

第1回 医療等情報の二次利用に関する技術作業班

日本の医療情報データ 二次利用の現状と課題

2024.2.8

東京大学情報基盤センター

データ科学研究部門 客員研究員

清水 央子



リアルワールドデータ（RWD）と医療等情報データ

リアルワールドデータ（RWD）とは、日常の実臨床の中で得られる医療データの総称。レセプトデータや電子カルテのデータなど、様々な情報源から収集される患者さんの健康状態や医療行為のデータが含まれる。臨床試験とは異なり、複数の疾患や診断名のない症状も記録されており、医療ビッグデータとして活用が期待されている
リアルワールドデータの例としては、以下のようなものがある

- レセプトデータ
- 電子カルテのデータ
- 患者レジストリデータ
- 健診データ
- ウェラブルデバイスから得られるバイタルデータ

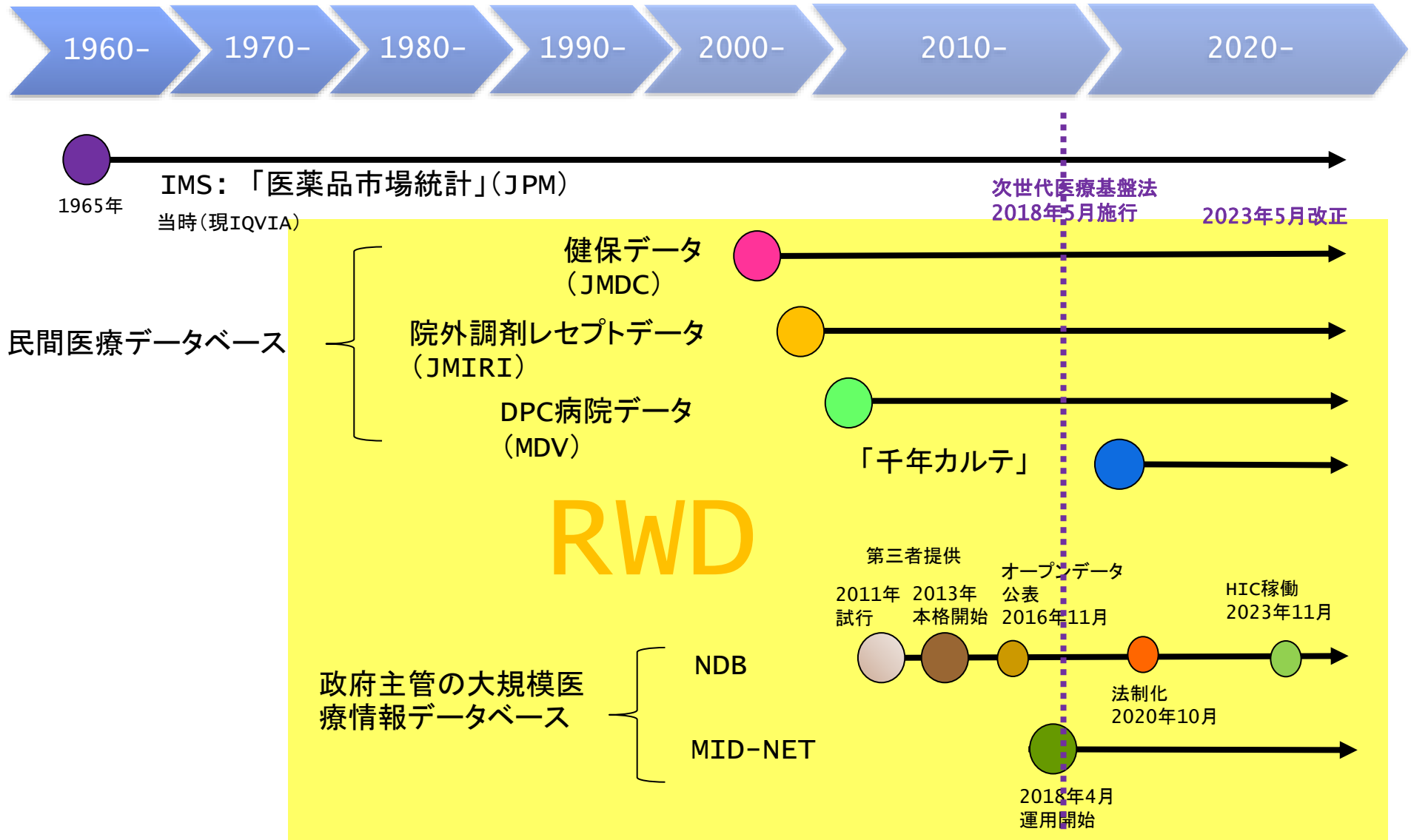
※MS Bingが「リアルワールドデータとは」で作成した資料（清水が一部改訂）

※逆に「臨床試験データ」は、実臨床の場ではなくコントロールされている下での治療結果なのでRWDには含まれない

👉 今回のWGでの議論の対象は「RWD（他医療関連情報）」の「二次利用」



日本における医療情報データの歴史





「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」

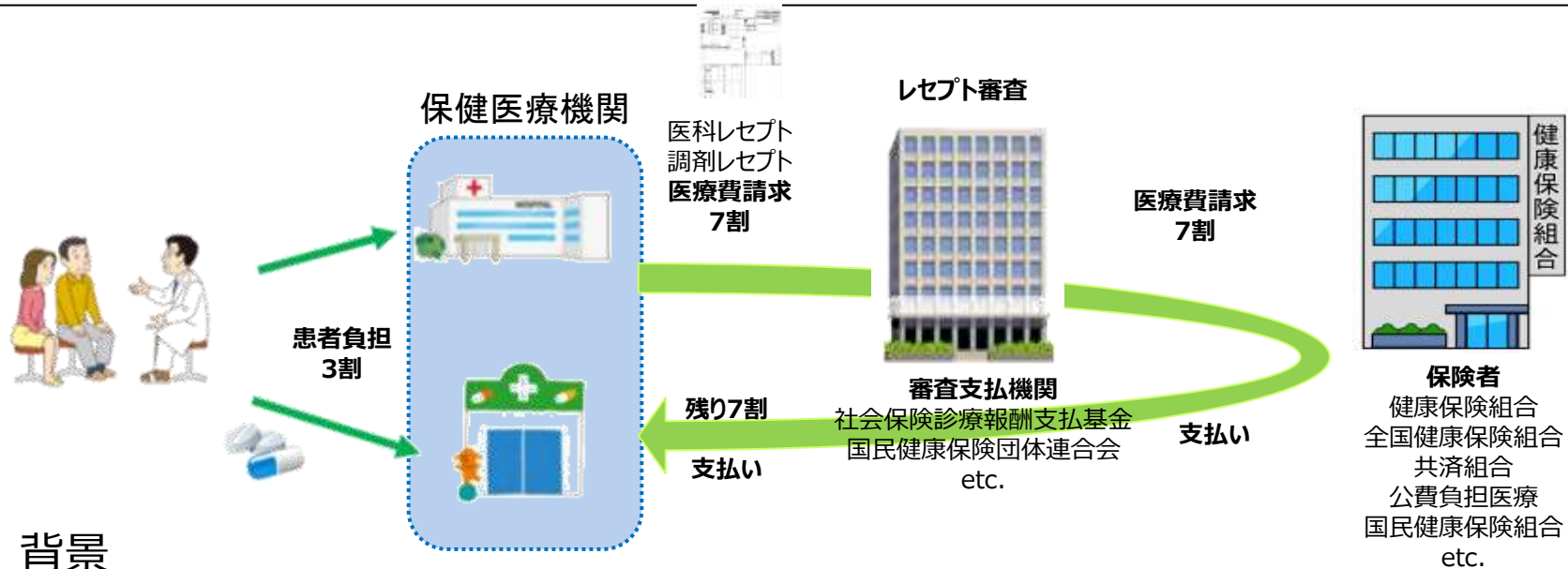
日本薬剤疫学会ホームページに掲載されたデータベース(他のリンクにないものはすべて日本薬剤疫学会ホームページに掲載されている)

データベース名	医療機関データベース										製薬者データベース						保険(調剤)薬局データベース				その他	
	NHO NCD/MIA	4DR	HCD/RWD	LDI	MID-NET	値観値	MDV	JMDC	NCD	NDB	JMDC	Cross Fact	MDV	DxC	IQVIA Claims data	Cross Fact	IQVIA NPA data	日本製薬	JMDC	HPWS		PatentMap
所有者(組織や会社など)の名称	国立病院機構	4DR	リアルワールドデータ株式会社	一般社団法人ライフデザインシステムズ	医療法人社団 医療情報センター	株式会社 JAMDC	一般社団法人 National Clinical Database	厚生労働省	株式会社 JAMDC	株式会社 インターナショナルワールド	JamNet	メディカル・データ・ソリューション株式会社	DxC	IQVIA ソリューションズ ジャパン株式会社	株式会社 インターナショナルワールド	株式会社 日本製薬協会	株式会社 JAMDC	5SRH Center Envis	株式会社 研経研			
データベース名称	NHO NCD/MIA	4DR Research Network	RWD データベース	千年カルテ	Medical Information Database Network (MED-IN)	厚生メディカルデータセンター	EMM Prossard	JMDC 医療情報データベース	レセプト情報・特定診察情報データベース	JMDC 調剤者データベース	Cross Fact	JamNet データベース	EMM Prossard	DeD Database	IQVIA Claims data	Cross Fact	IQVIA NPA data	調剤者データベース	JMDC 調剤者データベース	HPWS (National Health and Survey Database)	PatentMap	
