

➤ MID-NETにおけるデータ品質管理



中島 直樹
九州大学病院メディカル・インフォメーションセンター
日本医療情報学会・代表理事

発表の目的

- データ駆動型の解析事業は、データ品質管理が鍵になる
- MID-NETは継続的・組織的にバリデーションを実施した代表的な事業である
 - 本格稼働まで5年間バリデーションを行い、本格稼働後も1年以上継続している
- MID-NETをその限界を含めて知ることで、今後のデータ駆動型事業のデータ品質管理を議論することができる



前向き臨床研究とデータ駆動型臨床研究（後向き）の比較

この両者を駆使して、正確に対象全体を俯瞰する情報を得る！

	前向き臨床研究	データ駆動型臨床研究 (DDMS、後向き)	DDMSの 改善法
項目の設定	自由度高い	自由度が無い	Phenotyping
バイアス	管理可能	DB構築目的に依存	matching
データ品質	高い	低い	データ品質管理
コスト	高い	低い	✕
データ規模	小さい	大きい	
リアルタイム性	弱い	強い	

データ駆動型研究事業における課題解決の具体的な手段

◆項目の自由度が無い

- ePhenotyping (正しい病名・病態を推測する)

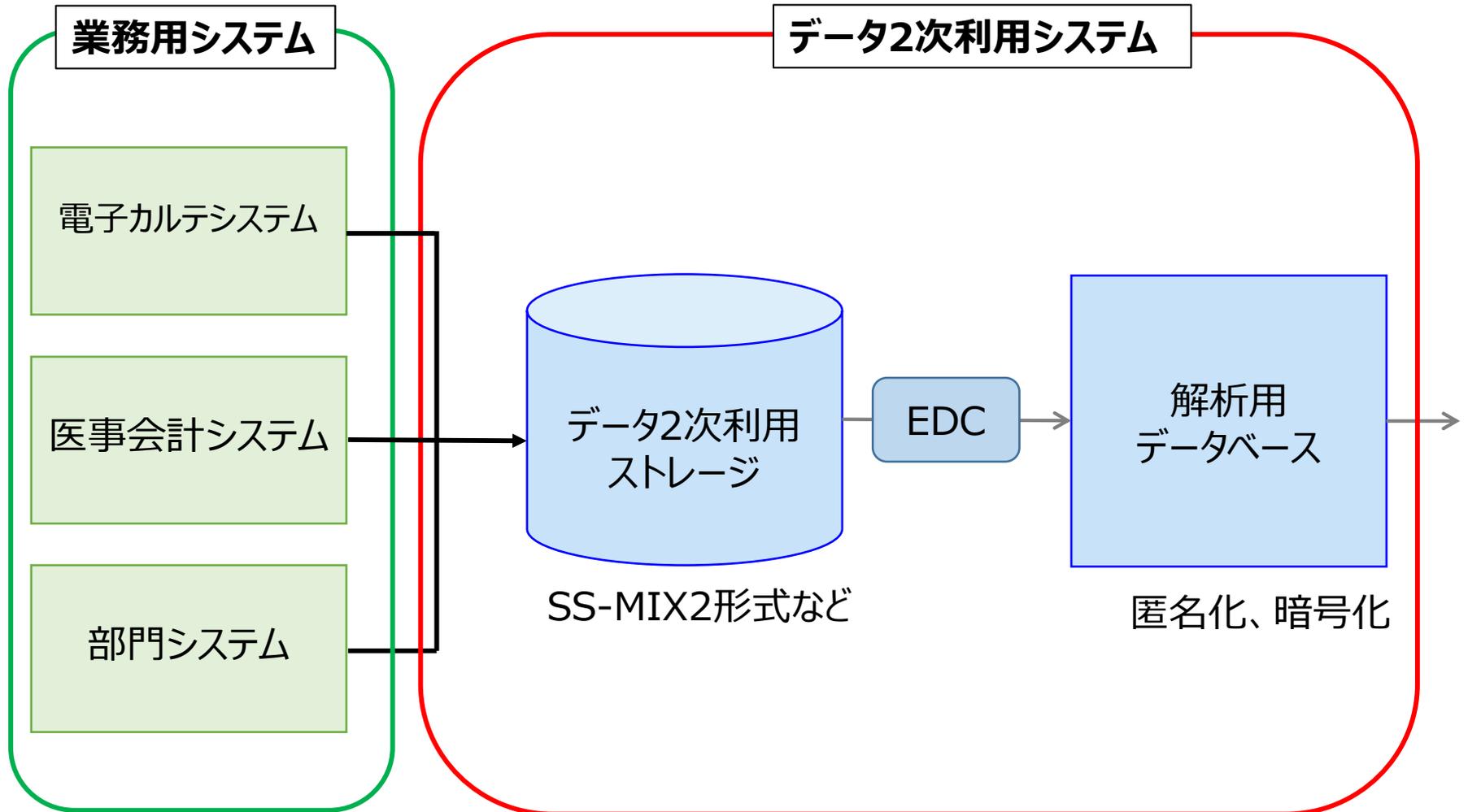
◆バイアス

- ネステッドケースコントロール研究, ケースコホート研究

◆データ品質管理

- EDCシステム導入マニュアル・運用手順
- バリデーション
- 継続的なガバナンス

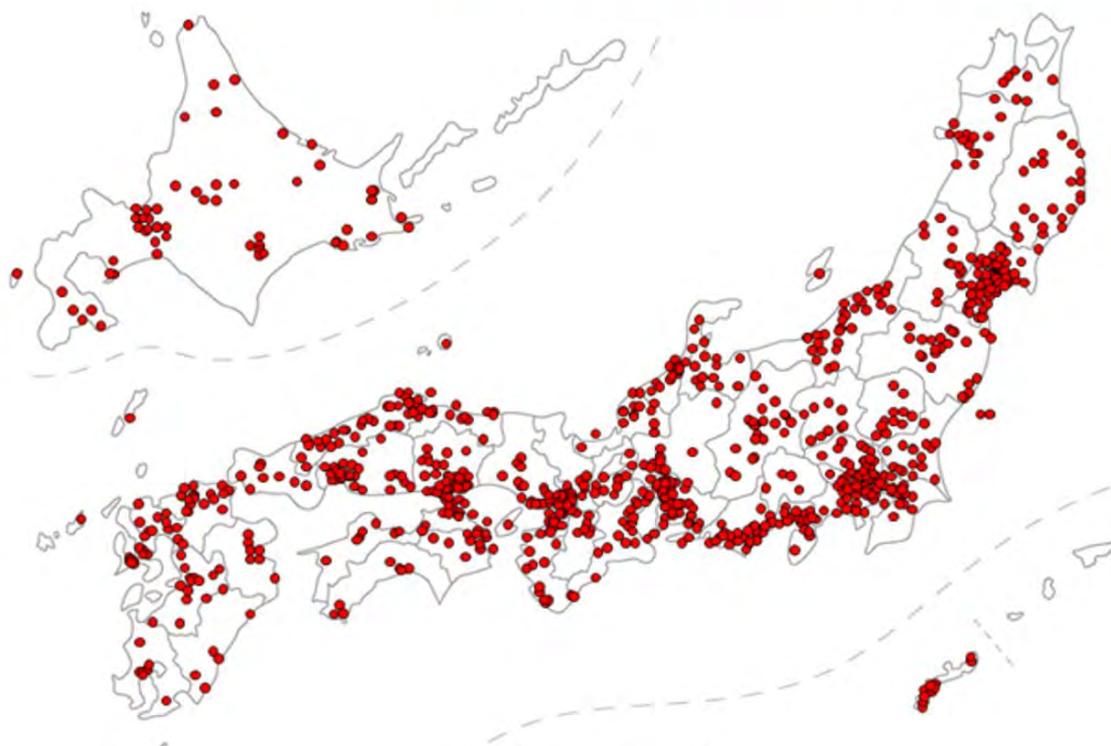
病院情報システムからのEDCによるデータ2次利用



■ SS-MIX2 普及状況

SS-MIX普及推進コンソーシアム ホームページより (2018.03末現在)

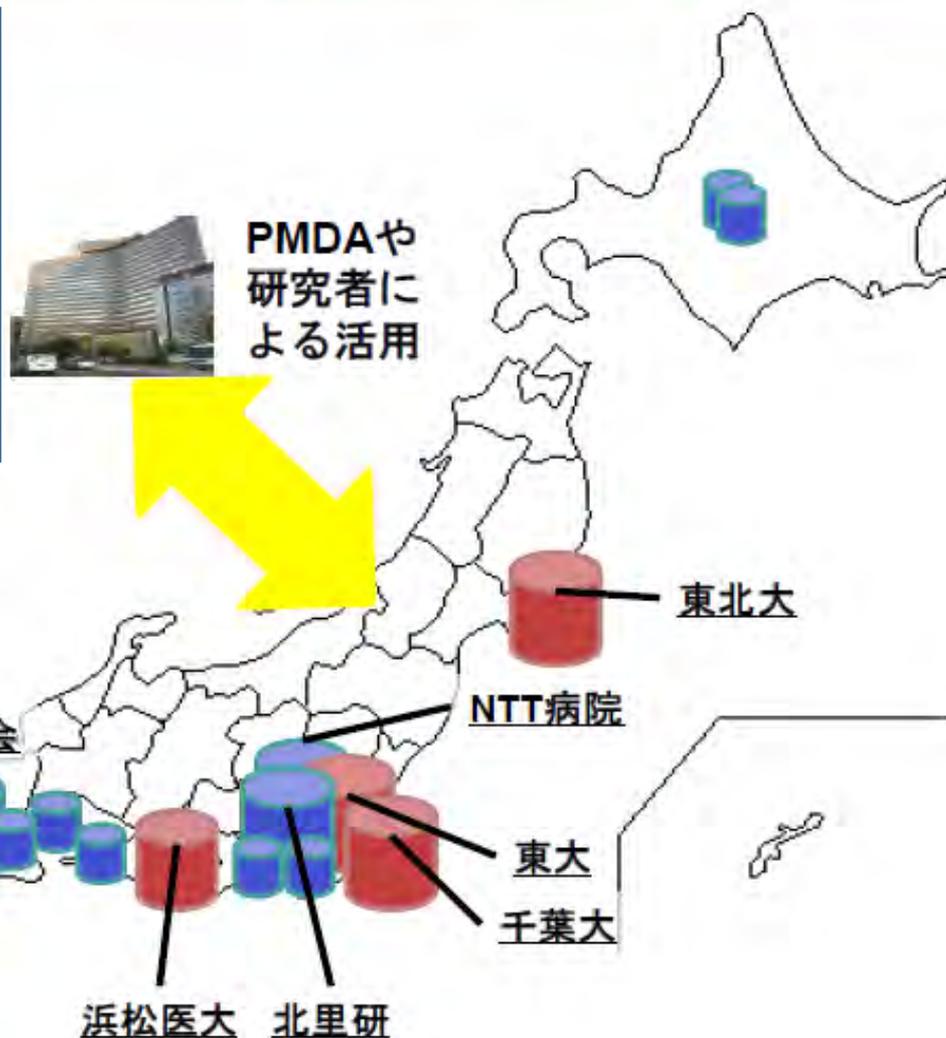
2018年3月末 標準化ストレージ：1,360施設 (処方・検体結果を出力：845施設)
2017年3月末 標準化ストレージ：1,114施設 (処方・検体結果を出力：754施設)
2016年3月末 標準化ストレージ： 996 施設 (処方・検体結果を出力：630施設)



