

日本薬剤疫学会

薬剤疫学とデータベースタスクフォース主催

「実務者のためのデータベース研究講座 その1」

『JMDCのレセプトデータベースの構築と品質管理』

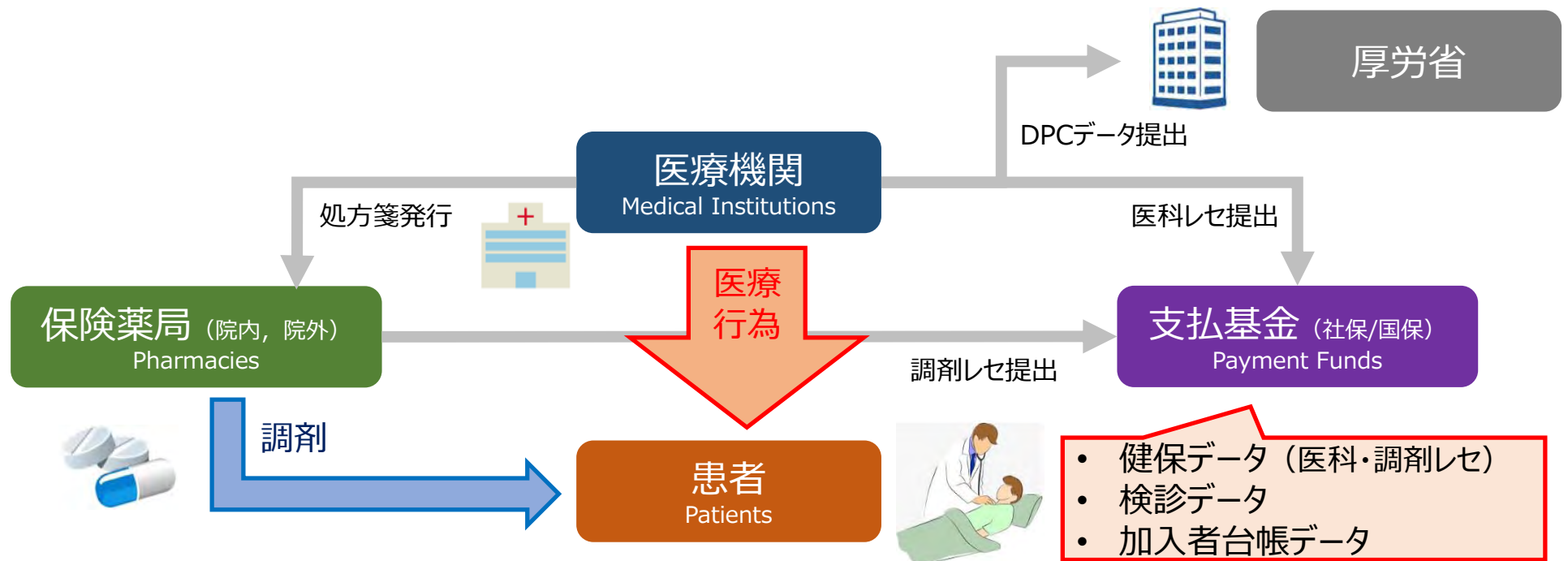
株式会社JMDC

メディカルファーママネジメント部

清田浩康

本演題発表に関連して、開示すべきCOI (Conflict of Interest)関係にある企業等はありません。

データソースの概要

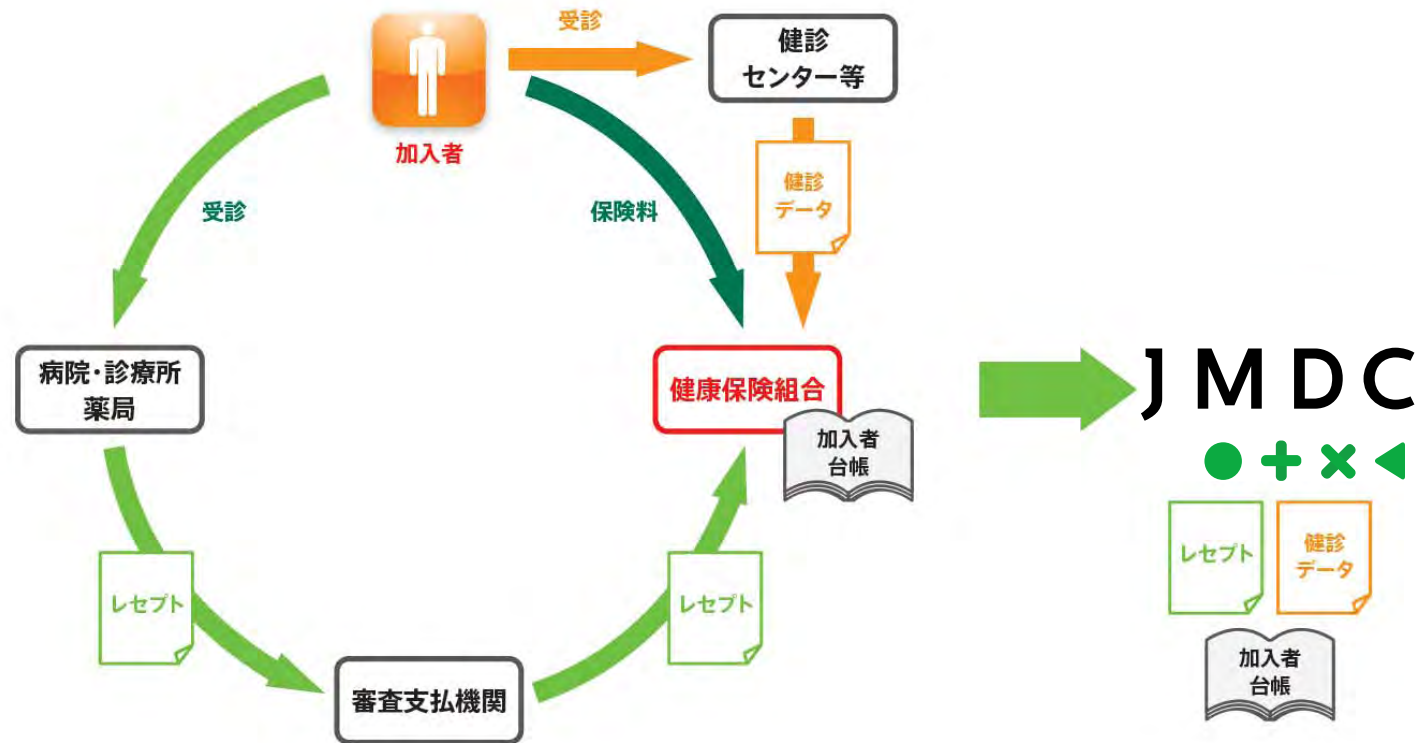


【ポピュレーションベースの医療情報データベース】

- 患者追跡性が高く、選択バイアスの少ないデータであり、一般人口における有病、イベントの発生の検討に適しています。
- 後期高齢者のデータが欠落しており、認知症、前立腺がん等の高齢で発症する疾患については注意が必要。

