

# レセプトデータ分析 のための前処理

岡本悦司  
(国立保健医療科学院)

# レセプトデータの特徴

- 通常のデータでは分析対象となるデータが同一行に含まれているが
- レセプトデータでは異なる行にまたがっている
- ナショナルデータベースでは通番(ID1)とレセプトID(ID2)が含まれている

# レコード識別

- 目的変数
  - IR: provider record
  - RE: patient record (name, DOB, SEX, etc)
  - HO: summary records (charges, N of visits)
- 説明変数
  - SY: diagnoses (date of diagnoses, prognoses, principal Dx)
  - IY: drugs (drug code, therapeutic class)
  - SI: clinical procedures (code, classification, N times provided)

# 2つの手法

- 1) 求めるデータを含むID2を抽出し(サブクエリ), そのID2が含まれるデータを抽出する
- 2) あらかじめ目的変数にあたるIR, RE, HOレコードを説明変数にあたる(SY, IY, SI, TO)レコードに挿入しておく

# 1) サブクエリ法

田中のデータだけどうやって抽出するか？

ID1	ID2	レコード識別	F1
1	レセ1	RE	田中
2	レセ1	SY	
3	レセ1	SI	
4	レセ1	IY	
5	レセ1	RC	
6	レセ2	RE	山本
7	レセ2	SY	
8	レセ2	SI	
9	レセ2	IY	
10	レセ2	RC	

まず田中のID2を取得し(**サブクエリ**), そのID2を含む行を抽出する

SELECT \* FROM MED where ID2 in

(**select ID2 from MED where レコード識別='RE' and F1='田中'**)

ID1	ID2	レコー	F1
1	レセ1	RE	田中
2	レセ1	SY	
3	レセ1	SI	
4	レセ1	IY	
5	レセ1	RC	

# 2) 目的変数を全説明変数に挿入

① 補正前レセプト

データ識別	レセプトごと連番	同枝番	レコード識別情報	診療識別	負担区分	コード	使用量	点数	回数
2	10	0	MN			レセプト管理レコード			
1	20	0	IR			医療機関情報			
1	30	0	RE			レセプト共通レコード			
1	40	0	HO			保険者レコード			
1	50	0	SI	21	1	診療行為1	1		
1	60	0	SI		1	診療行為2(加算)	2		
1	70	0	SI		1	診療行為3(加算)	22	118	1
1	80	0	IY		1	医薬品1(2回服用)	100		
1	90	0	IY		1	医薬品2(2回服用)	2	67	14
1	100	0	TO		1	特定器材1	2		
1	110	0	TO		1	特定器材2	3	20	1
1	120	0	CO		1	コメント1	文字		
1	130	0	CO		1	コメント2	文字		
1	140	0	SI	21	1	診療行為1	1		
1	150	0	SI		1	診療行為3(加算)	22	118	1
1	160	0	IY		1	医薬品1(2回服用)	100		
1	170	0	IY		1	医薬品2(2回服用)	2	67	14
1	180	0	TO		1	特定器材1	2		
1	190	0	TO		1	特定器材2	3	20	1
1	200	0	CO		1	コメント1	文字		
1	210	0	CO		1	コメント2	文字		
1	220	0	:						
2	230	0	RC			34sjdb456eivy59030cvkdhk45d25kse			