施設所在地を非公開とした304施設は含んでいない

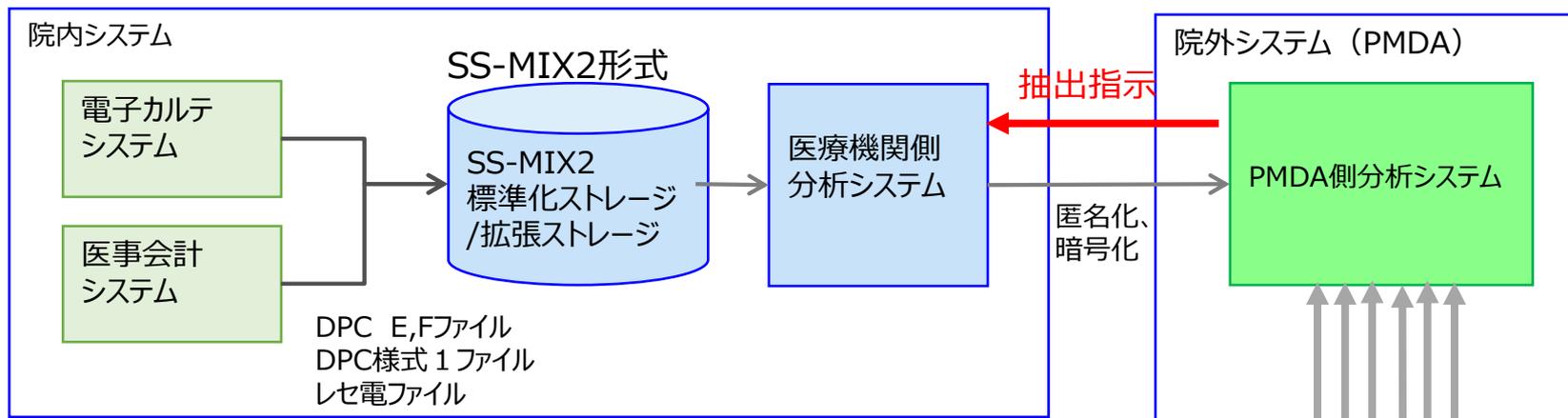
医療情報データベース基盤整備事業 (MID-NET)

- 薬の副作用の検知システム
- 2018年度から10医療機関を拠点としてデータの検索・調査を行い、副作用を分析・評価
(2018年末現在470万人分)
- 電子カルテ**全件**の処方、病名、検査結果情報を利用

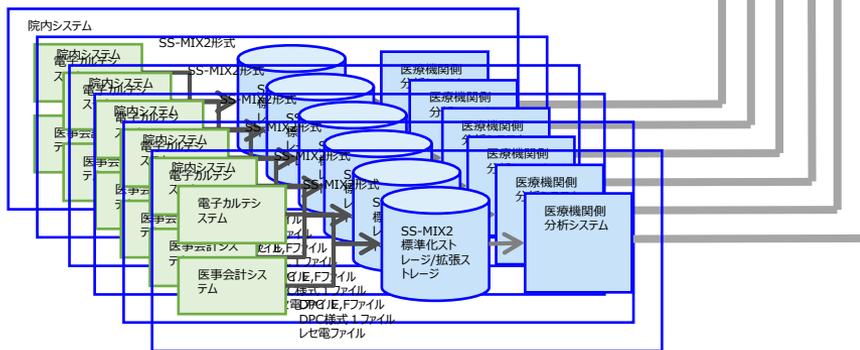


MID-NET事業のシステム概要

PMDA: (独)医薬品医療機器総合機構
複合施設統合データ処理センター



470万人の
リアルタイム
データベース
(2018年末)



データ駆動型研究事業における課題解決の具体的な手段

◆項目の自由度が無い

- ePhenotyping (正しい病名・病態を推測する)

◆バイアス

- ネステッドケースコントロール研究, ケースコホート研究

◆データ品質管理

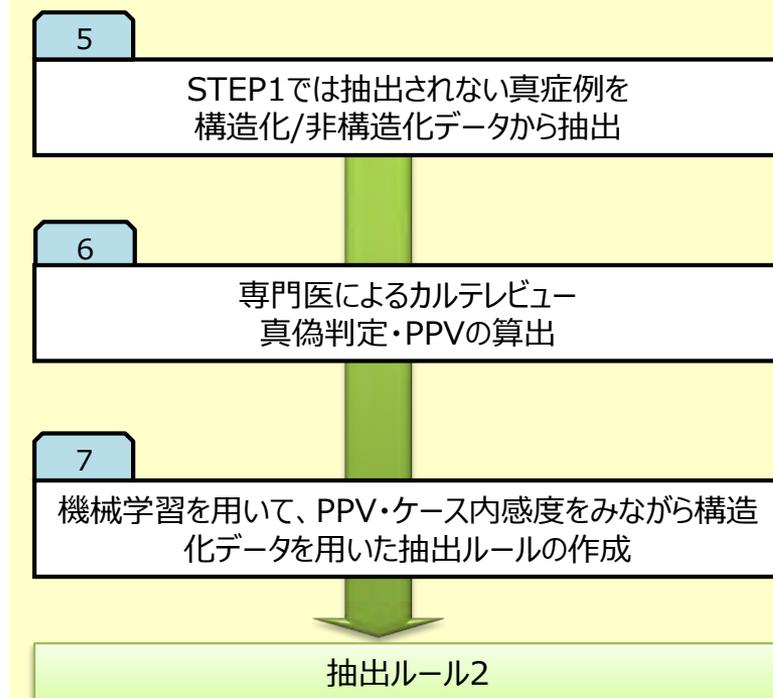
- EDCシステム導入マニュアル・運用手順
- バリデーション
- 継続的なガバナンス

MID-NETにおけるPhenotyping開発の流れ（間質性肺炎の例）

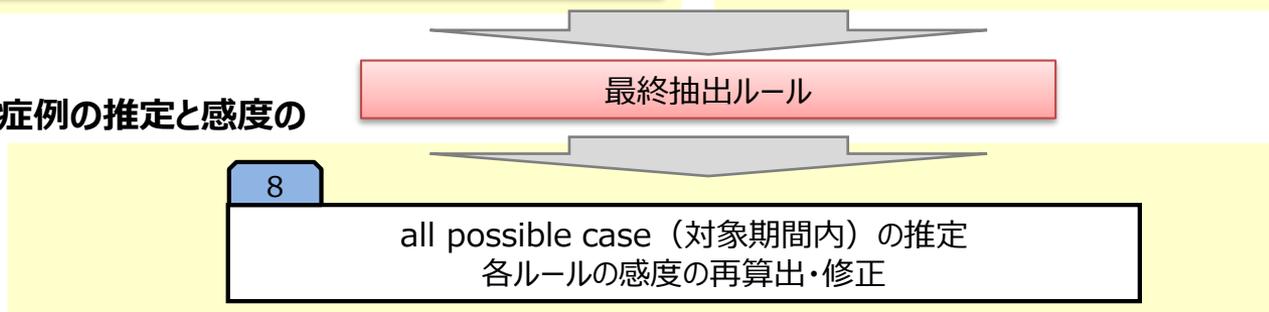
STEP1: 症例抽出ルールの作成



STEP2: STEP1で非抽出の症例抽出ルールの作成



STEP3: 把握された真症例の推定と感度の精緻化



初期ルール作成・PPV算出 STEP1-1~3

	間質性肺炎	
初期抽出ルールの作成	[病名(確定のみ)] or [KL-6値 \geq 430U/mL]	CT読影レポートに スリガラスなどの <u>テキストキーワード</u> を含む(専門医による選定)
対象期間	2014/1/1~2015/12/31における 本院の入外患者117,401名	2014/1/1~2015/12/31における 本院の入外患者117,401名
症例抽出	1,424件	5,141件
ランダム・サンプリング	200件	84件
専門医による真偽 判定 (PPV算出)	2名/症例 5段階評価 1: 確実に違う 2: たぶん違う 3: どちらでもない 4: たぶんそう 5: 確実にそう	
PPV	スコア5: 36.5% スコア4: 50.0%	スコア5: 46.4% スコア4: 60.7%

間質性肺炎 抽出ルール 1

初期ルール



検索式 (a or b) and (c or d)

