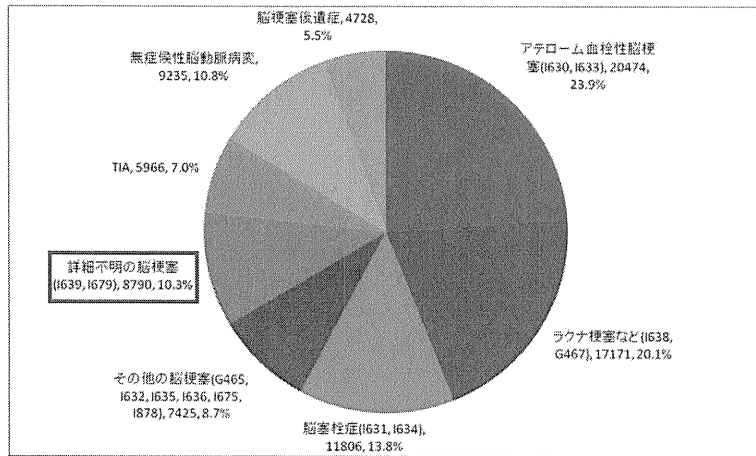


DPC研究班調査H22年7-12月における脳梗塞病型



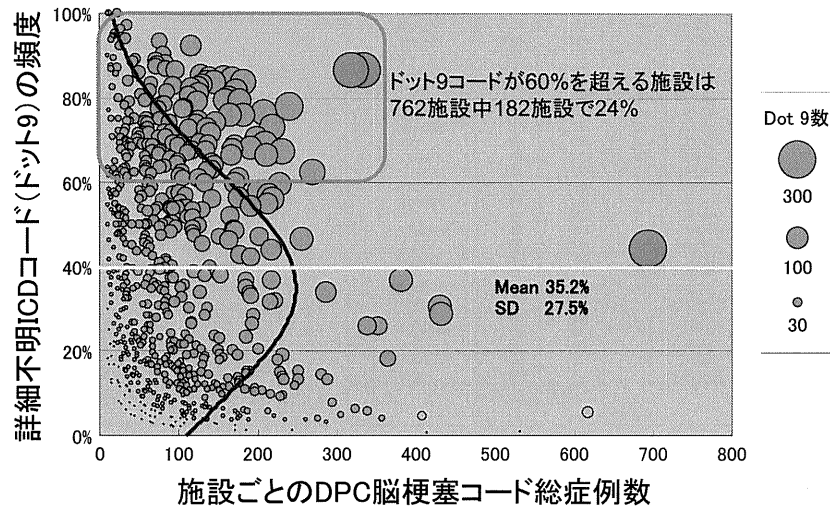
(DPC分類010060, N=85,595)

虚血性脳血管障害の主なICD-10 Coding

- G 45 一過性脳虚血発作および関連症候群
 - G 45.0 椎骨脳底動脈症候群
 - G 45.1 頸動脈症候群 (半球性)
 - G 45.2 多発性および両側性脳 (実質) 外動脈症候群
 - G 45.3 一過性黒内障
 - G 45.4 一過性全健忘
 - G 45.8 その他の一過性脳虚血発作関連症候群
 - G 45.9 一過性脳虚血発作, 詳細不明
- I 65 脳実質外動脈の閉塞・狭窄, 脳梗塞に至らなかったもの
 - I 65.0 椎骨動脈の閉塞・狭窄
 - I 65.1 脳底動脈の閉塞・狭窄
 - I 65.2 頸動脈の閉塞・狭窄
 - I 65.3 多発性および両側性の脳実質外動脈の閉塞・狭窄
 - I 65.8 その他の脳実質外動脈の閉塞・狭窄
 - I 65.9 詳細不明の脳実質外動脈の閉塞・狭窄
- I 63 脳梗塞
 - I 63.0 脳実質外動脈の血栓症による脳梗塞
 - I 63.1 脳実質外動脈の塞栓症による脳梗塞
 - I 63.2 脳実質外動脈の詳細不明の閉塞・狭窄による脳梗塞
 - I 63.3 脳動脈の血栓症による脳梗塞
 - I 63.4 脳動脈の塞栓症による脳梗塞
 - I 63.5 脳動脈の詳細不明の閉塞・狭窄による脳梗塞
 - I 63.6 脳静脈血栓症による脳梗塞, 非化膿性
 - I 63.8 その他の脳梗塞 (ラクナ梗塞など)
 - I 63.9 脳梗塞, 詳細不明
- I 64 脳卒中, 脳出血または脳梗塞と明示されないもの (疾病、傷害および死因統計分類提要)
- I 66 脳動脈の閉塞・狭窄, 脳梗塞に至らなかったもの
 - I 66.0 中大脳動脈の閉塞・狭窄
 - I 66.1 前大脳動脈の閉塞・狭窄
 - I 66.2 後大脳動脈の閉塞・狭窄
 - I 66.3 小脳動脈の閉塞・狭窄
 - I 66.4 多発性および両側性の脳動脈の閉塞・狭窄
 - I 66.8 その他の脳動脈の閉塞・狭窄
 - I 66.9 詳細不明の脳動脈の閉塞・狭窄
- I 69 脳血管疾患の続発・後遺症
 - I 69.0 クモ膜下出血の続発・後遺症
 - I 69.1 脳内出血の続発・後遺症
 - I 69.2 その他の非外傷性頭蓋内出血の続発・後遺症
 - I 69.3 脳梗塞の続発・後遺症
 - I 69.4 出血・梗塞と明示されない脳卒中の続発・後遺症
 - I 69.8 その他および詳細不明の脳血管疾患の続発・後遺症

236

施設別脳梗塞症例数と詳細不明コード頻度



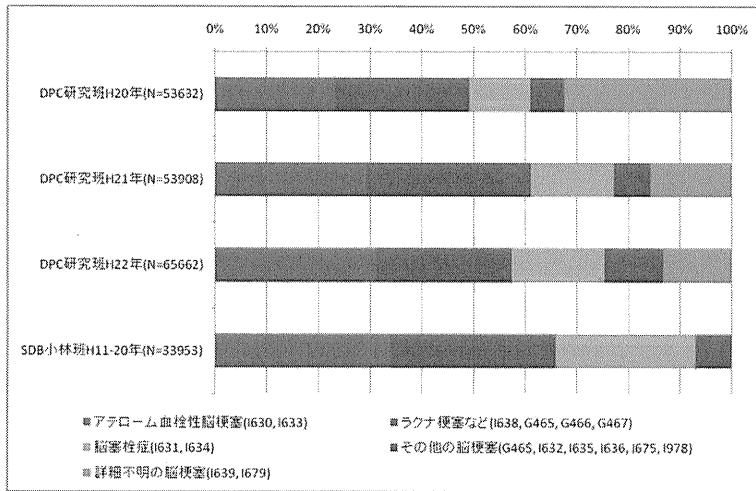
入江克実, 松田晋哉: 第□回日本脳卒中学会総会, □□□

DPC 機能評価係数 II (新たな機能評価係数)

項目	名称	評価の考え方	評価方法
1	○データ提出係数	対象病院における詳細な診療データの作成・提出に要する体制と、そのデータが活用されることで、医療全体の標準化や透明化等に貢献することを評価	【指数】 ①「データ提出の遅滞」については、翌々月に当該評価を50%・1ヶ月の間、減じる。 ②「部位不明・詳細不明のコード使用割合が40%以上」については当該評価を5%・1年の間、減じる。 (平成23年4月より評価) ※非必須項目の入力状況による評価は行わない。
2	○効率性係数	平均在院日数の変動に伴う病棟業務量の増減について、患者の疾病構造の違いを補正した在院日数の相対値により評価	【指数】 =〔全DPC対象病院の平均在院日数〕 ／〔当該医療機関の患者構成が、全DPC対象病院と同じと仮定した場合の平均在院日数〕 ※当該医療機関において、10症例(10か月)以上ある診断群分類のみを計算対象とする。 ※包括評価の対象となっている診断群分類のみを計算対象とする。
3	○複雑性係数	対象病院における診療の複雑さについて、当該病院における一入院当たり包括点数の相対値により評価	【指数】 =〔当該医療機関の包括点数(一入院当たり)を、診断群分類ごとに全病院の平均包括点数に置き換えた点数〕 ／〔全病院の平均一入院あたり包括点数〕 ※当該医療機関において、10症例(10か月)以上ある診断群分類のみを計算対象とする。 ※包括評価の対象となっている診断群分類のみを計算対象とする。
4	○カバー率係数	様々な疾患に対応できる総合的な体制について、当該病院で算定している診断群分類の広がり(種類の多さ)により評価	【指数】 =〔当該医療機関で一定症例数以上算定している診断群分類数〕 ／〔全診断群分類数〕 ※当該医療機関において、10症例(10か月)以上ある診断群分類のみを計算対象とする。 ※すべて(包括評価の対象・対象外の両方を含む)の診断群分類を計算対象とする。

※設定方法: [項目2-4]の指標から係数を設定する際には、各指標の平均±2SDを上限、下限とする。(出典: 中協22年1月29日)

DPC病型診断率の経年的改善



入江克実, 松田晋哉: 第□回日本脳卒中学会総会, □□□

脳梗塞の臨床病型を考慮したICD-10コーディング

臨床病型	ICD-10 (検索病名)
一過性脳虚血発作 G 45	
椎骨脳底動脈系TIA	G 45.0 椎骨脳底動脈症候群(椎骨動脈血行不全)
内頸動脈系TIA	G 45.1 頸動脈症候群(内頸動脈不全症)
その他のTIA	G 45.2~8
脳梗塞 I 63	
ラクナ梗塞	I 63.8 その他の脳梗塞(ラクナ梗塞など)
アテローム血栓性脳梗塞	
椎骨動脈・脳底動脈・頸動脈の血栓症	I 63.0 脳実質外動脈の血栓症による脳梗塞(脳外主幹動脈血栓症脳梗塞)
中大脳動脈・前大脳動脈・後大脳動脈・小脳動脈の血栓症	I 63.3 脳動脈の血栓症による脳梗塞(アテローム血栓性脳梗塞)
脳塞栓症	
椎骨動脈・脳底動脈・頸動脈の塞栓症	I 63.1 脳実質外動脈の塞栓症による脳梗塞(脳外主幹動脈塞栓症脳梗塞)
中大脳動脈・前大脳動脈・後大脳動脈・小脳動脈の塞栓症	I 63.4 脳動脈の塞栓症による脳梗塞(心原性脳塞栓症)
分類困難な脳梗塞	I 63.2, I 63.5
特定の病因による脳梗塞	I 63.8, I 67.0~8, I 68.0~8
無症候性脳動脈閉塞・狭窄 I 65, I 66	
椎骨動脈・脳底動脈・頸動脈の閉塞・狭窄	I 65.0~8 脳実質外動脈の閉塞・狭窄, 脳梗塞に至らなかったもの
中大脳動脈・前大脳動脈・後大脳動脈・小脳動脈の閉塞・狭窄	I 66.0~8 脳動脈の閉塞・狭窄, 脳梗塞に至らなかったもの
脳梗塞後遺症 I 69	I 69.3 脳梗塞の続発・後遺症