連絡先(代表メールアドレス等)	730-dip@ncd.mhl.go.jp	info@4dr.com	info@rwd.com	contact@rdp.or.jp	info@medinfo.jp	em_mails@mdc.co.jp	mdshp@mdc.co.jp	info@ncd-cdb.jp	www.ncdb.jp	mdshp@mdc.co.jp	www.ncdb.jp	support@jam-net.co.jp	em_mails@mdc.co.jp	new-protein@doc-inc.co.jp	www.rwd.com	www.jamdc.com	www.jamdc.com	info@jamdc.co.jp	mdshp@mdc.co.jp	www.hpws.com	patentmap@patentmap.com	
ウェブサイトURL	http://www.ncd.mhl.go.jp/ncd/000001.html	http://4dr.com	http://www.rwd.com	http://rdp.or.jp	http://www.medinfo.jp	http://www.mdc.co.jp	http://www.mdc.co.jp	http://www.ncd-cdb.jp	http://www.ncdb.jp	http://www.mdc.co.jp	http://www.ncdb.jp	http://www.jam-net.co.jp	http://www.mdc.co.jp	http://www.doc-inc.com	http://www.rwd.com	http://www.jamdc.com	http://www.jamdc.com	http://www.jamdc.co.jp	http://www.hpws.com	http://www.patentmap.com		
紹介論文DOI	DOI: 10.21467/CLPEP.5339072				10.1002/jpm.4777 10.1002/jpm.4670			DOI: 10.1007/s10095-013-0823-3 DOI: 10.5069/kjca.2014.47.3.437														
データベースの概要	国立病院機構の施設の電子カルテ情報と全施設のDPCレセプトデータを、連携体制に構築、集約したデータベース	医療機関内の臨床データを集約、匿名化し、活用することを可能にするネットワーク	電子診療録、DPCレセプト等の医療情報を集約したデータベース	特定機能病院、大規模病院の電子診療録 (EMM) が他医療機関との連携を可能にするデータベース	独立行政法人医薬品医療総合機構 (PMDA) が他医療機関との連携を可能にするデータベース	厚生労働省が推進する医療情報データベース	全国480の医療機関を対象とした、大規模な診療データベース	全国500以上のDPC非DPC病院から収集したレセプト・電子カルテ臨床検査情報	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース	厚生労働省が推進する医療情報データベース
データベースの構築	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年	2010年