※以下、指定期間内での有無とする。

- a. 間質性肺炎に関連した（確定）病名がある
- b. KL-6が施設基準値以上（ ≥ 430 U/mL）がある

追加ルール



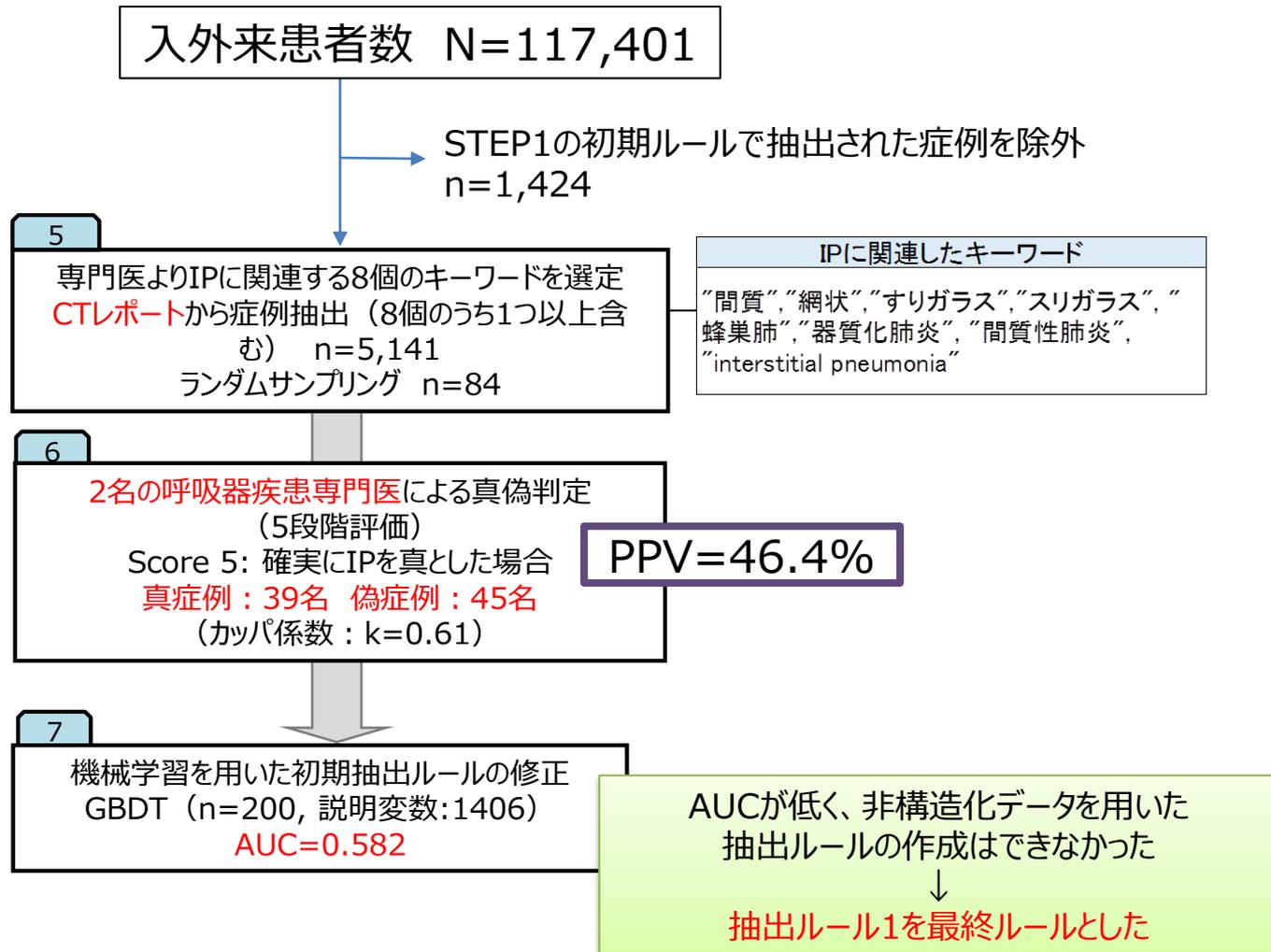
- c. SP-Dが $145 \mu\text{g/dL}$ 以上がある
- d. “間質性肺炎”のテキストを含むDPC病名がある

PPV : 80.0%
Recall : 60.3%

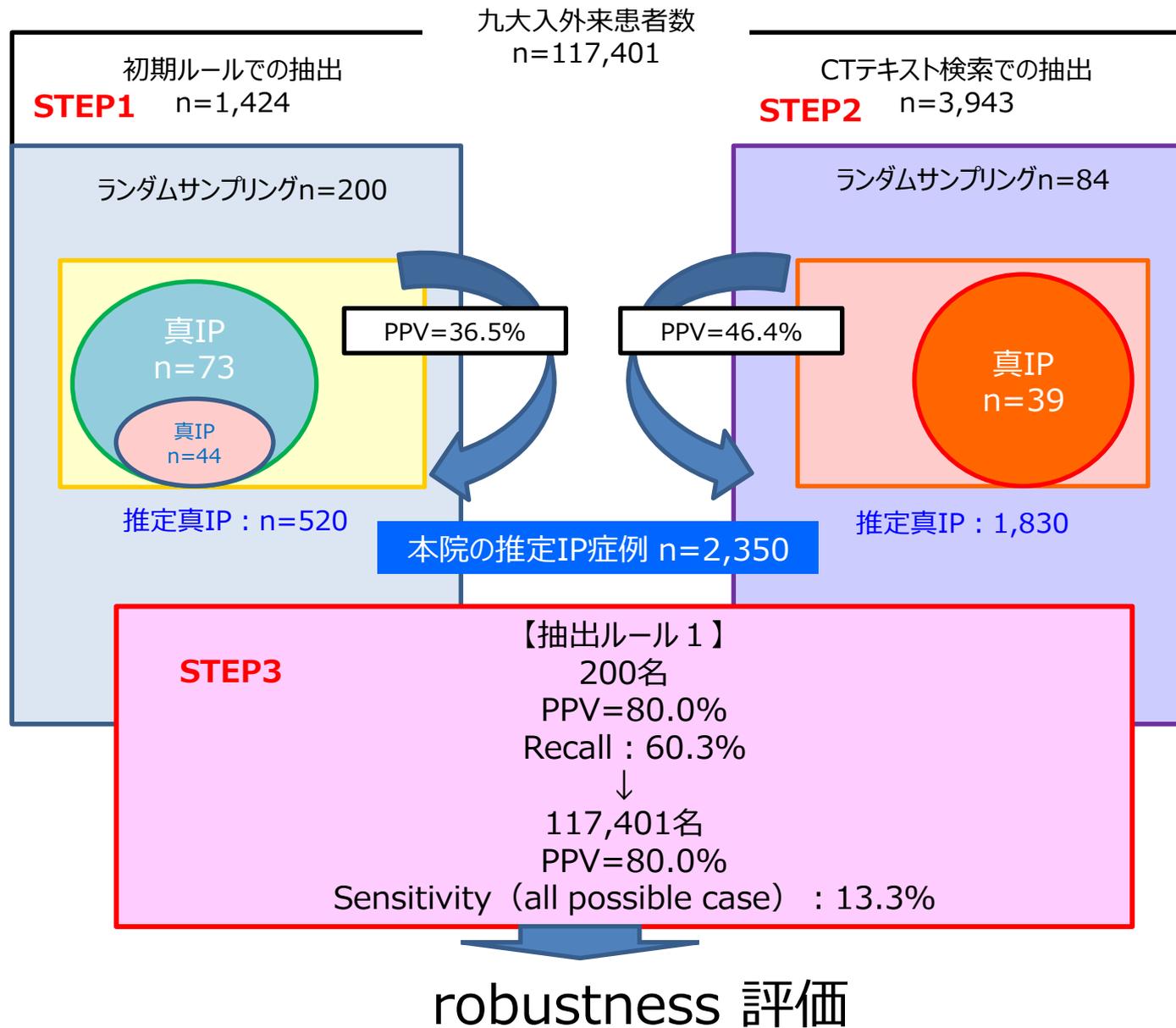
初期ルール作成・PPV算出 STEP2-5~7

	間質性肺炎	
初期抽出ルールの作成	[病名(確定のみ)] or [KL-6値 \geq 430U/mL]	CT読影レポートに スリガラスなどの <u>テキストキーワード</u> を含む(専門医による選定)
対象期間	2014/1/1~2015/12/31における 本院の入外患者117,401名	2014/1/1~2015/12/31における 本院の入外患者117,401名
症例抽出	1,424件	5,141件
ランダム・サンプリング	200件	84件
専門医による真偽判定 (PPV算出)	2名/症例 5段階評価 1: 確実に違う 2: たぶん違う 3: どちらでもない 4: たぶんそう 5: 確実にそう	
PPV	スコア5: 36.5% スコア4: 50.0%	スコア5: 46.4% スコア4: 60.7%

