

平成30年6月14日	資料4-2
第41回レセプト情報等の 提供に関する有識者会議	

レセプト情報等 オンサイトリサーチセンター(京都) 運用および活動に関する報告



medical informatics

大寺祥佑、酒井未知、朝野美穂、加藤源太、黒田知宏
京都大学医学部附属病院 レセプト情報等オンサイトリサーチセンター(京都)運用部





1. 運用および活動の概要
2. 模擬申出課題の進捗
3. 考察
4. 結語



レセプト情報等オンサイトリサーチセンター(京都) 運用および活動の履歴

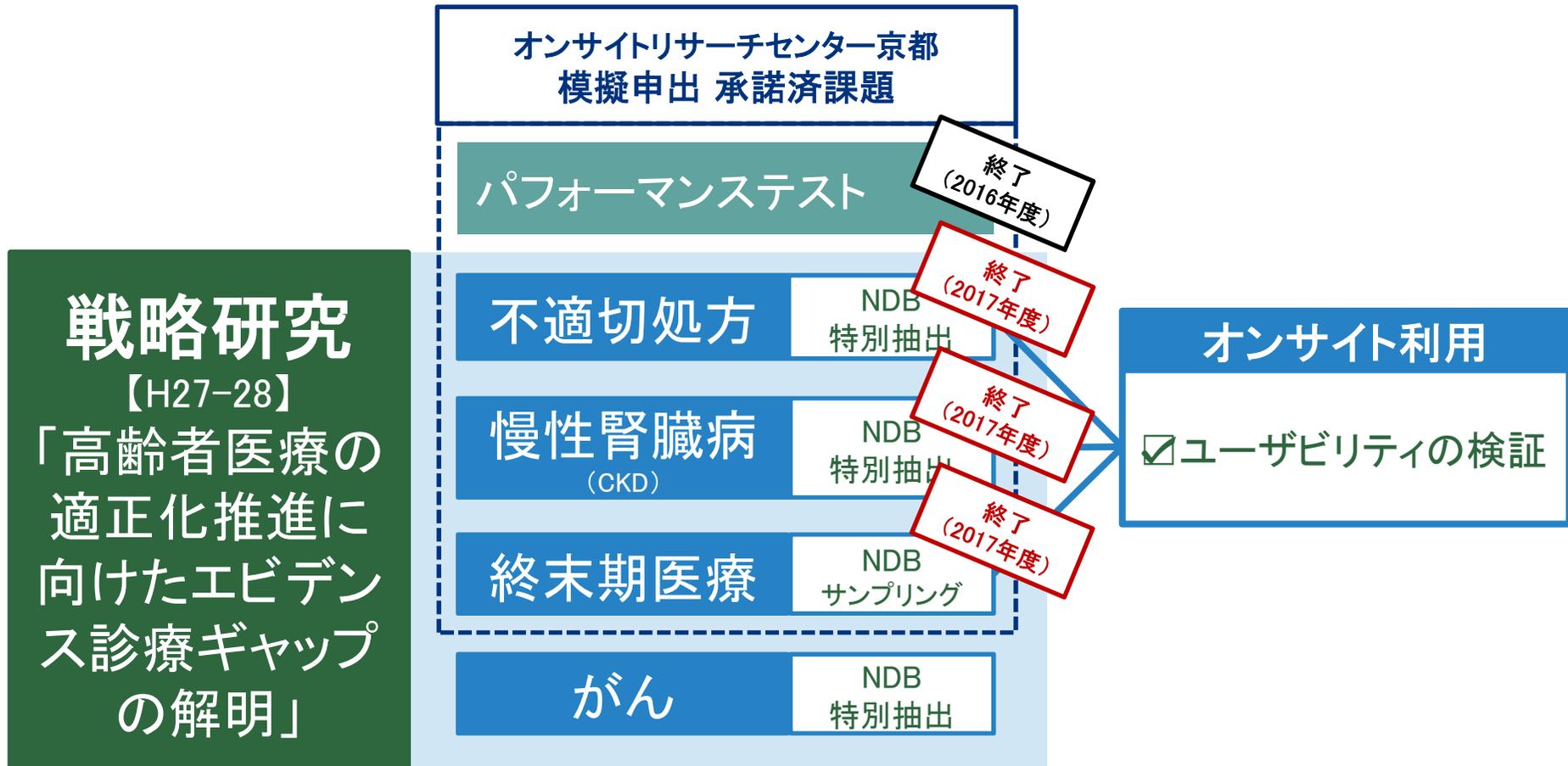
3

- レセプト情報等オンサイトリサーチセンター(京都)にて、
2016年2月17日より試行利用を開始
- レセプト情報等オンサイトリサーチセンター(京都)運用部運営委員会の開催
 - 計3回
 - 2016年3月、2017年4月、2018年4月
- 連携協力機関としての会議への参加
 - 連携協力に関する基本協定書に基づく連絡協議会
 - 第1回(2015年8月24日)より計5回参加
 - オンサイトリサーチセンター実務者会議
 - 第1回(2016年5月12日)より計11回参加
- レセプト情報等の提供に関する有識者会議での進捗報告
 - 第29回(2016年3月)、第32回(2016年7月)
- 医療情報学連合大会でのシンポジウム開催
 - 計3回
 - 2015年11月(宜野湾)、2016年11月(横浜)、2017年11月(大阪)



オンサイトリサーチセンター京都における 模擬申出課題の概要

4



* 課題名: レセプト情報等オンサイトリサーチセンター利用普及に向けたセンター機能と性能の実証的研究に関する研究

● 承諾済みの模擬申出をすべて終了した



模擬申出課題の終了について