次々と新しいデータベースが構築されています



日本の医療情報の中心は「レセプト」



背景

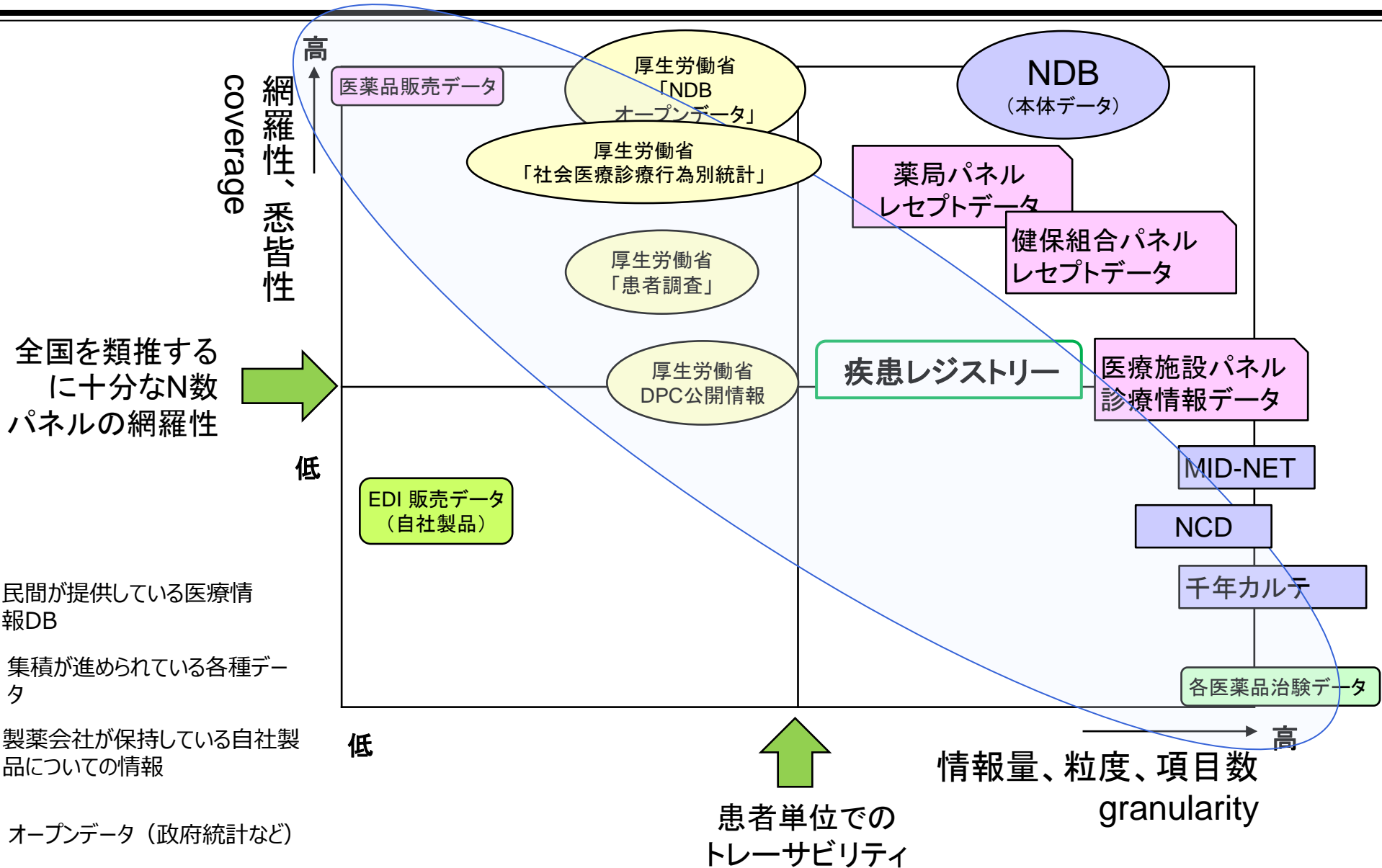
- 国民皆保険制度の下、全国民の保険医療についての情報がレセプトに反映されている → 複数の医療機関における治療履歴が追える
- 医療費の患者負担 3 割の残り 7 割の「請求書」なので、最低限必要な情報が記載されている → データが“きれい” (vs カルテは自由記載)
- (カルテ等に比べて) 電子化が進んでいる

ただし

- レセプト病名
- 治療内容はわかるが、そのアウトカムや転帰については明快でない場合も
例：検査をしたことはわかるが結果は記載されていない
- 保険組合を移動したら追うことが難しい → 長期にわたる治療記録が取れないことも



日本の医療情報データ相関図





日本のRWDの中心は民間のデータ

民間の医療情報データは

※データベース、ベンダーにより差がありますので、詳細は各社にお問い合わせください

- データが「きれい」
 - 集積したデータのクレンジングがされている
 - 主なマスターデータを裏に有している
- 入手までの期間が短く、データの速報性が高い
- (データを良く知る) ベンダー側に解析まで依頼することも可能
(※ただし追加料金必要)

一方で

- 代表性をよく理解しないと落とし穴も...
- とても高価... 



👉 「高かろう良かろう」の民間データ (vs 「安かろう悪かろう」)