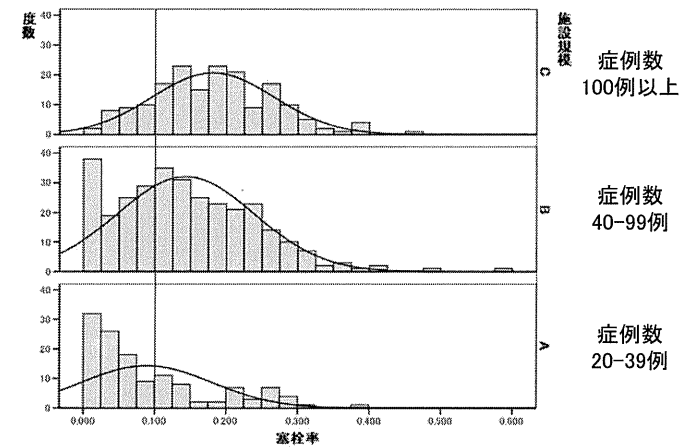
脳梗塞病名のICD-10変換における問題点

(疾病、傷害および死因統計分類提要)

(病名検索ソフト:病名くん2.0)

G 45 一過性脳虚血発作および関連症候群	
G 45.0 椎骨脳底動脈症候群	G 45.0⇔椎骨動脈血行不全、椎骨脳底動脈循環不全
G 45.1 頸動脈症候群(半球性)	G 45.1⇔内頸動脈不全症
G 45.9 一過性脳虚血発作, 詳細不明	G 45.9⇔一過性脳虚血発作
I 63 脳梗塞	
I 63.0 脳実質外動脈の血栓症による脳梗塞	I 63.0⇔脳外主幹動脈血栓症脳梗塞
I 63.1 脳実質外動脈の塞栓症による脳梗塞	I 63.1⇔脳外主幹動脈塞栓症脳梗塞
I 63.2 脳実質外動脈の詳細不明の閉塞・狭窄による脳梗塞	I 63.1⇔脳外主幹動脈閉塞脳梗塞
I 63.3 脳動脈の血栓症による脳梗塞	I 63.3⇔アテローム血栓性脳梗塞
I 63.4 脳動脈の塞栓症による脳梗塞	I 63.4⇔心原性脳塞栓症
I 63.5 脳動脈の詳細不明の閉塞・狭窄による脳梗塞	I 63.5⇔脳幹梗塞、小脳梗塞、穿通枝梗塞
I 63.8 その他の脳梗塞(ラクナ梗塞など)	I 63.8⇔ラクナ梗塞、無症候性脳梗塞、分水界梗塞
I 63.9 脳梗塞, 詳細不明	I 63.9⇔脳梗塞
I 65 脳実質外動脈の閉塞・狭窄, 脳梗塞に至らなかったもの	
I 65.0 椎骨動脈の閉塞・狭窄	
I 65.1 脳底動脈の閉塞・狭窄	
I 65.2 頸動脈の閉塞・狭窄	I 65.2⇔内頸動脈閉塞症、内頸動脈狭窄
I 66 脳動脈の閉塞・狭窄, 脳梗塞に至らなかったもの	
I 66.0 中大脳動脈の閉塞・狭窄	I 66.0⇔中大脳動脈閉塞症、中大脳動脈狭窄
I 66.1 前大脳動脈の閉塞・狭窄	
I 66.2 後大脳動脈の閉塞・狭窄	

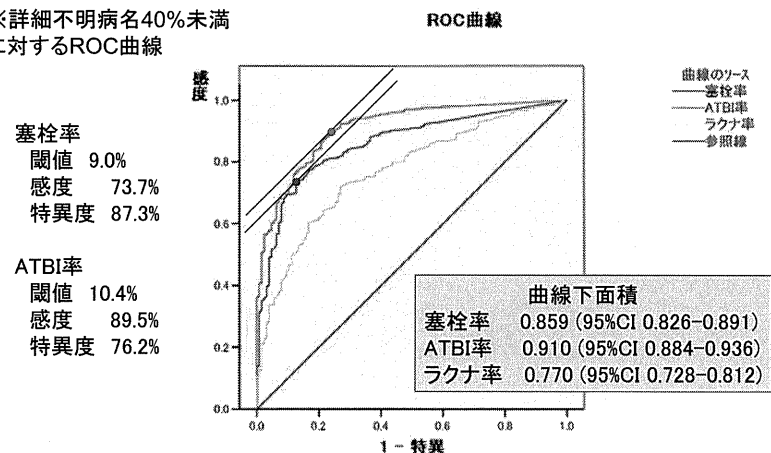
施設規模別にみたDPC脳塞栓症診断率



入江克実, 松田晋哉: 第36回日本脳卒中学会総会, 2011

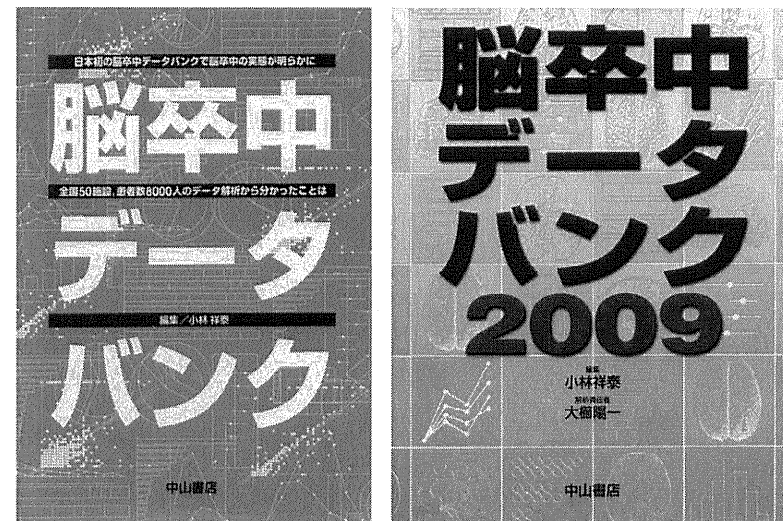
病型診断率による詳細不明病名高頻度施設の判別

※詳細不明病名40%未満
に対するROC曲線

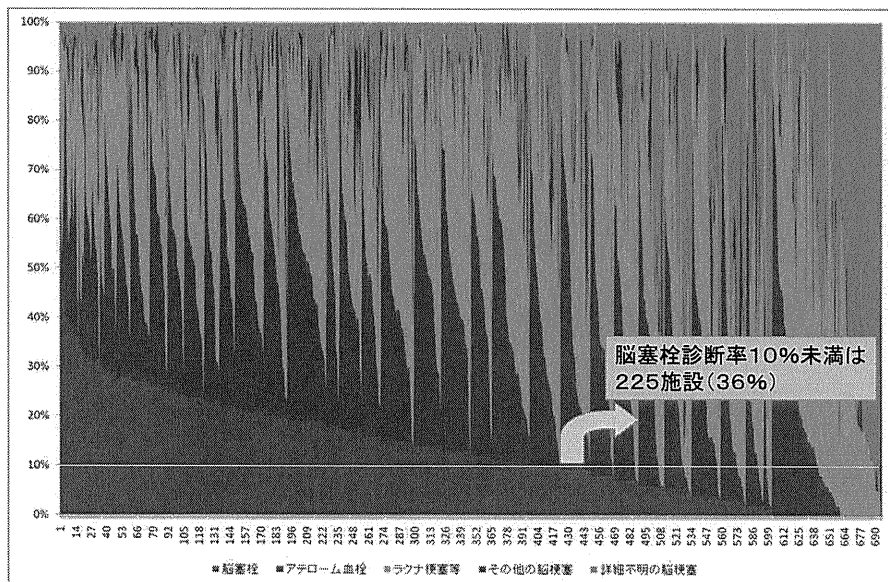


対角セグメントは同一値により生成されます。

入江克実, 松田晋哉: 第36回日本脳卒中学会総会, 2011



脳塞栓診断率でみた施設ごとの脳梗塞分類頻度



DPCデータの強みと弱み

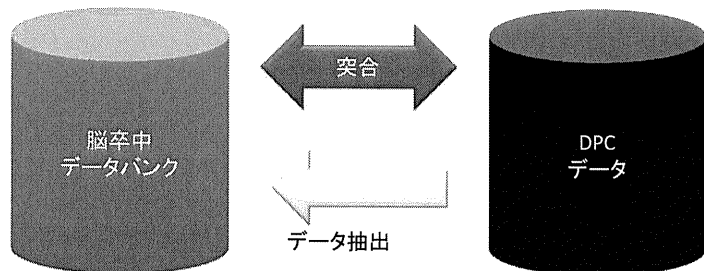
DPCデータが持たないもの

- 発症時間、NIHSSなど詳細な臨床情報
- 行為・薬剤の実施時刻
- 画像所見、検査所見、検査値
- 入院途中の進行・再発によるADLの経時的推移
- 退院後の長期的予後

DPCデータの強み

- データ抽出が容易で漏れがない(開始日、延べ日数・数量)
- 全ての薬剤、行為を抽出できる
- 転科前後のデータもある
- レセプトレベルのコスト分析ができる

DPCデータと脳卒中データバンクの連携



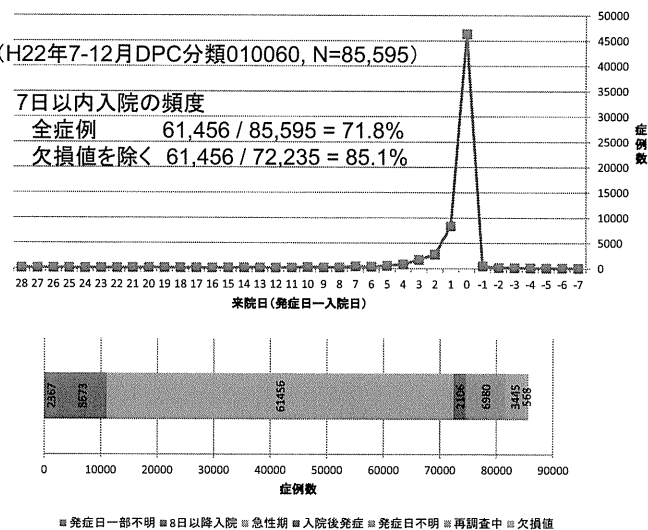
脳梗塞急性期入院(発症7日以内)は約72%

(H22年7-12月DPC分類010060, N=85,595)

7日以内入院の頻度

全症例 61,456 / 85,595 = 71.8%

欠損値を除く 61,456 / 72,235 = 85.1%



心筋梗塞データベース・脳卒中データベースを用いた
救急搬送情報と診療情報の連結のための研究(H23-心筋一般-001)
小林班H23年度第1回班会議 2011.6.18.

