で相変わらず解消されていない

### ■ 施設ごとの検査値の分布

基準値が揃っていても分布が施設で異なる場合も

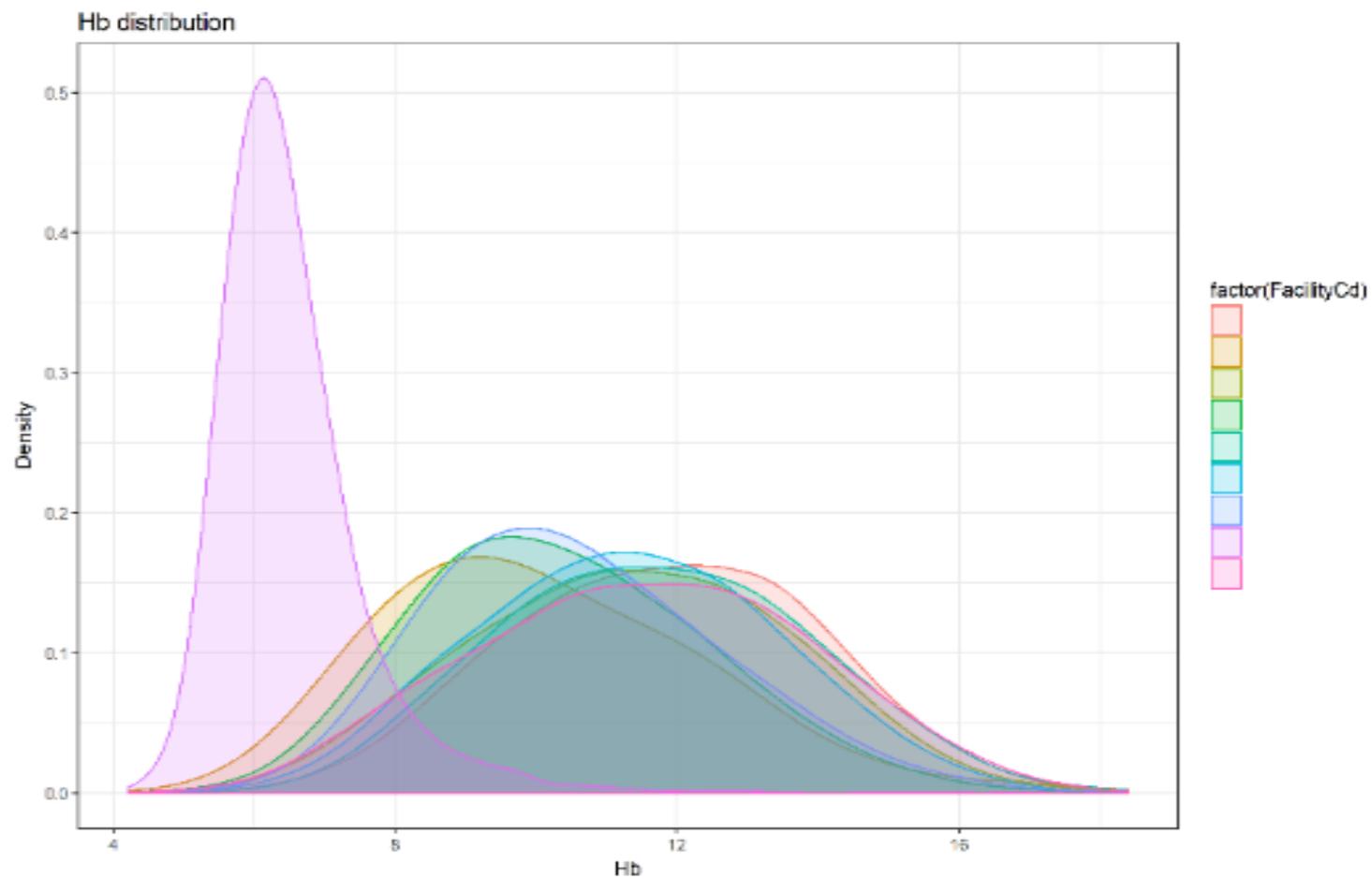
### ■ 検体検査の標準コード

- 標準をJLAC10コードとしているが、院内マスターでは採用されていない
- 厚労省事業によるDB構築では施設ごとのJLAC10の割当が必要
- 割当が不完全であるとDB上で、JLAC10とローカルコードが混在してしまう
- JLAC10では完全に確定できない場合がある  
→ JLAC11への移行が進んでいる

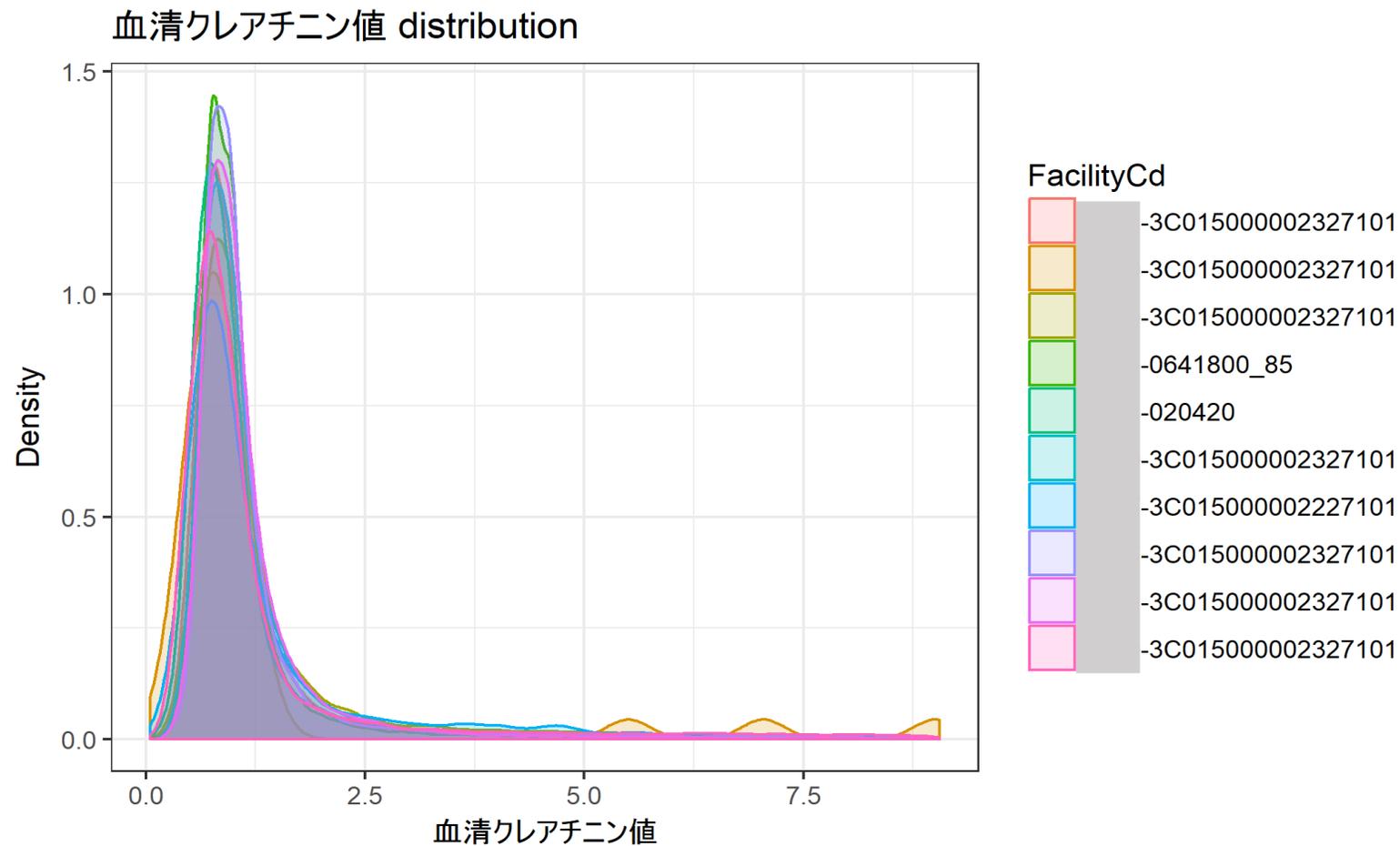
# 蓄積された電子カルテデータの研究活用上の課題(例)