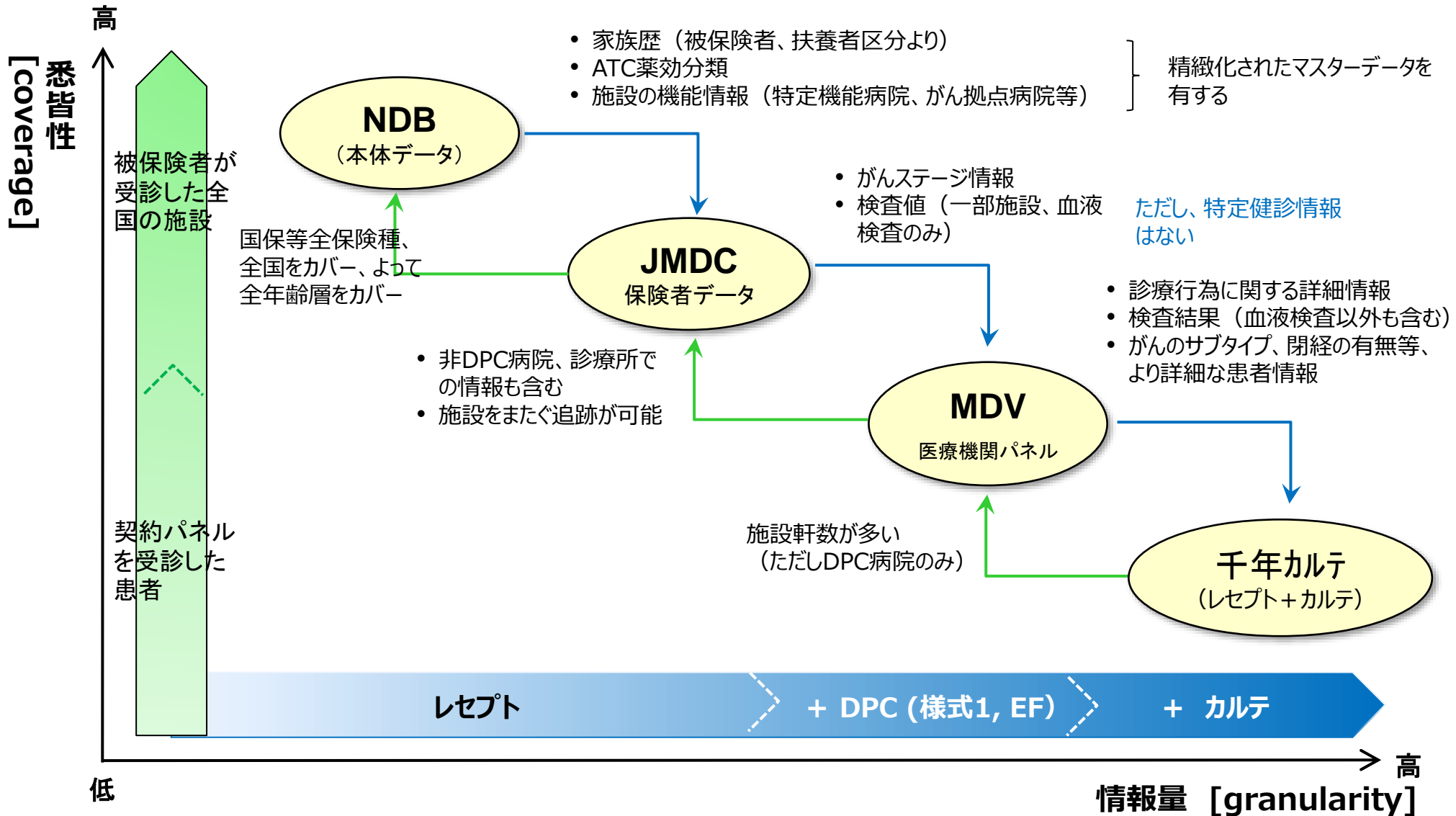


レセプトデータの比較 ~NDB vs JMDC/MDV~

MDV		NDB		JMDC
約450施設、年間約1100万人の診療記録	<<	N数は国内最大(保険医療のほぼ全数が対象)	>>	約850万人の組合員のレセプトデータ
データ集積期間:2010.4-	≐	データ集積期間:2009.4-	<	データ集積期間:2005.1-
すべての保険組合が対象 保険外も一部カバー	>	すべての保険組合が対象	>>	組合健保のみ → 高齢者のデータが少ない
患者IDは施設毎に精緻化されている	>	患者ID1+ID2の突合プロセスが必要	<<	患者IDは保険組合毎に精緻化されている
当該施設への転院前・後は追えない	<<	長期の履歴が追える	>	保険組合の転入前、転出後は追えない
マスターデータが整備されている	>>	マスターデータは解析者が作成	<<	マスターデータが整備されている
高価 ※解析まで依頼することも可能	<<	安価	>>	高価 ※解析まで依頼することも可能
契約すれば数週間程度でデータの取得が可能	>>	申請後データ取得までの1年以上かかる	<<	契約すれば数週間程度でデータの取得が可能
直近データは3-4か月遅れ	>>	直近データは5-6か月遅れ。ただし申請内容やタイミングによりもっと遅れることも	<	直近データは5-6か月遅れ
個人単位	=	個人単位	<	家族(被保険者-扶養者)関係がわかる



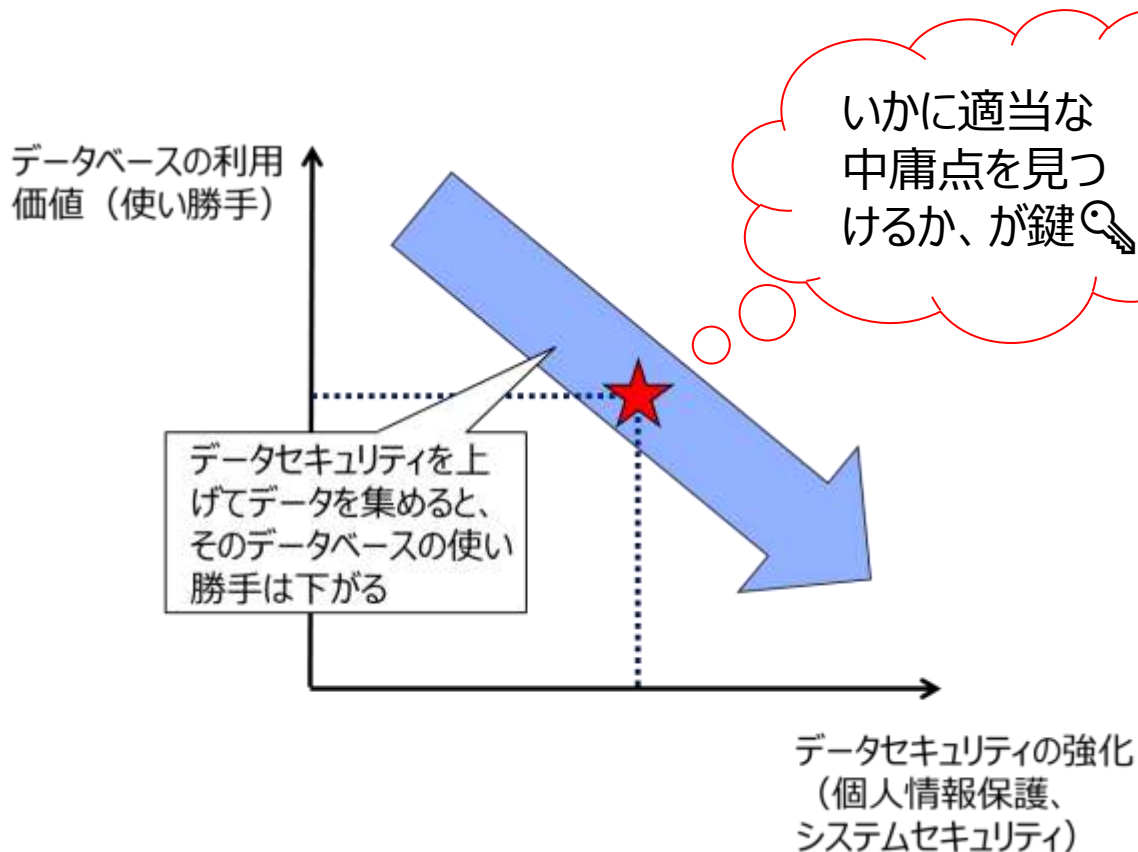
「カルテ」データ含めた各種RWDの優位性比較





セキュリティの強化 vs 使い勝手

データを集積（DB化）すれば、その利用が進むわけではない！
一方で、より質の良い（粒度・悉皆性の高い）データを提供するために、セキュリティ要件を強化しすぎ使い勝手が著しく低下しては利用の推進にはつながりにくい



