

申出者から見た レセプト利用手続きと解析

国立がん研究センターがん対策情報センター
がん政策科学研究部

東 尚弘

平成25年6月26日

本日のトピック

(準備編)

- 必要な物理的セキュリティの前提
- 入退室管理、ログ管理

(実務編)

- スケジュールについて
- データの処理の実際について

(その他)

- 実地監査と指摘事項について

物理セキュリティの前提

- 入退室が管理される部屋
 - 物理的に隔離が原則
 - 申出者・利用者以外は立ち入らないスペース
 - 他のプロジェクトでは使わない
- データの保管は、パソコンからさらに施錠
- パソコンは通常はネットから隔絶、
 - ただし、定期的にはアップデートが可能

本申出については

- 医療情報管理室を設置
- カードキーシステムを導入

実際の入退室管理、ログ管理

- 入退室管理

カードキーではカード解錠の記録のみ・・・

紙台帳で管理することにした

PCのログイン管理

①紙管理で開始

②Windowsのイベントビューアーを使ってみた

監査で①、②は改ざんが可能であると指摘

③専用ソフト「アクセスログ2」(Life Boat社)
に切り替え

私の申し出の概要

- 悪性新生物の診断のある患者
- がん関連の診療行為・薬剤
 - それぞれについて、対象コードのリストをcsvファイルで提出
- 2年間(2009年4月～2011年3月)

(参考)

- REのレコード数

医科レセプト	107616162	レコード
調剤レセプト	13264743	レコード
DPCレセプト	7549991	レコード

データ提供までのスケジュール

- ・平成24年

5月 申出

9月 有識者会議

10月 承諾書

抽出テンプレートを記入(数回のやりとり)

- ・平成25年

1月25日 データ受領

解析体制:

パソコン

- ・DELL Precision 490 Xeon X5535 2.66GHz 32GB
Windows Vista Professional 64 bit

外付けHD

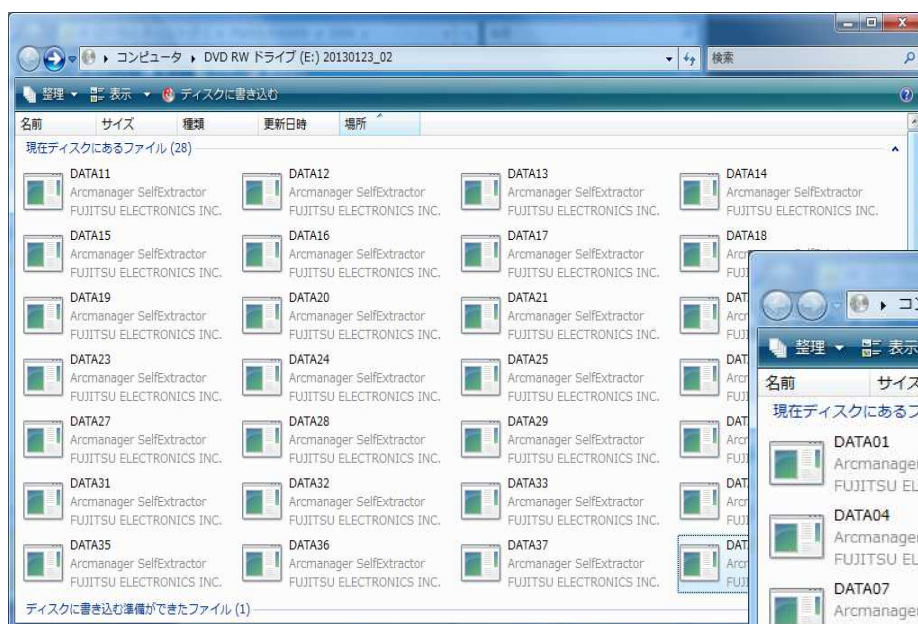
- ・ハードウェア暗号化・USB3.0

ソフト

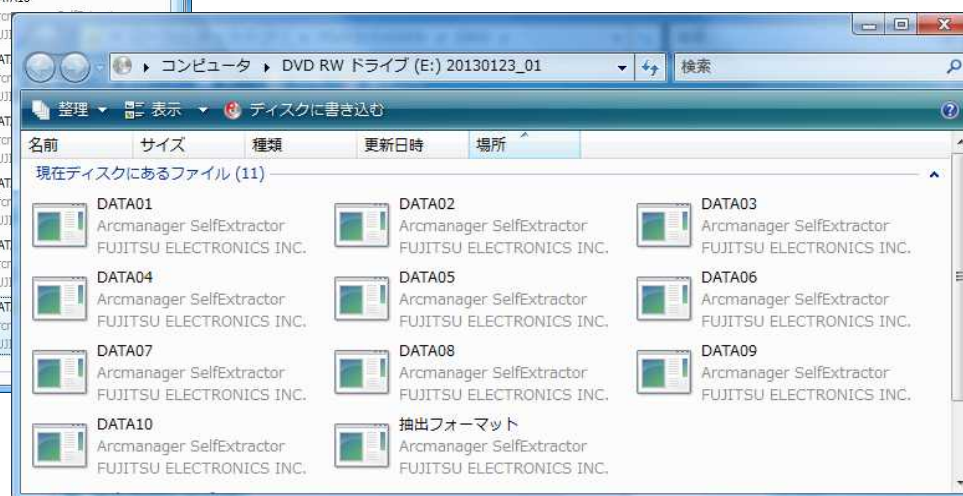
- ・MySQL 5.6 でデータ加工
- ・Stata11／MP4 で詳細解析(予定)