- 薬剤に関するデータの問題
  - ・ 院内ルールがあり、標準コードか否かとは別の問題がある
  - ・ 用法の標準が普及していない。蓄積データを用いた観察研究・疫学研究で重要課題
- 必須項目であり電子カルテにあるが、データ取得には通常、手入力を要する項目
  - 例) 体重、血圧
- 疫学研究で頻用されるものの、患者プロフィールに通常あって標準化されていない項目
  - 例) アレルギー、予防接種、手術歴
  - 例) 家族歴、社会歴、喫煙、等
  - 交絡要因や交互作用要因となる他の説明要因に関する情報、リスクに関する情報
- 観察研究に重要であるが、用語の意味さえ標準化されていない項目
  - 例) イベント、転帰情報
- DPCにあるが電子カルテデータと統合し標準的に管理されていない項目
  - 例) ADL、手術情報
- 研究目的では死亡情報が入手できない
- 薬剤投与のリーズニングは電カルからは難しい(なぜそうしたか、しなかったか、中止したか等)

# Hbの分布

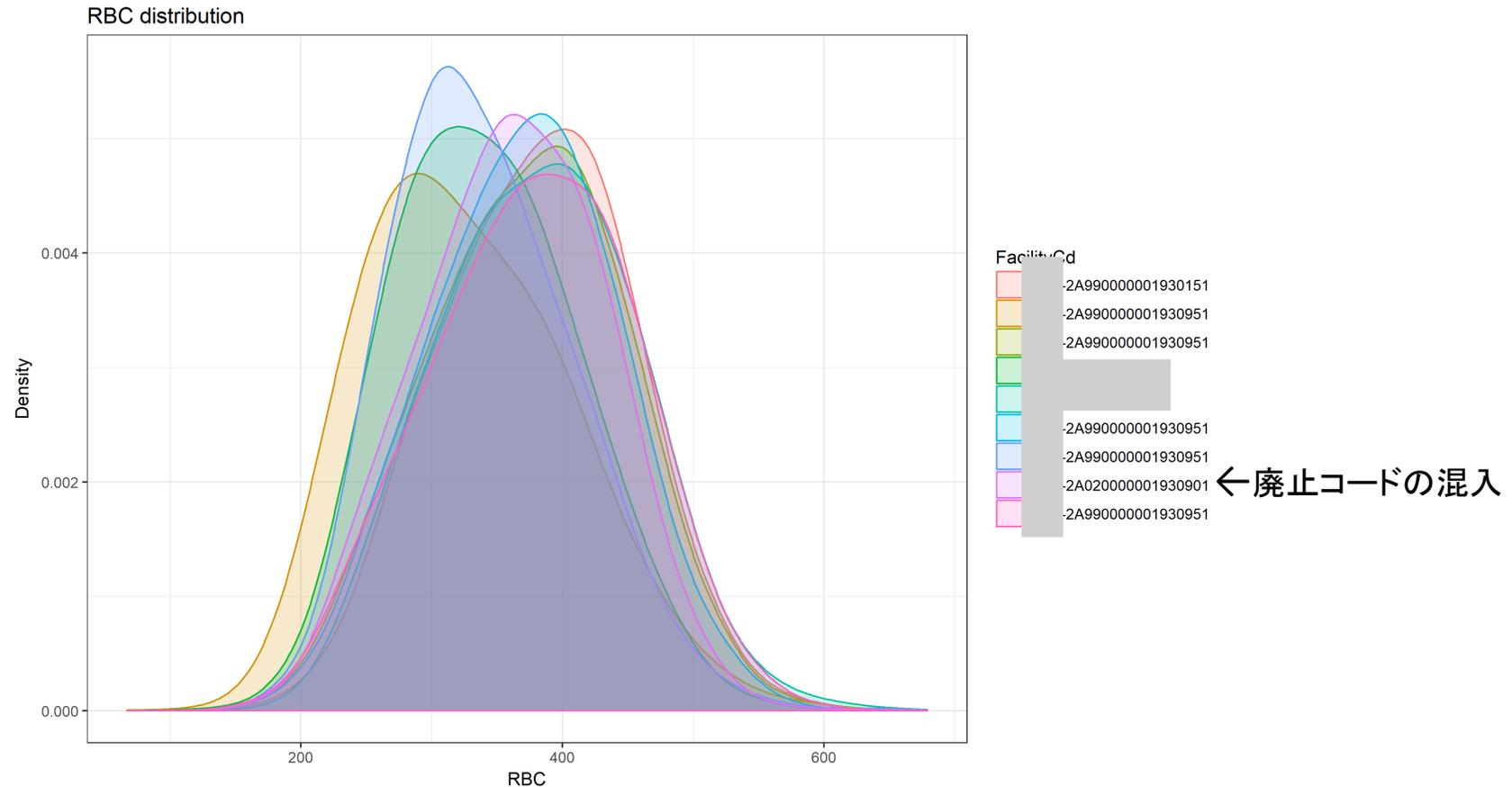


# 血清クレアチニンの分布



- ・ クレアチニンは、ほぼ補正無しで使えるレベル  
施設別分布を可視化することで検証可能

# RBCの分布



- CBC関連は標準物質の安定性の問題あり
- 残念ながらJLAC10レベルでは測定法の違いは判断不能  
(恐らくすべてインピーダンス法と考えられるが会社が異なる)

# 薬剤マスターにみられる事例

院内コード	名称(半角スペース削除後に全角変換)
VFE01110	・ブイフェンド錠50mg【禁運転】
ARI0432	アリクストラ皮下注1. 5mg(0. 3mL)【予防】
1026166	アムロジピンOD錠5mg
1031340	アムロジピンOD錠5mg
1025980	アムロジピンOD錠5mg
1026129	アムロジピンOD錠5mg
1026176	アムロジピンOD錠5mg
1026742	アムロジピンOD錠5mg
1026178	アムロジピンOD錠5mg
1025995	アムロジピンOD錠5mg
1025904	アムロジピンOD錠5mg
1026160	アムロジピンOD錠5mg
1025902	アムロジピンOD錠5mg
1032034	アムロジピンOD錠5mg
1026210	アムロジピンOD錠5mg
1027159	アムロジピンOD錠5mg
1026759	アムロジピンOD錠5mg
1026219	アムロジピンOD錠5mg
1026013	アムロジピンOD錠5mg
1026081	アムロジピンOD錠5mg
1025896	アムロジピンOD錠5mg
1023193	アムロジピンOD錠5mg
1026045	アムロジピンOD錠5mg
MIN12110	・ミノサイクリン塩酸塩錠50mg(ミノマイシン)
AMI2131	アミカシン注射液200mg/2mL(AMK)
AMI0632	硫酸アミカシン200mg(粉末バイアル)(AMK)
110706702	アミカシン硫酸塩注用100mg
110706702	アミカシン硫酸塩注射用100mg
EPI0832	エピルビシン塩酸塩50mg 末梢ポンプ禁
EPI0031	エピルビシン塩酸塩注 末梢ポンプ禁
EPI0031	エピルビシン塩酸塩注(エピルビシン)

名称に注意書きを含める?

これだけ異なるコードあり?

名称に補足、追記、編集?

注用と注射用で同じコード?

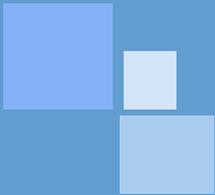
?

- 従来の医療データ標準はヘルスケアと規制領域では別々に 例えの話し

	RWD	臨床試験
担当部門の例	医療情報部門	ARO
関係組織の例	病院情報システムベンダー	製薬企業、規制当局
情報システムの例	電子カルテ	EDC (electronic data capture)
国際標準化団体の例	HL7	CDISC
標準用語等の例	ICD10対応標準病名	MedDRA
主な薬剤コード	YJコード、レセ電コード、GTIN、HOT、等	再審査コード、経済課コード、WHODrug、等
主な人材	医療情報担当者	ARO人材

- RWDのデータソースは診療現場に
- RWD/RWEのデータ標準はヘルスケアと臨床研究・規制領域にわたる
  - ➡ 両者の連携がこれまでになく求められている

RWD/RWEはどう価値があるのか



# RWD/RWEはどう価値があるのか？

## ■ 臨床試験における期待

- 観察研究で極めて高額なRCTに置き換えることが可能な試験があり得る
- 臨床試験におけるコントロールとしてのリアルワールドデータの活用
- 欧米ではPragmatic Clinical Trial\*を通じて、医薬品による疾患の重篤化リスク低減や予防を評価する取り組みが行われている

\*広範な患者集団を対象とし日常診療下における治療介入の有効性を示すためにデザインされた試験。日常診療で治療を受ける患者を対象に、一般的経験を持つスタッフが通常の機器を用いて治療を行い、状態に応じた柔軟な治療変更が許容される日常診療と変わらない状況で死亡や入院などの評価項目が評価される