👉 利用推進のための条件

- ①データ自体の質の向上
- ②適正なセキュリティ要件に基づく利用環境の整備（データ提供体制、解析環境）



①データ自体の質の向上

「データを集める」だけでは利用は進まない

利活用者から見たデータの標準化、Visiting環境での利用環境のあり方等について

課題：データ自体の質の向上

- 粒度 and/or 悉皆性を上げる↑
複数DBの連結（例：NDB+介護）→ そのためにはデータの標準化が必要
- マスターの整備（例：NDBの診療報酬/薬剤/医療機関マスター等、JLAC10）
- 情報の質（例：検査結果@カルテ）→ インput情報（データ入力時）の質の向上 → 医療現場からの協力も必要なのでは？（vs 現場の負担増）
- 精緻化された患者IDによる同一患者の特定と連結

👉 現時点では、利用者側にとって本来の解析の前の下準備の段階での負荷が大きく、誤った解析をしてしまう可能性も



NDB解析に必要な「マスターデータ」

例示

主要マスターデータ	一般公開されているマスターの種類
傷病名マスター (約26,000レセ電処理用傷病名)	「診療報酬情報提供サービス(厚労省保険局)」1) 「社会保険診療報酬支払基金」1) 2) 「医療情報システム開発センター[MEDIS]」2) 3)
医薬品マスター (約21,000の医療用医薬品)	
診療報酬マスター (医科診療行為: 約7,500)	
医療機関マスター (全国約18万医療施設)	全国版のマスターは存在しない 研究者各自が各地方厚生局のサイトより47都道府県分の医療機関リストをダウンロードし統合

共通課題: いずれもある時点でのマスターであり
過去に遡った累積情報を収載していない

データ解析という“二次利用”を想定したマスターデータが整備されていない

- 1) 別サイトで運用されているがソースは同じ
- 2) 「傷病名」においては共同で更新
- 3) 「診療報酬」においては、手術・検査など一部の項目のみ



「グランドマスター」の必要性

例示

RQ「DPC病院とそれ以外で〇〇の治療方法の差異を調べたい」

RQ「抗インフルエンザ薬の各シーズンの投与患者数の推移を知りたい」

『医療機関マスター』によるフラグ付け

『医薬品マスター』より過去のコードも含めて抽出



1. 地方厚生局URLより各都道府県の医療施設のリストと付随情報をD/LLし、
2. DPC病院のリストと照合してリストを作成

* 解析対象期間中に存在したすべてのコードを抽出することが必要

年度	都道府県番号	点数表	医療機関番号	医療機関名	DPC対象施設フラグ
2020	01	1	1234567	〇〇〇大学病院	1
2020	01	1	4567890	△△△	
2020	02	1	3456789	●●●	
2020	02	3	7891234	▲▲▲	

医薬品コード (レセプトコード)	医薬品名	後発品フラグ
660443018	リレンザ	0
610443074	タミフルカプセル	
610462002	タミフルドライシロ	
622012101	イナビル吸入粉	
622622501	ゾフルーザ錠	
622622601	ゾフルーザ	
621972202	ラビアクタ点滴静注液バイアル150mg	0
621972102	ラビアクタ点滴静注液バッグ300mg	0
621972101	ラビアクタ点滴用バッグ300mg	0
621972201	ラビアクタ点滴用バイアル150mg	0
622638801	オセルタミビルカプセル75mg「サワイ」	1
622638901	オセルタミビルDS3%「サワイ」	1

医療機関コードが変更
※移転、経営母体の変更等に伴う

2013.3.31付で廃止された古いコード
※販売会社、名称変更等に伴う

年度	都道府県番号	点数表	医療機関番号	医療機関名	DPC対象施設フラグ
2019	01	1	5678901	〇〇〇大学病院	1
2019	01	1	4567890	△△△クリニック	0
2019	02	1	3456789	●●●市立病院	1
2019	02	3	7891234	▲▲▲デンタルクリニック	0



「カルテ」データは万能か？

例示

例：「カルテ」で検査結果が判明できたのは75%



HR/HER2サブタイプの特定



サブタイプ「HR/HER2」の特定ができた割合は、A,B,Cの3施設の中で差が見られた。
また当該施設で検査を行った場合は8割以上でサブタイプ「HR/HER2」の判定ができたが、
当該施設で検査がなかった場合は半分以下となっており、大きな差が見られた

	対象患者	HR/HER2 判定可能患者数*		当該施設での 検査あり**	検査なし	合計
A病院	113	93 (82%)	HR/HER2 判定可能*	627 (81%)	72 (46%)	699 (75%)
B病院	421	320 (76%)	HR/HER2 判定不可	147 (19%)	84 (54%)	231 (25%)
C病院	396	286 (72%)	合計	774 (100%)	156 (100%)	930 (100%)
合計	930	699 (75%)				

* HR/HER2両者が判定可能なケースのみ

** レセプトで確認



②解析環境の整備

セキュリティ要件が厳しすぎると使い勝手が悪くなり、利用が進まない

🏠 情報の価値を最大限保ちつつセキュリティ要件を満たすための“落としどころ”を探る必要

「セキュリティ」管理の対象：

- A) 悪意ある第三者による情報の不正取得
 - ✓ システムに侵入（ハッカー）
 - ✓ 管理部屋に物理的な侵入
 - ✓ タコバヤでの窃視、聞き耳による → これで情報が盗めるのか？
- B) 悪意ある利用者による目的外使用 → 出口管理で対応（案）
- C) データから個人が特定できてしまう可能性
 - ✓ 希少疾病（N数が少ない）
 - ✓ 最小集計単位の設定
 - ✓ それでも特定できることはある（専門医の専門分野の場合等）