データソースの概要

保険加入者が医療機関を受診することでレセプトが発行され、そのレセプトは審査支払機関での審査を経て、各保険者のもとへ送付されます。JMDCでは観察集団としての固定の保険者よりレセプトをお預かりすることで、保険加入者の自由な受診行動を反映した全数調査を可能にしています。



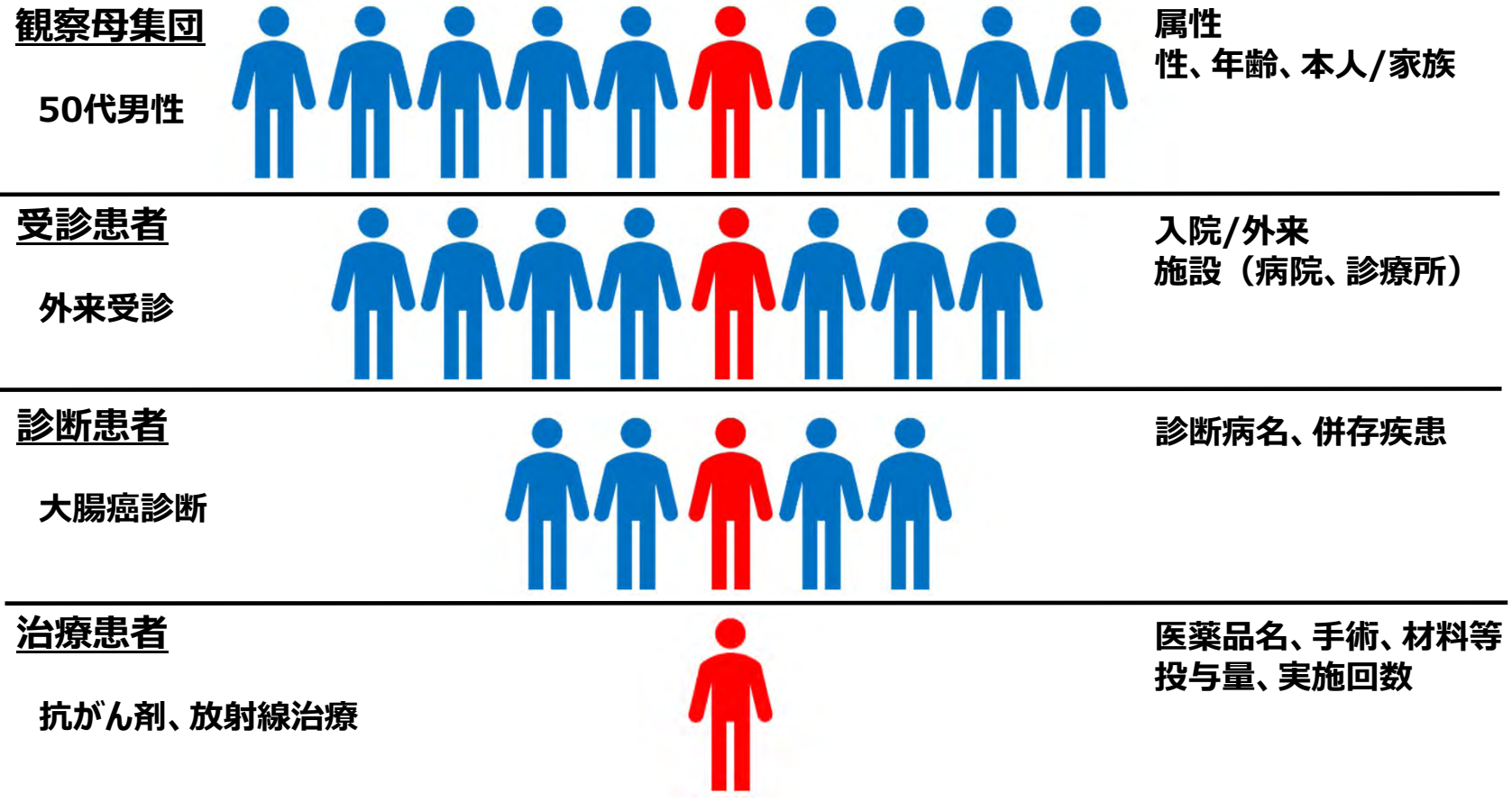
データソースの概要

全国に約2,900万人いる健康保険組合加入者のうち530万人が観察集団となります。
健康保険組合であるため、観察集団の年齢は0-74歳となります。

このうち530万人

保険名	保険者	保険者数	加入者種類	加入者数	年齢
組管掌健康保険	健康保険組合	1,431	大手企業等の従業員 およびその扶養家族	2,900万人	0-74 歳
全国保険協会 (協会けんぽ)	全国健康保険協会	1	中小企業等の従業員 およびその扶養家族	3,500万人	
共済組合保険	共済組合	85	国家公務員、地方公務員、 私学の教職員およびその 扶養家族	900万人	
国民健康保険	地方自治体	1,771	自営業者、年金生活者、 非正規雇用者等	3,500万人	75 歳以上
後期高齢者 医療制度	地方自治体	47	75歳以上および65歳～ 74歳で一定の障害の状態 にあることにつき認定を 受けた人	1,500万人	

データソースの概要



データソースの概要

JMDCでは2005年1月診療分よりデータを蓄積しています。
 現在、直近1年で約530万人、累計で約730万人の観察集団を保持しています。

- 直近1年 2018年1月～2018年12月 ●直近5年 2014年1月～2018年12月
- 直近3年 2016年1月～2018年12月 ●全期間 2005年1月～2018年12月