## ■ リアルワールドデータによる実臨床における真の効果の評価への期待

- 実臨床によってのみ可能な、真の効果の評価がリアルワールドデータに期待される

## ■ 臨床試験以外の観察研究、疫学研究の広がり

# RWD/RWEは本当に価値があるのか？

## RWD/RWEはどう価値があるのか？

### Feasibility of Using Real-World Data to Replicate Clinical Trial Evidence

VB Bartlett, SS Dhruva, ND Shah, et al. JAMA Network Open, 2019; 2(10): e1912869

RCTがゴールドスタンダードと考えられる中、RWEにより有効性、安全性の評価に関心が高まっているが、RCT同様のクリニカルクエスチョンに観察データで答えられるか明らかでない

【方法】 2017年1/1～12/31の間のトップインパクト医学雑誌7誌対象

New England Journal of Medicine, Lancet, JAMA, The BMJ, Annals of Internal Medicine, JAMA Internal Medicine, and PLoS Medicine

Clinical Trialの論文のうち医療費請求データおよび/またはEHRの観察データを用いて

Clinical Trialの置き換えができると判断される論文を調査

【結果】 220件のClinical Trialの論文を特定

- ・ 観察データから評価可能と考えられる介入試験に関する研究 86件
- ・ 86件のうち評価可能なもの 62件
- ・ 62件中inclusion、exclusion criteria評価が可能と考えられたもの 45件
- ・ 45件中少なくとも1primary end pointを評価可能と考えられたもの 33件

【結論】 220件中33件(15%)のclinical trialが観察データの分析で置き換えることができると考えられた

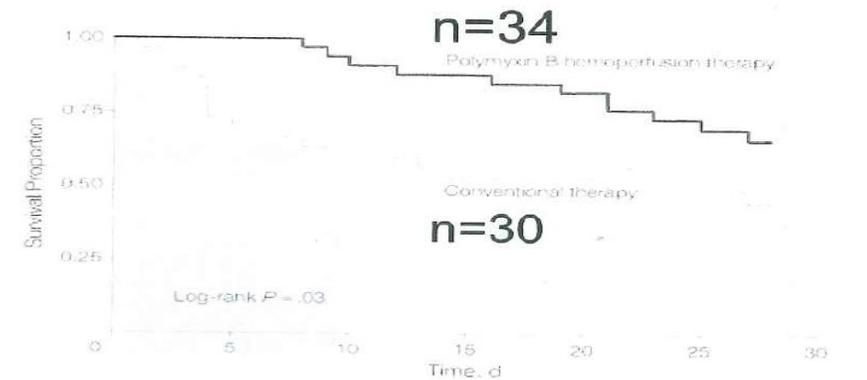
日本で1994年より肺血症性ショックに対してPMXの使用が保険適応となった。(しかし、この根拠となった試験は数十人のRMX使用患者の使用前後の比較で血圧上昇を認めたのみ。)

2014年、康永等のDPC研究(n=1,180)  
DPCデータベース 2007年~2011年の計34カ月、  
分析手法 Propensity score matching

その後、2015年フランスでのRCT (PMX群119人、対照群113人)にて、PMX群の方がむしろ死亡率が高い傾向にあった。  
2018年、北米でのRCTでも同様の結果

## 2009年、イタリアでの世界初RCT (n=64)

Early Use of Polymyxin B Hemoperfusion in Abdominal Septic Shock The EUPHAS Randomized Controlled Trial. JAMA. 2009;301(23):2445-2452

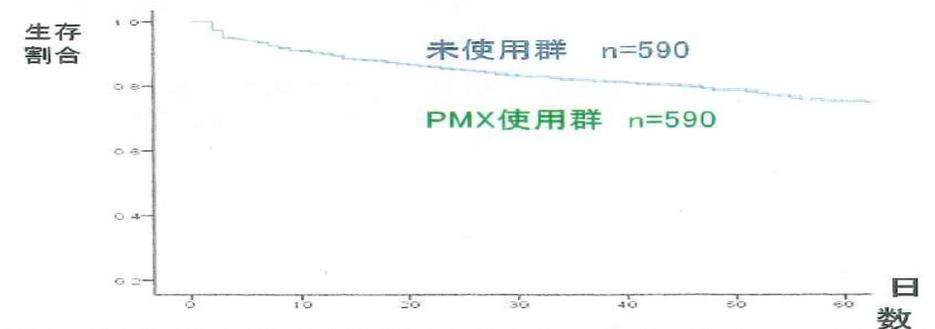


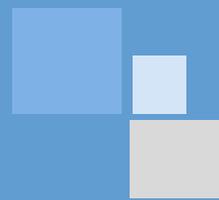
### 28日死亡率

PMX使用群 17.1 % (101人/590人)

未使用群 16.3 % (96人/590人)

P=0.696





何を満たさなければならないのか？

# RWDを用いた研究論文の撤回事例

2020年6月初めに著名学術誌で相次ぎ2件のCOVID-19 患者に関する研究論文が撤回される

## 【The Lancet論文の撤回】

**Retraction** - Mandeep R Mehra, Frank Ruschitzka, Amit N Patel. Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis, LANCET Vol. 395, Issue 10240, P1820, June 13, 2020

クロロキンやヒドロキシクロロキンを投与された患者の死亡率減少効果は見られず、心室性不整脈のリスクは有意に増加

- 多数の科学者から研究方法と利用したデータに懸念を表明する公開質問状
- Surgisphere Corporation(共著者の1人が設立)のデータセットを利用
- 独立のピアレビュアーにはSurgisphereは守秘義務契約より提供できないとし、情報提供が得られなかった
- これにより論文を撤回

## 【NEJMの論文の撤回】

**Retraction** - Mandeep R. Mehra, Sapan S. Desai, SreyRam Kuy, Timothy D. Henry, Amit N. Patel. Cardiovascular Disease, Drug Therapy, and Mortality in Covid-19. The New England Journal of Medicine, June 25, 2020 doi: 10.1056/NEJMoa2007621

ACE阻害薬やARBの影響について、心血管疾患の既往は死亡率の増加と関連があったが、治療薬の使用は死亡率増加と関連が見られなかった

- Surgisphere Corporationのデータセットを利用
- Lancet論文撤回のすぐ後、著者等はraw data へのアクセスができず第三者への提供もできなかった
- これにより論文を撤回

# RWD研究 - 何が満たされなければならないか?

## 観察研究におけるガイダンスの例

### STROBE ガイドライン (STrengthening the Reporting of OBservational studies in Epidemiology)の拡張

<https://www.strobe-statement.org/>

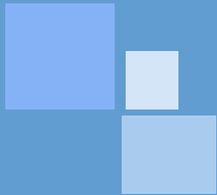
- *The REporting of studies Conducted using Observational Routinely collected health Data (RECORD) statement, 2015*
  - 日常的に収集される医療データを用いた観察研究の報告(RECORD声明)、報告にあたってのチェックリスト
- *The Reporting of Studies Conducted Using Observational Routinely Collected Health Data Statement for Pharmacoepidemiology (RECORD-PE), 2018*
  - RECORDを拡張した薬剤疫学研究に固有のガイドライン(RECORD-PE)、報告にあたってのチェックリスト

## RECORDチェックリストの一部 ※データのReliabilityとRelevanceの観点

	STROBEにはなくRECORD追加された項目の概要
タイトルとアブストラクト	1.1: データのタイプをタイトルかアブストラクトに示す 1.2: 地理的、時間帯範囲を(該当する場合は)アブストラクトに記載する 1.3: DB間のリンケージを行った場合はアブストラクトに明記する
方法 - 参加者	6.1: 研究対象の集団の選択の方法を記載する 6.2: 集団の選択に用いたコード/アルゴリズムを参照する バリデーションを行い、他で公表していない場合は方法と結果を詳述する 6.3: DBの連結を行った場合はプロセスを図示する方法を検討する
方法 - 変数	7.1: 暴露因子、アウトカム、交絡因子、効果修飾因子の分類に用いた完全なコードとリストを示す
方法 - データアクセスとクリーニングの方法	12.1: 対象集団の生成に研究者はどの程度、DBへのアクセスができたかを説明する 12.2: データクリーニングの方法に関する情報を記載する
方法 - リンケージ	12.3: 2つ以上のDBにまたがる個人レベル、組織レベル、その他のデータリンケージを行った場合、リンケージの方法、その質の評価方法を提示する
結果 - 参加者	13.1: 登録した個人の選定の詳細を説明する。データの質に基づいたフィルタリング、データの取得可能性とリンケージ
考察 - 制限事項	19.1: リサーチクエストに答える上でのデータ利用における制限事項を説明する 例. 誤分類バイアス、計測されない交絡、欠損データ、時間を追って変わり得る選択基準への適合
考察 - プロトコル、生データ、プログラミングコード	22.1: 研究プロトコル、生データ、プログラミングコード等、補足情報に、どのようにアクセスできるかに関する情報を提示する