- 申請手続きの簡素化と審査の迅速化
（例：NDBでは申請からデータ提供までのリードタイム1年以上、申請は研究機関単位＝学長が申請者という体）
- タコバヤの過度なセキュリティ要件の見直し
 - セキュアなネットワークとは？
 - タコバヤの規制
 - ✓ 入退室管理はどこまで必要か
 - ✓ 窃視、聞き耳対策は必要か？
- 個人が特定できてしまう可能性を排除するための規制
 - 最小集計単位の原則（K匿名化）
 - 罰則はどこまで？
 - 特定できる可能性があってはNG、では解析の幅が広がらない可能性も



NDBの解析は「三密」の環境で行うことが原則

例示

三密？

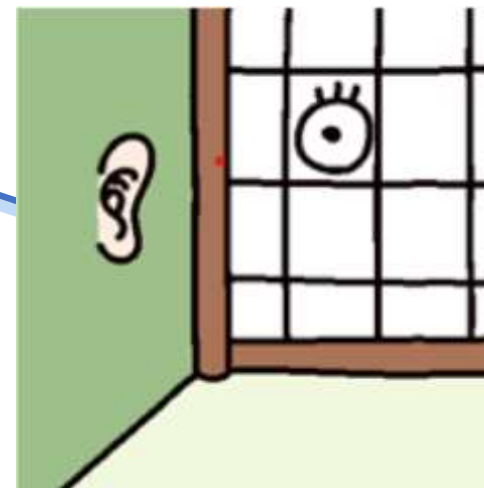
NDB専用室

解析用ソフトウェア

SQL、SAS、Python、R、JMP、
SPSS... も必要



バックアップ用HDDも必要



©ことわざ・慣用句の百科事典

※外部とネットワークでつながって
てはいけない → ソフトウェアのア
プデートもままならない

リモートワークでは解析作業は行うことができない...



ご参考：当研究室のNDBでの経験

例示

セキュアなネットワークでのアクセスが不承諾となった例：

申請したシステム構成図（この他にセキュリティに関する説明を加えて提出）



I型認定に近いシステム構成

不承諾の理由：

プライベートクラウドの利活用を必ずしも否定するものではないが、**取扱者自らが設置したものではない**うえ、設置主体、保守運営事業者、情報基盤センターおよび利用者の**責任分解点が明らか**ではなく、セキュリティインシデント等の**不備の事態が発生した場合における体制や対応についても不明な点が多数存在**することを指摘せざるを得ない。

このような状況下において「匿名レセプト情報・匿名特定健診等情報の提供に関するガイドライン」で定めるセキュリティ要件を満たしていると認められないため、本件申し出について不承諾が相当と判断したもの。

厚生労働省発保0705第2号より



不承諾の理由について納得がいかなかったが、以下の条文により**不服申し立ても断念**

匿名レセプト情報等の提供は厚生労働大臣と提供申出者（利用者）及び取扱者の双方の合意の基づく契約上の行為であり、行政手続法（平成5年法律第88号）第2条第2号の処分にならないことから、**行政不服審査法（昭和37年法律第160号）の対象外**となりますことご了承ください



「K匿名化」の例

例示

「NDBオープンデータ」 医薬品の使用量（例）

1000単位（錠剤の場合は「錠数」）未満をK匿名化（最小集計単位の原則）

→ 「980か355か」がわかるとどうやって個人の特定ができるのか？

薬効分類	薬効分類名称	医薬品コード	医薬品名	単位	薬価基準収載 医薬品コード	薬価	後発品 区分	総計	01			02			03					
									県別合計	北海道	青森県	岩手	岩手	岩手	岩手					
235	下剤, 浣腸剤	620480601	センノシド錠 1.2mg 「サワイ」	錠	2354003F2367	5.0	1	228,215,579	228,215,579	10,984,513	4,691,918	3,5								
		620481301	ブルゼニド錠 1.2mg	錠	2354003F2316	5.6	0	144,253,041	144,253,041	7,739,994	2,500,621	1.5								
		622182901	アミティーザカプセル 2.4μg	カプセル	2359006M1025	123.0	0	140,085,864	140,085,864	9,045,733	1,937,566	1.6								
		620491801	アローゼン顆粒	g	2359102X1072	7.2	0	83,457,982	83,457,982	12,607,997	1,037,764	1.3								
		620006963	センナリド錠 1.2mg	錠	2354003F2251	5.0	1	78,667,155	78,667,155	3,411,686	1,202,166	1.1								
		620008638	センノサイド錠 1.2mg	錠	2354003F2278	5.0	1	64,176,581	64,176,581	4,291,427	1,200,831	4.1								
		621278001	センノシド錠 1.2mg 「トーワ」	錠	2354003F2375	5.0	1	62,800,082	62,800,082	5,535,207	999,143	9.9								
		612350042	ヨーデルS糖衣錠-80 80mg	錠	2354002F2028	6.5	1	62,096,053	62,096,053	590,726	50,051	1.1								
		622564801	グーフイス錠 5mg	錠	2359008F1025	105.8	0	57,566,025	57,566,025	4,280,332	573,429	1.1								
		620480205	センノシド錠 1.2mg 「セイコー」	錠	2354003F2340	5.0	1	55,339,486	55,339,486	4,348,619	1,686,154	1.1								
		620480707	センノシド錠 1.2mg 「武田テバ」	錠	2354003F2448	5.0	1	44,620,761	44,620,761	2,719,457	865,331	3.3								
		620481202	センノシド錠 1.2mg 「ファイザー」	錠	2354003F2413	5.0	1	42,906,246	42,906,246	3,401,734	1,077,759	1.1								
		622625001	モビコール配合内用剤 6.8523g	包	2359110B1029	83.9	0	41,930,890	41,930,890	1,733,828	390,699	1.1								
		620006956	ヤチロ配合錠	錠	2359103X1034	5.6	0	36,670,635	36,670,635	2,964,931	65,170	1.1								
						錠	2354003F2332	5.0	1	34,995,144	34,995,144	2,820,024	1,402,261	1.1						
				622263901	センノシド錠 1.2mg 「クニヒロ」	錠	2356002X1000	1.2	0	17,013	15,467	-	-	-	-	-	-	-	-	
		620006886	加香ヒマシ油「ケンコー」	mL	2359001X2017	7.2	0	17,013	15,467	-	-	-	-	-	-	-	-			
		612350146	カルメロースナトリウム	g	235900501110	9.6	1	16,111	12,246	4,424	-	-	-	-	-	-	-			
		620488301	ビコスルファートナトリウム顆粒 1% 「イセイ」	g	2355002X1016	0.7	0	12,736	10,134	-	-	-	-	-	-	-	-			
		620481711	硫酸マグネシウム「東海」	g	2356001X1019	1.2	0	12,383	9,132	-	-	-	-	-	-	-	-			
		612350033	ヒマシ油	mL	2354003F2294	5.0	1	11,843	9,896	-	-	-	-	-	-	-	-			
		620480901	ソルダナ錠 1.2mg	錠	2359005F1170	5.8	1	9,818	9,526	-	-	-	-	-	-	-	-			
		620007051	フェニスルー錠 2.5mg	錠	2356001X1193	1.4	0	4,840	1,680	-	-	-	-	-	-	-	-			
		620006895	ヒマシ油（山善）	mL	235910981077	0.7	1	4,000	4,000	-	-	-	-	-	-	-	-			
		612350062	人工カルルス塩「コゾカイ・M」	g	2356001X1124	1.4	0	2,320	1,700	-	-	-	-	-	-	-	-			
		620006883	ヒマシ油「東海」	mL	2356002X1013	1.3	0	1,610	1,000	-	-	-	-	-	-	-	-			
		612350048	加香ヒマシ油	mL	235910981115	0.7	1	1,270	0	-	-	-	-	-	-	-	-			
		612350065	人工カルルス塩「ヤマゼン」	g	2355002X1016	0.7	0	-	0	-	-	-	-	-	-	-	-			
		612350119	硫酸マグネシウム「カナダ」	g	2356002X1013	1.2	0	-	0	-	-	-	-	-	-	-	-			
		620482319	加香ヒマシ油「東海」	mL	2359005M1098	5.0	1	-	0	-	-	-	-	-	-	-	-			
		621278301	ビコスルファートナトリウムカプセル 2.5mg 「T.C」	カプセル	2356002X1188	1.7	0	-	0	-	-	-	-	-	-	-	-			
		620006893	加香ヒマシ油（山善）	mL	2356002X1013	1.3	0	-	0	-	-	-	-	-	-	-	-			
		620482310	加香ヒマシ油「ニコッ」	mL				-	0	-	-	-	-	-	-	-	-			