H22年度DPC様式1改定による脳梗塞急性期評価 — Stroke Data Bankとの比較 —

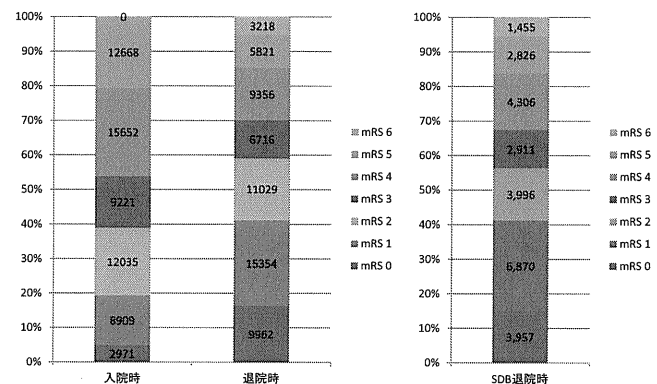
入江克実¹⁾, 林田みどり¹⁾, 桑原一彰²⁾,
松田晋哉³⁾, 伏見清秀⁴⁾

- 1) 社会医療法人白十字会白十字病院
- 2) 九州大学大学院医学研究院医療経営・管理学講座
- 3) 産業医科大学公衆衛生学
- 4) 東京医科歯科大学大学院医療政策学

急性期脳梗塞におけるADLの改善

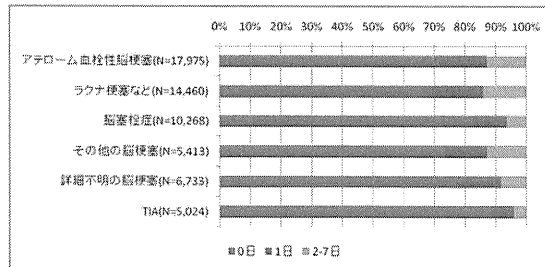
DPC研究班H22年7-12月
(N=61,456)

Stroke Data Bank H11-H20年
(N=26,299)

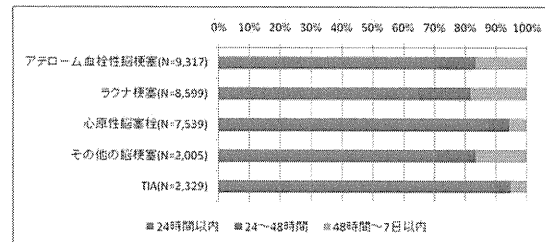


脳梗塞急性期の約75%は24時間以内に入院

DPC研究班
H22年7-12月

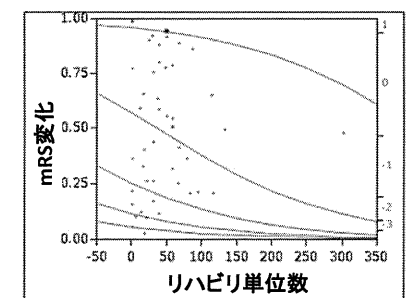
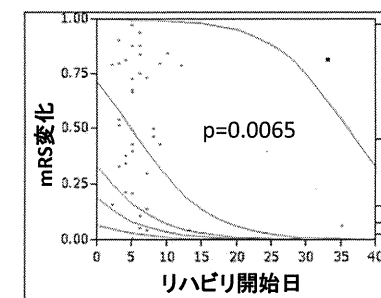
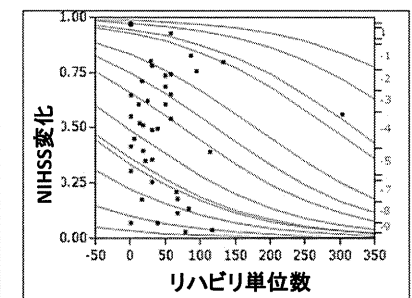
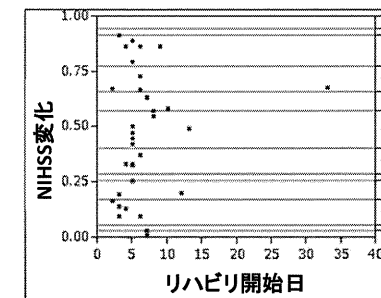
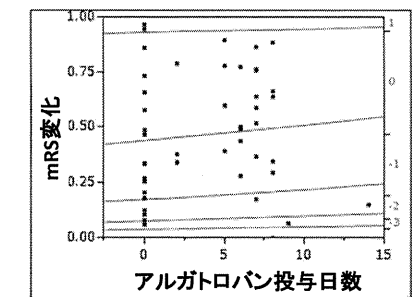
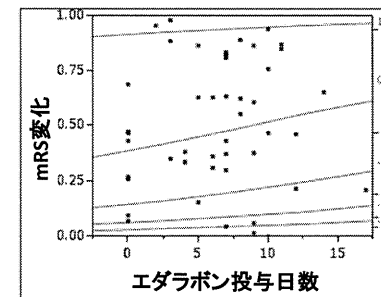
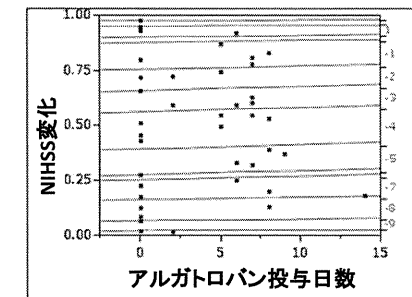
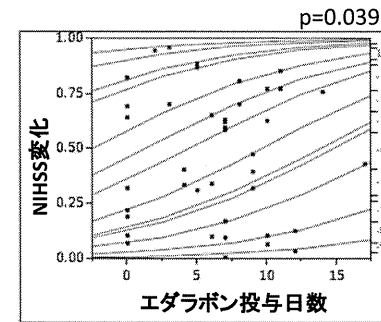


Stroke Data Bank
H11-H20年



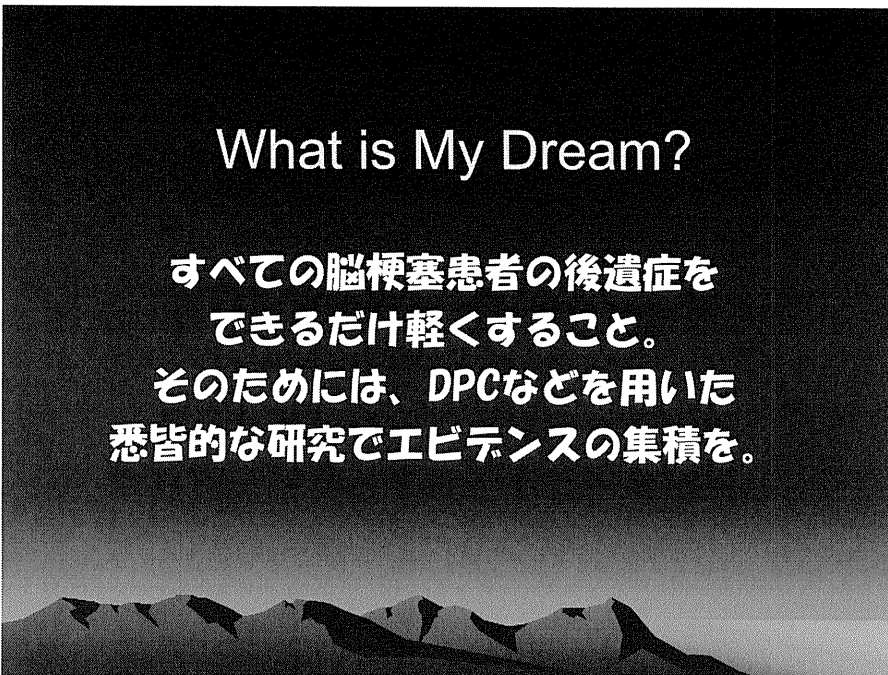
島根大におけるDPCデータとStroke Databankの突合

- 対象: H20および21年度脳梗塞100例中入院時NIHSSが5~22であった47例
- 突合の注意点:
 - 院内発症では入院日と異なる(DPCは自院の入院日)
 - 転科した場合に治療情報と行為情報の不一致が発生する(転科先でのリハビリ状況や退院時スコア評価項目など)
- 評価項目:
 - エダラボン・アルガトロバンの投与日数
 - 入院からのリハ開始日数、リハ単位数
 - NIHSSおよびmRSの変化



What is My Dream?