# リアルワールドデータ

## 品質管理と標準

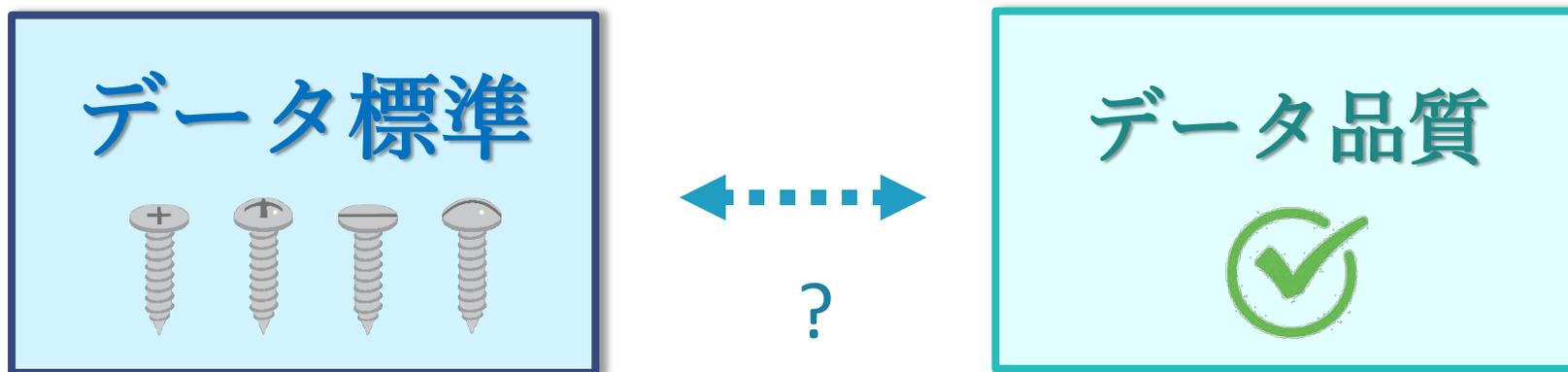


# リアルワールドデータの質と標準

## RWDのデータソースは診療現場に

- RWD/RWEにおけるデータ標準の議論はヘルスケアと臨床研究・規制領域にわたる
- ヘルスケア領域と臨床研究・薬事領域にわたる取り組みが必要

「データの標準」と「データの品質」はどう関わるのか？



Nahm M, Hammond WE. : **Data Standards  $\neq$  Data Quality**, Stud Health Technol Inform. 2013; 192: 1208.

「データの品質」のいかなる点を論じるのかを明確にする必要がある

# データの品質保証 (Data Quality Assurance)

## ベリフィケーション(verification) と バリデーション(validation)

環境Data verification及びData validationに関するガイダンス (米国環境保護庁,US Environmental Protection Agency)

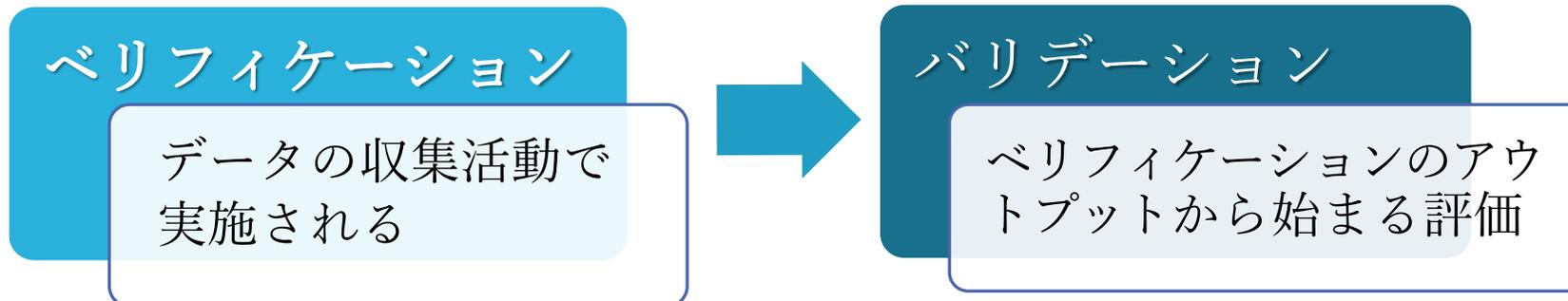
### データ・ベリフィケーション:

データセットの完全性、正確性、適合性 (conformance/compliance)を方法、手続き的要件、契約要求事項に照らして評価するプロセス

### データ・バリデーション:

データ・ベリフィケーションを超えて特定のデータセットの分析的なクォリティを評価するプロセス  
「プロジェクトに特有のデータニーズ」に焦点

一般にデータのベリフィケーションとバリデーションはシーケンシャルなステップ



※ほとんど常にデータ収集者やデータ利用者とは独立のパーティにより実施される

# FDA RWEのフレームワーク

## レギュラトリ・ディシジョンの利用におけるRWD適合性評価(Assessing Fitness of RWD)

**RWEの強度:** 臨床試験の方法とData Reliabilityおよび根底にあるData Relevanceに依存

※ Sentinel SystemではRWEの生成のためのRWD評価にData ReliabilityとRelevanceを検討

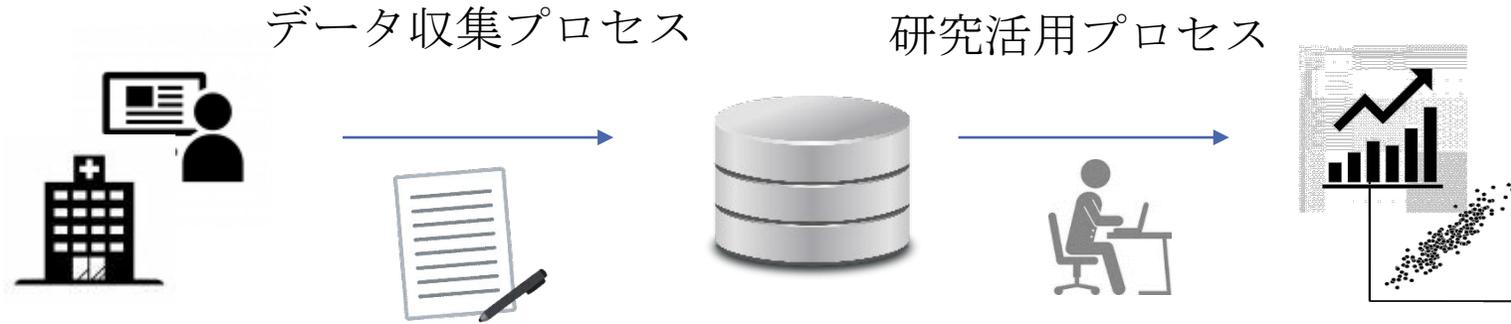
データの信頼性(Reliability)	データの適切性(Relevancy)
データ収集プロセス( <u>accrual</u> )と データ品質管理/品質保証( <u>quality control/assurance</u> )	リサーチ・クエスチョンごと
<b>Data Accrual:</b> RWDのソースはどうデータ要素が収集されるか、データ要素の定義、データ集積・統合の方法、適切な時間窓(relevant time windows)などが予め定められていること	<b>Dataのfit-for-purpose :</b> Fit-for-purposeは利用に照らしての適合性 この議論は複雑
<b>Reliability Assessment</b> コードやコードの組み合わせが適切に根底にある意図する医療コンセプトを表現できているか	<b>Relevance Assessment</b> データはfit-for-purpose であるか、暴露、アウトカム、共変量に関するデータがキャプチャされているかのアセスメントを含む

- 安全性評価に関してEHRのRWDを用いるにあたっての企業とスタッフへのガイダンス発行

*Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data*

# FDA RWEのフレームワーク

レギュラトリ・ディシジョンの利用におけるRWD適合性評価(Assessing Fitness of RWD)



## Reliability Assessment

コードやコードの組み合わせが適切に根底にある意図する医療コンセプトを表現できているか

## Relevance Assessment

データはfit-for-purposeであるか、暴露、アウトカム、共変量に関するデータがキャプチャされているかのアセスメントを含む

データ品質管理/品質保証(quality control/assurance) :