提供データの様子

- DVD上は2枚、請求月38個のファイルに分かれて保存



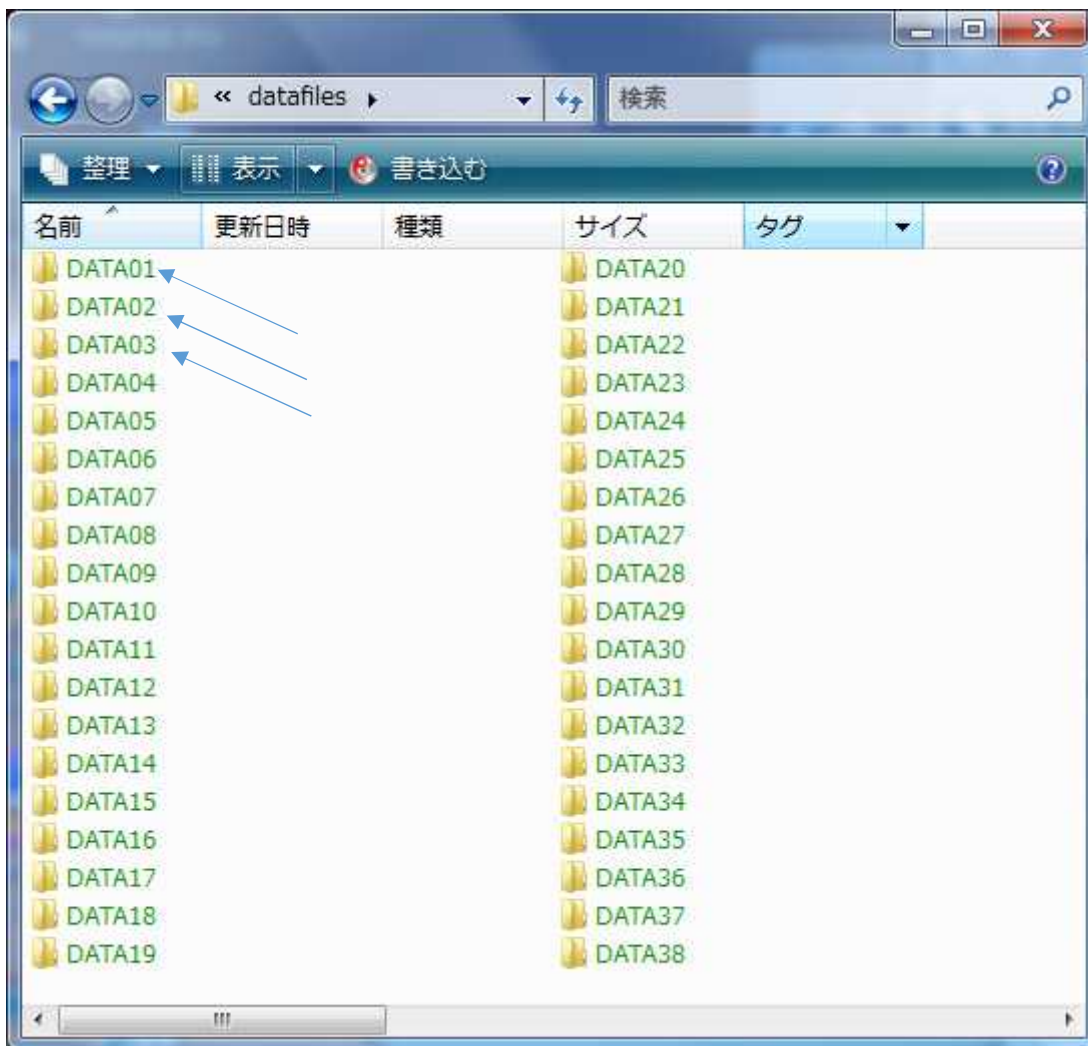
3.1 GB



3.67GB

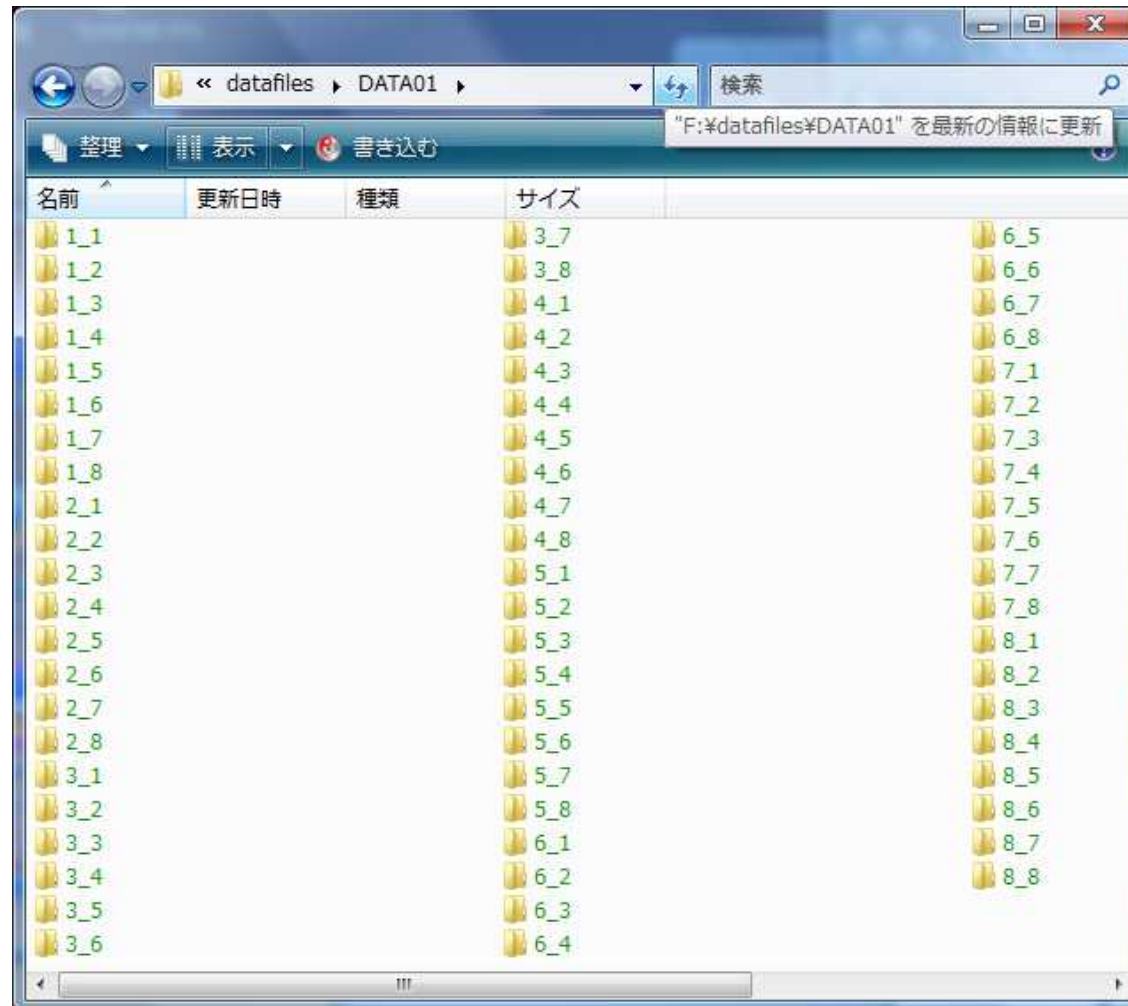
全てのファイルにパスワードがついていた

すべてを展開すると38フォルダ