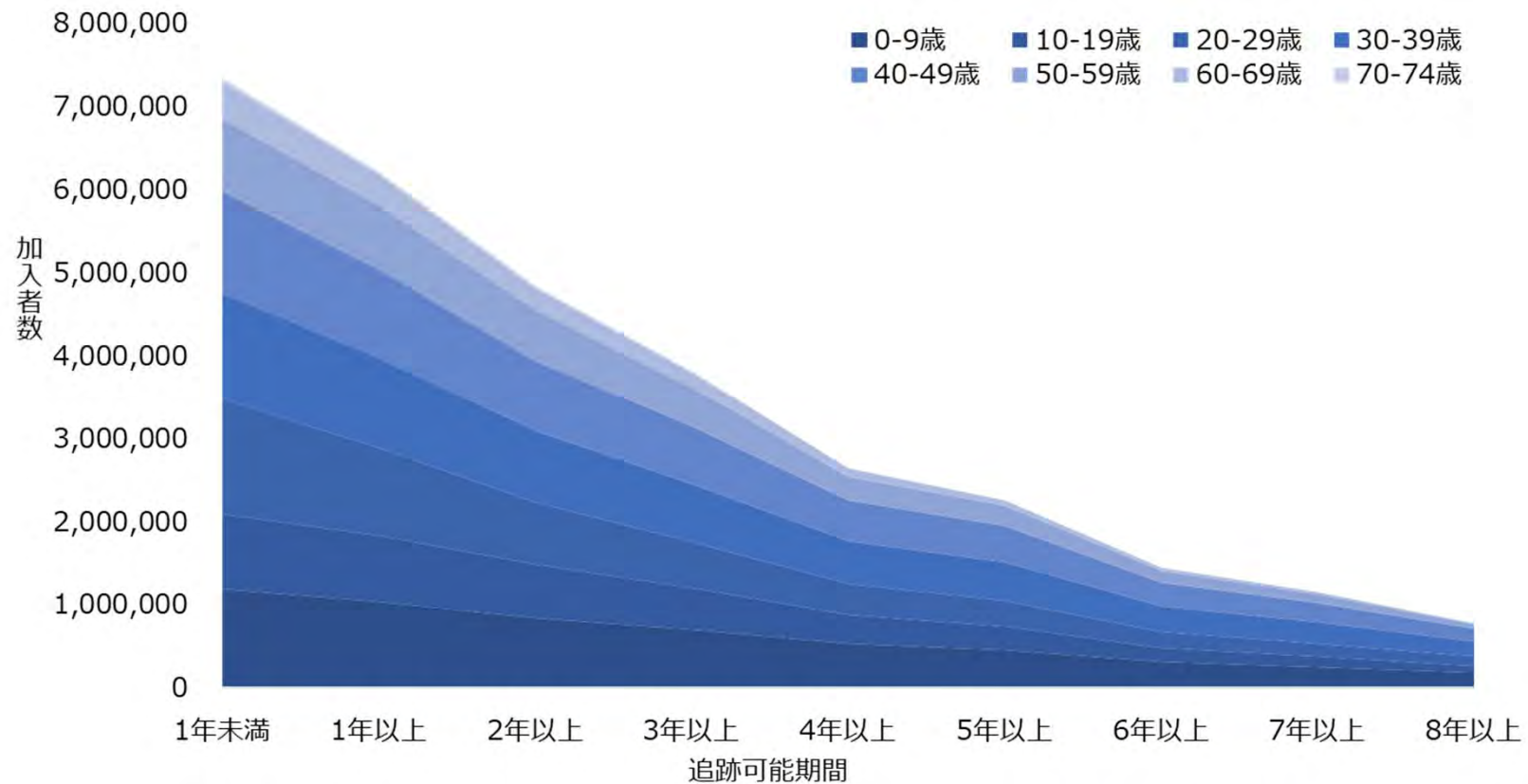
● 年齢階級×性別構成比

・年齢は期間の midpoint の年齢

性別	年齢階級	直近1年		直近3年		直近5年		全期間	
		母数	構成比	母数	構成比	母数	構成比	母数	構成比
男性	00-09歳	358,168	6.7%	417,493	6.8%	470,842	7.1%	663,426	9.1%
	10-19歳	373,039	7.0%	416,494	6.8%	451,111	6.8%	622,623	8.5%
	20-29歳	471,229	8.8%	581,077	9.5%	656,216	9.8%	705,500	9.6%
	30-39歳	475,323	8.9%	535,392	8.7%	580,018	8.7%	667,053	9.1%
	40-49歳	546,235	10.2%	592,192	9.6%	623,563	9.4%	614,232	8.4%
	50-59歳	469,357	8.7%	477,050	7.8%	479,311	7.2%	436,879	6.0%
	60-69歳	215,906	4.0%	254,252	4.1%	272,252	4.1%	176,123	2.4%
	70-74歳	21,439	0.4%	24,208	0.4%	24,312	0.4%	11,874	0.2%
女性	00-09歳	340,215	6.3%	396,508	6.5%	447,027	6.7%	629,982	8.6%
	10-19歳	341,675	6.4%	373,924	6.1%	399,748	6.0%	505,107	6.9%
	20-29歳	329,691	6.1%	432,909	7.1%	498,855	7.5%	564,919	7.7%
	30-39歳	380,980	7.1%	457,096	7.4%	511,542	7.7%	629,505	8.6%
	40-49歳	490,251	9.1%	557,703	9.1%	596,493	9.0%	568,800	7.8%
	50-59歳	371,023	6.9%	401,774	6.5%	414,655	6.2%	370,475	5.1%
	60-69歳	155,988	2.9%	190,474	3.1%	206,661	3.1%	139,397	1.9%
	70-74歳	23,856	0.4%	29,604	0.5%	31,223	0.5%	18,448	0.3%
計	5,364,375	100.0%	6,138,150	100.0%	6,663,829	100.0%	7,324,343	100.0%	

データソースの概要

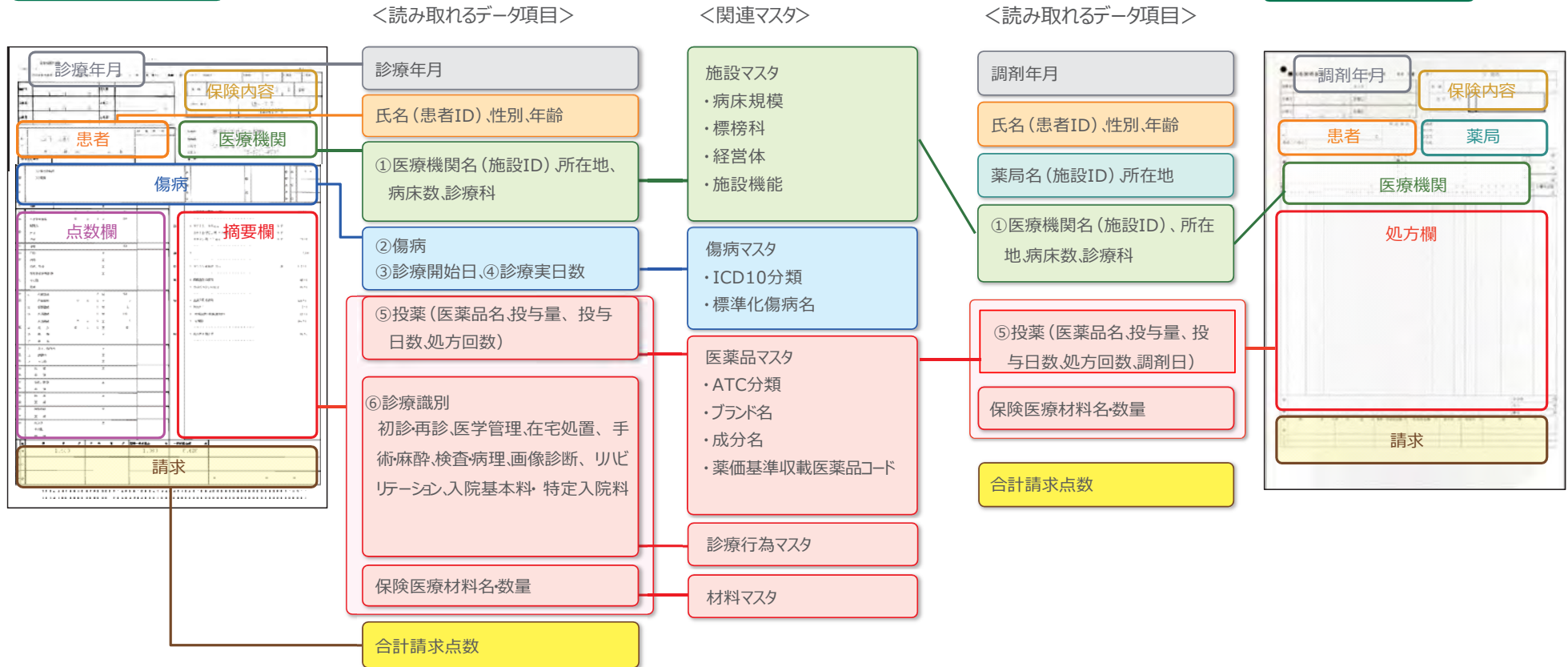
JMDCでは2005年1月診療分よりデータを蓄積しています。
現在、直近1年で約530万人、累計で約730万人の観察集団を保持しています。