… 摘要情報

② 補正後レセプト

レコードごと連番	データ識別	レセプトごと連番	同枝番	レコード識別情報	診療識別	診療行為番号	負担区分	コード	使用量	点数	回数
1	2	10	0	MN	34sjdb456eivy59030cvkdhk45d25k			レセプト管理レコード			
2	1	20	0	IR	34sjdb456eivy59030cvkdhk45d25k			医療機関情報			
3	1	30	0	RE	34sjdb456eivy59030cvkdhk45d25k			レセプト共通レコード			
4	1	40	0	HO	34sjdb456eivy59030cvkdhk45d25k			保険者レコード			
5	1	50	0	SI	34sjdb456eivy59030cvkdhk45d25k	21	1	診療行為1	1	118	1
6	1	60	0	SI	34sjdb456eivy59030cvkdhk45d25k	21	1	診療行為2(加算)	2	118	1
7	1	70	0	SI	34sjdb456eivy59030cvkdhk45d25k	21	1	診療行為3(加算)	22	118	1
8	1	80	0	IY	34sjdb456eivy59030cvkdhk45d25k	21	1	医薬品1(2回服用)	100	67	14
9	1	90	0	IY	34sjdb456eivy59030cvkdhk45d25k	21	1	医薬品2(2回服用)	2	67	14
10	1	100	0	TO	34sjdb456eivy59030cvkdhk45d25k	21	1	特定器材1	2	20	1
11	1	110	0	TO	34sjdb456eivy59030cvkdhk45d25k	21	1	特定器材2	3	20	1
12	1	120	0	CO	34sjdb456eivy59030cvkdhk45d25k	21	1	コメント1	文字		
13	1	130	0	CO	34sjdb456eivy59030cvkdhk45d25k	21	1	コメント2	文字		
14	2	140	1	SI	34sjdb456eivy59030cvkdhk45d25k	22	2	診療行為1	1	116	1
15	3	150	2	SI	34sjdb456eivy59030cvkdhk45d25k	23	2	診療行為3(加算)	22	116	1
16	4	160	3	IY	34sjdb456eivy59030cvkdhk45d25k	24	2	医薬品1(2回服用)	100	67	14
17	5	170	4	IY	34sjdb456eivy59030cvkdhk45d25k	25	2	医薬品2(2回服用)	2	67	14
18	6	180	5	TO	34sjdb456eivy59030cvkdhk45d25k	26	2	特定器材1	2	20	1
19	7	190	6	TO	34sjdb456eivy59030cvkdhk45d25k	27	2	特定器材2	3	20	1
20	8	200	7	CO	34sjdb456eivy59030cvkdhk45d25k	28	2	コメント1	文字		
21	9	210	8	CO	34sjdb456eivy59030cvkdhk45d25k	29	2	コメント2	文字		
22	1	220	0	:							
23	2	230	0	RC	34sjdb456eivy59030cvkdhk45d25kse						

… データ補正箇所

↑  
レセプトを一意に識別する16進数32ケタ


# 個人IDを挿入する

下記の個人IDは実際は64ケタのハッシュ関数に置き換えられている






ID	ID2	レコード	識別 F1		ID	ID2	個人ID	レコード	識別 F1
1	レセ1	RE	田中		1	レセ1	田中	RE	田中
2	レセ1	SY			2	レセ1	田中	SY	
3	レセ1	SI			3	レセ1	田中	SI	
4	レセ1	IY			4	レセ1	田中	IY	
5	レセ1	RC		→	5	レセ1	田中	RC	
6	レセ2	RE	山本		6	レセ2	山本	RE	山本
7	レセ2	SY			7	レセ2	山本	SY	
8	レセ2	SI			8	レセ2	山本	SI	
9	レセ2	IY			9	レセ2	山本	IY	
10	レセ2	RC			#	レセ2	山本	RC	



# デザインビューでフィールド挿入 SQLではALTER TABLE ADD 個人ID

	フィールド名	データ型
	ID1	数値型
	ID2	テキスト型
	個人ID	テキスト型
		テキスト型
		テキスト型
		テキスト型
		テキスト型
		テキスト型
		テキスト型
		テキスト型
		テキスト型

	主キー(K)
	切り取り(D)
	コピー(C)
	貼り付け(P)
	行の挿入(I)