間質性肺炎 STEP2

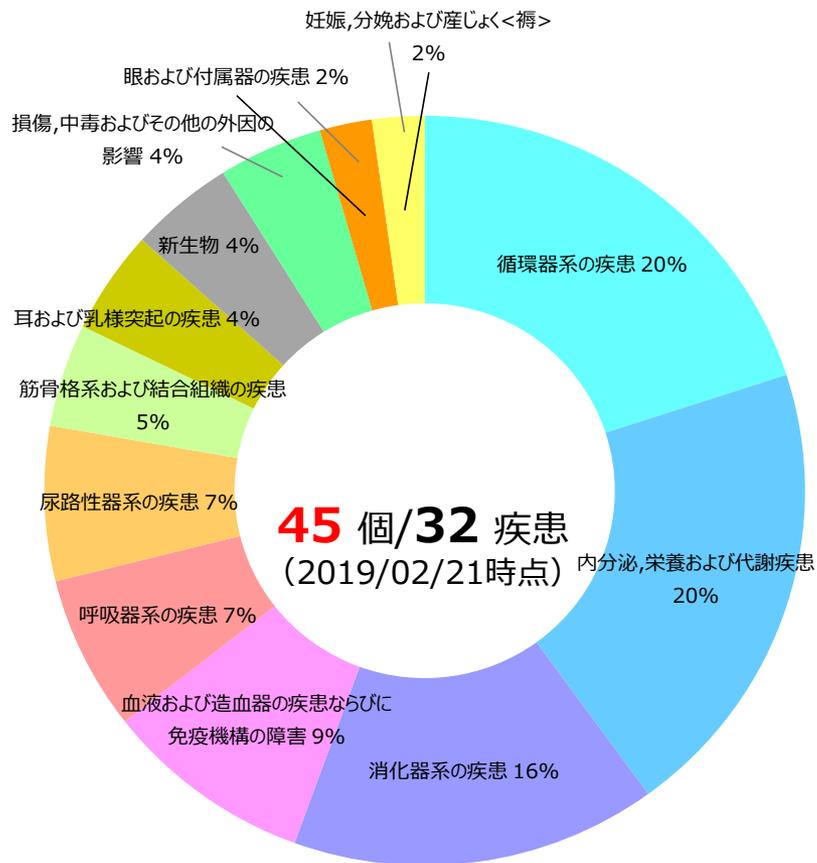


All possible caseによる間質性肺炎の感度計算



ePhenotyping（アウトカム定義）ライブラリ作成（AMED中島班）

A. ICD10分類別の集計

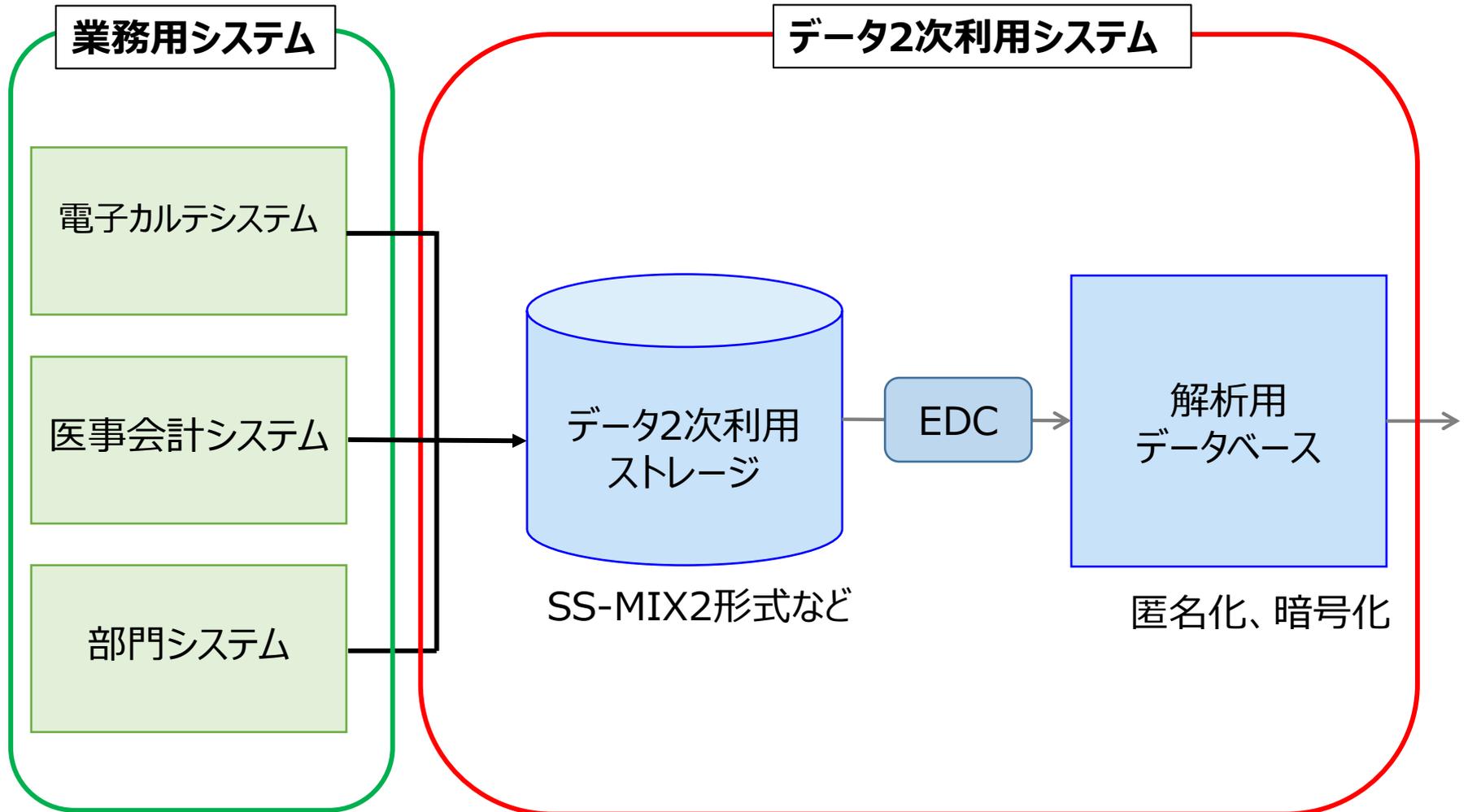


現在AMED・宇山班では、MID-NETに用いるためのアウトカム定義開発を行っている。今年度のターゲットに悪性腫瘍が含まれる。

B. 対象疾患/現象の内訳

国際疾病分類名	病名/現象	個数
新生物	乳癌	2
血液および造血器の疾患ならびに免疫機構の障害	DIC	1
	無顆粒球症	2
	血小板減少症	1
内分泌, 栄養および代謝疾患	I型糖尿病	1
	II型糖尿病	1
	甲状腺機能亢進症	1
	高血糖	2
	脂質異常症	1
	低カルシウム血症	1
	糖尿病	1
	薬剤性糖尿病	1
耳および乳様突起の疾患	聴覚障害	1
	難聴	1
眼および付属器の疾患	眼内出血	1
循環器系の疾患	急性心筋梗塞	2
	心不全	1
	高血圧	1
	頭蓋内出血	2
	VTE	1
	PTE	1
	脳梗塞	1
呼吸器系の疾患	間質性肺炎	3
消化器系の疾患	急性膵炎	1
	消化管穿孔	1
	消化管出血	4
	薬剤性肝機能障害	1
筋骨格系および結合組織の疾患	横紋筋融解症	2
泌尿器系の疾患	急性腎不全	2
	薬剤性腎障害	1
妊娠, 分娩および産じょく<褥>	妊娠・分娩	1
損傷, 中毒およびその他の外因の影響	アナフィラキシー	2
計		32疾患 45個

