

レセプト情報等 オンサイトリサーチセンター（京都） パフォーマンステスト結果報告



medical informatics

酒井未知¹⁾、大寺祥佑¹⁾、岩尾友秀²⁾、岡本和也²⁾、加藤源太³⁾、
黒田知宏²⁾

- 1) 京都大学大学院医学研究科 社会健康医学系専攻 健康情報学分野
- 2) 京都大学医学部附属病院 医療情報企画部
- 3) 京都大学医学部附属病院 診療報酬センター



本日お話する内容

1. 試行的利用開始後の模擬申出実施状況
2. オンサイトリサーチセンター（京都）
の機能とパフォーマンス








試行的利用開始後の模擬申出実施状況

- **2016.02.17 京都大学にて試行的利用開始**
 - ・ オンサイトリサーチセンター（京都）運用部の設置
- **2016.03.16 有識者会議にて意見提供**
 - ・ 教育プログラム、公開演習環境等の整備、情報環境の見直し（ソフト導入、データマート整備）を提案
- **2016.05.12 第1回実務者会議**
 - ・ 各種アプリケーションを用い、試行的に件数集計を実施
 - ・ 傷病名「前立腺癌 or 前立腺肥大症」
 - ・ 対象診療年月は「平成23年1月～12月」
- **2016.06.09 第2回実務者会議**
 - ・ 抽出件数や作業負荷量を増加させ、件数集計、個票ダウンロード、統計解析を実施
 - ・ 「高血圧」に関連する傷病名
 - ・ 対象診療年月は「平成22年1月～27年12月」



オンサイトリサーチセンター データ抽出・分析アプリケーション

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
マウス操作			コマンド入力		
プルダウンメニューから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 
定形分析	非定形分析				
レセプトの行単位	レセプトの行単位・レセプト単位・患者単位				
サーバー			ローカル		



定形帳票

プルダウンで分析対象レセプトの条件を指定し
件数集計や作図を行う

The screenshot shows the Oracle BI Business Intelligence interface. The top part displays search criteria for medical data analysis. A red box highlights the '診療年月(RE)' dropdown menu, which is set to '2015年12月'. A red arrow points to this dropdown with the text '診療年月を選択' (Select treatment year). Another red box highlights the '傷病(SY)' dropdown menu, with the text '傷病名を選択' (Select injury name). Below the search criteria, a table titled '【医科】年齢層別・男女区分別集計' (Medical: Age Group and Gender Distinction Summary) is shown. The table has columns for '男女区分' (Gender Distinction), '年齢層' (Age Group), '件数' (Number of Cases), '合計点数' (Total Points), and '診療日数' (Treatment Days). The table is divided into two main sections: '1:男' (Male) and '2:女' (Female). The '件数' column is highlighted in red. To the right of the table, a '月別集計グラフ (件数)' (Monthly Summary Graph (Number of Cases)) is displayed, showing a single blue bar representing the total number of cases for the selected criteria.

件数集計の結果



自由分析

マウスで分析対象レセプト条件、集計方法
を選択し、件数集計や作図を行う

標準 結果 プロンプト 詳細

サブジェクト・エリア

- A-02.医科(オンサイトセンター)
 - 医科コメントレコード(CO)ビュー
 - 医科保険者レコード(HO)ビュー
 - 診療実日数
 - 合計点数
 - 回数(食事療養・生活療養)
 - 合計金額(食事療養・生活療養)
 - 医療保険(負担金額)
 - 減額割合(負担金額)

選択された列

「サブジェクト・エリア」ペインで列名をダブルクリックして、分析に追加します。追加した列は、このペインにソールを合せます。

医科レセプト管理レコード(MND)ビュー	医科保険者レコード(HO)ビュー
診療年月	合計点数 診療実日数

テーブルと列を
ドラック&ドロップ

件数集計の結果

複合レイアウト

ピボット表

男女区分名称	西暦名称	診療日数	合計点数
1:男	2014年11月	1,830	2,471,354
	2014年09月	22	22
	2014年08月	20	102
	2014年07月	315	718
	2014年05月	10	10
	2014年04月	10	10
	2013年12月	24	24
2:女	2013年10月	377	19
	2014年11月	26	24,691,356

分析対象レセプトの条件
をフィルタで指定

フィルタ

削除

新規フィルタ

列 診療年月 **診療年月を選択**

演算子 次で始まる

値 422

詳細オプションの追加 すべてクリア

フィルタの保護

このフィルタをSQLに変換する



SQL Plus

SQLで分析対象レセプトの条件と集計方法を指定し、
件数集計を行う

```
SQL Plus
select count (distinct sy.SEQ2_NO)
from CMDV_T_RCP_MED_SY sy
where sy. SKWD_NAME_CD in (1859003,
8836591) ;
```