品質保証の原則については国際的に合意された一定の基準があるわけではない

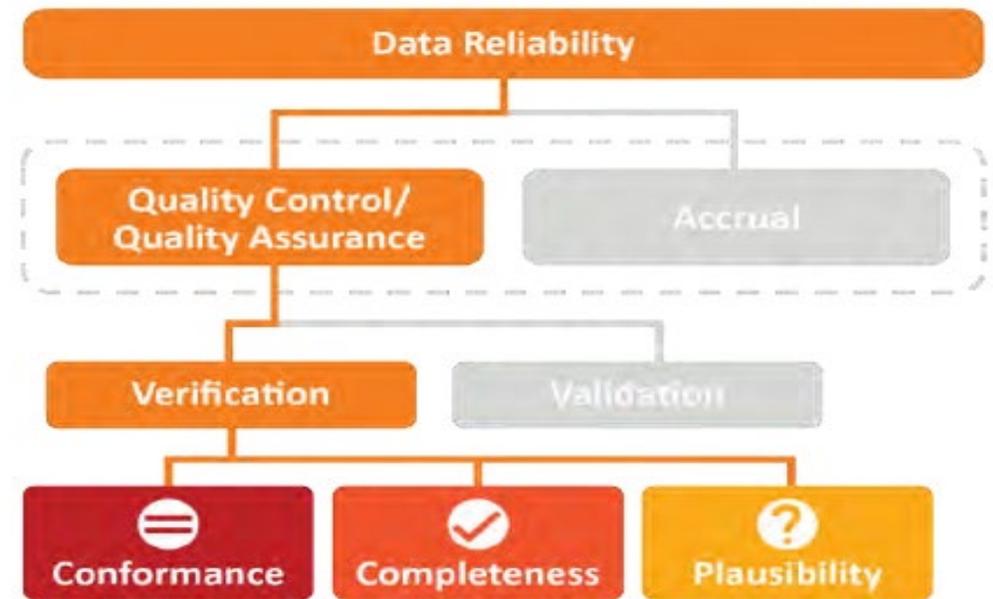
# Determining Real-World Data's Fitness for Use and the Role of Reliability, Duke Margolis Center for Health Policy, September 26, 2019

Figure 2. Examples of Quality Check Principles



Duke      Sentinel      PCORnet      FDA

Figure 3. How Verification Checks for Completeness, Conformance, and Plausibility Contribute to Assessing Reliability to Support Fit-for-Use RWD\*



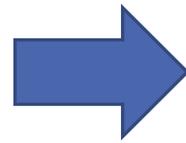
Conformance, Completeness, Plausibility

NOTE: FDA has identified issuing guidance on assessments of the reliability and relevance of RWD in generating RWE on drug product effectiveness as a program item in its 2018 RWE Framework. This paper does not reflect the views, guidance, or recommendations of FDA.

# データ品質アセスメント(Data Quality Assessment: DQA)

Kahn等はBiomedical Informaticsにおけるデータ品質評価に関する既存のターミノロジーを単一化させ(unify)、調和されたDQAターミノロジーを整理・提案  
Kahn M, Callahan T, Barnard J, et al. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. EGEMS (Wash DC). 2016;4(1):1244. doi: 10.13063/2327-9214.1244. eCollection 2016

調和されたDQAターミノロジー  
3つのカテゴリ  
(このしたに8つのサブカテゴリを定義)



## Conformance (適合性)

データが以下に適合する度合い：

- ① あらかじめ定義された内的または外的なformatting制約への適合 (Value Conformance)
- ② 構造的な制約への適合
- ③ 既存の値に基づいて計算して導出された値の正確性 (Calculation Conformance)

## Completeness (完全性)

データの存在が以下の2つの観点からアセスされること：

- ① 時間軸の一点 (Atemporal Completeness)
- ② 複数の時点 (Temporal Completeness)

## Plausibility (質が合理的で真であるらしい度合い)

The quality of being reasonable and likely to be true :  
信じるに足る程度が次のように評価(assess)されること：

- ① 論理的に制約された測定や分布、同一の事実についての独立な複数の測定が一致すること (Atemporal Plausibility)
- ② 時間軸の属性や順序、測定値の状態遷移 (Temporal Plausibility)
- ③ 重複する測定値、変数の値または記録の存在 (Uniqueness Plausibility)

※RWDの完全性(Completeness)はRCTで前向きに収集されたデータに比べて低い

# データ品質アセスメント(Data Quality Assessment: DQA)

## 調和されたデータ品質アセスメント (Data Quality Assessment: DQA) ターミノロジー

- データに内在する※DQAカテゴリに限定 ※「データの値そのものに属する次元」
- 外来性 (extrinsic)※のカテゴリは含まない ※「データが利用されるコンテクスに依存する次元、fitness-for-use 基準のデータの利用可能性、タイムリネス等」

データ品質の議論は用語自体、単純ではないが、調和されたDQAターミノロジーによるとデータ標準の議論も整理し易い

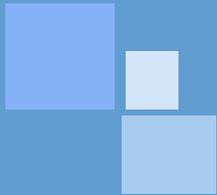
適合性(Conformance)はデータの相互利用可能性(Interoperability)に強く関連

RWDの完全性(Completeness)は、RCTで前向きに収集されたデータに比べて低い

## データ標準の観点から

- データ・バリデーションのコンテクストからのデータ品質チェックは重要であるが、データ・ベリフィケーションのコンテクストにおけるチェックに比べ、定義、実施とも複雑
- データ・ベリフィケーションにおけるデータ品質、特に適合性(Conformance)の観点から data interoperability を中心にデータ標準について考える

# データ品質と医療情報標準



# Data Interoperabilityをコアとするデータ標準の取り組み

## HL7 FHIR(Fast Healthcare Interoperability Resource)の展開

- 臨床試験、規制領域ではCDISCにおけるHL7 FHIRの取組
- 米国医薬品食品局(FDA)、欧州医薬品庁(EMA)から参加しての医薬品あるいは臨床研究のFHIRリソース、product and substance identificationのためのFHIR implementation guideの開発
- FHIRリソースでは臨床研究のためのResearchStudy, ResearchSubjectなどの開発

## 研究ネットワークとコモンデータモデル

- 臨床研究へのRWD活用の一環としてコモンデータモデル(Common Data Model: CDM)
- 様々な研究ネットワークで、それぞれのコモンデータモデルの策定：
  - 1) PCORNet(national patient-centered clinical research network) コモンデータモデル
  - 2) i2b2 (Informatics for Integration Biology and Bedside) コモンデータモデル
  - 3) OMOP(Observational Medical Outcomes Partnership) コモンデータモデル
  - 4) Sentinel model

## PCORnet (全国患者中心臨床研究ネットワーク)

- 2014年、患者中心アウトカム研究所(Patient Centered Outcomes Research Institute: PCORI)が2億5千万ドル以上を投資して開発
- 国に渡る代表性の高いパートナー・ネットワークからなる臨床データ研究ネットワーク(Clinical Data Research Networks: CDRNs)と患者が強化する研究ネットワーク(Patient Powered Research Network: PPRN)
- 臨床研究を研究者指向型から患者中心型へとシフトさせていく
- 患者、ケア提供者、その他ステークホルダーが研究のデザイン、実施、分析から結果の普及に至るまで関与することを支援