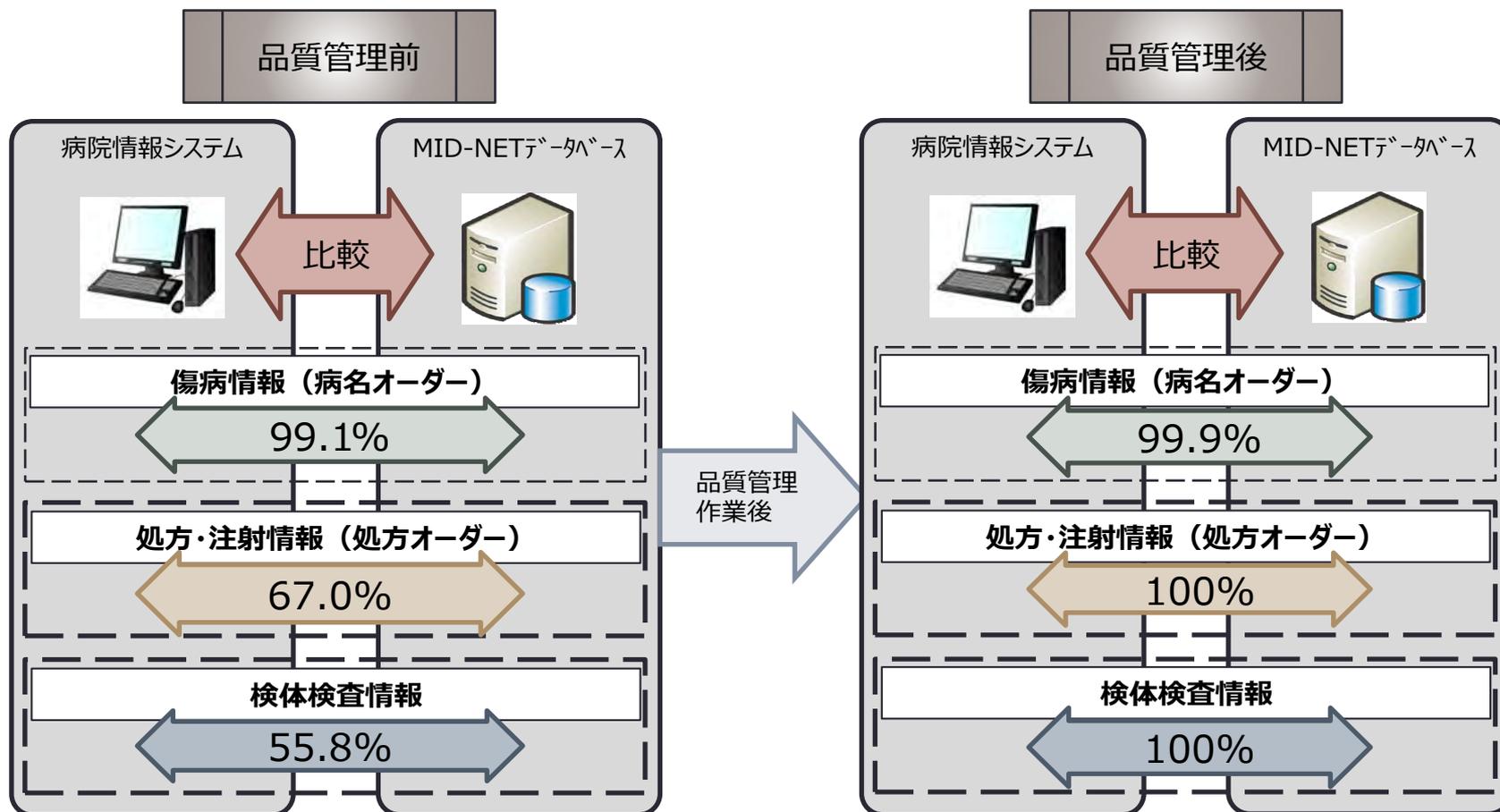
病院情報システムからのEDCによるデータ2次利用



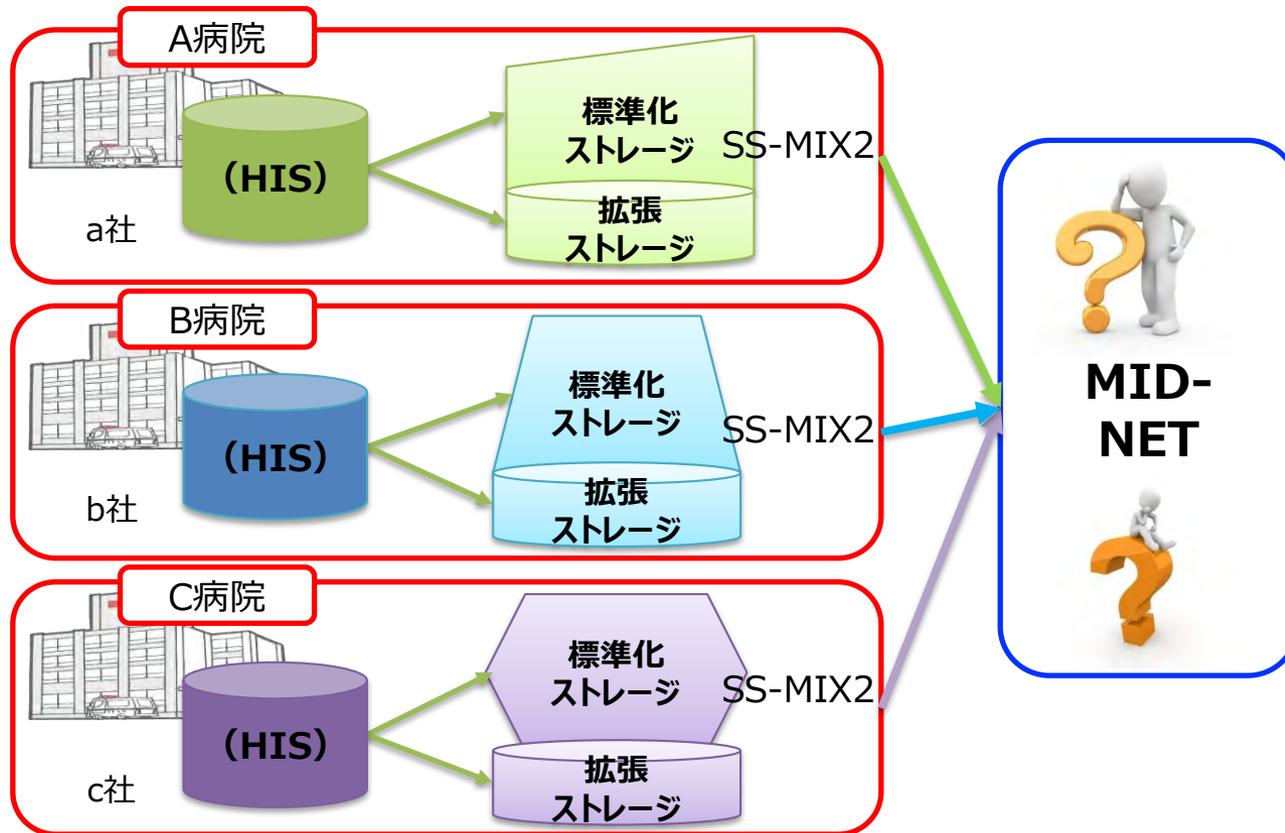
MID-NETのデータの品質管理について

PMDA資料より引用・改変

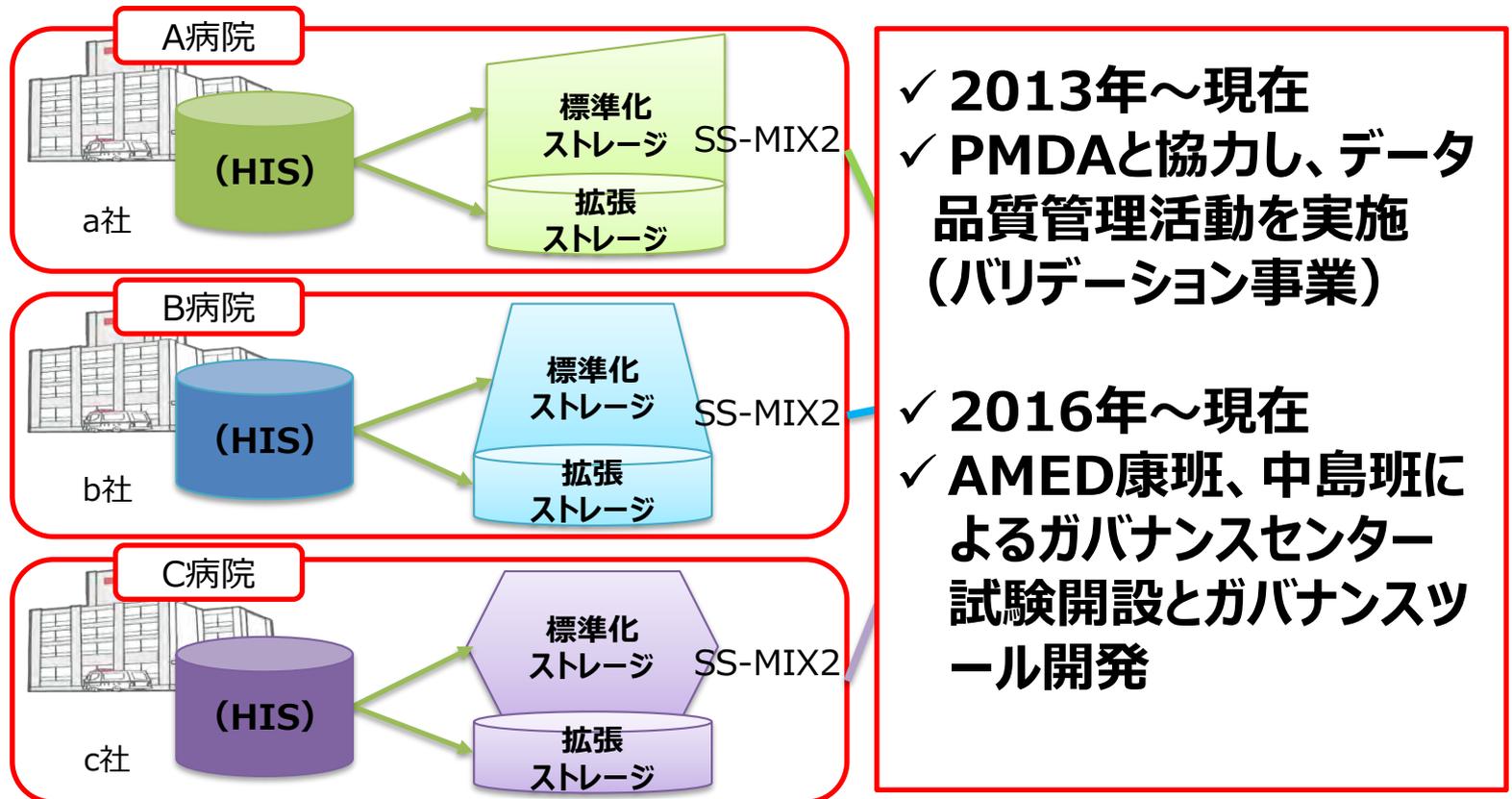
- PMDAが各医療機関に設置したデータベースのバリデーションを順次実施し、病院で保存されている実際の情報が正確にデータベースへ転送されていることを確認した