220,006,818

分析対象レセプトの条件と
集計方法を指定するSQL

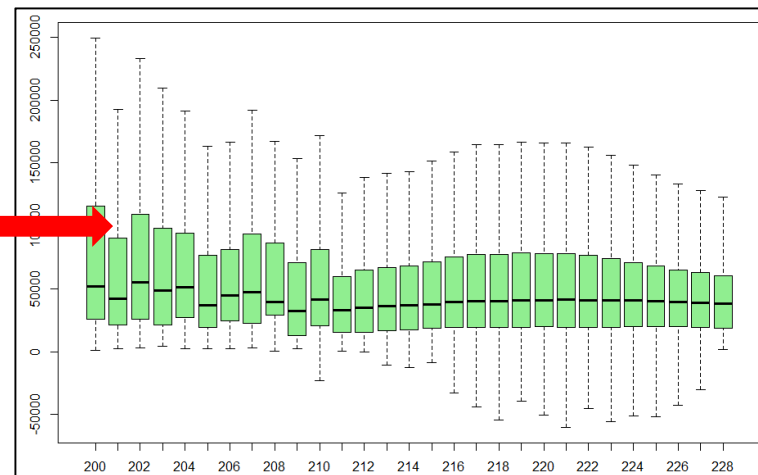
件数集計の結果

統計ソフト コマンドで分析対象レセプトの条件と解析方法を指定し、解析を行う

```

R
Rコマンド入力
R: Y:\log\r_test160711.R - R Editor
start <- proc.time()
q <- "
create table DATA1 as
select count(distinct re.ID1N) N
from CMDV_T_RCP_MED_SY sy
inner join CMDV_T_RCP_MED_RE re
on sy.SEQ2_NO=re.SEQ2_NO
where sy.SKWD_NAME_CD in (
'8830212','8842687','8844144','8832479',
'8842709','8848335','8848337','8833421',
'4372001','8833422','3621009','8842178',
'8833423','8833424','8844713','4029010',
'8833425','8833426','4039033','4039005',
'4039006','8833427','4372003','4372002',
'4319003','3621010','8842488','8842089',
'8834071','8842500','8834893','8835022',
'8844234','4019016','4019017','8842765',
'8848357','8835367','8835586','8835605',
'8835614','8842791','8842804','8842094',
'8844086','8846195','8838336','8838398',

```



```

SAS
SASコマンド入力
SAS
NOTE: Copyright (c) 2002-2012 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) Proprietary Software 9.4 (TS1M2 6553170)
Licensed to MINISTRY OF HEALTH, LABOUR AND WELFARE -BSC, Site 10207599.
NOTE: このセッションはX64 7PRO プラットフォーム上で実行されています。

NOTE: Updated analytical products:
SAS/STAT 19.2

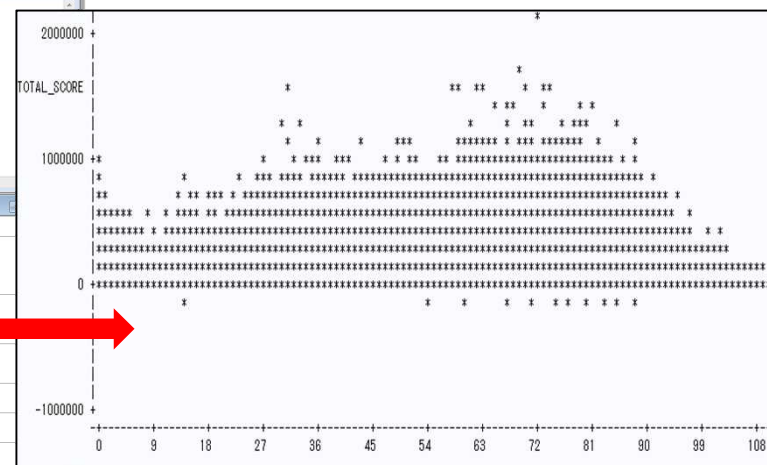
NOTE: Additional host information:
X64 7PRO MIN 6.1.7601 Service Pack 1 Workstat!

NOTE: SASシステム初期化処理
処理時間 2.59 秒
CPU時間 0.71 秒

sas160706
proc univariate data=Data;
run;
proc univariate data=Data normal plot;
var AGE;
histogram;
output out=wk1 n= mean=mean std=std median=median
mi=mi min=mi max=mi;
run;
proc plot data=Data;
plot AGE^TOTAL_SCORE;
run;
quit;
proc plot data=Data;
plot TOTAL_SCORE^AGE;
run;
quit;
proc corr data=Data Pearson;
var AGE TOTAL_SCORE;
run;
proc reg data=Data;
model TOTAL_SCORE=AGE;
run;
proc reg data=Data;
model TOTAL_SCORE=AGE SEX_DIV;
run;

```

統計解析の結果





パフォーマンステスト

1. レセプト・患者の件数集計
2. ローカルでのデータ解析
3. サーバーでのデータ解析

オンサイトの各アプリケーションを使用し
データ件数、実行所要時間、操作感を検証



パフォーマンステストの全体像

		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	テスト 1					
	患者数		テスト 2				
データダウンロード				テスト3			
ローカルでの統計解析						テスト4	
サーバーでの統計解析					テスト5		