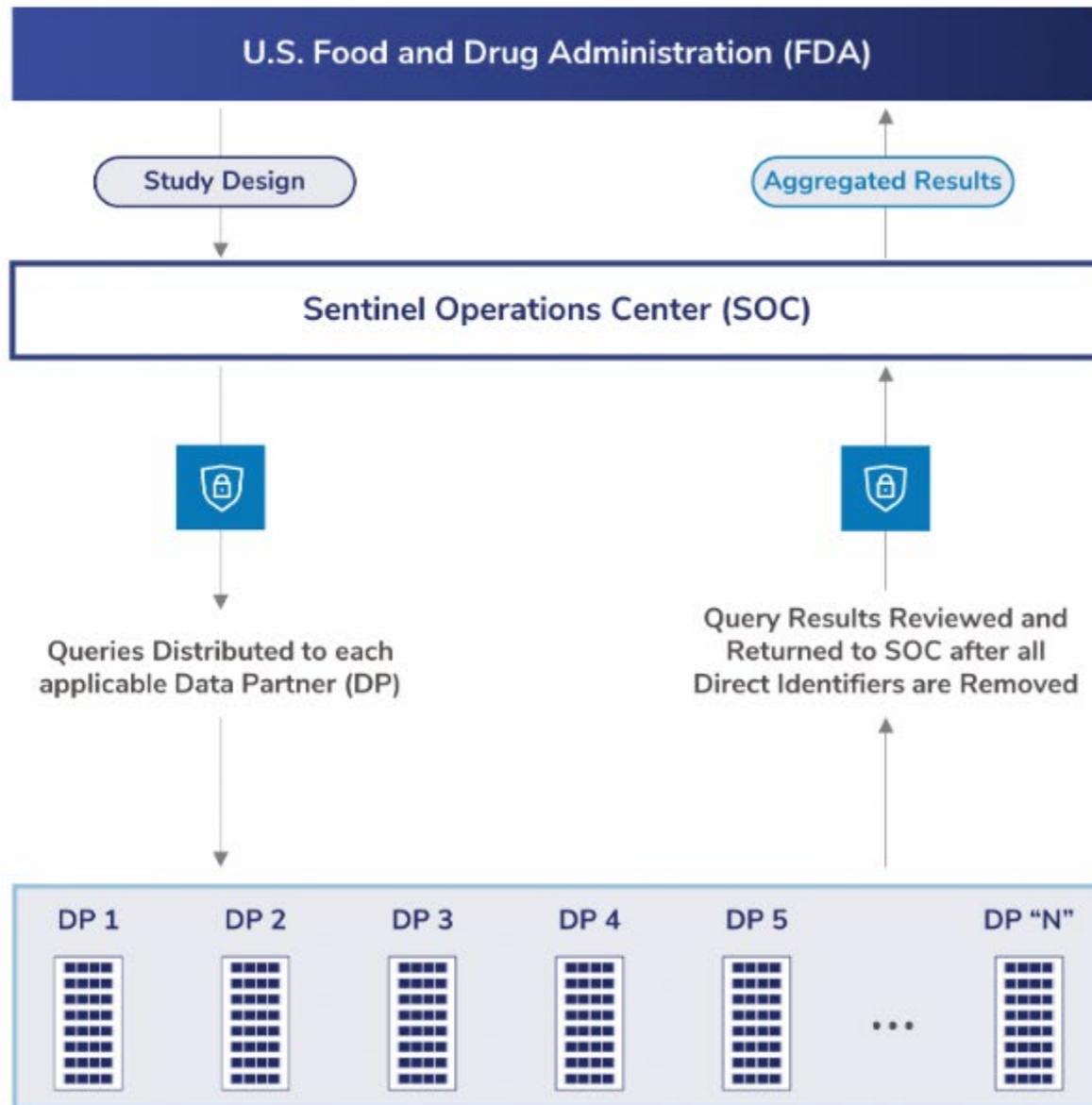
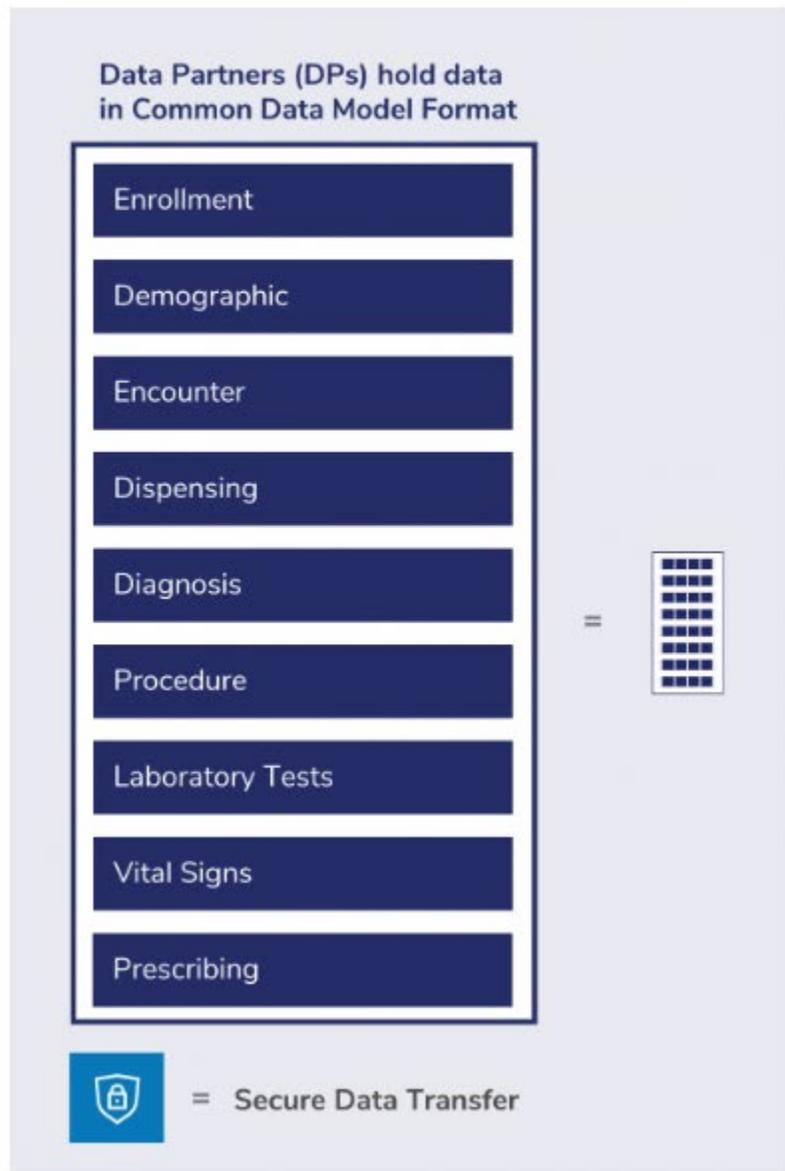
## Informatics for Integration Biology and Bedside (i2b2)

- 2004年、NIHのファンドによるBiomedical Computingのためのナショナルセンターに設立
- ハーバード連携病院、ハーバード公衆衛生大学院、ハーバード医学大学院、ハーバード/MITヘルスサイエンス&テクノロジー部門の研究者等が参加
- 国のサイエンス・コミュニティへのi2b2 ツールの提供

## Sentinel

- FDA Sentinel Initiativeは2007のFDA Amendments Act of 2007にて開始。初期デザインからCDMを用いる

illustrates Sentinel's distributed data approach.



# Common Data Models Harmonization (CDMH)

## 連邦機関の主導によるエビデンス創出のための 各種CDMとオープンスタンダードのハーモナイゼーション

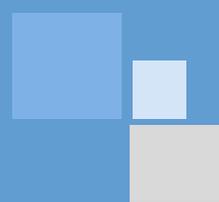
- 日々の医療提供で生成されるRWDを用いて患者中心のアウトカムリサーチを行うためのデータ基盤の構築とRWD/RWE促進のためのデータ標準のハーモナイゼーションの取り組み
- FDAは、Harvard Pilgrim Health Care Instituteとハーバード医学部と協力し、複数データソースにわたるEHRデータの質、完全性、安定性を評価・記述するための統一されたメタデータ標準の開発
- 複数連邦機関、研究者が患者中心のアウトカムリサーチへのRWD活用を可能とする複数データモデルのハーモナイゼーションのためのメタ・コモンデータモデルを開発  
FDA、国立医学図書館(NLM)、国立がん研究所(NCI)、国立先端トランスレーショナル科学センター(NCATS)、米国国家医療IT調整官室(Office of the National Coordinator for Health Information Technology: ONC)が参加
- 連邦機関と研究ネットワークによるCommon Data Models Harmonization : CDMH プロジェクトでは、データクエリーや、クエリーから結果を得るなど様々な観点からの取り組み

# Common Data Models Harmonization (CDMH)

## Common Data Models Harmonization (CDMH) FHIR IG

- HL7 BR&R (Biomedical Research and Regulation) Work Groupにより開発、現在version 0.1.0
- 複数の研究ネットワークから抽出される研究データは、それぞれ異なるデータモデルによるデータ表現がなされ得る
- PCORNet CDM、i2b2 CMD、OMOP CDM、SENTINEL modelに焦点
- 患者中心のアウトカムリサーチ(PCOR)から抽出された観察データ、クエリーから得られる結果のFHIRフォーマットへのマッピングと変換に焦点
- 以下への投資とともに米国のデータ基盤の向上を目指す
  - Learning Health System (LHS)やPCOR Trust FundによるDAF (Data Access Framework) - Research IG
  - 構造化データキャプチャ(Structured Data Capture: SDC) IG、
  - 患者報告アウトカムFHIR IG
  - Women's Health Technologies Coordinated Registry Network (CRN) FHIR IG
- 将来クエリーやワークフローが自動化可能となる段階でDAF-Research IGを向上させることが期待