レセプトから読み取れるデータ項目

医科レセプト

調剤レセプト



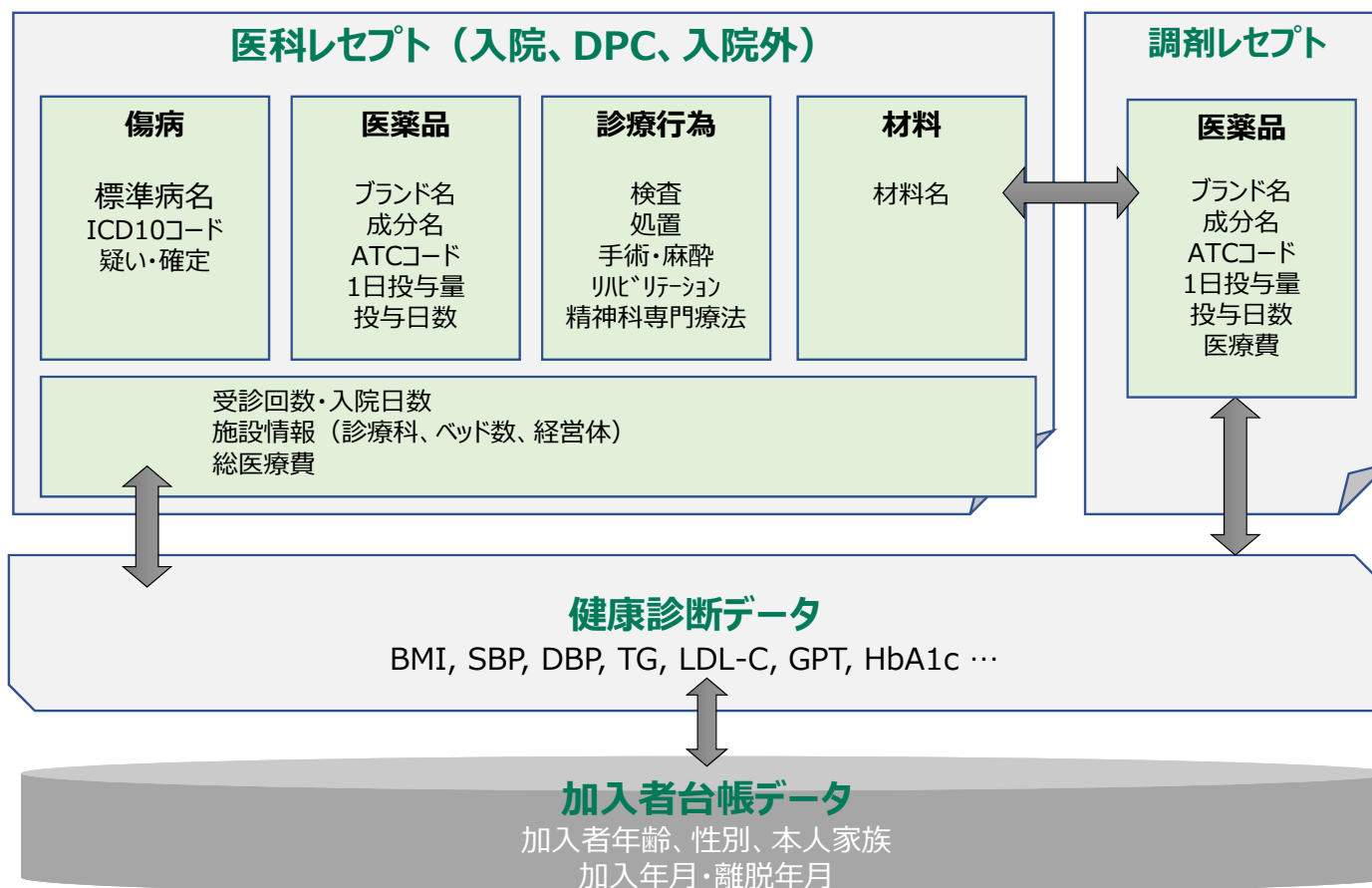
レセプトから読み取れるデータ項目

JMDCでは、DPCLレセプトの包括評価による算定を行った期間の診療行為、医薬品及び特定器材の情報が記載されたコーディングデータについて、提出が義務付けられた2009年1月診療分よりデータ化を行っております。

診療行為等	計	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
【注射その他】																																
ソルテム3 A 輸液 500mL	7	3	3	1																												
【検査・病理】																																
キシロカインゼリー 2%	30			30																												
ブスコバン注射液 2% 1mL	1			1																												
末梢血液一般	1			1																												
Alb	1			1																												
BUN	1			1																												
クレアチニン	1			1																												
UA	1			1																												
グルコース	1			1																												
LDH	1			1																												
ALP	1			1																												
CPK	1			1																												
GOT	1			1																												
GPT	1			1																												
K	1			1																												
Ca	1			1																												
CRP (定量)	1			1																												

JMDCでは、レセプト、健康診断、加入者台帳データを提供しています

レセプト、健康診断、加入者台帳データは、全て加入者一意のIDで連結します。



情報源から収集した医療データの品質管理



取り込み

- 頻度：毎月
- 授受方法：オンラインシステム、セキュリティ便で受領

データクリーニング

- 対象となるデータ項目：傷病データ、健診データ
- データクリーニングの具体的な内容：999コード（その他の傷病）に対してコーディング（標準化）
健診値について学会等の公式基準に基づき点検

