

# レセプトデータ分析の手法

岡 本 悦 司

国立保健医療科学院

医療・福祉サービス研究部

# データ受領における暗号化処理

- 保険組合(加入者約6万人)に毎月支払基金より納入されるDVDよりデータを受領
- 支払基金からのDVDを復号処理
- 組合事務所のPCを利用しACCESSにより匿名化処理を施す
- 7か月分の医科, 調剤レセプトが約1.3GB(ACCESSは2GB限度)
- 指紋認証USBで運搬

# ACCESSによる匿名化処理後

すべてのテーブル	DS	TSUBAN	EDABAN	RS	F2	F3	F4	F5	F6
MED201107	2	1	0 MN	91000335	北海道札	01142308			
MED201108	1	2	0 IR	1	01	1		0116258	
MED201108 : テーブル	1	3	0 RE	2247	1122	42307		9665	1
MED201109	1	4	0 HO	06137384	KIGO	BANGO	1		44323
MED201109 : テーブル	1	5	0 KO	51096022	1711142			1	44323
MED201110	1	6	0 SY	5559001	4151120	1			
MED201110 : テーブル	1	7	0 SY	2809009	4180120	1			
MED201111	1	8	0 SY	8840786	4180120	1			
MED201111 : テーブル	1	9	0 SY	1579002	4230728	1		8002	
MED201112	1	10	0 SY	5770014	4230728	1		8002	
MED20112	1	11	0 SI	12	2	11201131			70
MED201201	1	12	0 SI	13	2	11300291			250
MED201201 : テーブル	1	13	0 SI	33	2	13000931			47
PHA201201	1	14	0 IY		2	64331018	1		
PHA201201 : テーブル	1	15	0 IY		2	64331018	1		
PHA201112	1	16	0 IY		2	64046200	4		40134
PHA201112 : テーブル	1	17	0 CO	33	2	81000000			レミケード
PHA201111	1	18	0 CO		2	81000000			患者体重&
PHA201110	1	19	0 CO	60	2	84000009			28
PHA201110 : テーブル	1	20	0 SI		2	16016025			48
PHA201109	1	21	0 SI	60	2	16009571			13
PHA201109 : テーブル	1	22	0 SI	60	2	16017017			40
PHA201108	1	23	0 SI	60	2	16006181			125
PHA201108 : テーブル	1	24	0 SI	60	2	16006191			144
PHA201107	1	25	0 SI	60	2	16006211			144
PHA201107 : テーブル	1	26	0 CO	60	2	84000009			28

氏名が番号に置き換わった

記番号を置き換え

# 研究室での処理(通番挿入)

- 元のレセプトデータには通番すらない
- ACCESSのオートナンバーを挿入(1から順)
- しかし複数月のデータでは重複してしまう

SERIAL	DS	TSUBAN	EDABAN	RS
1	2	1	0 MN	
2	1	2	0 IR	
3	1	3	0 RE	
4	1	4	0 HO	
5	1	5	0 SY	
6	1	6	0 SY	
7	1	7	0 SY	
8	1	8	0 SI	
9	1	9	0 SI	
10	1	10	0 SI	
11	1	11	0 SI	

# 複数月レセプトにユニークIDを挿入

- たとえば2011年7月分なら
- UPDATE MED set ID=1107000000+SERIAL
- で複数月にまたがるレセプトにユニークな通番を挿入する

SERIAL	ID	DS	TSUBAN	EDABAN	RS
1	1107000001	2	1	0 MN	
2	1107000002	1	2	0 IR	
3	1107000003	1	3	0 RE	
4	1107000004	1	4	0 HO	
5	1107000005	1	5	0 SY	
6	1107000006	1	6	0 SY	
7	1107000007	1	7	0 SY	
8	1107000008	1	8	0 SI	
9	1107000009	1	9	0 SI	
10	1107000010	1	10	0 SI	
11	1107000011	1	11	0 SI	
12	1107000012	1	12	0 TY	

# SQL Serverへのインポート

SQL Server インポートおよびエクスポート ウィザード

コピー元のテーブルおよびビューを選択  
コピーするテーブルおよびビューを1つ以上選択します。

テーブルとビュー(T):

変換元	変換先
✓ `MED`	[dbo].[MED]

列マッピング

変換元: `MED`  
変換先: [dbo].[MED]

変換先テーブルを作成する(R)  変換先テーブル内の行を削除する(W)  変換先テーブルに行を追加する(P)