# 氏名があるのはレコード識別RE

SELECT \* FROM MED where レコード識別='RE'  
上のクエリを次頁では☆と表す

ID	レセコ	個人I	レコー	F1
1	レセ1		RE	田中
6	レセ2		RE	山本

# サブクエリとjoin

SELECT \* FROM

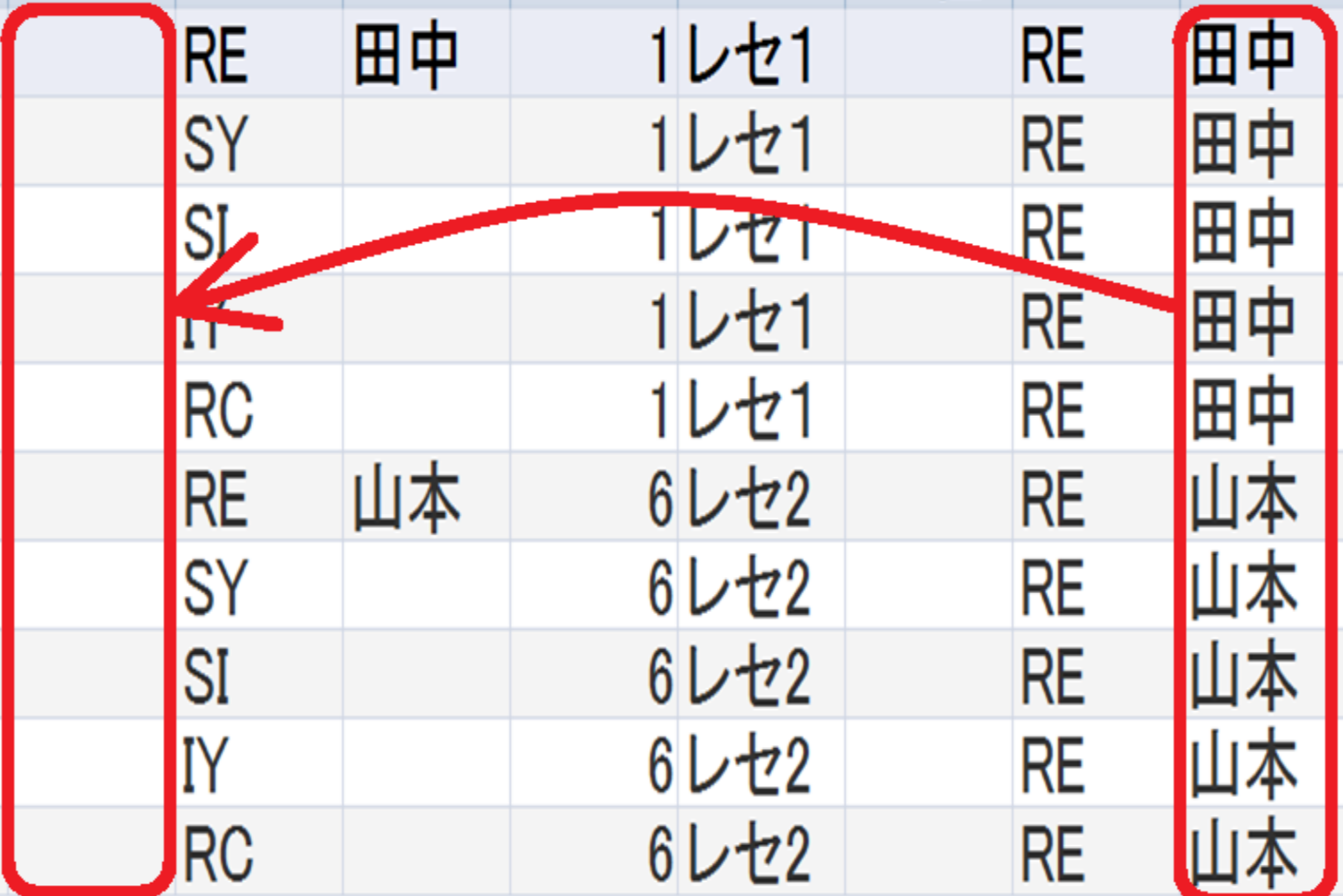
**MED inner join (★) as RE on MED.ID2=RE.ID2;**

赤字部分を次頁では★とする

MED.ID1	MED.ID2	MED.ID3	MED.ID4	MED.ID5	RE.ID	RE.レ	RE.個	RE.レ	RE.F1
1	レセ1		RE	田中	1	レセ1		RE	田中
2	レセ1		SY		1	レセ1		RE	田中
3	レセ1		SI		1	レセ1		RE	田中
4	レセ1		IY		1	レセ1		RE	田中
5	レセ1		RC		1	レセ1		RE	田中
6	レセ2		RE	山本	6	レセ2		RE	山本
7	レセ2		SY		6	レセ2		RE	山本
8	レセ2		SI		6	レセ2		RE	山本
9	レセ2		IY		6	レセ2		RE	山本
10	レセ2		RC		6	レセ2		RE	山本