No.	研究名称 ※目的	申出者
1	高齢患者における不適切処方薬の処方実態に関する研究 ※65歳以上の不適切処方の回数、投与量、併用薬を調査し、有害事象との関連、医療経済評価を行う	黒田知宏 (医療情報企画部)
検証内容を本日このあと報告		
2	終末期高齢患者において提供される医療に関する実態評価 ※終末期高齢者同定アルゴリズムの構築および終末期高齢者医療のEvidence Practice Gapの解明	黒田知宏 (医療情報企画部)
3	高齢慢性腎臓病患者における診療の質の定量的評価 ※高齢慢性腎臓病患者の診療の質の解明および診療の質と慢性腎不全進行の関連を検証	黒田知宏 (医療情報企画部)

- 現行の模擬申出課題は戦略研究（平成26-28）で設定した3研究
- 特別抽出データで実施した内容と同等の結果を得たので終了
 - レセプト情報等オンサイトリサーチセンター（京都）運用部
平成30年第1回 運営委員会（平成30年4月24日）にて承認



本日の話題

1. 運用および活動の概要
2. **模擬申出課題の進捗**
3. 考察
4. 結語



大規模データ解析のパフォーマンス(2016度)

		データ抽出・件数集計ツール			統計解析ツール		
		定形帳票	自由分析	SQL Plus	Oracle R	R	SAS
検証事項		プルダウンで検索条件を選択する	マウスでテーブルや列を選択する	SQL構文を使う	Rコマンドを使う	Rコマンドを使う	SASコマンドを使う
件数集計	レセプト数	行数集計は可能	出来る				
	患者数		出来る				
データダウンロード				長時間かかる	メモリの制約が大		
ローカルでの統計解析					メモリの制約が大	出来る	
サーバーでの統計解析					出来る		



高齢患者における不適切処方薬の 使用実態に関する研究(2017年度)