SQLの編集(S)...  変換先テーブルを削除した後、再作成する(D)  ID挿入を許可する(I)

マッピング(M):

変換元	変換先	型	NULL...	サイズ	有効...	小...
SERIAL	SERIAL	int	<input type="checkbox"/>			
ID	ID	int	<input checked="" type="checkbox"/>			
DS	DS	char	<input checked="" type="checkbox"/>	1		
TSUBAN	TSUBAN	smallint	<input checked="" type="checkbox"/>			
EDABAN	EDABAN	tinyint	<input checked="" type="checkbox"/>			
RS	RS	nvarchar	<input checked="" type="checkbox"/>	255		
F2	F2	nvarchar	<input checked="" type="checkbox"/>	255		
F3	F3	nvarchar	<input checked="" type="checkbox"/>	255		
F4	F4	nvarchar	<input checked="" type="checkbox"/>	255		
F5	F5	nvarchar	<input checked="" type="checkbox"/>	255		

変換元列: F4 VarChar (255)

ヘルプ

OK キャンセル

# インポート完了 → RSPTIDを追加

プロジェクト エクスプローラー

SQLQuery1.sql - ...-ASUS¥AtoZ (52))\*

```
select * from MED
```

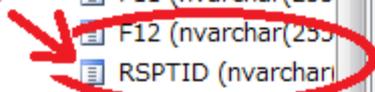
結果

	ID	DS	TSUBAN	EDABAN	RS	F2	F3	F4
1	1107000001	2	1	0	MN	910000213	東京都港区新橋	1314220591000021
2	1107000002	1	2	0	IR	1	13	1
3	1107000003	1	3	0	RE	23	1118	42204
4	1107000004	1	4	0	HO	06132013	1234567	79
5	1107000005	1	5	0	SY	2500013	4140619	1
6	1107000006	1	6	0	SY	5739014	4140619	1
7	1107000007	1	7	0	SY	8833421	4150716	1
8	1107000008	1	8	0	SI	12	1	112007410
9	1107000009	1	9	0	SI	12	1	112011010
10	1107000010	1	10	0	SI	13	1	113001810
11	1107000011	1	11	0	SI	21	1	120000710
12	1107000012	1	12	0	IY	21	1	610443044
13	1107000013	1	13	0	IY	NULL	1	620008041
14	1107000014	1	14	0	IY	NULL	1	613960041
15	1107000015	1	15	0	IY	21	1	620004502
16	1107000016	1	16	0	SI	25	1	120001210
17	1107000017	1	17	0	SI	25	1	120003170
18	1107000018	1	18	0	SI	27	1	120001810
19	1107000019	1	19	0	SI	60	1	160000310
20	1107000020	1	20	0	SI	60	1	160010010
21	1107000021	1	21	0	SI	60	1	160022510
22	1107000022	1	22	0	SI	NULL	1	160022610
23	1107000023	1	23	0	SI	NULL	1	160020410
24	1107000024	1	24	0	SI	NULL	1	160019410
25	1107000025	1	25	0	SI	60	1	160095710

データベース

- システム データベース
- データベース スナップショット
- ReportServer
- ReportServerTempDB
- sampleDB
  - データベース ダイアグラム
  - テーブル
    - システム テーブル
    - dbo.MED
      - 列
        - ID (int, NULL)
        - DS (char(1), NULL)
        - TSUBAN (smallint)
        - EDABAN (tinyint)
        - RS (nvarchar(255))
        - F2 (nvarchar(255))
        - F3 (nvarchar(255))
        - F4 (nvarchar(255))
        - F5 (nvarchar(255))
        - F6 (nvarchar(255))
        - F7 (nvarchar(255))
        - F8 (nvarchar(255))
        - F9 (nvarchar(255))
        - F10 (nvarchar(255))
        - F11 (nvarchar(255))
        - F12 (nvarchar(255))
        - RSPTID (nvarchar(255))

RSPTID  
を追加



# レセプトを一意に識別するIDを挿入

- 1レセプトはRレコードで終わるのでその行のIDを同一レセプトの行にRSPTIDとして挿入
- UPDATE MED set RSPTID='MED'+RC.ID from MED, (select ID, TSUBAN from MED where RS='RC' and EDABAN=1) as RC where MED.ID>(RC.ID-RC.TSUBAN) and MED.ID<=RC.ID
  - 2011年4月分より医科Rレコードに枝番1が追加されているので注意
- 上記の処理が最も時間がかかる(数日かかることも)。1万件ごとに分割する等して迅速化をはかる。