コード化

- 対象となるデータ項目：傷病、医薬品、診療行為、材料
- コード化の具体的な内容，手順：毎月メンテナンスされたマスタの付与

継続した品質管理

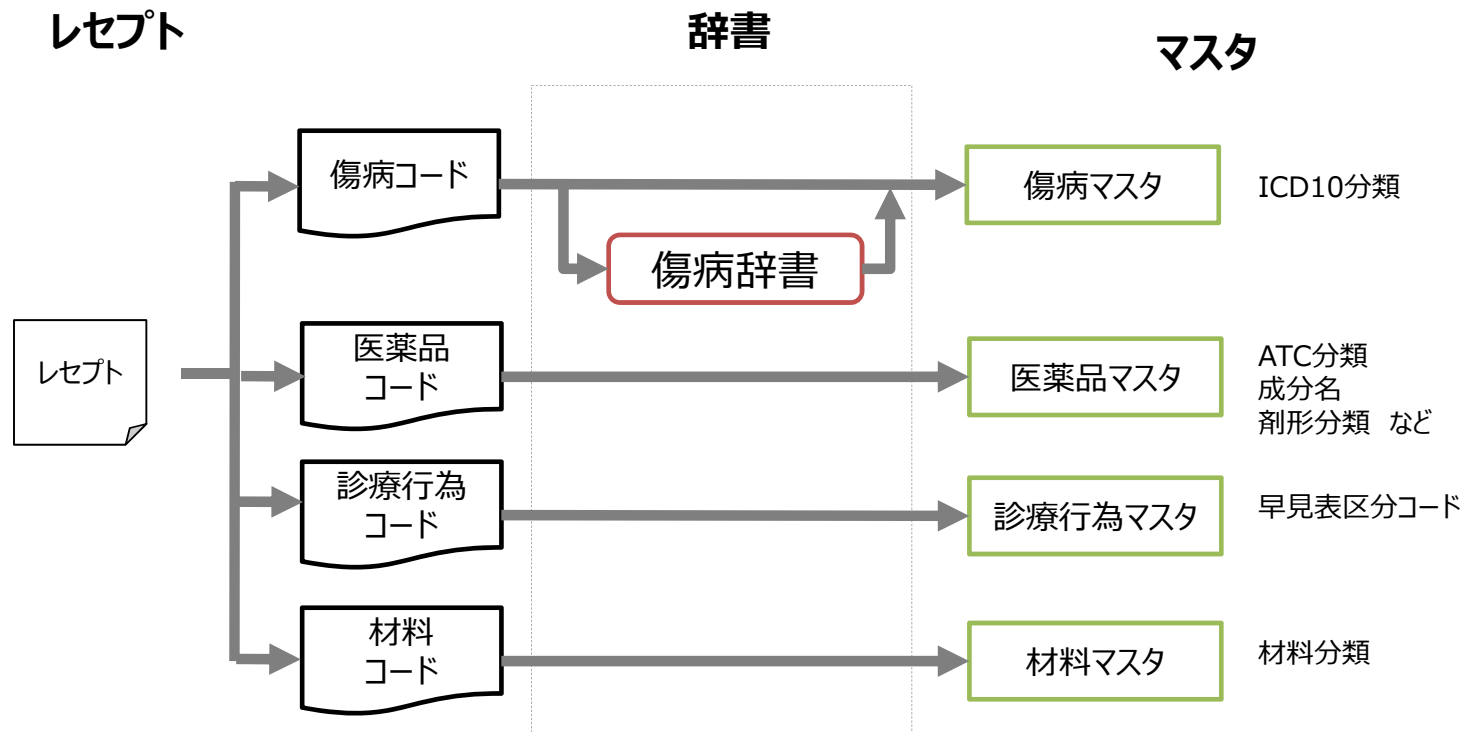
- 対象となるデータ項目，品質管理の具体的な内容：定期的なマスタ更新

2019/2/26



情報源から収集した医療データの品質管理

紙レセプトを電子化、医薬品、診療行為、材料データは、コードマッチングでJMDCマスタを付与します。
傷病のみ、コードマッチングしなかったもの（999コード）はJMDC傷病辞書にあてて、標準化されたデータに変換します。



J M D C の使用マスタとデータ化率

傷病マスタ

診療報酬情報提供サービスが電子レセプト用として提供している傷病名マスタの「傷病名基本名称」に準拠して作成しています。1年間に2回、傷病名マスタの更新時に合わせて更新しています。

医薬品マスタ

「薬価基準収載医薬品」をデータベース化したものです。レセプトに記録されている多様な表現の医薬品名、規格、投与量などを標準化しています。ATC分類（WHO、EphMRA）や成分名、剤形分類など、レセプト記載以外の情報を付与することで、体系的な分析に対応できるよう開発しています。厚労省ホームページの情報に準拠し、随時更新しています。

診療行為マスタ*

診療報酬情報提供サービスの医科診療行為マスタを基に作成しており、医科レセプトの摘要欄に記載されている様々な表現の診療行為名称を標準化し、さらにこれらを診療点数早見表区分コード等に紐付けすることで、体系的な分析に対応できるよう開発しています。

材料マスタ*

診療報酬情報提供サービスの特定器材マスタを基に作成されており、レセプトに記載されている様々な表現の医療材料名称を標準化し、さらにこれらを診療点数早見表区分コード等に紐付けすることで、体系的な分析に対応できるよう開発しています。

*2年に一度の厚生労働省による診療報酬「大」改定ごとにバージョンを持たせています。改定とは別に新規収載された診療行為に関しては、診療報酬情報提供サービスの「告示・通知に関する案内」などを確認し、変更の都度随時反映しています。

項目	明細数（件）	マッチング数（件）	データ化率（%）
傷病	12,084,505	12,047,834	99.70%
医薬品	10,756,521	10,752,600	99.96%
診療行為	29,728,156	29,720,524	99.97%
材料	197,185	196,982	99.90%

2018年8月診療分

JMDCのデータ化フロー

お預かりしたデータは万全なセキュリティの下で管理され、二次利用されるデータは標準化済み、個人を特定しうる情報を排除した匿名加工情報です。



データエントリー

保険者より送付された台帳、レセプト、健診データをエントリーします。
紙、画像レセプトも全て入力します。

匿名加工処理①

JMDCプライバシーポリシーで定められている、直接的に特定の個人を識別できる情報について削除します。

標準化

傷病、医薬品、診療行為、材料について、医療機関ごとに異なる医療方言をJMDC独自の医療辞書で自動変換によって標準化し、マスタを付与します。
さらに投与量単位の統一、数値等の標準化を実施します。