8

目的

- ・ 高齢者の不適切処方の実態を記述
- ・ 不適切処方が起こる要因を検討
- ・ 不適切処方と有害事象発生に関連を検証

対象者

- ・ 65歳以上(2010年10月1日)
- ・ 2010年4月 - 2015年3月診療分

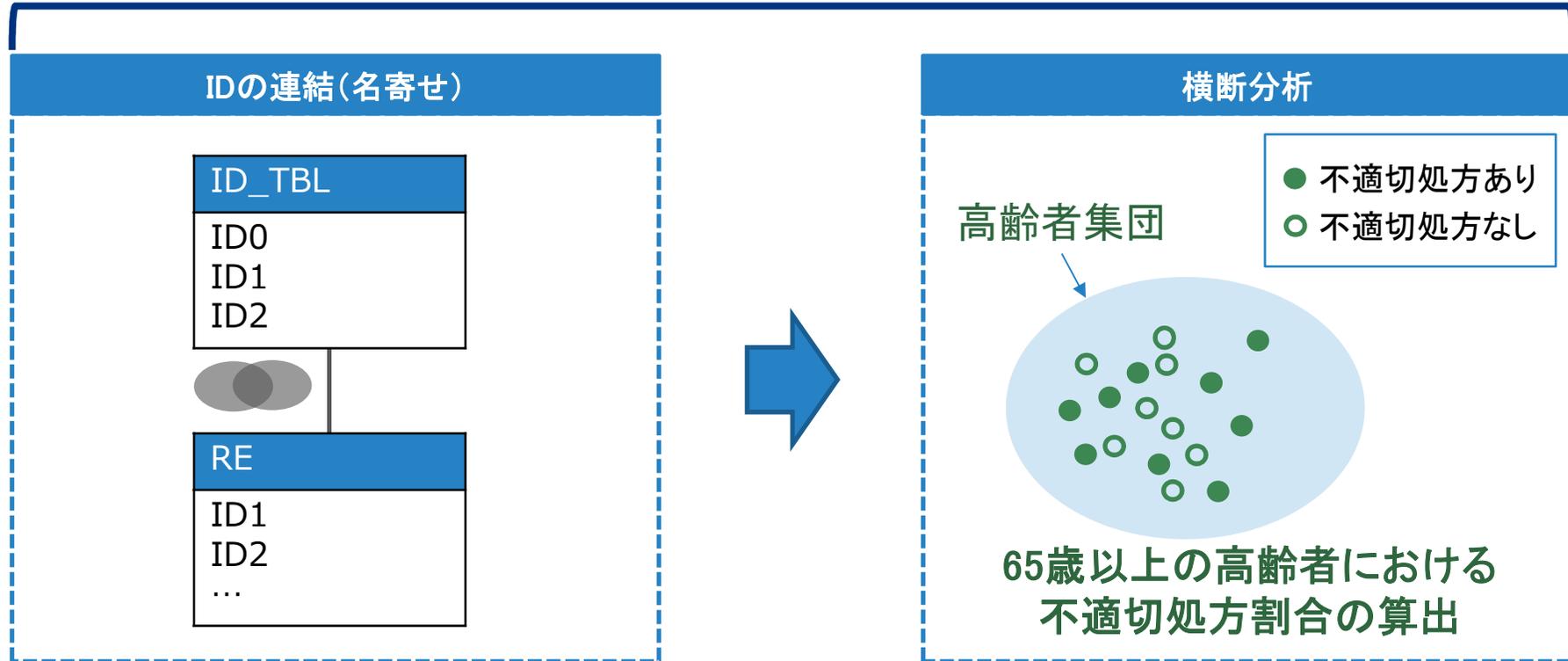
平成27年度戦略研究 分担研究報告より



不適切処方研究における オンサイト利用のねらい

9

特別抽出データ解析環境とオンサイトのパフォーマンス比較



※図はすべてイメージであり、実際のデータではありません。

- ID連結、横断的分析
 - **【POINT】特別抽出データの解析環境と比較して、同等もしくはそれ以上のパフォーマンスが得られるか**



不適切処方研究 データハンドリングの経過

対象

- ・ 65歳以上の内科、DPC、調剤レセプト(約3千万ID)

IDの連結(名寄せ)

- ・ 奈良医大を中心に開発されたアルゴリズム
(野田、久保、明神 他. 厚生指標 2017)
- ・ ID1が同じなら同一人物
- ・ ID1が異なってもID2が同じなら同一人物
- ・ 他人同士を紐付けないための独自ロジックあり

ID1 (健康保険証記号番号由来)

yL5eFiGRHF2iSk9dfJZY
5fXKPCeeU3VUtz8Kvj7i
m4JAbiXTxTW2qN3QaJ4Q
m4JAbiXTxTW2qN3QaJ4Q
HxzxxCgKbu6pcpd958dK

ID2 (姓名由来)

m4JAbiXTxTW2qN3QaJ4Q
gXZQnLXqSX6BJ9696PRU
gXZQnLXqSX6BJ9696PRU
TEHUXLM3uALS5HcH3fBq
vnrTbtCAWuDLYzSbV5Kc



ID0

m4JAbiXTxTW2qN3QaJ4Q
5fXKPCeeU3VUtz8Kvj7i
5fXKPCeeU3VUtz8Kvj7i
5fXKPCeeU3VUtz8Kvj7i
HxzxxCgKbu6pcpd958dK



名寄せの処理時間

	1年度分 (H22)	2年度分 (H22-23)
ID数	2900万ID	3300万ID
処理時間	5時間	7時間



不適切処方研究 データハンドリングの経過

横断研究用データ抽出

- 抽出データ: 患者ID、年齢、性別、医薬品(コード、使用情報)、傷病名
- 医薬品コード840種類
 - 定型抗精神病薬(ハロペリドール、クロルプロマジンなど)
 - 非定型抗精神病薬(リスペリドン、オランザピンなど)
 - ベンゾジアゼピン系睡眠薬
- 傷病名コード40種類
 - 認知症