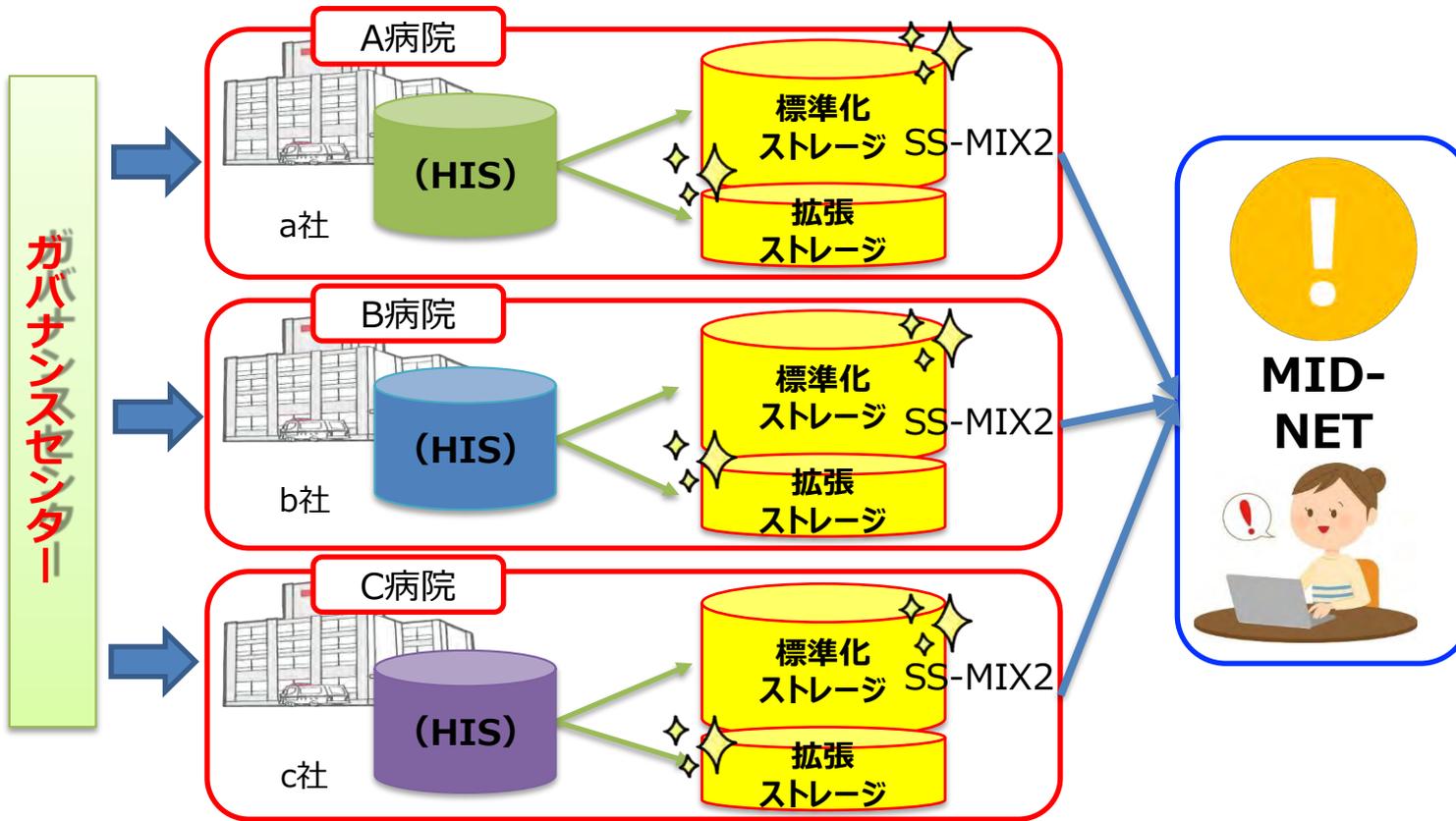
■ MID-NETのデータ品質の課題



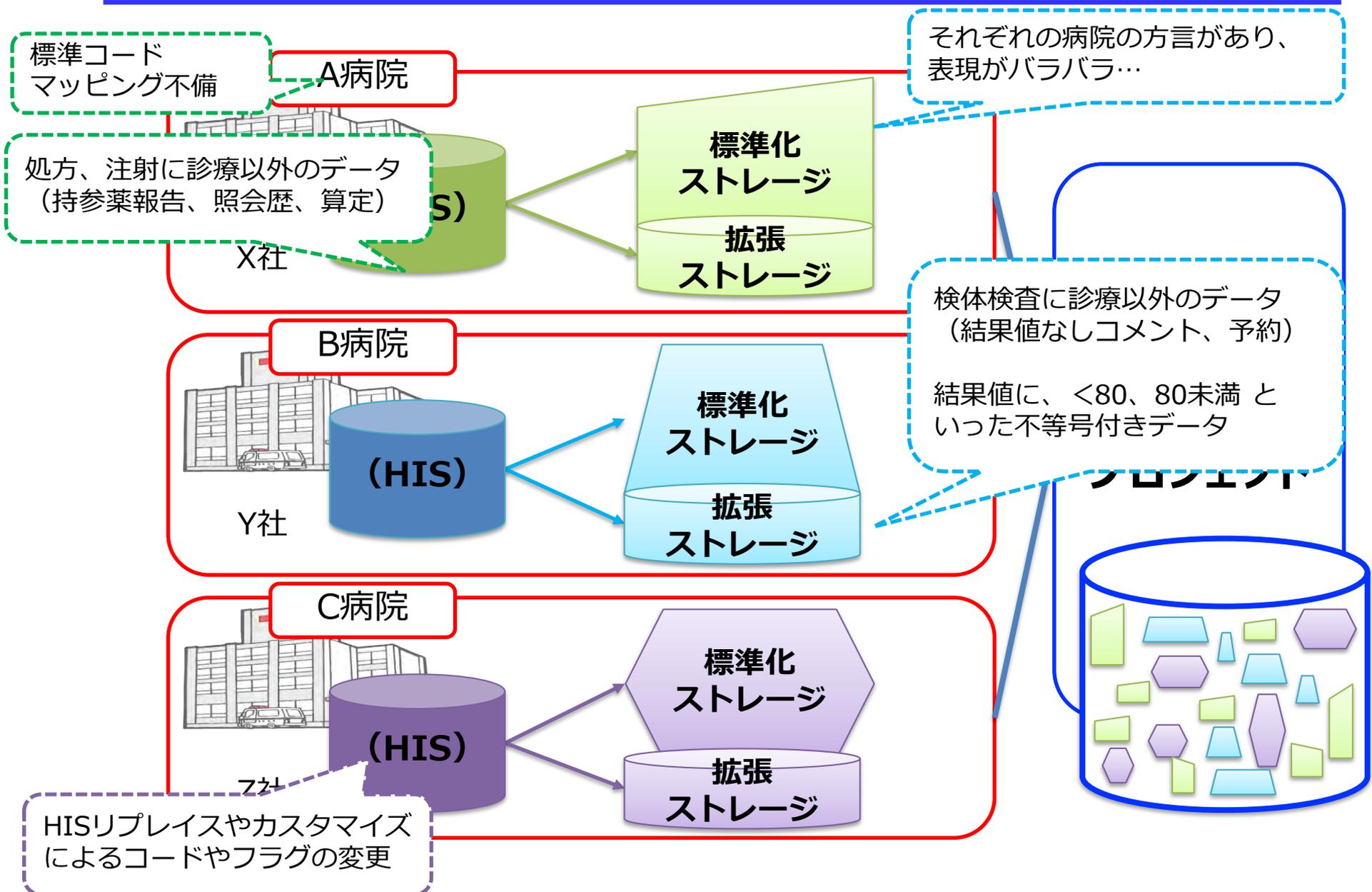
■ MID-NETのデータ品質の課題解決 (MID-NET)



■ MID-NETのデータ品質の課題解決 (MID-NET)



MID-NETで噴出したデータ品質の課題



MID-NETで生じたシステム運用・データ品質課題の洗い出しと整理

① マッピング不備、未設定

- 標準コード未設定
- 院内製剤、器材、治験薬、持参薬 →マッピング留意
- 単位、用法、投与経路 →標準コード見直し
- ワープ口病名

② HISリプレイス、カスタマイズ、運用など制限事項

- 注射オーダ → 旧・新HISで削除フラグの違い
- 退院転帰コードの変更
- 医薬品、検査項目の追加
- 細菌検査) 抗酸菌感受性の結果が、外注検査へ移行のためデータ無し