# 医薬品規制領域におけるHL7 FHIRの開発動向



IDMP Project	<ul style="list-style-type: none"><li>FDAと欧州規制当局者ネットワーク、HL7の協力によるISO IDMP標準をFHIR仕様で記述する取り組み。IDMP関係のFHIRリソースの開発がMaturity Level 1まで進んでいる Maturity Level：0～5のスケールで表される各リソースの成熟度(大きい方が成熟度が高い)</li></ul>
Profiles for ICSR Transfusion and Vaccination Adverse Event Detection and Reporting	<ul style="list-style-type: none"><li>FDA CBER Biologics Effectiveness and Safety (BEST) イニシャティブのファンドによる(FDA CBER BEST IBM プロジェクトチーム)による。AE検出と報告の二つの要素</li><li>FHIR データで駆動するClinical Quality Language (CQL) アルゴリズムを用いたAE 検出と、FHIRベースのAE報告生成とICH ICSR仕様のマッピングのためのプロファイル集合等</li></ul>
FDA Quality/ Chemistry, Manufacturing, and Controls Data Standardization	<ul style="list-style-type: none"><li>FDA CDERが進める効率的、効果的な規制意思決定支援を目的とするプロジェクト</li><li>Quality (PQ)/ Chemistry, Manufacturing, and Controls (CMC)データの主要分析の自動化を可能するためデータ要素、用語、データ構造の特定と標準化を進めている</li><li>Quality SpecificationでFHIRによるデータ交換のテストを実施</li></ul>
FHIR SPL Implementation Guide	<ul style="list-style-type: none"><li>FDAのSPL(Structured Package Labeling)提出をFHIRで行うための実装ガイド</li><li>HL7 V3 SPLからFHIRへの移行を可能にするため、SPL HL7 V3仕様とFHIRの間のマッピングと、両仕様で受信できるdual submission architectureを含む</li></ul>
EU ePI (electronic Product Information) Common Standard	<ul style="list-style-type: none"><li>ePI はEU ePI Common Standardを用いたEU医薬品電子製品情報</li><li>Common Standardは2021年にEMA、NCA、ECによって実施されたePIプロジェクトの主要成果物の1つ</li><li>ePIには患者向けの添付文書と医療専門家向けの製品特性 (SmPC) の要約が含まれる</li><li>EUで認可されたすべての医薬品に添付され、その使用方法と処方方法を説明</li></ul>
Vulcan Accelerator Project	<ul style="list-style-type: none"><li>Vulcanでは臨床研究にフォーカスをあてた各種ユースケースのためのFHIR開発を推進</li><li>SDOとしてCDISCとHL7の連携</li><li>ユースケースとしてDCT Real time collection, CDISC mapping, sIRB (Single IRB), Phenopackets in EHR, RWD, AE from HER, Vital signs, Demographics, Patient Cohort Identification, Pragmatic Clinical Trial, Protocol Design/Feasibilityなど</li></ul>

VulcanはHL7 FHIR Accelerator Projectsの中の一つ。2019年9月発足

## 【課題】

- 臨床研究のための EHR データの取得、利用は日常的ケアとは大きくかけ離れている

## 【目的】

- 医療ケアと臨床研究の間のギャップを埋め、トランスレーショナル及び臨床研究コミュニティ全体の利害関係者を結び付ける
- 臨床研究の設計・実施・報告を最適化するための標準の開発・改良・使用を加速
- 世界の規制当局に包括的推奨事項を提供
- 必要な FHIR 研究リソースを成熟するまで開発

## 【運営・組織】

- 行政、学界、テクノロジー企業、標準開発団体、患者、業界コンソーシアムからの投資を受けた代表者グループによるリード
- SDOとしてCDISCの参画

## Clinical and Translational Clinical Research Community

### Current Member Organizations of Vulcan As of December 2022

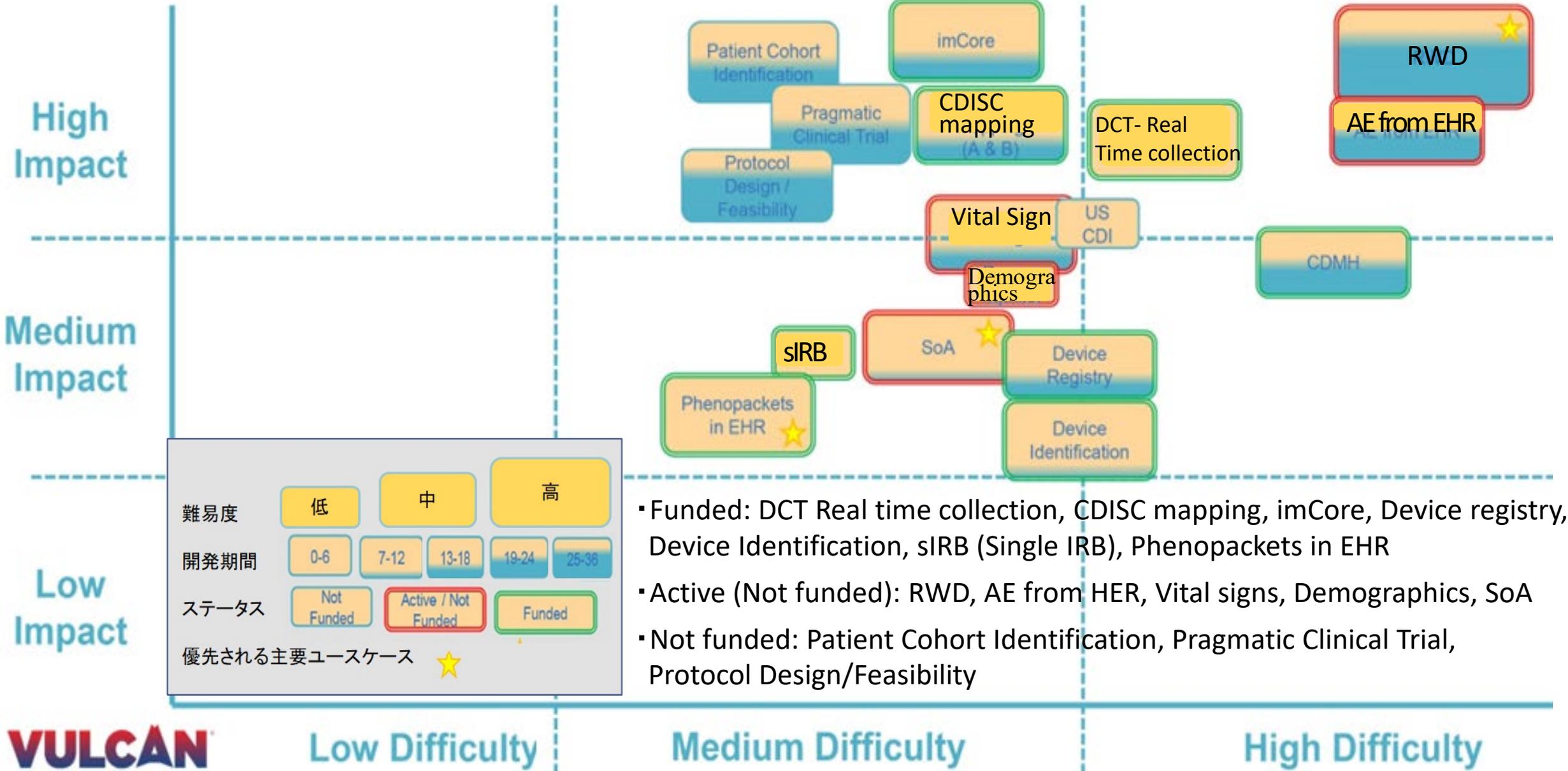
Academia	 Cedars Sinai  Duke University School of Medicine  Johns Hopkins School of Medicine  国立がん研究センター 東病院 National Cancer Center Hospital East  OREGON HEALTH & SCIENCE UNIVERSITY  UAMS
Consortia	 phuse  Society for Clinical Data Management SCDM  TransCelerate BIOPHARMA INC.
Government Agencies	 LÆGEMIDDELSTYRELSEN DANISH MEDICINES AGENCY  FDA  NIH National Center for Advancing Translational Sciences  NIH U.S. National Library of Medicine
Implementers	 accenture  BioVeras BLOCKCHAIN FOR LIFE SCIENCE  Epic  FUJITSU  IgniteData  infor  InterSystems Health   Business   Government
Pharma	 GSK  Johnson & Johnson  Pfizer  Roche
SDOs	 cdisc  HL7 International
Others (e.g., thought leaders, SMEs, CROs, Patient Advocates)	 ACRO  CROHN'S & COLITIS FOUNDATION  FELLESKATALOGEN

VULCAN

メンバー組織 (2022年12月現在)

\* indicates a convening member of Vulcan

# Vulcan Use Case Assessment Prioritization

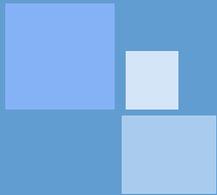


難易度	低	中	高		
開発期間	0-6	7-12	13-18	19-24	25-36
ステータス	Not Funded	Active / Not Funded	Funded		
優先される主要ユースケース	★				

- **Funded:** DCT Real time collection, CDISC mapping, imCore, Device registry, Device Identification, sIRB (Single IRB), Phenopackets in EHR
- **Active (Not funded):** RWD, AE from HER, Vital signs, Demographics, SoA
- **Not funded:** Patient Cohort Identification, Pragmatic Clinical Trial, Protocol Design/Feasibility