仮名加工の取り扱いを可能とする解析環境 次世代医療基盤法（内閣府）案

IV. ③ 認定利用事業者の安全管理措置に関する基本的な考え方（1）

○ 「I型認定」「II型認定」及び「管理区域」「取扱区域」

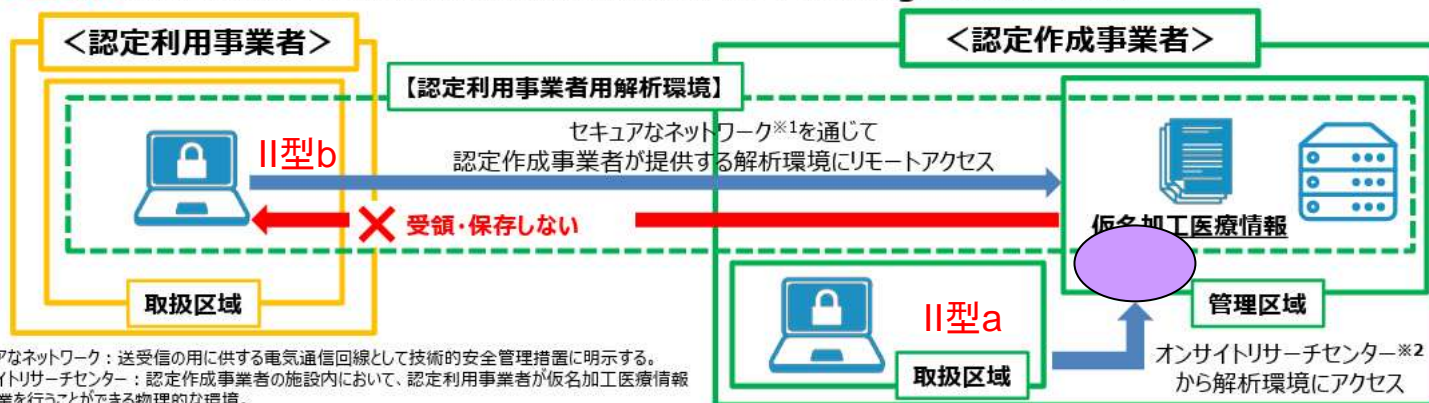
- ・ 確実な安全管理措置の確保と仮名加工医療情報の利活用促進の両立の観点から、利用事業者が自らの整備した環境下に仮名加工医療情報を保存することが可能な **I型認定**に加え、認定作成事業者等が整備したVisiting環境での利用に限定し、当該環境を前提とした安全管理措置の要件を定める **II型認定**の2種類を設ける（新GLIV-5-1-1-2参照）。
- ・ また、提供仮名加工医療情報※を取り扱う施設設備として、仮名加工医療情報を保存する機器が設置される**管理区域**と、提供仮名加工医療情報の操作・解析等を行う**取扱区域**を特定し、安全管理のために必要かつ適切な措置を講じることを求める（新GLIV-14-3-1-1参照）。 ※認定作成事業者から認定利用事業者に提供された仮名加工医療情報

【I型認定】仮名加工医療情報を認定利用事業者の管理区域に保存する



個別の解析のために切り出したデータセット

【II型認定】認定利用事業者は仮名加工医療情報を保存せず、Visiting環境限定で利用



※1 セキュアなネットワーク：送受信の用に供する電気通信回線として技術的安全管理措置に明示する。
 ※2 オンサイトリサーチセンター：認定作成事業者の施設内において、認定利用事業者が仮名加工医療情報を用いた作業を行うことができる物理的環境。

次世代医療基盤法検討WG資料より



解析環境準備の負荷を軽減するために

ご提案

前頁の仕組みにおける各事業者の負荷

	作成事業者	利用事業者
I型	データを切り出したらそれを利用者に渡す仕組み（アップロードあるいは郵送）までを整備	「管理区域」準備の負担最大
II型a	「取扱区域」（visiting作業現場）準備の負担最大	visitingの場所まで物理的に移動しないと作業ができない
II型b	リモートアクセスの受け側のシステム環境整備が必要	Thinkクライアントとセキュアなネットワーク環境の用意が必要