	1年度分(H22)	2年度分(H22-23)
行数	6億行	13億行
データサイズ	24GB	50GB
処理時間	4時間	13時間

※ダウンロードは300MB/24時間で中止



解析用データセット作成

- ・ データ形式を縦長から横長に
- ・ 解析や集計に使える1行1ID化

患者ID	年齢	性別	医薬品	傷病名
p6Fi2bRc	75	M	610461209	NULL
p6Fi2bRc	75	M	NULL	8842618
Kc8hWzra	84	F	620002517	NULL
Kc8hWzra	84	F	620005594	NULL



患者ID	年齢	性別	認知症	抗精神	BZD
p6Fi2bRc	75	M	1	1	0
Kc8hWzra	84	F	0	0	1

	1年度分(H22)	2年度分(H22-23)
行数	2900万ID	3300万ID
データサイズ	1.1GB	1.3GB
処理時間	2時間	4時間



不適切処方研究 解析用データ作成過程

	特別抽出	オンサイト
データ入手		
承諾からデータ操作まで	14ヶ月	承諾後ただちに操作可
解析用データ作成		
月別データファイル統合	1年分:24時間	不要
名寄せ	5年分:72時間	2年分:7時間
データの妥当性検証	可能	可能
抽出条件の再考	再審査、再抽出	可能

手戻りの発生



本日の話題

15

1. 運用および活動の概要
2. 模擬申出課題の進捗
3. **考察**
4. 結語



NDBを利用した研究: 4つの段階

16

4つの段階	知識、技術の種類
1. 研究計画	<ul style="list-style-type: none">研究デザインNDB、レセプト
2. 利用準備 (データ入手)	<ul style="list-style-type: none">NDB利用申出倫理申請
3. 解析用データの作成	<ul style="list-style-type: none">データベース言語への理解データハンドリング全般へのリテラシー
4. 実際の解析	<ul style="list-style-type: none">統計解析手法統計解析ソフトウェア

• NDBに特有の知識、技術
• 工学的要素が強い
• この段階以外については支援を要さない利用者が一定程度存在する



- データベース言語への理解
 - SQLの工夫(パーティション、TRUNCATE、DLI…)
 - 各種SQL言語(MySQL、PostgreSQL、SQL Serverなど)のOracle SQLへの翻訳
- データハンドリング全般へのリテラシー
 - データの成型やクレンジングの技術
 - 割り当て表領域上限150GBに対する配慮
 - 不要なオブジェクトの削除
 - データの規模などを考慮した処理時間の把握



「解析用データの作成」におけるPDCA

18

■ P: SQLの開発

■ D: 解析用データの作成

■ C: データの妥当性の検証

■ A: 抽出条件の試行錯誤

オンサイトによる改善ポイント①

- 承諾から利用までの期間短縮化
- 利用者個人では用意が困難な環境の提供

オンサイトによる改善ポイント②

- 特別抽出では非常に困難
- 再審査や再抽出に多大な労力



効率的なNDBオンサイトの利用に向けて

19

レセプト、DPCなどの使用経験があった者もNDBでは苦勞…



ツール

- ・ データ構造、データハンドリングなどの知識、技術の標準化
- ・ ノウハウ共有、学習教材の開発



ひと

- ・ 利用者間のコミュニティづくり
- ・ サポートデスク設置(支援人材育成)





本日の話題

20

1. 運用体制の概要
2. 模擬申出課題の進捗
3. 考察
4. **結語**



- オンサイト京都の運用および活動について報告
 - 連携協力機関として関連する会議、学会に参加
- パフォーマンステストの振り返り、及び模擬申出課題の進捗について報告
 - 昨年度パフォーマンス検証を終了し、今年度より個別の模擬申出課題である不適切処方に関する疫学研究に着手
 - 高齢者の1年分のデータを対象とした検証では、IDの結合(名寄せ)、データ抽出が実用的な時間内に実施可能



- 模擬申出課題の実施を通じた考察を報告
 - 解析データの作成
 - 現実的な時間内にデータ作成できる環境を実現
 - 手戻りなくデータ抽出をくり返すことが可能
 - オンサイトの機能ではカバーできない部分を今後、ツールの開発や人材充実を補う必要あり