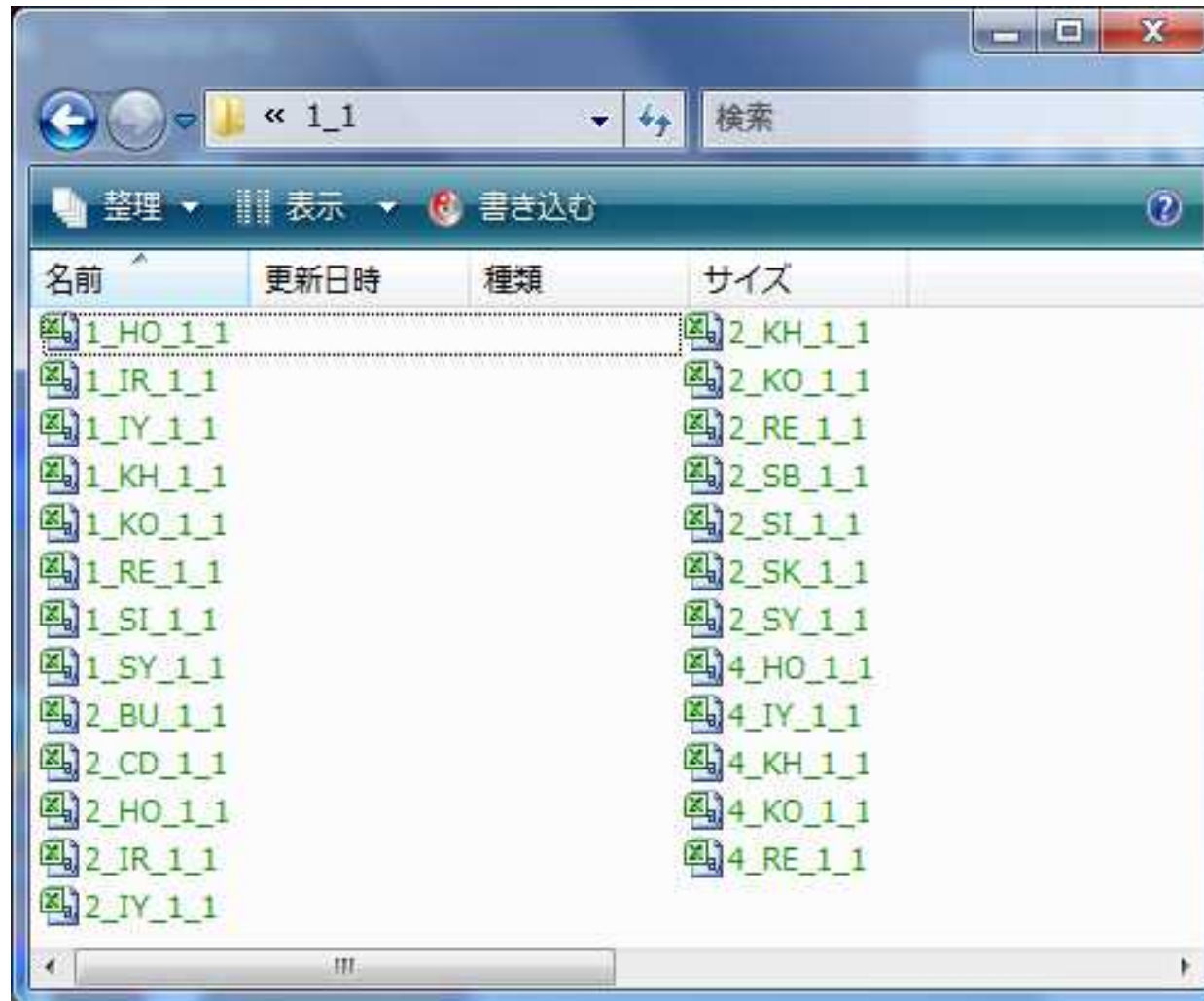


ファイル全体 で36.1 GB

これらフォルダの下に64フォルダ



レセ識別子毎にファイルが構成



ファイルの統合・・・

- 一つの識別子(情報タイプ)で
64処理系×38月=2432個のファイル
- 識別子は医科8種、調剤5種、DPC12種、
- 全部で・・・ $2432 \times 25 = 60800$ ファイル