すべての脳梗塞患者の後遺症を
できるだけ軽くすること。
そのためには、DPCなどを用いた
悉皆的な研究でエビデンスの集積を。



平成22年度[様式1]改定で追加された項目

【脳卒中の発症時期】

- (入力ようになった理由) 発症直後の急性期診療の入院かその後の入院かを区分する必要があるため¹

医療資源を最も投入した傷病名が010020(くも膜下出血、破裂脳動脈瘤)~010070(脳血管障害(その他))に該当する場合は入力する。

YYYY(西暦4桁)MM(月2桁)DD(日2桁)の値(計8桁)を入力する。例 2010年7月10日 → 20100710
発症時期が明らかでない場合について、臨床所見及び病歴等から最も妥当な推定発症日を入力するものとする。なお、過去の情報等を参考にしても不明な場合は“00000000”を入力する。一部不明の場合は、不明な部分のみ“0”とする(例:1998年5月に発症したことがわかっている場合は“19980500”と入力する)。²

【入院時および退院時 modified Rankin Scale (mRS)】

- (入力ようになった理由) 脳神経疾患の診療において国際的にもよく利用されている。比較的簡便であり、妥当性・信頼性も実証されている。この指標を用いることにより、脳神経疾患の機能分類別の治療内容・早期アウトカムや医療費が分析可能となる¹

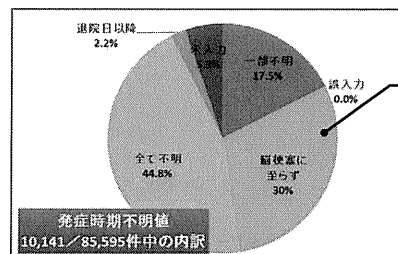
入院の契機となった傷病名がMDC01(神経系疾患)に該当する場合に入力する。
入院した時点での“0”~“6”までの値を入力する。不明の場合は“9”を入力する。²

¹ 平成22年度 第1回 診療報酬調査専門組織・DPC評価分科会(H22.4.22) 資料(D-2-1)「平成22年度調査に向けた様式1の見直し(案)」についてより

² 平成22年度「DPC導入の影響評価に係る調査」実施説明資料より ※黒文字・斜体文字部分は、平成23年度様式1改定で変更された

脳卒中の発症時期の入力状況について ~①入力不明値~

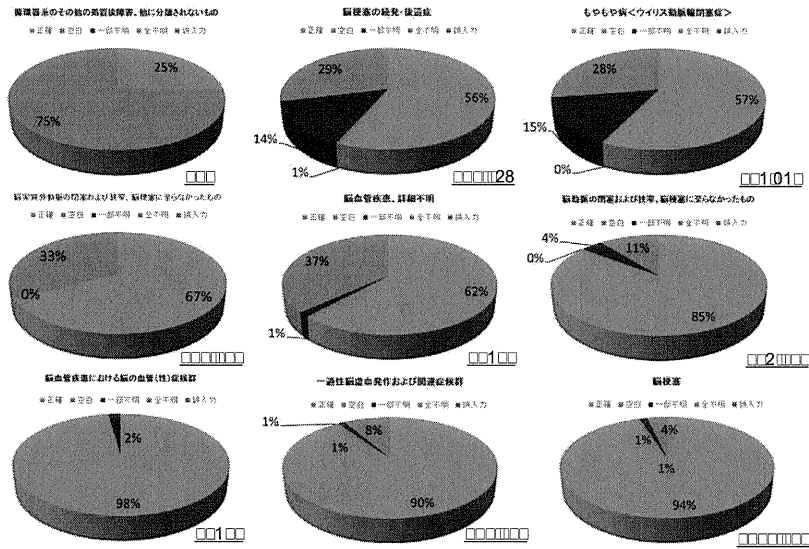
- ・ 脳卒中の発症時期が...
 - 年、月、日の全てが不明のデータ
入力例) 00000000 もしくは、0
 - 一部不明のデータ
入力例) 0000/01/01 や 2010/10/00 等
 - 退院日以降のデータ
入力例) 退院日:2010/11/26
発症日:2022/11/02
 - 年であれば「0000」、月日であれば「00」以外の日付と認識できない数値データ
入力例) 2010/07/77



ICDのうち『165\$ 脳実質外動脈の閉塞および狭窄、脳梗塞に至らなかったもの』、『166\$ 脳動脈の閉塞および狭窄、脳梗塞に至らなかったもの』のうち、入力値が、「全て不明」もしくは、「一部不明」は「脳梗塞に至らず」に分類

※生年月日より前の日付データはなかった。

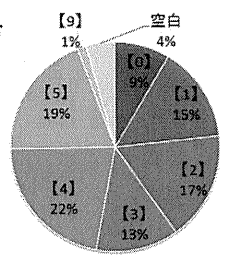
脳卒中の発症時期の入力状況について ~②別々の不明値の割合~



入院時および退院時mRsの入力状況について

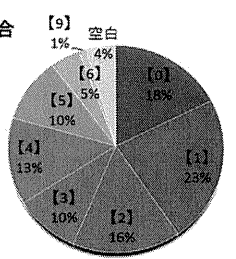
入院時mRs判定別割合

入院時mRs	件数
[0]	7,312
[1]	12,639
[2]	14,427
[3]	10,984
[4]	18,629
[5]	16,678
[9]	1,163
空白	3,763
総数	85,595



退院時mRs判定別割合

退院時mRs	件数
[0]	15,189
[1]	19,538
[2]	13,501
[3]	8,215
[4]	11,513
[5]	8,812
[6]	3,960
[9]	969
空白	3,898
総数	85,595

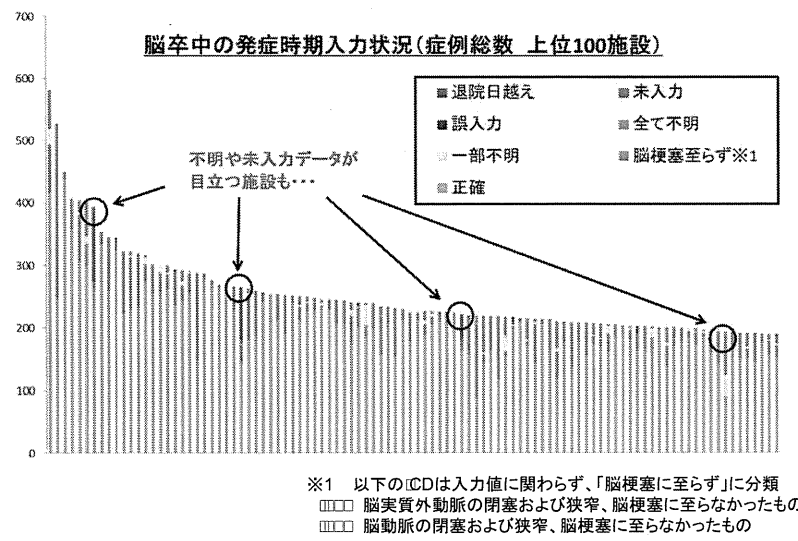


◆不明値について...

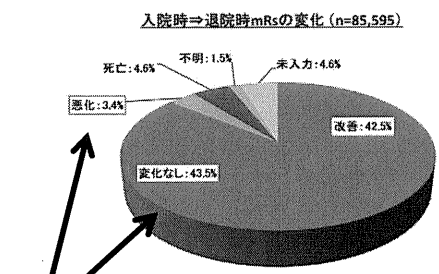
- 指定された数値以外が入力されたデータは、0件であった。
- 退院時転帰が、【□=最も医療資源を投入した傷病による死亡】□R【□=最も医療資源を投入した傷病以外の死亡】であるにも関わらず、退院時□R□が【□死亡】以外の判定値が入力されていたデータは、8□□□件中210件あった。
- 逆に、退院時□R□が【□死亡】であるにも関わらず、【□=最も医療資源を投入した傷病による死亡】□R【□=最も医療資源を投入した傷病以外による死亡】以外の退院時転帰が入力されていたデータは、8□□□件中2件あった。

242

脳卒中の発症時期の入力状況について ~③施設別不明値の割合~



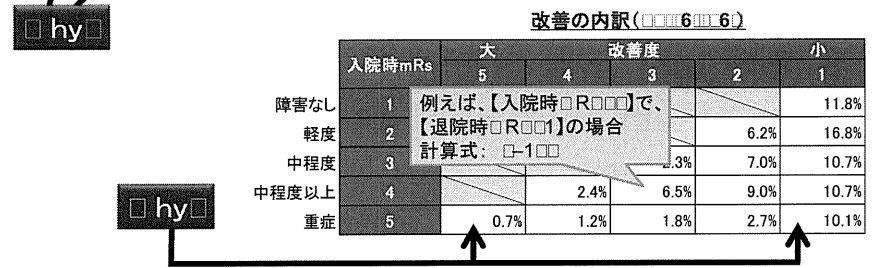
分析対象 ~まずは何を明らかにするのか?~



①様式1を□□□□で開いて、以下の条件で分類してグラフ化してみる

種別	定義
改善	(入院時mRs - 退院時mRs) > 0
変化なし	(入院時mRs - 退院時mRs) = 0
悪化	(入院時mRs - 退院時mRs) < 0
死亡	退院時mRs = 6
不明	入院時mRs = 9 AND/OR 退院時mRs = 9
未入力	入院時mRs = 空白 AND/OR 退院時mRs = 空白

②どのくらい改善しているのかを計算してみる



分析の準備 ~分析の資料の一つとしてガイドラインを参考にする~

脳卒中治療ガイドライン 2009



<http://www.jsts.gr.jp/jss08.html>

1. 脳卒中リハビリテーションの進め方

1-4. 急性期リハビリテーション

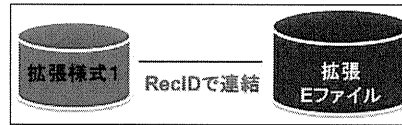
推奨

1. 療養症候群を予防し、早期のADL向上と社会復帰を図るために、十分なリスク管理のもとにできるだけ発症後早期から積極的なリハビリテーションを行うことが強く勧められる(グレードA)。その内容には、早期座位・立位、嚥食を用いた早期歩行訓練、摂食・嚥下訓練、セルフケア訓練などが含まれる。
2. 脳卒中ユニット、脳卒中リハビリテーションユニットなどの増強化された場で、リハビリテーションチームによる集中的なリハビリテーションを行い、早期の退院に向けた積極的な指導を行うことが強く勧められる(グレードA)。

分析の目標...