# リアルワールドデータ 医療施設の観点から



# リアルワールドデータ利活用のための取組みの事例

## AMED医療技術実用化総合促進事業

### 臨床研究中核病院 RWE創出のための取組み(通称: 臨中ネット)

- 2018年より開始(当初12病院、現在14病院)
- 病院情報システムからRWDを抽出して研究利活用するための課題抽出、院内RWD整備
  - SWG1: データ項目・各種ユースケースによる研究
  - SWG2: RWD生成のための人材育成
  - SWG3: 標準化すべき情報項目・標準化の方法
  - SWG4: RWDの品質 (バリデーション等)
  - SWG5: RWDの出力
  - SWG6: 研究倫理・ユースケース(研究テーマ)

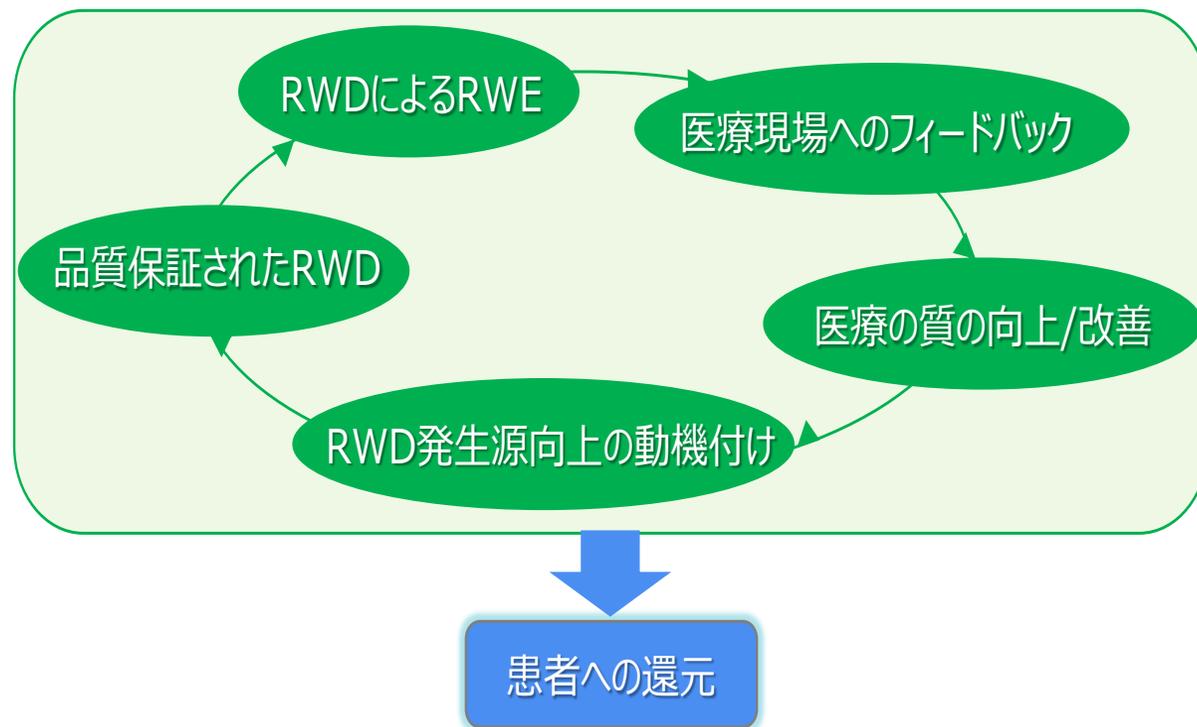
## 高品質RWD生成の良循環から患者への還元へ

日々蓄積されるRWD

→ RWDの活用・解析結果が、医療現場にフィードバック

→ このフィードバック機構の形成が医療の質向上に繋がる

→ それが質の高い診療データ生成の動機へと繋がる良循環に



## 医療機関にしかできないこと

- 蓄積された医療データから医療現場にフィードバックし得る結果が生まれるためにはリアルワールドデータの質が問われる

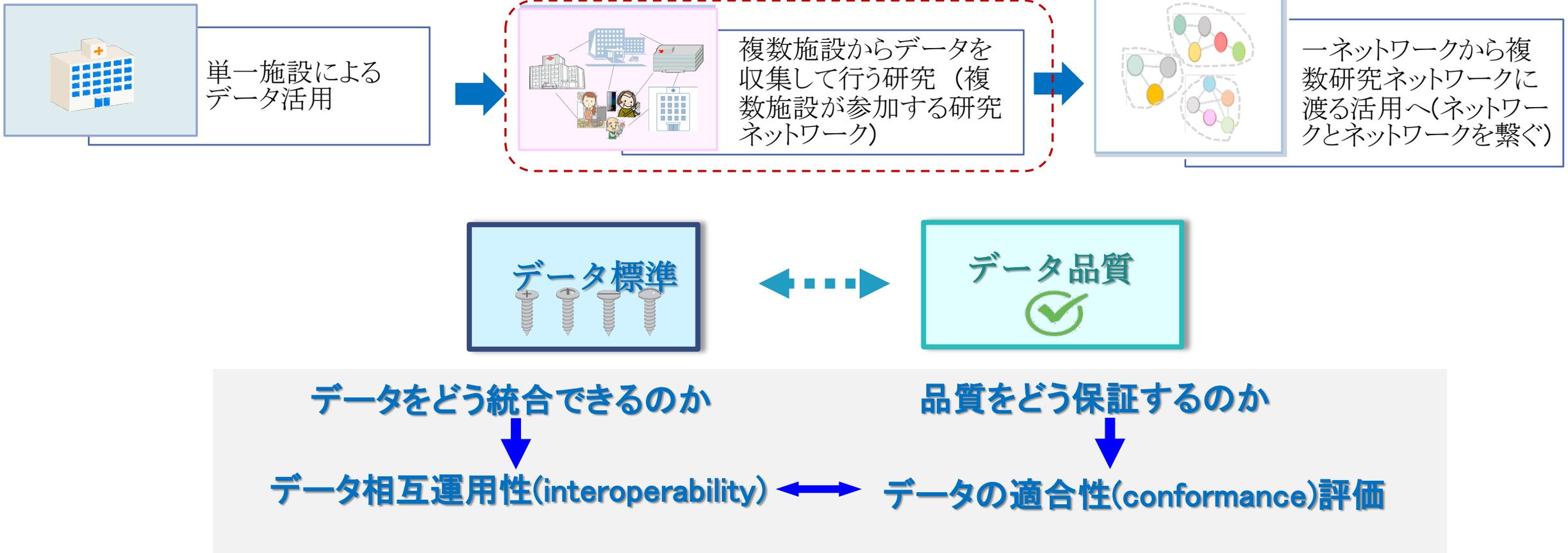
- RWDの質：

- ①生成、②格納、③収集、④集積 の各ステップ  
クレンジング処理である程度の対応ができるのは③収集、④集積の段階だけ

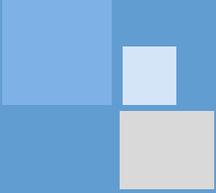
- データが①生成され、②格納されるステップは施設内でしか対応できない

高品質診療データの発生源が施設内に整備されそれが最適な医療に、患者への還元につながる  
これが「患者に還元される二次利用」の姿では

# おわりに



医療情報の標準がデータ品質保証を支える



# リアルワールドデータの利活用と標準化

## －現状と今後への展望－

### 応えるのは誰？

# CAPS

*Council for Accelerating Pharmaceutical  
Information Standards*



[TOP](#)

[ご挨拶](#)

[設立趣旨](#)

[組織概要](#)

[事務局\(問い合わせ先\)](#)

## 医薬品情報標準化協議会(CAPS)

### 医薬品情報標準化協議会(CAPS)が設立されました

医療における医薬品のライフサイクルに係る情報の  
横断的な交換、共有、活用を可能とするため  
関係者の協力により標準化の実現を目指します

[TOP](#)[ご挨拶](#)[設立趣旨](#)[組織概要](#)[事務局\(問い合わせ先\)](#)

## 組織概要

名称	医薬品情報標準化推進協議会 Council for Accelerating Pharmaceutical Information Standards (CAPS)
代表	折井 孝男 医薬品情報標準化推進協議会
顧問	森田 朗 東京大学名誉教授 次世代基盤政策研究所代表理事 中村 陽子 元厚生労働省 村上 貴久 一般財団法人 日本医薬情報センター会長
特別会員	赤川 治郎 一般財団法人 日本医薬情報センター理事長 武田 泰生 一般社団法人 日本病院薬剤師会会長 俵木 登美子 一般社団法人 くすりの適正使用協議会理事長 前田 秀 GS 1 Japan (一般財団法人流通システム開発センター) 常務理事 山本 信夫 一般社団法人 日本薬剤師会会長
事務局	一般社団法人 医療データ活用基盤整備機構 Institute of Health Data Infrastructure for all (IDIAL)

ご清聴ありがとうございました



謝辞本報告の一部はAMEDの課題番号JP20mk0101103の支援を受けた