# 各行にレセプトIDが入った (前処理ほぼ完了)

SQLQuery1.sql - ...-ASUS¥AtoZ (52))\*

```
select ID, RSPTID, TSUBAN, RS from MED
```

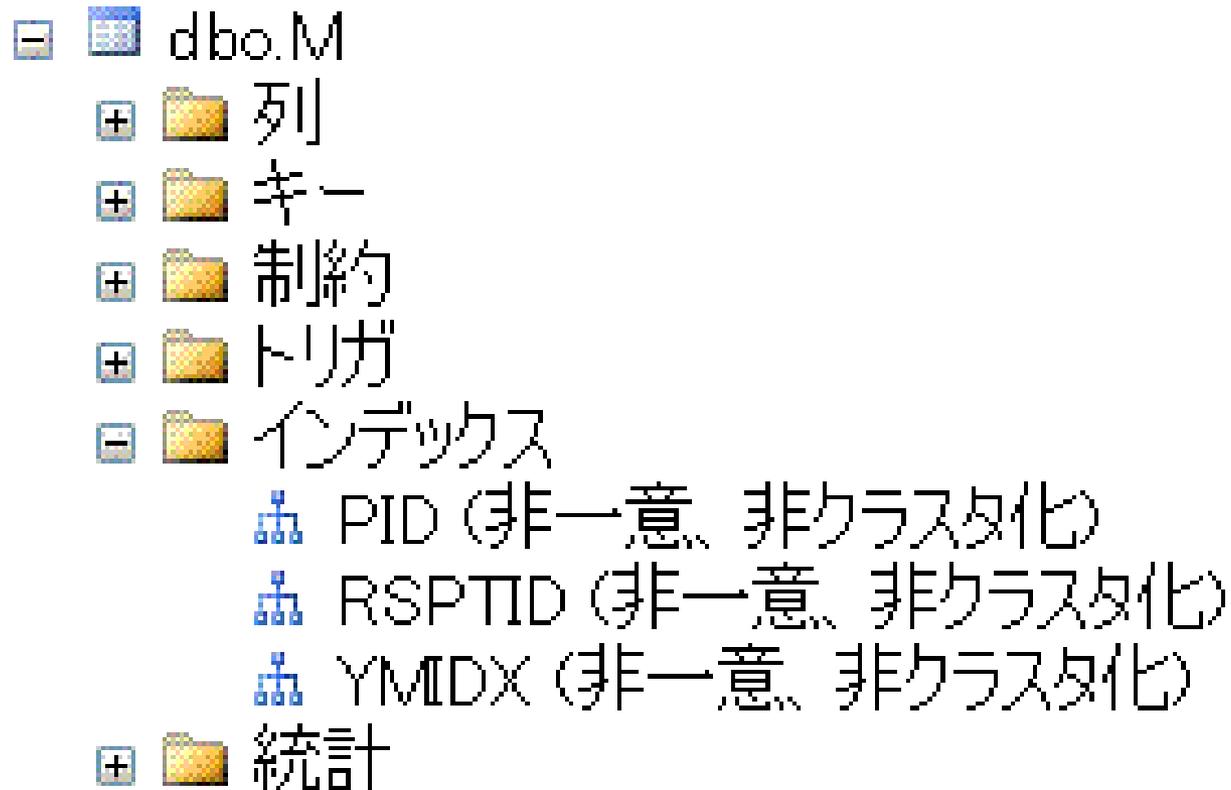
結果 メッセージ

	ID	RSPTID	TSUBAN	RS
1	110700001	MED1107000029	1	MN
2	110700002	MED1107000029	2	IR
3	110700003	MED1107000029	3	RE
4	110700004	MED1107000029	4	HO
5	110700005	MED1107000029	5	SY
6	110700006	MED1107000029	6	SY
7	110700007	MED1107000029	7	SY
8	110700008	MED1107000029	8	SI
9	110700009	MED1107000029	9	SI
10	110700010	MED1107000029	10	SI
11	110700011	MED1107000029	11	SI
12	110700012	MED1107000029	12	IY
13	110700013	MED1107000029	13	IY
14	110700014	MED1107000029	14	IY
15	110700015	MED1107000029	15	IY
16	110700016	MED1107000029	16	SI
17	110700017	MED1107000029	17	SI
18	110700018	MED1107000029	18	SI
19	110700019	MED1107000029	19	SI
20	110700020	MED1107000029	20	SI
21	110700021	MED1107000029	21	SI
22	110700022	MED1107000029	22	SI
23	110700023	MED1107000029	23	SI
24	110700024	MED1107000029	24	SI
25	110700025	MED1107000029	25	SI
26	110700026	MED1107000029	26	SI
27	110700027	MED1107000029	27	SI
28	110700028	MED1107000029	28	EX
29	110700029	MED1107000029	29	RC
30	110700030	MED1107000044	1	MN
31	110700031	MED1107000044	2	IR
32	110700032	MED1107000044	3	RE
33	110700033	MED1107000044	4	HO