医療資源を最も投入した傷病名が脳梗塞(DPC6桁:01000)で、かつ入院日が脳卒中の発症時期から一週間以内の患者を入院時Rごとに分類し、診療プロセス(今回は、リハビリ)を可視化し、そのアウトカムへの影響を調査する。
(※赤線...様式1からデータを取得、青太文字...Eファイルからデータを取得)

分析の準備 ~診療明細クエリを作成する~



レセ電コード	内容
180027610	脳血管疾患等リハビリテーション料(Ⅰ)
180027710	脳血管疾患等リハビリテーション料(Ⅱ)
180030810	脳血管疾患等リハビリテーション料(Ⅲ)

①DPC6桁=010060 (脳梗塞) で抽出条件を設定

②経過日数 >=0 and <=7 で抽出条件を設定

③抽出条件を設定

④入院から何日目に行為が実施されたかを計算: 相対日=実施日-入院日

⑤在院日数を計算: 在院日数=(退院日-入院日)+1

分析の準備 ~Accessを使用して、分析用のデータを作成する~



拡張様式1をさらに加工する(Accessを使用した場合)

- ①脳卒中中の発症時期(以下、発症日)、入院日、退院日を日付型に変更する
日付型に変更できないデータは分析対象からIsDate関数を使用して除外する。*1
・Mid([日付],1,4) & '-' & Mid([日付],5,2) & '-' & Mid([日付],7,2)
・IsDate(Mid([日付],1,4) & '-' & Mid([日付],5,2) & '-' & Mid([日付],7,2))=True
- ②発症日と入院日から【発症から入院までの経過日数】を計算する
・DateDiff('d',CDate(発症日),CDate(入院日))

*1 不明値が多い場合...以下の処理を加えるとn数を増やすことができる
脳卒中中の発症時期の日のみ不明なデータの00を01に置き換える
・Mid([発症日],1,4) & '-' & Mid([発症日],5,2) & '-' & If(Mid([発症日],7,2)='00','01',Mid([発症日],7,2))

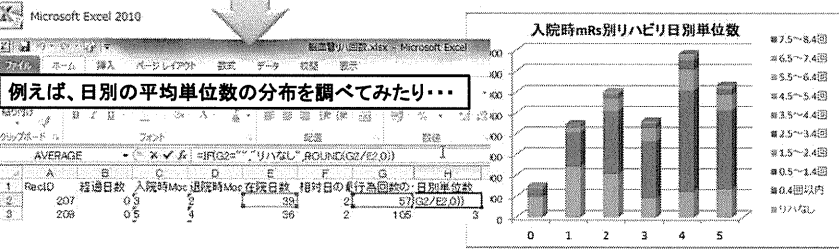
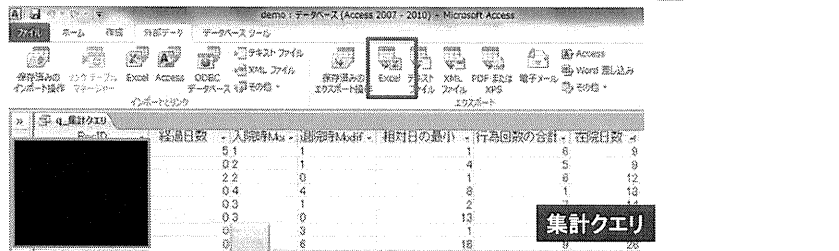
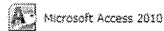
分析の準備 ~診療明細クエリからデータを集計する~

『診療明細クエリ』から『集計クエリ』を作成

集計の設定をそれぞれ、以下のよう
に設定する。
・相対日は「最小」
(リハビリの開始日を取得するため)
・行為回数は「合計」
(入院期間中の合計単位数を取得
するため)

分析 集計クエリ結果をE□CE□でグラフにする①

『集計クエリ』を□□□□にエクスポートして、見やすい形で表現する



参考資料

- 今回、説明した分析方法は特に新しい手法ではなく、これまでの開催されたセミナーやDPC研究班の先生方の書籍で説明されている手法です。



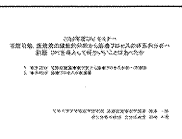
藤森先生の本・・・
 「DPCデータ分析 アクセス・SQL活用編」
 拡張様式1、EFファイルの作成方法、それを使用した分析の解説があります。
 また、付属のCD-ROMには、今回、取り上げた脳梗塞リハビリ状況のサンプルSQLもあるので、とても参考になる一冊です。

- Q310_在院日数分析.sql
- Q320_術前後後日数.sql
- Q410_療養室退院データ.sql
- Q420_EFファイル区分集計.sql
- Q430_EFファイル区分別集計.sql
- Q440_EFファイル区分集計.sql
- Q450_EFファイル区分集計.sql
- Q460_脳梗塞リハビリ状況.sql
- 拡張EFファイル集計.sql
- 拡張EFファイル集計.sql
- 拡張EFファイル集計.sql
- 拡張EFファイル集計.sql

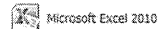


病院に届いているはず・・・
 「平成22年度 総括・分担研究報告書」
 先生方の論文以外にも昨年のセミナーの資料や各種分析の説明が載っています。
 冊子は白黒ですが、付属のDVDにカラーのPDFがあります。分析のヒントがあるはず。

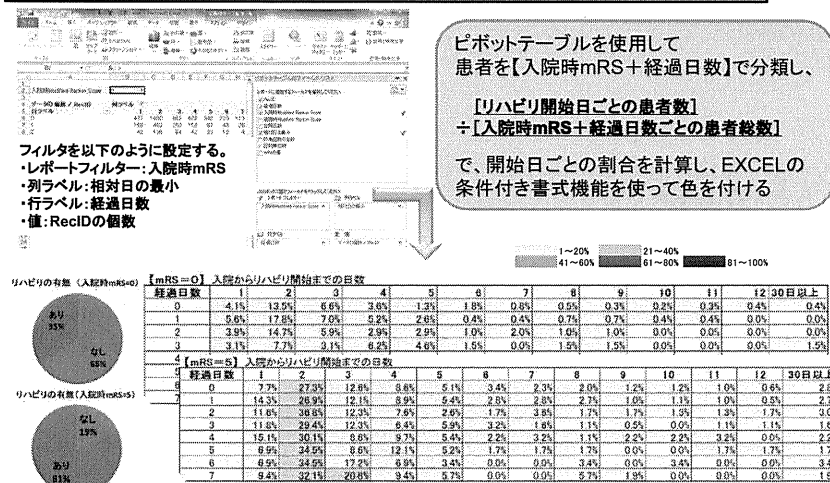
P.321～
 診療プロセス分析演習



分析 集計クエリ結果をEXCELでグラフにする②



例えば、入院時mrs・経過日数別のリハビリの開始日の分布を調べてみる・・・



D・E・Fファイル、損益計算書を活用した 病院活動基準原価計算の試み

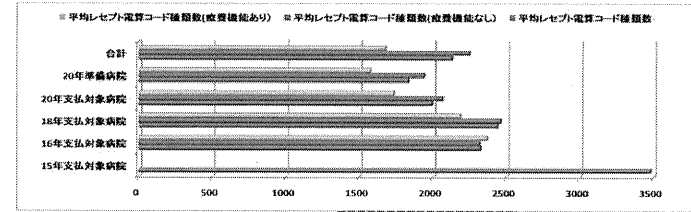
診療活動量測定
診断群分類別、経過日別、費用・収益分析

九州大学大学院医学研究院 医療経営・管理学専攻
桑原 一彰
済生会熊本病院
田崎 年晃

ファイルの説明

- PDFファイル 本日のプレゼン資料
 - プレゼンパワーポイントは配布資料を若干の変更しています
- アクセスファイル(使用します)
 - デモ演習用
- テキストファイル(使用します)
 - クエリ文章、クエリの名前
- エクセル(使用します)
 - 費用集計、利益の可視化、収益費用の関係可視化(散布図化、回帰分析)
 - 途中脱落された方が、再参加できるようにしています
- 注意 今回のファイルは基礎的作業を省略しています
 - 勘定科目とレセプト電算コードの対応関係(ここがポイント)
 - 病院独自の考えがあって良い
 - 入院相対日の計算 = [実施日] - [入院日]
 - 入院外来合わせたFファイルの件数の集計
 - Dファイル円・点区分の整理

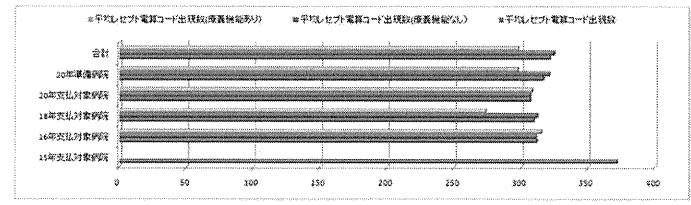
レセプト電算コードの質 レセプト電算コード施設別出現種類数 平成19年度DPC調査965施設 DPC参加年度別



DPC参加年度	平均レセプト電算コード種類数	平均レセプト電算コード種類数範囲
15年支払対象病院	3479	[2180 - 4371]
16年支払対象病院	2823	[1144 - 4000]
18年支払対象病院	2439	[831 - 3513]
20年支払対象病院	1995	[306 - 3399]
20年準備病院	1833	[294 - 3852]
合計	2131	

DPC参加年度	平均レセプト電算コード種類数(原費集計あり)	平均レセプト電算コード種類数(原費集計なし)
15年支払対象病院	2369	[1144 - 3415]
16年支払対象病院	2185	[863 - 3083]
18年支払対象病院	1733	[565 - 3031]
20年支払対象病院	1573	[516 - 3442]
20年準備病院	1678	
合計	1678	

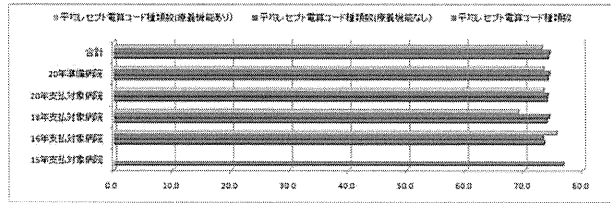
レセプト電算コードの質 レセプト電算コード患者一人当たり平均出現件数 平成19年度DPC調査965施設 DPC参加年度別



DPC参加年度	レセプト電算コード合計数	患者数	平均レセプト電算コード出現数	平均レセプト電算コード出現数範囲
15年支払対象病院	160973538	492772	327	[235 - 655]
16年支払対象病院	89308868	191962	312	[143 - 505]
18年支払対象病院	205740024	654273	316	[123 - 501]
20年支払対象病院	151717713	495825	307	[113 - 529]
20年準備病院	295866692	922142	317	[102 - 677]
合計	873575625	2714974	322	