③ SS-MIX2への送信エラー

- 「来院等情報（退院）」の死亡日の欠落、手技が入力されていない注射オーダの欠落
- HL7メッセージ変換エラー

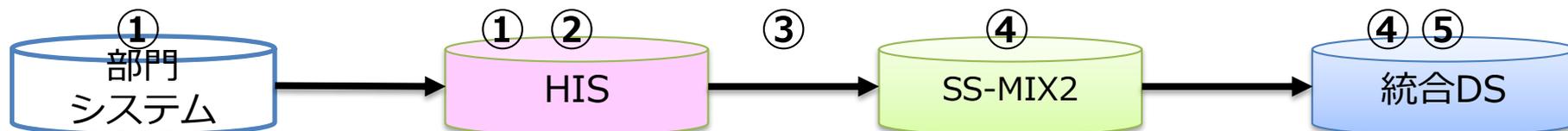
④ 診療以外のデータが存在

- 検体検査値なし、コメントなど
- 持参薬報告、中止処方、指示変更など
- 放射線オーダとして検査室予約

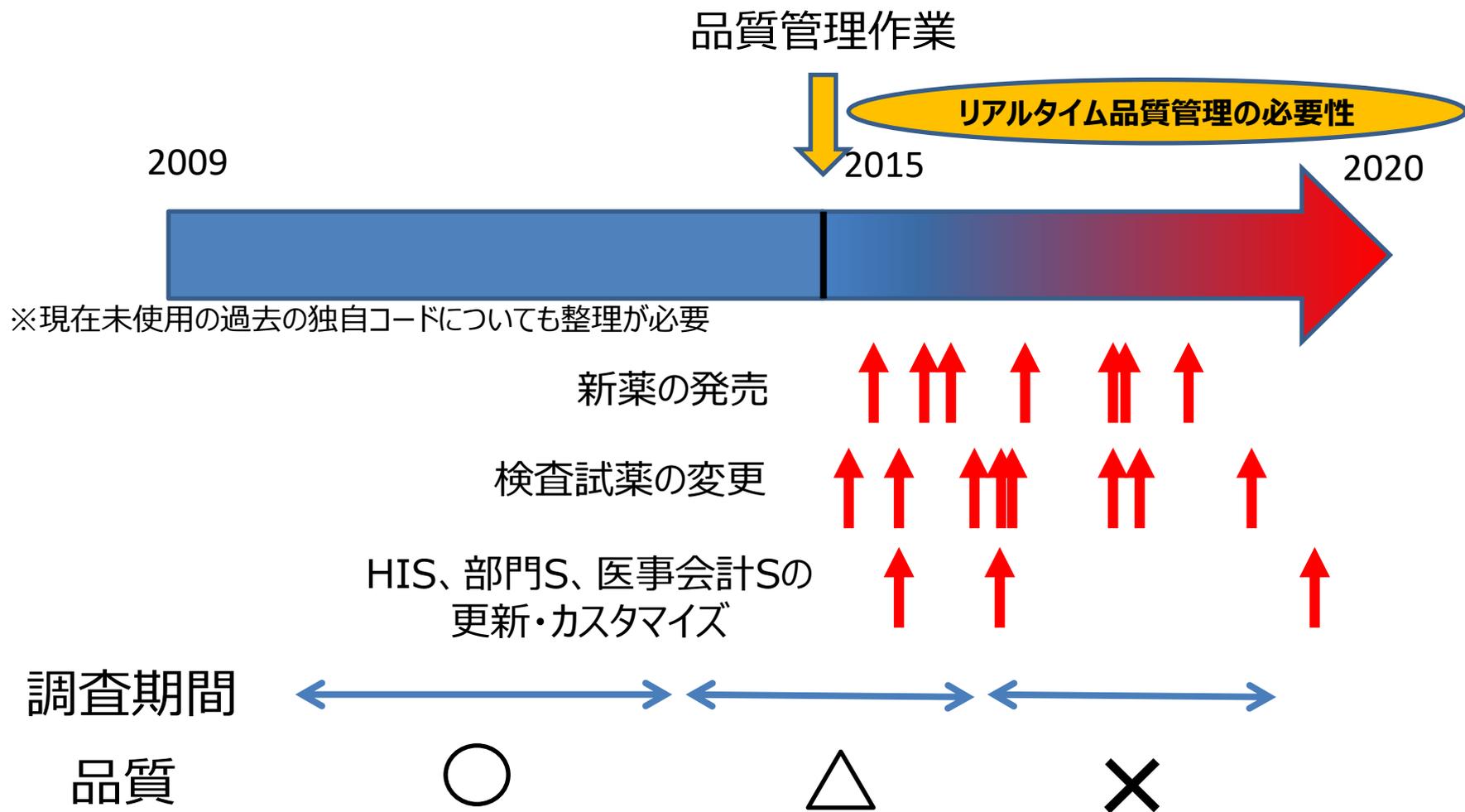
⑤ 二次利用データベース取込

- 退院サマリ傷病情報の取込欠落
→退院サマリ病名開始日未入力の場合
- 日付の認識（オーダ日、実施日、服用開始日など）

約300件



過去のデータ品質管理と リアルタイム品質管理

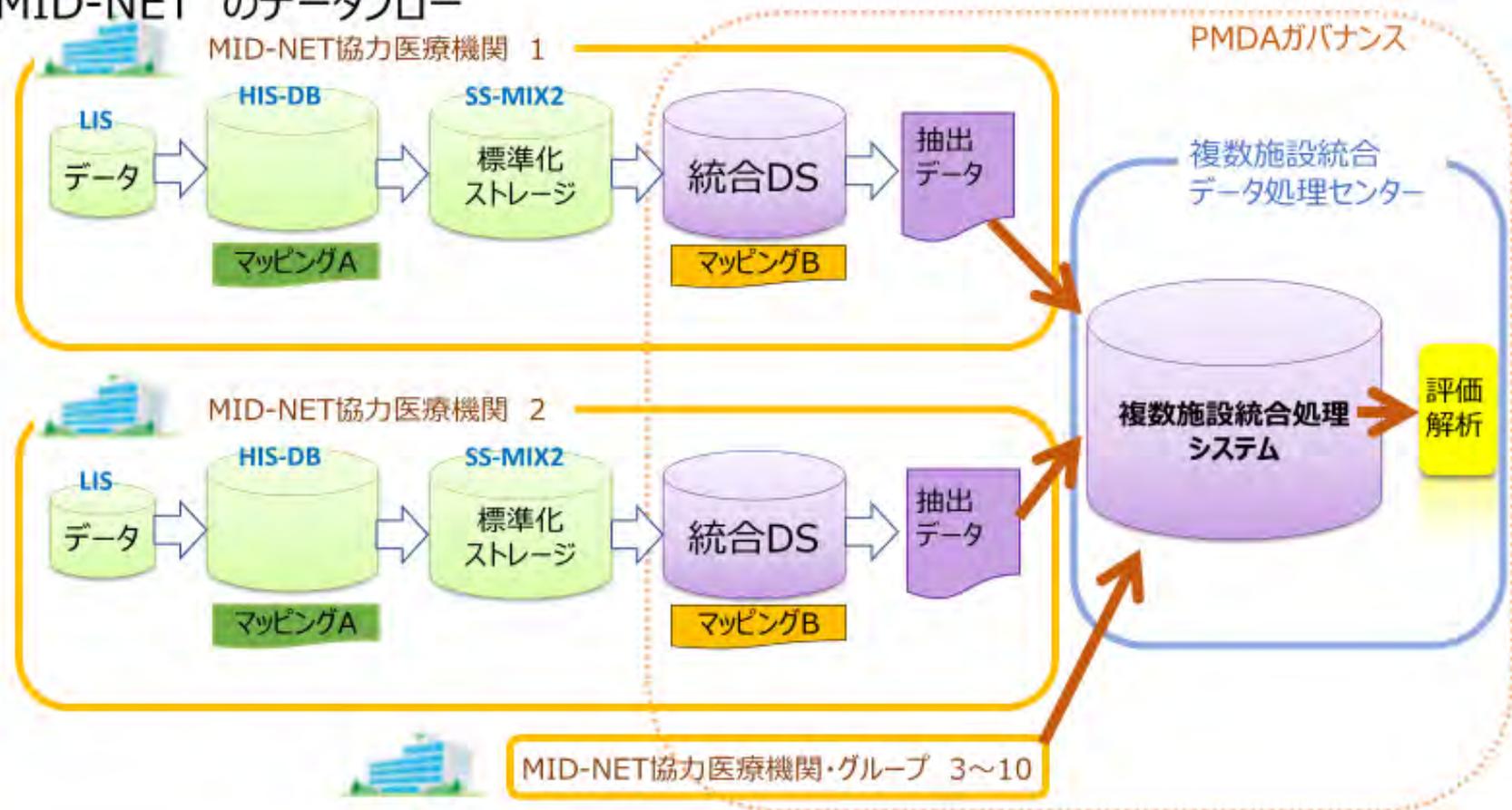


2019年1月の九州大学病院HISリプレイスに伴う MID-NET問題発生

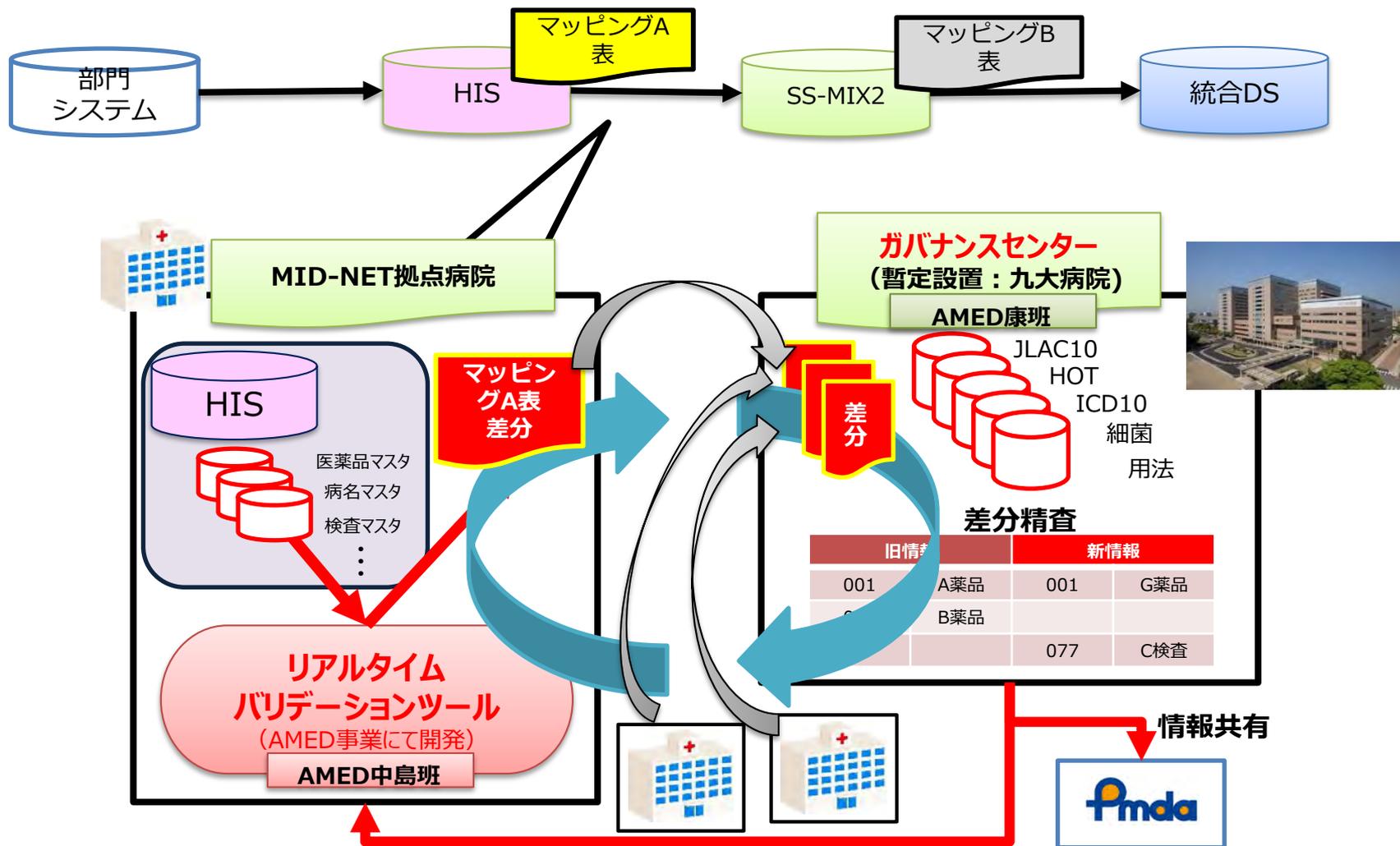
No	エンティティ	内容
1	G : 傷病情報 (病名オーダ)	標準コード (病名管理番号) で管理されていると思われるコード体系のデータがローカルコードとして (PRB-3のコーディングシステム名が99Z21) 送信されている。
2	K : 処方・注射情報 (注射実施)	ID、日付、オーダ番号を用いて比較したところ、HISのみのデータが存在した。 (SSMIXの最終メッセージがDelとなっている。)
3	その他 (転科転棟実施)	2018/12/6に発生したデータ誤送信に対するリカバリの確認を行ったところ、リカバリが行われていない誤送信データが1件存在。
4	その他 (外来診察受付)	2018/12/6に発生したデータ誤送信に対するリカバリの確認を行ったところ、誤送信が発生していないデータがリカバリ実行日(2019/1/24)に新たに4件送信されていた。
5	I : 処方・注射情報 (処方実施)	ID、日付、オーダ番号を用いてHISとDSを比較したところ、DSのみに存在するデータが認められた。
6	B : 来院等情報 (外来)	ID、日付、オーダ番号は同じだが診療科コードが異なるデータ (片方が診療科コードNULL) が存在した。HIS側は2レコードとなるが、DS側には1レコードとなり差分となっていた。
7	G : 傷病情報 (病名オーダ)	ローカルコードで管理されていると思われるコード体系 (Z始まり) のデータが標準コード (病名管理番号) として (PRB-3のコーディングシステム名がMDCDX2) で送信されている。
8	K : 処方・注射情報 (注射実施)	ID、日付、オーダ番号、医薬品ローカルコードをキーとして比較したところ、キー情報は一致するものの診療科がHISとDSで異なるデータが認められた。(HIS : NULL、DS : 06循環器内科)
9	K : 処方・注射情報 (注射実施)	RXA-4 (投薬終了日時) の編集ルールが事前提供資料と実際の送信内容で異なる。

MID-NETにおける標準コード管理手順

MID-NET のデータフロー



リアルタイムバリデーションツールによるMID-NETガバナンスのトライアル (AMED康班、中島班)