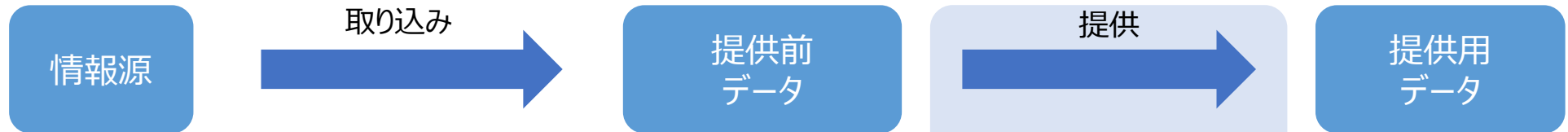
匿名加工処理②

JMDCプライバシーポリシーで定められている、他の情報と容易に照合することができ、それにより特定の個人を識別できる可能性のある情報について削除します。

二次利用DB

標準化済み、**対応表を保持しない**匿名加工情報DBが生成されます。

情報源から収集した医療データの品質管理



提供用データの作成

- 全体？切り出し：ご要望の条件（傷病、医薬品、診療行為、材料）に応じてデータを切り出し
- 切り出しの場合、確認方法：ダブルプログラミングを行いデータ抽出
- 二次利用DBは毎月すべての加入者IDが洗い替えられる
- GPSP基準に準拠した体制で解析



提供可能な項目（レセプトデータ）

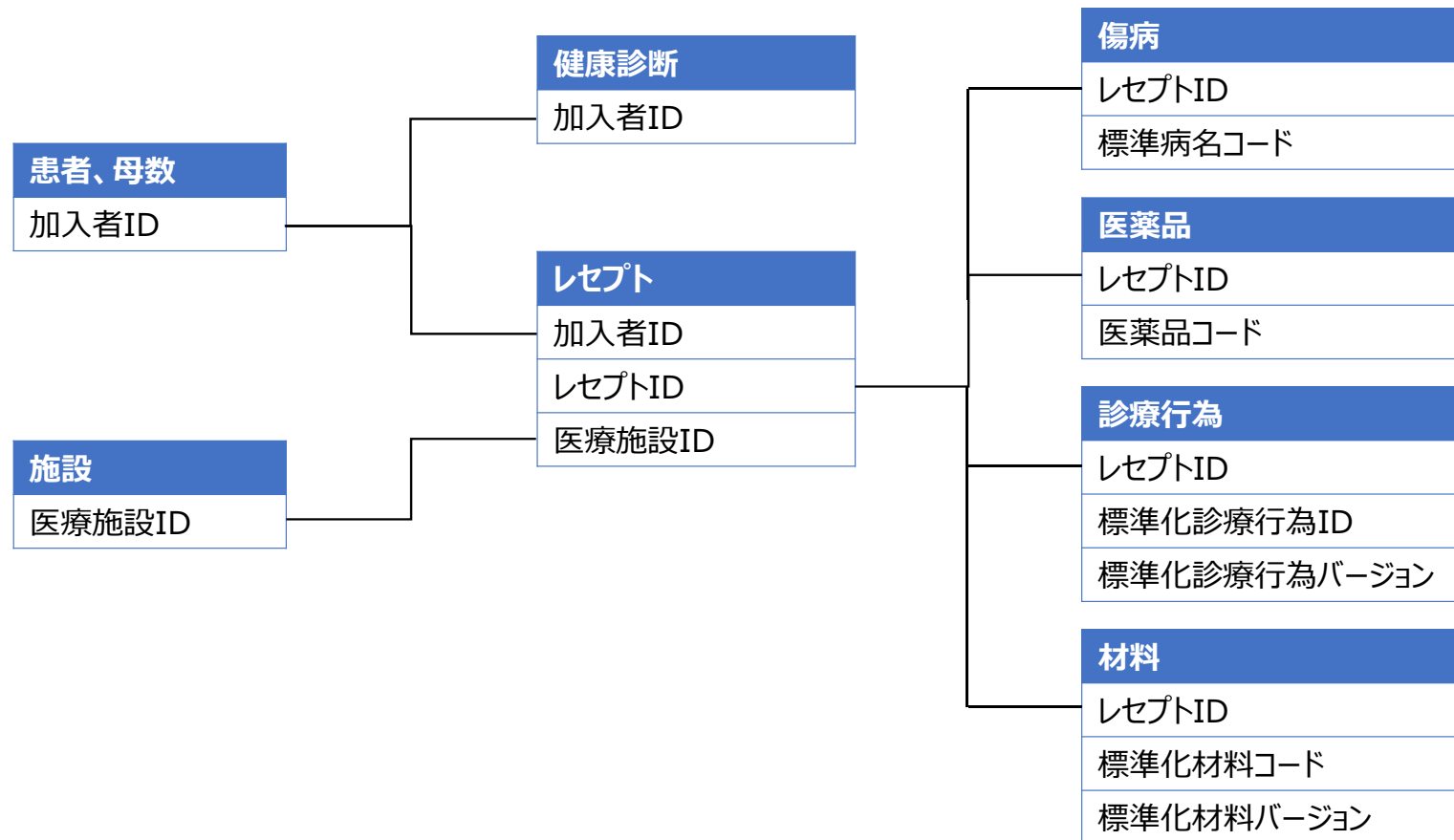
母数・患者	施設	レセプト	傷病	医薬品	診療行為	材料
加入者ID	医療施設ID	加入者ID	加入者ID	加入者ID	加入者ID	加入者ID
加入者生年月	ベッド数	レセID	レセID	レセID	レセID	レセID
加入者性別	HPGP区分	レセ種別	レセ種別	明細ID	明細ID	明細ID
本人家族	診療科大分類	診療年月	診療年月	レセ種別	レセ種別	レセ種別
観察開始年月	診療科中分類	医療施設ID	医療施設ID	診療年月	診療年月	診療年月
観察終了理由(死亡)フラグ	経営体	診療科大分類	ICD10大分類コード	医療施設ID	医療施設ID	医療施設ID
	在宅療養支援診療所	診療科中分類	ICD10大分類名	ATC大分類コード	診療区分中分類名	大分類名称
	地域医療支援病院	診療実日数	ICD10中分類コード	ATC大分類名	診療区分小分類名	中分類名称
	がん診療連携拠点病院	入院日	ICD10中分類名	ATC中分類コード	診療区分細分類名	診療点数早見表区分コード
	DPC導入医療機関	退院日	ICD10小分類コード	ATC中分類名	診療区分細々分類名	標準化材料コード
	特定機能病院	総点数	ICD10小分類名	ATC小分類コード	診療点数早見表区分コード	標準化材料名
			ICD10細分類コード	ATC小分類名	標準化診療行為ID	1回あたりの使用量
			ICD10細分類名	ATC細分類コード	標準化診療行為名	回数
			標準傷病コード	ATC細分類名	回数	使用年月日
			標準病名	WHO-ATCコード	数量	材料価格
			主傷病フラグ	WHO-ATC名	実施年月日	材料加算割合
			入院の契機となった傷病フラグ	薬価基準収載医薬品コード	1回あたり点数	診療区分
			疑いフラグ	個別12桁コード	1回あたり加算割合	同時ID
			診療開始年月日	成分名	1回あたり金額	同時IDごとの点数
			転帰	ブランド名	診療区分	
			死亡フラグ	医薬品名	同時ID	
			増悪フラグ	規格単位	同時IDごとの点数	
				薬価		
				処方日		
				調剤日		
				1処方あたりの1日投与量		
				1処方あたりの投与日数		
				剤形大分類名		
				剤形中分類名		
				剤形小分類名		
				頓服フラグ		
				診療区分		
				同時ID		
				同時IDごとの点数		