DPC参加年度	レセプト電算コード出現数(原費集計あり)	レセプト電算コード出現数(原費集計なし)	DPC参加年度	レセプト電算コード合計数	患者数	平均レセプト電算コード出現数(原費集計あり)	平均レセプト電算コード出現数(原費集計なし)
15年支払対象病院	160973538	427772	15年支払対象病院	703880	22297	315	[439 - 259]
16年支払対象病院	89308868	16565	16年支払対象病院	221650	4485	274	[318 - 123]
18年支払対象病院	19650424	81968	18年支払対象病院	2202474	73774	309	[484 - 149]
20年支払対象病院	13865239	42850	20年準備病院	5219847	170652	299	[677 - 151]
20年準備病院	24223345	75160	合計	9421592	315452	293	
合計	77514843	209408					

レセプト電算コードの質 レセプト電算コード1患者当たり平均出現種類数 平成19年度DPC調査965施設 DPC参加年度別



DPC参加年度	平均レセプト電算コード種類数	平均レセプト電算コード種類数(範囲)
15年度参加病院	70.6	(2 - 501)
16年度参加病院	70.4	(1 - 421)
18年度参加病院	70.8	(2 - 458)
20年度参加病院	70.8	(1 - 427)
20年度参加病院	74.0	(1 - 430)
合計	74.1	

DPC参加年度	平均レセプト電算コード種類数(療養機能別)	平均レセプト電算コード種類数(療養機能別) (範囲)
15年度参加病院	70.2	(1 - 414)
16年度参加病院	74.2	(2 - 458)
18年度参加病院	74.0	(1 - 427)
20年度参加病院	74.0	(1 - 430)
合計	74.4	

DPC参加年度	平均レセプト電算コード種類数(療養機能別)	レセプト電算コード種類数(療養機能別) (範囲)
15年度参加病院	70.6	(2 - 501)
16年度参加病院	70.4	(1 - 421)
18年度参加病院	70.8	(2 - 458)
20年度参加病院	70.8	(1 - 427)
20年度参加病院	74.0	(1 - 430)
合計	74.1	

導入

病院事業の損益を明確化、事業計画に活用させるツール
準拠すべき統一された方法は存在しない

「病院原価計算要綱」 (昭和29年日本病院会)

- 目的
- 原価管理
 - 価格決定
 - 損益管理
 - 予算編成
 - 基本計画策定

診断群分類(DPC)を活用した 診療報酬包括支払制度

- 原価適減を図るインセンティブ
 - 診療報酬に原価を反映させる必要性
- 平成15年度 中央社会保険医療協議会
コスト調査分科会 医療機関のコスト調査

議論 サービス別原価(=分類別原価) > 入院基本料 のはず

原価管理目的の意識向上
価格決定目的達成のための、DPC別原価計算の必要性

病院原価計算の問題点

- ▶ 標準化された間接費配賦方法がない
- ▶ 費用収益把握を内部管理に適時活用出来ていない
- ▶ 複数病院間の経営ベンチマークが出来ない

原価計算とDPCデータ



先行研究 林田(2004):DPC別患者別原価計算の研究

- ▶ DPC毎のコストと診療報酬の相関関係を分析
- ▶ 在院日数・診療報酬点数比等を用いて配賦
- ▶ 診療活動量を考慮した配賦方法の必要性を唱えている

病院内部管理への活用可能性、科学的根拠に基づいた価格決定

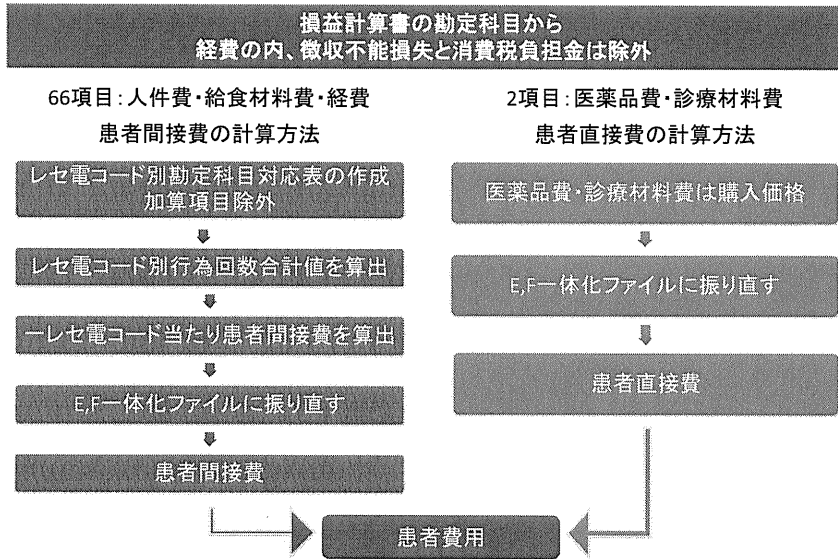
演習のねらい

- ▶ 全国共通形式のDPC調査データ(D,E,Fファイル)を活用
- ▶ レセプト電算処理システムコード(レセ電コード)を診療活動量単位
- ▶ 診療活動量を考慮した患者別原価計算手法の開発
- ▶ 主要診断群(MDC: Major Diagnostic Category)別手術処置有無別の費用収益関係を分析

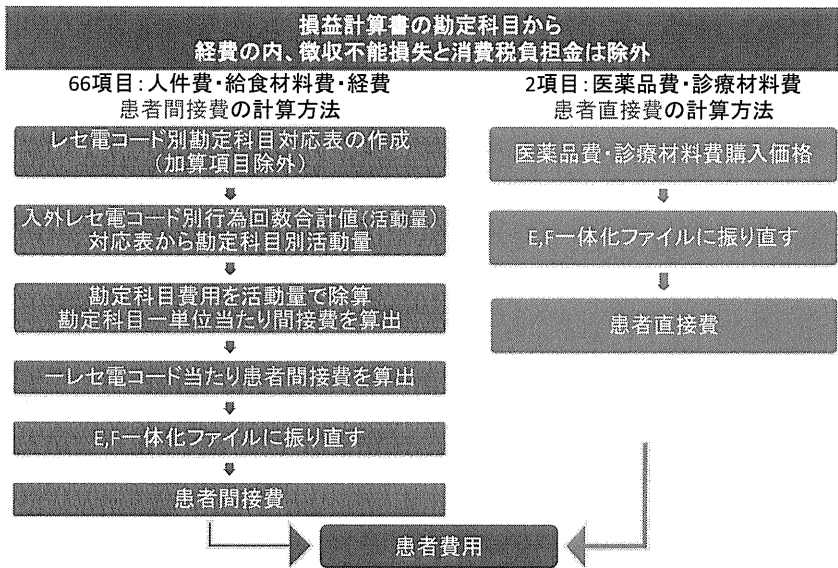
何を素材とするか

患者	対象期間に入退院した患者	
データ	診療録情報(様式1)	患者基本属性
	E,F一体化ファイル	レセ電コード、実施年月日 使用量、行為回数 診療活動量の測定
	包括診療明細情報 (Dファイル)	出来高支払収益 包括支払収益
	損益計算書勘定科目	勘定科目別費用金額
	診療行為・医薬品・器材マスタ	レセ電コード別勘定科目対応表

費用集計の流れ



費用集計の流れ



流れ1 勘定科目別活動量合計を算出

患者間接費配賦方法



流れ2 レセプト電算コード別間接費の算出



流れ3 患者別入院相対日別直接費用の集計

5. レセ電コード別患者間接費

レセ電コード	レセ電コード別患者間接費(円)
111000110	3,000
111011810	3,200
112006470	2,000
112006570	2,000
112006670	2,000
112006770	3,000
112011310	2,000
112011710	2,500
113000310	2,000
113000410	3,000
113000510	3,000
113001010	

6. E,F一体化ファイル

患者ID	経過日	レセ電コード	行為回数	使用量
1234567890	1	111000110	1	1
1234567890	2	111011810	1	1
1234567890	3	112006470	1	1
1234567890	4	112006570	1	1
1234567891	1	112006670	1	1
1234567891	2	112006770	1	1
1234567891	3	112011310	1	1
1234567890	1	610406001	1	10
1234567890	2	610406002	1	20
1234567890	3	610406003	1	5
1234567890	4	610406004	1	10
1234567891	1	610406005	1	7
1234567891	2	610406006	1	10
1234567891	3	610406008	1	10



患者ID	経過日	患者間接費	患者直接費	出来高収益	包括収益
1234567890	1	11100	3,000	1,900	1,000
1234567890	2	11100	3,200	2,200	1,400
1234567890	3	11200	2,000	7,000	4,500
1234567890	4	11200	2,000	2,000	2,000
1234567891	1	11200	2,000	10	1,500

患者ID	在院日数	患者間接費	患者直接費	収益	診断群分類番号	年齢
1234567890	4	11200	10,200	2,200	100003X01	100
1234567891	1	11200	7,000	169	100003X09	100
1234567890	1	11200	3,000	40		
1234567891	1	11200	2,000	10		

患者ID	在院日数	610406001	610406002	610406003	610406004	610406005	610406006	610406008
1234567890	4	1	10					1,900
1234567890	1		20					100
1234567890	1			5				50
1234567890	4				10			150
1234567891	1					7		119
1234567891	2						10	40
1234567891	3							10

在庫物流管理、電子カルテの応用 後述

248

演習用Accessテーブルの説明

- T0_レセプト電算コードと勘定科目との対応表_1
- T1_年間レセプト電算コード入外総件数
 - 考え方
 - 活動量をレセプト電算コードで測定
 - 加算は省く
- T2_損益計算書費用_間接費用
 - 労務費、材料費(医薬品、材料費は除く)、経費(消費税負担金は除く)
- T3_レセプト電算コード別材料平均価格_直接費用
- T3_医薬品平均購入価格_直接費用
- T4_EFファイル
 - 生産(種々の疾患の患者の治療)における活動をレセプト電算コードで明確に区別し、生産にどれだけの活動を投入(レセプト電算コードの件数合計)したのかを識別する
 - 医薬品、材料のレセプト電算コード=患者に直課+Comedicalの活動量としても収集
- T4_Dファイル
 - 収益と分類番号
 - 円・点を統一