JLAC 改訂スケジュール（予定）

JLAC10

公表されたコード表のメンテナンス

更新
停止

JLAC10並用期間 → JLAC11

JLAC11

開発（公開
/パブリックコ
メント）

準備

公表されたコード表のメンテナンス

2017

2018

2019

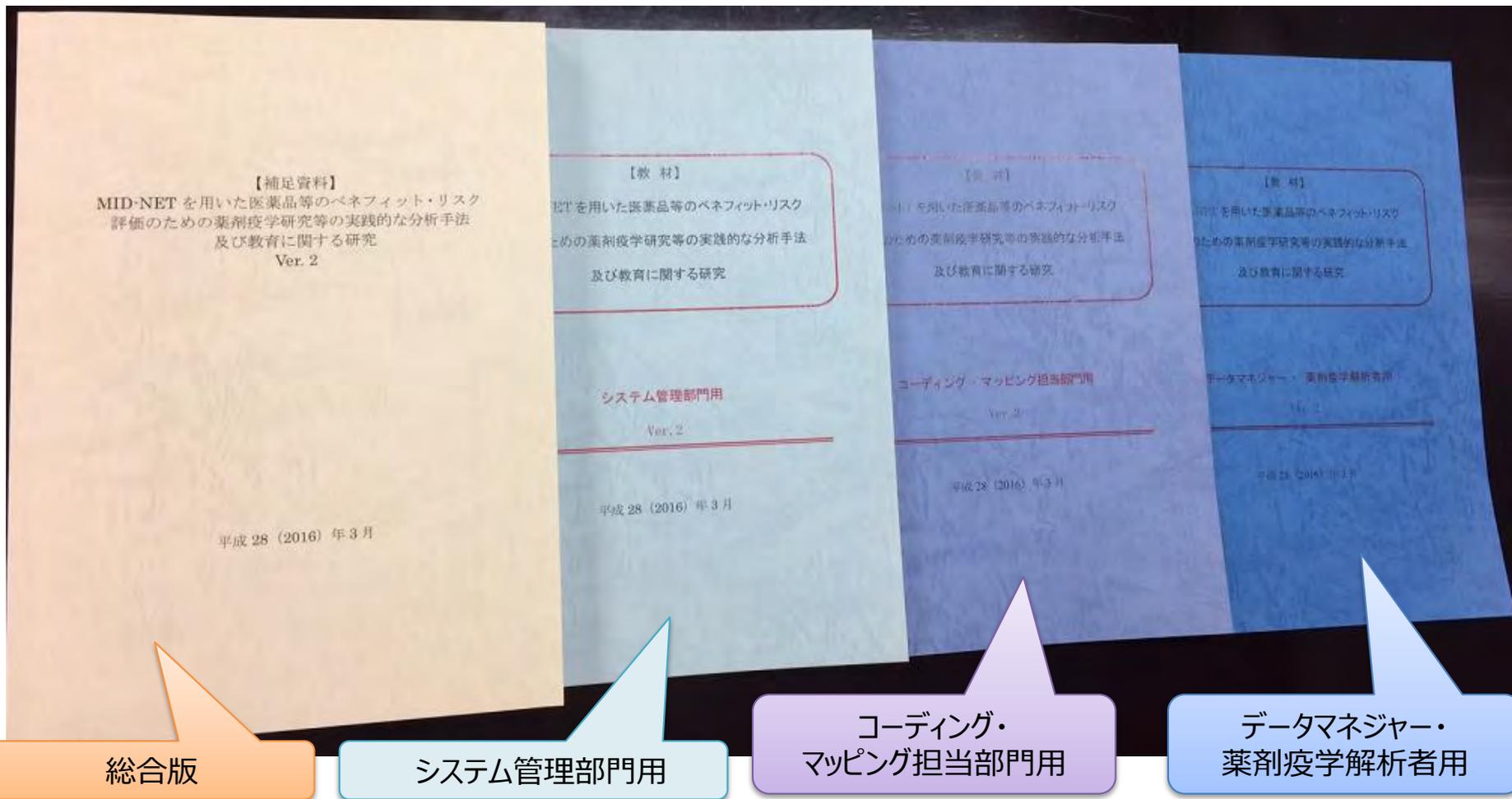
2020

2023

2024

MID-NETのバリデーション事業、厚労科研/AMED事業での経験を教材化

AMED研究事業「MID-NETを用いた医薬品等のベネフィット・リスク評価のためのデータ標準化の普及に関する研究」（代表：中島直樹）

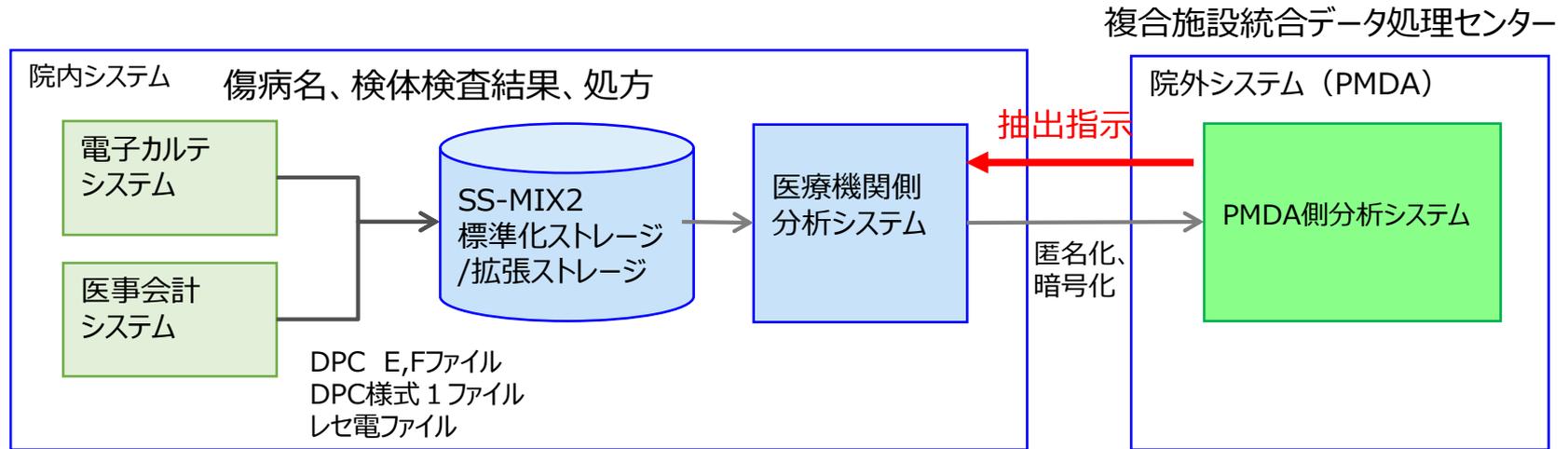


↑ 2016年度AMED研究にて開発した教材

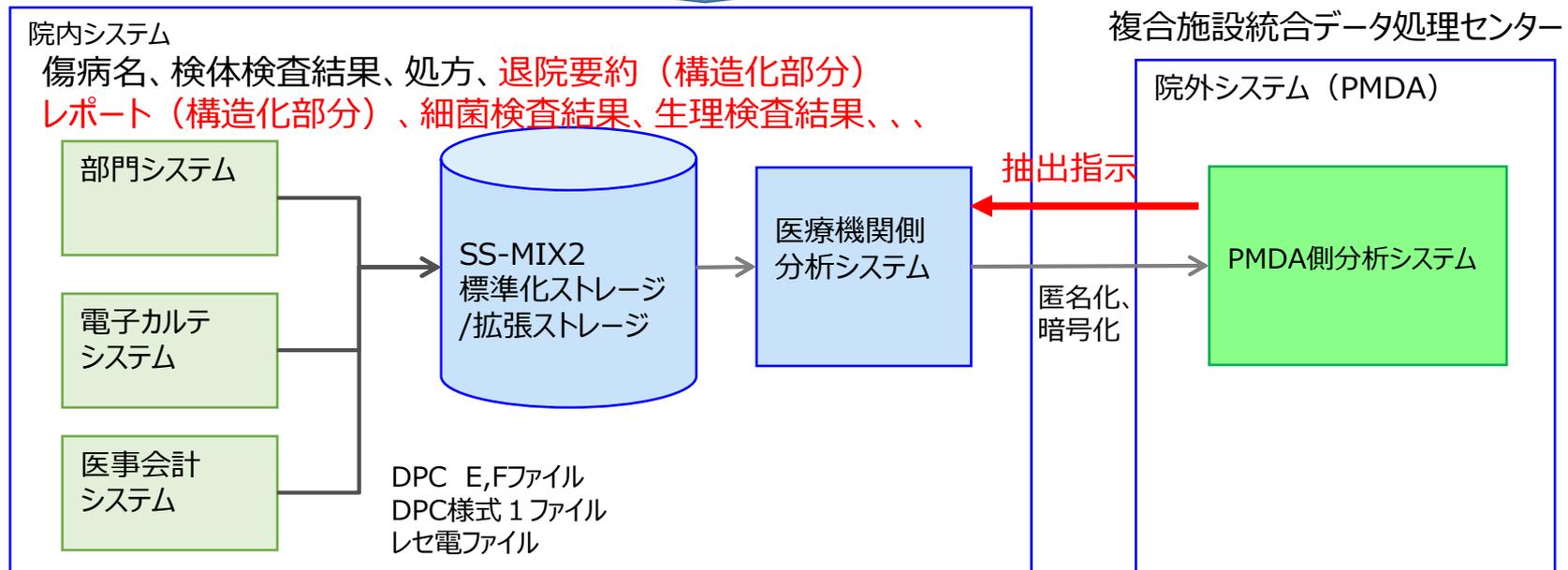
**2015年度に初版、2016年度、2018年度にバージョンアップし、
2019年度から3年間の新規のAMED事業²⁹（中島班）でさらにバージョンアップする予定**

MID-NET事業の利活用データ種の「今」と「期待」

現在



期待



AI時代に向けた医療情報次世代標準化シンポジウムのご案内

－SS-MIX2からHL7FHIRに向けて－

主催：日本医療情報学会 NeXEHRs研究会
(次世代健康医療記録システム共通プラットフォーム研究会)

共催：東京大学COI「自分で守る健康社会」健康医療情報ICT標準化チーム

日時：2019年7月8日(月) 13:30～15:30

場所：東京大学伊藤国際学術研究センター内「伊藤国際謝恩ホール」
https://www.u-tokyo.ac.jp/campusmap/cam01_00_33_j.html

プログラム：

主催者挨拶

日本医療情報学会NeXEHRs研究会代表 大江和彦(東京大学)

講演 題未定 厚生労働省(交渉中)

講演 医療RWD共通プラットフォームに向けて

日本医療情報学会代表理事 中島直樹(九州大学)

講演 医療情報標準の国際動向

日本医療情報学会副代表理事 澤 智博(帝京大学)

講演 電子カルテの標準化－日本医師会からの期待－

日本医師会理事 石川広己

講演 SS-MIXとHL7FHIR－それぞれの役割－(仮題)

日本HL7協会 木村通男(浜松医科大学)

講演 次世代健康医療記録システムのための共通プラットフォームの構築へ

日本医療情報学会NeXEHRs研究会代表 大江和彦(東京大学)

質疑応答

参加資格：なし(日本医療情報学会の会員でない方もご参加いただけます。)

参加費：無料

定員：360名(先着順)

医療情報技師資格更新ポイント：2ポイント

告知！

<https://www.jami.jp/?mode=seminar&no=37>

データ駆動型研究事業では、データ品質管理を効率的・継続的に行いましょう



ご清聴ありがとうございました

ご質問は、nnaoki@info.med.Kyushu-u.ac.jpへ