提供可能な項目（健康診断データ）

健康診断		
BMI	HbA1c	喫煙
腹囲	尿糖	運動習慣
収縮期血圧	尿蛋白（定性）	食べ方1(早食い等)
拡張期血圧	血清尿酸	食べ方2(就寝前)
採血時間(食後)	血清クレアチニン	食べ方3(夜食/間食)
中性脂肪（トリグリセリド）	ヘマトクリット値	食習慣
HDLコレステロール	血色素量(ヘモグロビン値)	飲酒習慣
LDLコレステロール	赤血球数	飲酒量
GOT (AST)	心電図(所見の有無)	睡眠
GPT (ALT)	眼底検査（キースワグナー分類）	
γ-GT(γ-GTP)	眼底検査（シェイエ分類：H）	
空腹時血糖	眼底検査（シェイエ分類：S）	
随時血糖	眼底検査（SCOTT分類）	

喫煙	現在、たばこを習慣的に吸っている。(1：はい 2：いいえ)
運動習慣	1日30分以上の軽く汗をかく運動を週2回以上、1年以上実施(1：はい 2：いいえ)
食べ方1(早食い等)	人と比較して食べる速度が速い。(1：速い、2：ふつう、3：遅い)
食べ方2(就寝前)	就寝前の2時間以内に夕食をとることが週に3回以上ある(1：はい 2：いいえ)
食べ方3(夜食/間食)	夕食後に間食(3食以外の夜食)をとることが週に3回以上ある。(1：はい 2：いいえ)
食習慣	朝食を抜くことが週に3回以上ある。(1：はい 2：いいえ)
飲酒習慣	お酒(清酒、焼酎、ビール、洋酒など)を飲む頻度(1：毎日、2：時々、3：ほとんど飲まない)
飲酒量	飲酒日の1日当たりの飲酒量(1：1合未満、2：1～2合未満、3：2～3合未満、4：3合以上)
睡眠	睡眠で休養が十分とれている。(1：はい 2：いいえ)

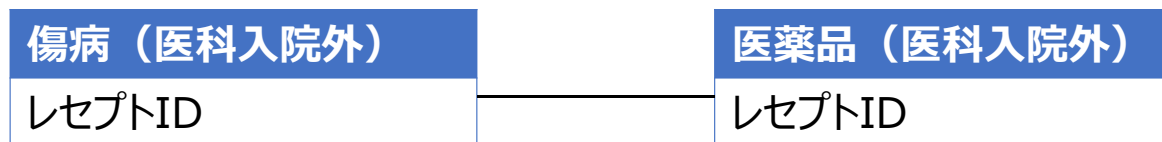
提供データのデータ構造



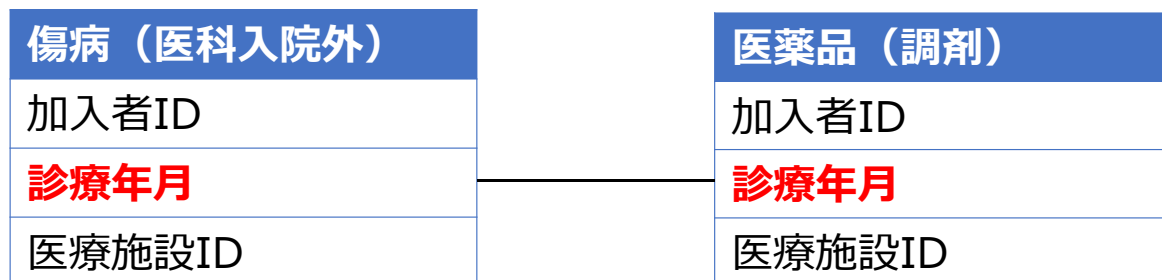
提供データのデータ構造

医薬品・材料の調剤レセプトを、傷病・診療行為の入院外の情報と結びつける場合は、注意が必要です。

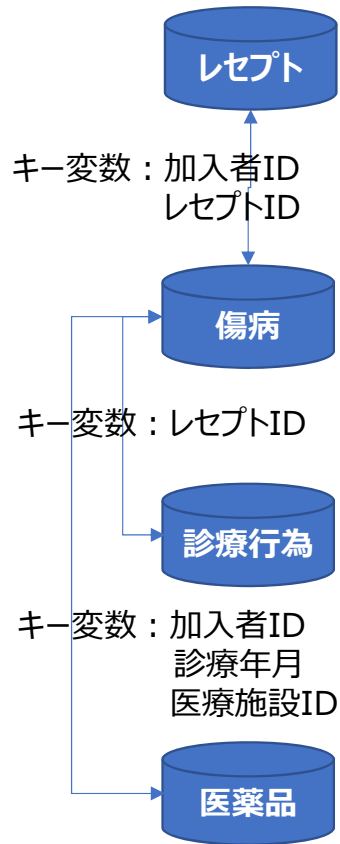
- 院内処方の場合
(医薬品、傷病とも同じレセプト種別：DPC/入院/入院外)



- 院外処方の場合で、レセプト単位で結合する場合
(医薬品はレセ種別：調剤、傷病はレセプト種別：入院外)



提供データのデータ構造



加入者ID	レセプトID	レセプト種別	診療年月	医療施設ID	診療実日数	総点数
000001	C000001	入院外	201801	F000001	1	1411
000001	C000002	調剤	201801	F000001		3252
000002	C000003	入院外	201801	F000002	1	394

加入者ID	レセプトID	レセプト種別	診療年月	医療施設ID	標準病名	疑い
000001	C000001	入院外	201801	F000001	糖尿病	
000001	C000001	入院外	201801	F000001	高血圧	
000001	C000001	入院外	201801	F000001	高尿酸血症	1

加入者ID	レセプトID	レセプト種別	診療年月	医療施設ID	診療行為名	回数
000001	C000001	入院外	201801	F000001	再診	1
000001	C000001	入院外	201801	F000001	HbA1c	1
000001	C000001	入院外	201801	F000001	グルコース	1

加入者ID	レセプトID	レセプト種別	診療年月	医療施設ID	医薬品名	投与日数	投与量	単位
000001	C000002	調剤	201801	F000001	アマリール0.5mg錠	30	60	錠
000001	C000002	調剤	201801	F000001	メトグルコ錠250mg	30	180	錠
000001	C000002	調剤	201801	F000001	ランタスXR注ソロスター	1	2	キット