Step1

Step1_Q_勘定科目別レセプト電算コード行為回数合計

解らなかったら、テキストファイルのクエリ文を貼り付けてね

Step2

Step2_Q_勘定科目別一単位当たり間接費用

解らなかったら、テキストファイルのクエリ文を貼り付けてね

Step3

Step3_Q_レセプト電算コード別間接費用合計

レセプト電算コード別間接費用合計

レセプト電算コード	間接費用合計
X0000	814845627250374
X0001	814845627250374
X0002	814845627250374
X0003	814845627250374
X0004	814845627250374
X0005	814845627250374
X0006	814845627250374
X0007	814845627250374
X0008	814845627250374
X0009	814845627250374
X0010	814845627250374
X0011	814845627250374
X0012	814845627250374
X0013	814845627250374
X0014	814845627250374
X0015	814845627250374
X0016	814845627250374
X0017	814845627250374
X0018	814845627250374
X0019	814845627250374
X0020	814845627250374
X0021	814845627250374
X0022	814845627250374
X0023	814845627250374
X0024	814845627250374
X0025	814845627250374
X0026	814845627250374
X0027	814845627250374
X0028	814845627250374
X0029	814845627250374
X0030	814845627250374
X0031	814845627250374

解らなかったら、テキストファイルのクエリ文を貼り付けてね

Step5

Step5_Q_患者別入院相対日別_医薬品直接費用

患者別入院相対日別_医薬品直接費用

RecID	入院経過日	医薬品費用
阿南	2	3191
阿南	3	559
阿南	4	6
阿南	5	509
阿南	6	54
森原	1	3802.4
今中	1	206
今中	2	17549
今中	3	1526
今中	4	802
今中	5	1471
今中	6	512
今中	7	512
今中	8	231
小林	1	3580.6
小林	2	81388.1
小林	3	9753.72
小林	4	1662.4
小林	5	1507.9
小林	6	372
小林	7	118
小林	8	118
小林	10	4547.4
小林	11	3892
小林	13	55
小林	18	3570
松田	2	61
松田	3	6880
松田	4	8280
石川	1	3786.6
池田	1	206
田崎	2	16376

解らなかったら、テキストファイルのクエリ文を貼り付けてね

Step4

Step4_Q_患者別入院相対日別_間接費用

患者別入院相対日別_間接費用

RecID	入院経過日	間接費用合計
阿南	1	2179.939290565656
阿南	2	3727.82451835766
阿南	3	4864.52768703564
阿南	4	7801.94950533506
阿南	5	10854.69223590827
阿南	6	4540.22062235121
森原	1	32853.195182567
森原	2	12239.693966368
森原	3	4540.22062235121
今中	1	2699.90938635984
今中	2	23861.6733449541
今中	3	30877.5846121766
今中	4	7301.75805167841
今中	5	25988.92270974946
今中	6	10504.6819414743
今中	7	8387.98341675872
今中	8	8125.00136306008
今中	9	20898.5031254882
今中	10	4786.91315606873
今中	11	4786.91315606873
今中	12	4540.22062235121
小林	1	10365.7516306769
小林	2	95650.4652779321
小林	3	67571.7227965579
小林	4	38344.9770975775
小林	5	24121.2057636727
小林	6	33611.4039862109
小林	7	10835.0852551492
小林	8	15726.8823783855
小林	9	26895.4362716374

解らなかったら、テキストファイルのクエリ文を貼り付けてね

Step6

Step6_Q_患者別入院相対日別_材料直接費用

患者別入院相対日別_材料直接費用

RecID	入院経過日	材料費用
阿南	2	35
森原	2	6910
今中	2	6735
今中	3	35
今中	5	35
今中	6	35
今中	7	35
小林	2	290050
小林	3	0
小林	4	0
小林	5	0
小林	6	35
小林	10	6910
松田	3	6910
石川	1	5281
池田	3	0
田崎	2	35
田崎	2	25478.96
田崎	3	35
田崎	5	35
田崎	7	35
田崎	9	35
田崎	10	35
田崎	11	35
田崎	12	35
藤森	2	64453
藤森	3	0
藤森	4	0
藤森	5	0
藤森	6	0
藤森	2	304234

解らなかったら、テキストファイルのクエリ文を貼り付けてね

Step7

Step7_Q_患者別入院相対日別_収益

Query Window: T4_Dファイル
RecID
行為即決
入院相対日
診療コード
MDC手術有無分類
行為点数_1

RecID	T4_Dファイル	MDC手術有無分類	入院相対日	行為点数_1
阿南	竹	1	6800	6800
阿南	竹	2	6800	6800
阿南	竹	3	6044	6044
阿南	竹	4	6789	6789
阿南	竹	5	6296	6296
阿南	竹	6	6351	6351
森原	松	1	6576	6576
森原	松	2	16725	16725
森原	松	3	7825	7825
今中	梅	1	6794	6794
今中	梅	2	17108	17108
今中	梅	3	4951	4951
今中	梅	4	4951	4951
今中	梅	5	4951	4951
今中	梅	6	6178	6178
今中	梅	7	5947	5947
今中	梅	8	5947	5947
今中	梅	9	5847	5847
今中	梅	10	6082	6082
今中	梅	11	6082	6082
今中	梅	12	6523	6523
小林	松	1	7102	7102
小林	松	2	407894	407894
小林	松	3	20066	20066
小林	松	4	6969	6969
小林	松	5	7767	7767
小林	松	6	6966	6966
小林	松	7	6868	6868
小林	松	8	6471	6471
小林	松	9	7831	7831
小林	松	10	16613	16613
小林	松	11	7913	7913

解らなかったら、テキストファイルのクエリ文を貼り付けてね

Step9

Step9_Q_患者別_費用収益

Query Window: Step9_Q_患者別_費用収益

RecID	MDC手術有無分類	取込	費用
阿南	竹	448078	445632641540821
森原	松	23226	60093605892682
今中	梅	235701	162157473975564
小林	松	536821	874742069322079
松田	松	38732	121712525697778
石川	松	26169	165187670250575
田中	梅	284316	627707394410216
藤村	竹	36104	959739775660506
沼崎	梅	260417	256474315322431
森原	松	483076	647139445896801
飯塚	松	636656	112495426164679
山口	竹	29446	

解らなかったら、テキストファイルのクエリ文を貼り付けてね

Step8

Step8_Q_患者別入院相対日別_費用収益

Query Window: Step8_Q_患者別入院相対日別_費用収益

RecID	入院相対日	MDC手術有無分類	行為点数_1	取込	費用
阿南	1	竹	6800	4778200000000	
阿南	2	竹	6800	4778200000000	
阿南	3	竹	6044	4645216000000	2131
阿南	4	竹	6789	4684900000000	539
阿南	5	竹	6296	4684900000000	6
阿南	6	竹	6351	4684900000000	309
森原	1	松	6576	2252156000000	54
森原	2	松	16725	12233556800000	36264
森原	3	松	7825	4549182000000	
今中	1	梅	6794	3593900000000	208
今中	2	梅	17108	18361939400000	17831
今中	3	梅	4951	3697330400000	1826
今中	4	梅	4951	3697330400000	420
今中	5	梅	4951	3697330400000	1471
今中	6	梅	6178	3697330400000	312
今中	7	梅	5947	3697330400000	512
今中	8	梅	5947	3697330400000	251
今中	9	梅	5847	3697330400000	
今中	10	梅	6082	3697330400000	
今中	11	梅	6082	3697330400000	
今中	12	梅	6523	3697330400000	
小林	1	松	7102	18361939400000	35836
小林	2	松	407894	8593482779401	613391
小林	3	松	20066	3717303600000	873372
小林	4	松	6969	3334497700000	16024
小林	5	松	7767	2417120000000	15076
小林	6	松	6966	3334497700000	779
小林	7	松	6868	1000580000000	462
小林	8	松	6471	15726822700000	119
小林	9	松	7831	26065436214000	
小林	10	松	16613	1275736000000	48474
小林	11	松	7913	35255806600000	2689

解らなかったら、テキストファイルのクエリ文を貼り付けてね

Step10

Step9_Q_患者別_費用収益_1

Query Window: Step9_Q_患者別_費用収益_1

RecID	MDC手術有無分類	取込	費用
阿南	竹	448078	445632641540821
森原	松	23226	60093605892682
今中	梅	235701	162157473975564
小林	松	536821	874742069322079
松田	松	38732	121712525697778
石川	松	26169	165187670250575
田中	梅	284316	627707394410216
藤村	竹	36104	959739775660506
沼崎	梅	260417	256474315322431
森原	松	483076	647139445896801
飯塚	松	636656	112495426164679
山口	竹	29446	

注意！
Accessでは
空欄の数値を合計すると
結果も空欄になる

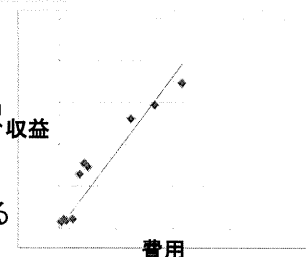
解らなかったら、テキストファイルのクエリ文を貼り付けてね

脱落しても、途中から参加できます

一円の費用でいくらの収益を上げるのか

Excelによる散布図の可視化

1. H2:H13, D2:D13を別々に選ぶ
2. 挿入→散布図→左上のシンプルな図を選択
3. 図の上で右クリック→近似曲線の追加
→切片0.0を選択
→費用ゼロでお金を貰うと財○省に怒られる



Excelによる単回帰分析【切片=0】

1. Excelファイル→オプション→アドイン→分析ツールを選択→OK
2. データの右に「データ分析」
3. データ分析→回帰分析
4. 入力Y範囲に収益(=行為点数_1)
5. 入力X範囲に費用
6. 定数に0を使用をクリック
7. OKをクリック
⇒ X1傾き=1円の費用で@円の収益を上げる

まとめ

- E,Fファイルから診療活動量を測定
 - ▶ タイムスタディを行わずに済む
 - ⇒ 意識させないタイムスタディ(山形済生館 岩淵談)
 - ▶ 原価計算の事務作業が軽減
 - ▶ 実現可能性を重視
- 診断群分類という病態別の、診療業務量を活動量とした原価計算
 - ▶ M.E.Porter(2009):病態別の診療業務を定義し、それを活動とした『活動基準原価計算』の重要性を示唆

限界点

- レセ電コードに定義のない医療行為(医師看護師の観察業務)は活動量として収集されていない
- ▶ 電子カルテ情報(アクセスログ)を用いれば、診察、看護の活動が収集でき(時間?クリック数、文字数)、より精緻な診療活動量が測定可能
岩淵勝好 電子カルテアクセスログと実施データによるDPC別人件費推定と原価計算 全自病協雑誌50(7)1126-33.
- 診療行為の重み付けの妥当性が未確認である
 - ▶ 重み付けの標準化は困難 意見百出