TSUBAN1~29行  
に同一の  
RSPTID  
(MED1107000029)  
が入った。

# インデックス化による処理迅速化

- レセプトID, 個人識別ID, 診療年月等しばしば分析する項目にインデックスを付与して処理の迅速化をはかる



# その他の処理

## 摘要欄の空白を埋める

RSPTID	D	T	EI	RS	ID1	ID	F2	F3	F4	F5	F6	F7
df061727	1	8	0SI	サンプル	7	12	1	112007410	69	1		
df061727	1	9	0SI	サンプル	7	12	1	112011010	52	1		
df061727	1	10	0SI	サンプル	7	13	1	113001810	225	1		
df061727	1	11	0SI	サンプル	7	21	1	120000710	9	1		
df061727	1	12	0IY	サンプル	7	21	1	610443044	1			
df061727	1	13	0IY	サンプル	7		1	620008041	1			
df061727	1	14	0IY	サンプル	7		1	613980041	0.5	19	35	
df061727	1	15	0IY	サンプル	7	21	1	620004502	2	2	35	
df061727	1	16	0SI	サンプル	7	25	1	120001210	42	1		
df061727	1	17	0SI	↑一連の行為単位				120003170	65	1		
df061727	1	18	0SI	↑				120001810	8	1		
df061727	1	19	0SI	サンプル	7	60	1	160000310	26	1		
df061727	1	20	0SI	サンプル	7	60	1	160010010	50	1		
df061727	1	21	0SI	サンプル	7	60	1	160022510				
df061727	1	22	0SI	サンプル	7		1	160022810				
df061727	1	23	0SI	サンプル	7		1	160020410				
df061727	1	24	0SI	サンプル	7		1	160019410	56	1		
df061727	1	25	0SI	サンプル	7	60	1	160095710	13	1		
df061727	1	26	0SI	サンプル	7	60	1	160061910	144	1		

RSPTID	D	T	EI	RS	ID1	ID	F2	F3	F4	F5	F6	F7
df061727	1	8	0SI	サンプル	7	12	1	112007410	69	1		
df061727	1	9	0SI	サンプル	7	12	1	112011010	52	1		
df061727	1	10	0SI	サンプル	7	13	1	113001810	225	1		
df061727	1	11	0SI	サンプル	7	21	1	120000710	9	1		
df061727	1	12	0IY	サンプル	7	21	1	610443044	1	19	35	
df061727	1	13	0IY	サンプル	7	21	1	620008041	1	19	35	
df061727	1	14	0IY	サンプル	7	21	1	613980041	0.5	19	35	
df061727	1	15	0IY	サンプル	7	21	1	620004502	2	2	35	
df061727	1	16	0SI	サンプル	7	25	1	120001210	42	1		
df061727	1	17	0SI	サンプル	7	25	1	120003170	65	1		
df061727	1	18	0SI	サンプル	7	27	1	120001810	8	1		
df061727	1	19	0SI	サンプル	7	60	1	160000310	26	1		
df061727	1	20	0SI	サンプル	7	60	1	160010010	50	1		
df061727	1	21	0SI	サンプル	7	60	1	160022510		56	1	
df061727	1	22	0SI	サンプル	7	60	1	160022810		56	1	
df061727	1	23	0SI	サンプル	7	60	1	160020410		56	1	
df061727	1	24	0SI	サンプル	7	60	1	160019410	56	1		
df061727	1	25	0SI	サンプル	7	60	1	160095710	13	1		
df061727	1	26	0SI	サンプル	7	60	1	160061910	144	1		

医科の摘要欄データはプリントアウトした時に見やすいよう左のように空白がある。右のように加工して空白を埋める