傷病を定義する際の注意点

病名はレセプトに記録される病名であるため、医療機関あるいは医師の記載の仕方に依存します。どのレベルで設定すべきか、はじめに探索的に確認することが必要です。

ICD10中分類	ICD10小分類	標準病名
E10-E14 糖尿病 14,071,195人 100%	E10 インスリン依存性糖尿病 155,392人(1.1%)	…1型糖尿病、不安定糖尿病など
	E11 インスリン非依存性糖尿病 4,824,921人(34.3%)	…2型糖尿病、インスリン抵抗性糖尿病など
	E12 栄養障害に関連する糖尿病 0人 (0%)	…栄養不良関連糖尿病
	E13 その他の明示された糖尿病 85,935人(0.6%)	…ステロイド性糖尿病、薬剤性糖尿病など
	E14 詳細不明の糖尿病 11,623,315人 (82.6%)	… 糖尿病

2型糖尿病の大半は単に「糖尿病」と記載されるケースが多い。この標準病名はICD10小分類では「E14詳細不明の糖尿病」に分類され、ICD10小分類の「E11インスリン非依存性糖尿病」のみで設定してしまうと、定義漏れの原因となります。

入院を分析する際の注意点

➤ 入院日

医科入院レセプトの場合

レセプトに記録されている入院日は、入院基本料の起算日であり、必ずしも実際の入院日と一致するとは限らない。退院日から3か月未満の間に、同一傷病で同一医療機関に再入院した場合、レセプトにおける入院日の欄には「今回の入院日」ではなく「前回の入院日」が記録されている。

DPCレセプトの場合

DPC算定対象となる病棟以外の病棟から、DPC算定対象となる病棟に転棟した場合、転棟年月日が記載され、必ずしも実際の入院年月日と一致するとは限らない。

➤ 退院日

医科入院レセプトの場合

フォーマットに退院日の項目がなく、コメントレコードの文字データとして記録される。

DPCレセプトの場合

診断群分類点数表等による診療報酬額の算定を終了する場合、当該終了日が記載され、必ずしも実際の退院年月日と一致するとは限らない。

入院回数の分析はレセプトでは限界があるが、入院期間については診療実日数、医薬品明細、診療行為明細の日付情報から定義することも可能

入院を分析する際の注意点

医科入院レセプト

診療報酬明細書 平成 30年 4月分 県番: 13 院番: 9999913

1 医科 1 社保 2 2 併 5 家入

保険者番号 06132013 助付割合

記号・番号 1 2 3 4 5 6 7 1

東京 東京都港区新橋
診療所 サンプル病院
所在地 TEL:03-9999-9999

氏名 サンプル 八
性別 女 年齢 4 年 19 月 5 日 6 生

病歴上の事由
① 呼吸 (主)
② 肺炎球菌感染症の疑い
③ アデノウイルス感染症の疑い
④ 急性心筋炎の疑い
⑤ 急性心筋炎の疑い

以下、概要欄

①	ネフローゼ症候群の疑い	平 28 年 5 月 4 日	私	1 6 日
②	急性尿毒症の疑い	平 28 年 5 月 4 日		
③	脱水症	平 28 年 5 月 4 日		
④	免疫グロブリン異常症の疑い	平 28 年 5 月 4 日		
⑤	低蛋白血症の疑い	平 28 年 5 月 4 日		
⑥	低酸素血症の疑い	平 28 年 5 月 4 日		
⑦	敗血症の疑い	平 28 年 5 月 5 日		
⑧	血液透析の疑い	平 28 年 5 月 5 日		
⑨	肺炎	平 28 年 5 月 5 日		
⑩	急性心筋炎 (主)	平 28 年 5 月 6 日		
⑪	マヨロウイルス感染症	平 28 年 5 月 6 日		
⑫	深層瘻の疑い	平 28 年 5 月 6 日		
⑬	アレルギー性鼻炎	平 28 年 5 月 6 日		
⑭	全身性エリテマトーゼスの疑い	平 28 年 5 月 6 日		
⑮	下咽頭機能低下症の疑い	平 28 年 5 月 6 日		
⑯	甲状腺機能低下症の疑い	平 28 年 5 月 6 日		
⑰	サイトメガロウイルス感染症の疑い	平 28 年 5 月 6 日		
⑱	腎性高血圧症	平 28 年 5 月 6 日		

入院年月日 30 年 4 月 4 日

入院基本料 17140 円

食事 17140 円

生活 39497 円

医療 39497 円

請求 41,100 円

決定 41,100 円

負担 0 円

請求 17,140 円

決定 17,140 円

負担 0 円

請求 41,100 円

決定 41,100 円

負担 0 円

DPCLレセプト

診療報酬明細書 平成 30年 4月分 県番: 13 院番: 9999913

1 医科 1 社保 1 単独 9 高入 7

保険者番号 06132013 助付割合

記号・番号 1 2 3 4 5 6 7 1

東京 東京都港区新橋
診療所 サンプル病院
所在地 TEL:03-9999-9999

氏名 サンプル 六
性別 男 年齢 1 年 3 月 19 日 5 日 16 生

病歴上の事由
① 呼吸 (主)
② 肺炎球菌感染症の疑い
③ アデノウイルス感染症の疑い
④ 急性心筋炎の疑い
⑤ 急性心筋炎の疑い

以下、概要欄

①	ネフローゼ症候群の疑い	平 28 年 5 月 4 日	私	1 6 日
②	急性尿毒症の疑い	平 28 年 5 月 4 日		
③	脱水症	平 28 年 5 月 4 日		
④	免疫グロブリン異常症の疑い	平 28 年 5 月 4 日		
⑤	低蛋白血症の疑い	平 28 年 5 月 4 日		
⑥	低酸素血症の疑い	平 28 年 5 月 4 日		
⑦	敗血症の疑い	平 28 年 5 月 5 日		
⑧	血液透析の疑い	平 28 年 5 月 5 日		
⑨	肺炎	平 28 年 5 月 5 日		
⑩	急性心筋炎 (主)	平 28 年 5 月 6 日		
⑪	マヨロウイルス感染症	平 28 年 5 月 6 日		
⑫	深層瘻の疑い	平 28 年 5 月 6 日		
⑬	アレルギー性鼻炎	平 28 年 5 月 6 日		
⑭	全身性エリテマトーゼスの疑い	平 28 年 5 月 6 日		
⑮	下咽頭機能低下症の疑い	平 28 年 5 月 6 日		
⑯	甲状腺機能低下症の疑い	平 28 年 5 月 6 日		
⑰	サイトメガロウイルス感染症の疑い	平 28 年 5 月 6 日		
⑱	腎性高血圧症	平 28 年 5 月 6 日		

入院年月日 30 年 4 月 4 日

入院基本料 17140 円

食事 17140 円

生活 39497 円

医療 39497 円

請求 41,100 円

決定 41,100 円

負担 0 円

請求 17,140 円

決定 17,140 円

負担 0 円

請求 41,100 円

決定 41,100 円

負担 0 円