まとめ

- 本手法 重み付けを部分的に考慮(・:加算を除くすべてのレセ電算コードを網羅的に勘定科目に対応させている)

ある処置の重症度が高いなら、その前後の診療行為が増えるはず

その処置に関連する診療活動量が多くなる

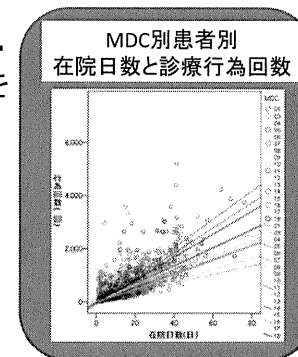
結果、その処置の重みは部分的にその前後のレセ電コードの件数で反映

- 重み付けと配賦基準の問題 私見
患者数、職員数、日数、面積、@@@...
何でやっても分類間のコストのばらつきを完全には解消しないであろう

▶ 何がよりましたか

- 本手法の可能性

- ▶ 分類別、入院経過日別の費用収益関係が分析可能
- ▶ どの日の費用に問題なのか、の把握
- ▶ 活動基準経営管理に有用



まとめ

- ◆ 本手法では、外科手術の患者の費用が手術日に低くなる
 - ⇒ 手術の考え方を精緻にする必要
 - ⇒ 麻酔時間はFファイルにある
 - ⇒ あとは医師、看護師の時間給が必要
 - ⇒ 外保連試算に時間給がある
 - ⇒ 病院が持つ(決める)時間給
 - ⇒ 時間=電子カルテのアクセスログなど(出勤してから閉じるまでなど)
- ◆ 医薬品、材料の購入価格の管理が必要
 - ⇒ 医薬品、材料の患者投入日、量、コードの管理
 - ⇒ 山形済生館 岩淵勝好先生がバーコードを活用した消費、在庫管理の研究
 - ⇒ 在庫・物流管理へ
- ◆ MDC別、入院経過日別の費用収益関係が把握出来、活動基準経営管理が可能
- ◆ E,Fファイルと電子カルテを結合=医師診療、看護観察が把握出来
 - ⇒ より精緻な診療活動量を測定する事が可能
- ◆ 診断群分類別(サービス別)原価計算が可能であり、根拠に基づいた診療報酬の決定(コスト分科会の役割、分類別入院基本料)に寄与できる

EFデータ抽出ツール

東京大学
堀口 裕正

DPCデータを分析に使うには

- HISによるシステム化
 - DPC分析ソフト(サービス)の利用
- ARROWS/EVE/ヒラソル.....
- ↑ 高
有償
低い
-
- 自分でDB化して分析
- ↑
研究班がセミナーで支援
- 無料

医療機関自らDBを作って分析する

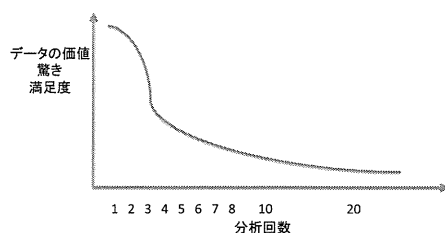
- メリット
 - コストがかからない
 - 分析に必要な考え方が身につく
 - 1度作ってしまえば、同じデータを毎回出すには手間がかからない
- とはいっても
 - 分析用データセットを作るには手間と時間がかかる

例えば・・・

- ACCESSでのデータ分析
- 毎月のDataのインポートにかかる手間
- 分析したいものに応じて作る中間テーブルの設計
- データを修正したときの差し替えの手間
- 毎月決まったものを出すOR自分で高度な分析をしたい(しできるスキルがある)のでないと継続しない??

分析の結果ほしいもの

- 毎月継続してほしいもの
- 1回か2回データを出せば十分なもの



結果として

- ちょっと1回データの分析結果を見てみたい
- あまり手間をかけたくない

といった希望が出てくる。そうすると

- DB化は面倒、DPCのデータファイルをそのまま使って分析できないかなあ・・・。

ここからは研究班側の話

- そもそも研究遂行のために必要なので、情報システム+分析システムを構築
- 自分たちが使いやすいよう、あるいはより多くの成果が出るように改良やノウハウの蓄積を進めてきた。

研究班での分析システムの歴史

- 初期 統計ソフト(SPSS/STATA)+ACCESS

これでは処理しきれなくなって

- 中期 SQLサーバー+統計ソフト+BIソフト
- BIソフト(QlikView/Spotfire・・・)

ここで得たノウハウやData Management手法をセミナーで医療機関の皆様へ還元

近年の研究班におけるIT的課題1

- SQLサーバーでは量をこなせない??
 - 2009年 Fファイル9億行
 - 2010年 E/F統合ファイル 16億行
 - 7年分で約1000万退院分のデータ
 - Fファイル換算約50億行分

処理に時間がかかる、たいへん……。今年もまた16億行分のデータ……。非常にうれしいけど……。

近年の研究班におけるIT的課題2

- 論文になったり、公表したりした成果を作るために利用した加工データの加工アイデアやロジックがそのあとお蔵入り……。
- (ありがたいことですが)年間数十本の論文成果、このデータの作成ロジック自体も世の中に貢献できないだろうか

課題1を解決するために

- 今「はやり」の分散処理システムを使うことにしました。
- Hadoop/Pig
- たくさんのコンピュータを並列で使ってデータが増えても機材の台数を増やすだけで解決。
- 処理をするためのスクリプト言語は統計ソフトのスクリプト並み

これはいいということで導入
(詳しい説明は後述)

分散処理システム

- これからはBigDataの時代・・
- スケールアップよりスケールアウト
- SQLなんて古い、これからはNOSQLの時代
- なんてもてはやされています
- なんといってもOSS(オープンソースソフトウェア)なので無償で使える

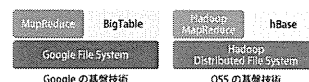
すばらしい、やはり時代はこちらに……
(By 東京のIT屋さん)

もちろん

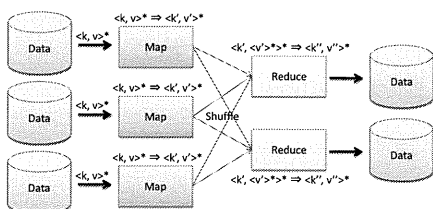
- SQLサーバーを使用し続けながらデータ量をこなす問題を解決する手法もあります。
 - SSD+高速RAIDを使う
 - データベースを分割する
 - 機械の性能を上げる
- こちらの手法も研究班内では利用してしままでのSQLで培った資産を生かす試みもやっています。
- (ただ、この手法は大量データを処理するためのノウハウなので、そこまでデータ量のない医療機関の皆様には還元できる成果が少ないです)

Hadoopとは？

- Googleの基盤ソフトウェアのクローン
 - Google File System, MapReduce
- Yahoo Research の Doug Cutting氏が開発
 - 元々はLuceneのサブプロジェクト
 - Apache Project
- Javaで記述!



MapReduceの実行フロー



Pig

- Hadoopを扱いやすくするミドルウェアの1つ
- Pig Latinという専用言語で簡単にDWH的な処理を書けるようにしたミドルウェア
 - Javaを使用せずに、SQL的な言語でMapReduce処理を記述する事が出来る
 - Googleでは、同様にSawzallというスクリプト言語で、MapReduce処理が簡単に書けるようになっている
- Yahoo!が開発
- データのロード・結合やフィルタ処理を楽に書ける

システム開発の前提条件

- 利用者は研究者
 - 分析は自分の手慣れたソフトで試行錯誤を繰り返しながら使いたい
 - SAS/SPSS/STATA/R等の統計ソフトを使う+シンタックスぐらいはかける(理解できる)
- 結果データは1患者1レコードのCSVデータとして作成、それ以上の加工はしない
- Pigというスクリプト言語を利用し、ほしいデータを抽出するスクリプトは利用者が自分で作成する(UIを作らない)

開発の概要

- 今回の抽出・加工を行うためのUDF(スクリプト言語で使う関数)を開発し、それを利用した作業が行えるようにする
- UDFで実現する内容
 - 抽出・加工内容を定義した定義ファイルがわかりやすくかける
 - 定義ファイルを使って必要なレコードを抽出する
 - 定義ファイルを使ってクロス表を作成する
 - (おまけ)日付データの取り扱いを楽にする

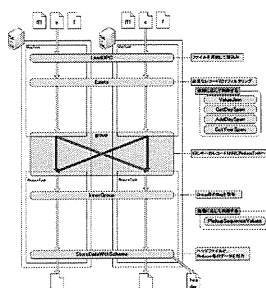
GroupFilterFormat (抽出・集計用定義)

- フォーマット
– <グループ名>(<項目名>:[<付加値>, ...], ...), ...
- サンプル
ボスミン(642450005[1,注],
642450164[2,注],
620517902[1,注]),
ワルファリン(613330003[1,錠],
621938101[0.2,細粒])

作成した関数

関数	定義
Exists	指定した文字が含まれるか判定する。
ValueJoin	指定した付加値を Bag として追加する。
InnerGroup	Bag を Group してタプルにする。
GetDaySpan	2つの日付から経過日数を算出する。
GetYearSpan	2つの日付から経過年数を算出する。
AddDaySpan	日付に値を加算して日付を得る。
PickupSequenceValues	Bag 中で値が連続しているタプルだけ取り出す。

処理イメージ



処理速度

- 本研究室内環境で2009年DPCデータから20種類程度の薬剤・処置の実施の有無を1患者1レコード型に書き出す加工
- 様式1 約300万・Fファイル 約9.5億行

台数	処理時間
2	6900sec
4	3650sec
8	2080sec
16	1010sec
48	380sec

本システムの今後

- このシステムは処理がDPCデータに特化していないため、多様な応用が可能
 - オーダリングのデータ
 - 電子レセプトデータ
 - Webのアクセスログ...
- スクリプトの記載が簡単であることを確保しながら、関数を増やしていく。
- 同じスクリプトを医療機関で簡単に(パソコン1台で)利用できるプログラムも開発中

新しい分析ソフトのご紹介

- 研究班で分散処理技術の開発を行ってきた副産物としてできたソフト
- パソコン1台で実行可能
- DPCデータと処理用のスクリプトを入れればその場で結果ファイルをはき出す。

DPCデータ抽出・整形プログラム(通称 Durok)