👉 解析環境の準備における両事業者の負担が大きすぎることにより、利用の推進を妨げる可能性も



共通の解析基盤を構築し、データ提供者、利用者の双方から利用料を徴収することにより運営する仕組みを整備

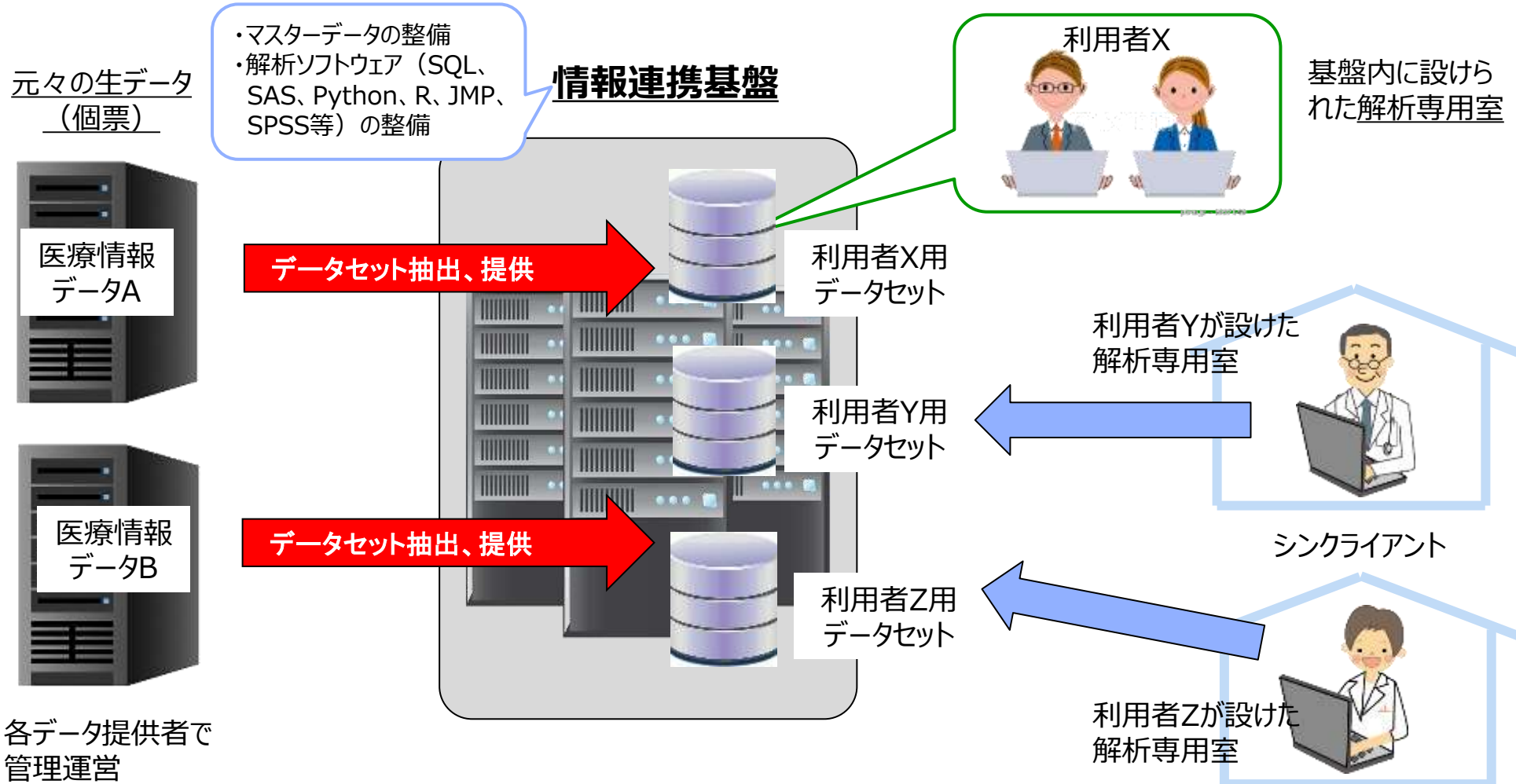
- ✓ 「セキュアなネットワーク」を定義し提供
 - ✓ 解析専用室はシンクライアントでデータを持たないため、窃視や聞き耳などの対策も不要
- 👉 結果として全体のコスト、負荷は軽減し、データ利用のハードルも下がる可能性



解析環境の整備：情報連携基盤（素案）

ご提案

解析環境の準備におけるデータ提供者、解析者双方の負担軽減のために**共通の解析基盤を整備**し、個別解析（利用）のために切り出したデータセットを格納。解析者はセキュアなネットワークでアクセスして個々に解析を行う





現在の稼働中サービス



Wisteria/BDEC-01

通常サービス中

システムの特徴

Wisteria/BDEC-01システムは、東京大学情報基盤センターが運用するスーパーコンピュータシステムです。Wisteria/BDEC-01はシミュレーションノード群（Odyssey（オデッセイ））とデータ・学習ノード群（Aquarius（アクエリアス））の2つの計算ノード群を有したシステムです。シミュレーションノード群（Odyssey）は、世界最高性能を有するスーパーコンピュータ「富岳」と同じ富士通株式会社の「FUJITSU Processor A64FX」を搭載した計算ノード7,680台から構成されます。データ・学習ノード群（Aquarius）はインテル社製第3世代Xeonスケーラブルプロセッサ（開発コード名Ice Lake）およびNVIDIA A100 Tensor コア GPUを搭載した計算ノード45台から構成されます。

東京大学情報基盤センター（以下、センター）では、現在Wisteria/BDEC-01 スーパーコンピュータシステムを運用しております。本スーパーコンピュータシステムでは、「通常利用（一般）」によるサービスを行っています。

🏠 大学・公共機関は、承認されれば安価（電気代程度）での利用が可能

データ提供者、利用者双方から相当の利用料を徴収し運営する仕組みを整備



ご清聴ありがとうございました

清水連絡先：

eshimizu@g.ecc.u-tokyo.ac.jp

eiko_shimizu@ds.itc.u-tokyo.ac.jp (情報基盤センター内)