# RE.F1をMED.個人IDにコピー

MED.	MED.	MED.	MED.	MED.	RE.ID	RE.レ	RE.個	RE.レ	RE.F1
1	レセ1		RE	田中	1	レセ1		RE	田中
2	レセ1		SY		1	レセ1		RE	田中
3	レセ1		SI		1	レセ1		RE	田中
4	レセ1		IY		1	レセ1		RE	田中
5	レセ1		RC		1	レセ1		RE	田中
6	レセ2		RE	山本	6	レセ2		RE	山本
7	レセ2		SY		6	レセ2		RE	山本
8	レセ2		SI		6	レセ2		RE	山本
9	レセ2		IY		6	レセ2		RE	山本
10	レセ2		RC		6	レセ2		RE	山本



# UPDATE ★

SET MED.個人ID=RE.F1.

ID1	ID2	個人ID	レコー	F1
1	レセ1	田中	RE	田中
2	レセ1	田中	SY	
3	レセ1	田中	SI	
4	レセ1	田中	IY	
5	レセ1	田中	RC	
6	レセ2	山本	RE	山本
7	レセ2	山本	SY	
8	レセ2	山本	SI	
9	レセ2	山本	IY	
10	レセ2	山本	RC	

```
select * FROM MED where 個人  
ID='田中'
```

ID1	ID2	個人I	レコー	F1
1	レセ1	田中	RE	田中
2	レセ1	田中	SY	
3	レセ1	田中	SI	
4	レセ1	田中	IY	
5	レセ1	田中	RC	

# 具体例

「NDBデータ抽出のための前処理」  
2頁参照

# <http://enkaku3.niph.go.jp/renandi/>

左上の「教材グループ検索」→検索KWに「レセプト」と入れて検索

インターネットによる遠隔教育 国立保健医療科学院  
National Institute of Public Health e-Learning

ログインID  日本語 English  
パスワード  Login

[お知らせ](#) | [科目検索](#) | [教材グループ検索](#) | [FAQ](#)

### サンプルレセプトACCESSファイル

検索 [\[戻る\]](#)

本システムにどのような教材グループが登録されているか検索することができます。  
→ [条件を指定して検索](#)

教材グループコード	レセプト分析法
教材グループ名	サンプルレセプトACCESSファイル
登録者名	岡本悦司
作成者名	岡本悦司
バージョン	
分野	
公開範囲	未登録を含む全ての利用者が利用可能
概要	支払基金が公表する医、歯、薬、DPCのサンプルデータを分析可能なように処理したもの。
教材 <a href="#">一括ダウンロード</a> (SCORM教材は含まれません)	<ul style="list-style-type: none"><li><a href="#">サンプルレセプトACCESSファイル</a> (20739 Kb) 登録日: 2011/04/17 20:29</li><li><a href="#">説明書</a> (1144 Kb) 登録日: 2011/04/17 20:29</li><li><a href="#">NULLを除去しがないファイル</a> (20739 Kb) 登録日: 2011/04/17 20:29</li><li><a href="#">ナショナルデータベースのフォーマット</a> (15467 Kb) 登録日: 2011/04/22 13:37</li></ul>



# 例1) 糖尿病(ICD10がE11)の患者のみを抽出

- `SELECT ID2 FROM SHOBYOJOHO  
where LEFT(ICD10,3)='E11' group by ID2;`

# 例2) DPCレセプトにおいて診療識別の診療行為の実施回数を実施日ごとに集計

- SELECT MID(CODE,2,2) FROM SHINRYOJOHO where LEFT(RSPTSHUBETSU,1)='2'...☆
- ☆ group by MID(CODE,2,2) ...★

TRANSFORM COUNT(\*)



PIVOT JISSHIYMD

# 例3) 薬効中分類ごとの投薬点数 を抽出

- ```
SELECT LEFT(Y.YJCODE,2),  
SUM(SHINRYOJOHO.TENSU) FROM  
SHINRYOJOHO inner join Y on  
SHINRYOJOHO.CODE=Y.CODE where  
SHINRYOJOHO.RS in ('IY','CD') group by  
LEFT(Y.YJCODE,2);
```