どうやって統合するか？

ファイル名の法則性

1__HO__2__3.csv



1: 医科
2: 歯科
3: DPC
4: 調剤

レセプト識別子

フォルダ
8×8の分散処理系

ファイル名の法則性を使って

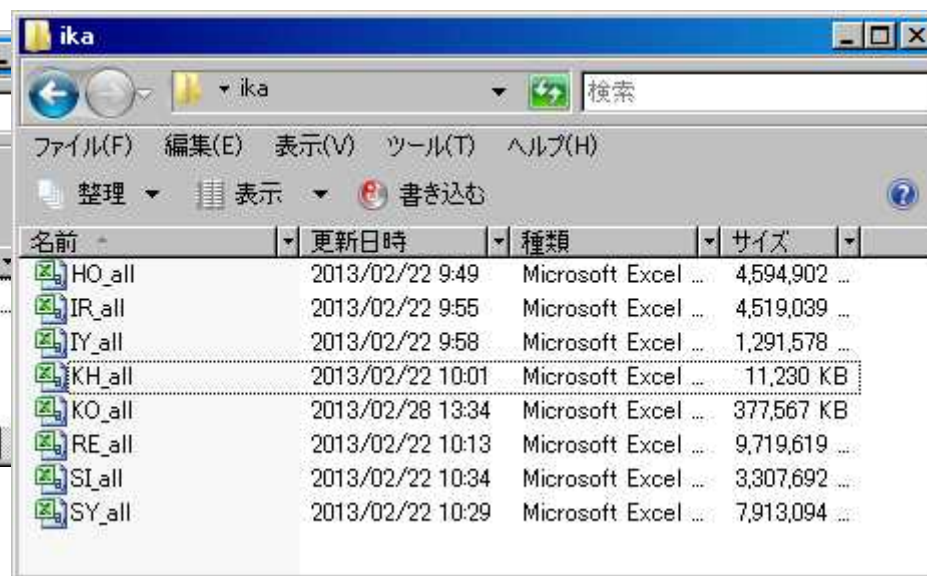
- Windowsのコマンドプロンプトを使用
 - ① 識別子毎にフォルダを作って、ファイルを移動
 - ② フォルダ毎に全部統合

①全データを識別子毎フォルダに分配

```
for /d %%d in (*) do (  
    cd g:¥datafiles¥%%d  
    for /d %%n in (*) do (  
        cd g:¥datafiles¥%%d¥%%n  
        for %%f in (*.csv) do (  
            set x=%%d  
            set y=%%f  
            set s=!y:~0,1!  
            set p=!x:~4,2!  
            set q=!y:~2,2!  
            copy %%f g:¥flat¥!s!¥!q!.!p!_%%f  
        )  
    )  
)
```

② 識別子フォルダ毎に統合

```
for %%i in (HO IR IY KH KO RE SI SY BU CD SI SK) do (  
copy g:\flat¥1¥%%i.* g:\assemble¥ika¥%%i_all.csv  
copy g:\flat¥2¥%%i.* g:\assemble¥dpc¥%%i_all.csv  
copy g:\flat¥4¥%%i.* g:\assemble¥rx¥%%i_all.csv  
);
```



MySQLに読み込み、解析開始

・時間がかかる！

診療行為(SI)に手術分類表を結合:21分24秒 ・・①

診療行為(SI)から放射線治療を抽出: 3分56秒・・②

調剤レセ(IY)に化学療法テーブルを結合:2時間30分

医薬品レセ(IY)に化学療法テーブルを結合:2時間37分

①と施設(IR)を結合 →朝～夕・終わらず

②とIRを結合 →朝～夕・終わらず

と、思ったら（最近発見・・・）

- MySQLでインデックス付与→大幅に処理速度Up
①の手術ファイルと施設(IR)を結合：数分
- 特に突合などの処理をする場合には、キーとなる変数にはインデックスを付けるのは必須

（参考） MySQLのコマンド

```
CREATE INDEX myindx ON tablename(colname);
```

教訓

「何事も基礎に忠実に」

- ・インデックスは付与した方が良い

(ただし、インデックスの付与そのものに結構時間がかかるので一定の覚悟は必要)

データについて知るには

- 診療報酬情報提供サービス(厚労省保健局運用)

<http://www.iryohoken.go.jp/shinryohoshu/>

- 各種マスター

医科診療行為マスター、医薬品マスターなど

ただし、順次更新され古いものが削除されるので注意
(検索はこのページから可能です)

- 記録条件仕様(医科、DPC、歯科、調剤)

「オンラインまたは光ディスク等による請求に係る記録条件仕様」などがわかりやすい

などが収載

最後に：実地監査

- 2人の監査者が来訪。3時間程度
 - インタビュー
 - 実際の現場の検証
 - パソコン環境のチェック

指摘事項

2. 一部要件を満たしていなかった事項

- ①アクセスログが事後的に内容変更が可能な状態。変更できないように設定を施すこと
- ②パスワードの複雑性に関する設定を施す

3. 対応が望まれる事項

- ①分析用IDとメンテ用IDを分離
- ②情報機器持ち込みルールの明確化
- ③外部接続ドライブのシステム上の無効化
- ④前室の入退室管理を整備すること

(祈) 有用な解析と活用を

- ご利用は計画的に
 - データの項目については密に事務局と協議するのが良いと思います。
 - データの範囲を絞るのは効率を確保するために重要
 - データベースの基礎に忠実に(インデックスなど)基礎知識のある人を巻き込むのが理想