同じ作業を各アプリケーションで実施し
データ件数・実行時間・操作感を比較



集計・解析対象レセプト

レセプト件数の集計

テスト1

- 病名に「前立腺癌」又は「前立腺肥大症」を含む内科レセプト
- 診療年月が平成23年1月～23年12月（1年分）

患者件数の集計

テスト2

- 病名に「高血圧」を含む内科レセプト
- 診療年月が平成22年1月～27年12月（6年分）

データダウンロード、統計解析

テスト3

テスト4

テスト5

- ICD10コードが高血圧症性疾患（I10-I15）のDPCレセプト
- 診療年月が平成22年1月～23年12月（2年分）



レセプト件数集計テスト

【病名に前立腺癌または前立腺肥大症を含む患者の
1年分の医科レセプト数を集計】

テスト1

	定形帳票	自由分析	SQL Plus	Oracle R
抽出数	34,964,611行	29,156,818件	29,156,818件	29,156,818件
実行時間① 5月6日 18:30	6分	1分	12分	2分
実行時間② 5月10日 14:20	40分以上	85分以上	4時間27分	10分

- 定形帳票は、レセプトの「行数」の集計結果が出力された
- 自由分析・SQL Plus・Oracle Rでレセプト数集計が一定時間内に可
- レセプト取込作業中は、実行時間が増加した

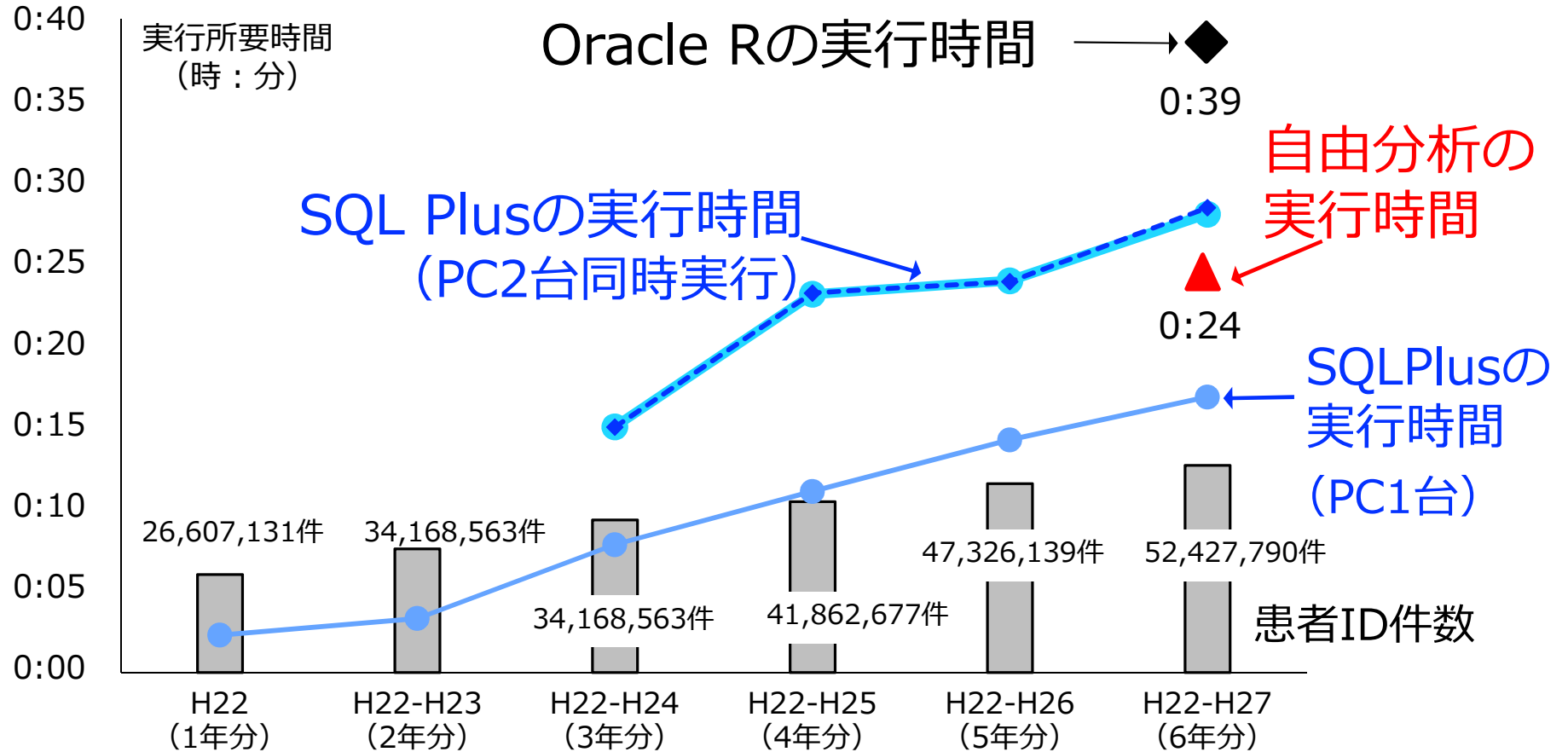
(注) 病名抽出条件：傷病名コードに「前立腺癌」または「前立腺肥大症」という語を含む



患者数集計テスト

テスト2

【病名に高血圧を含む医科レセプトが発行された患者数を集計】



- 自由分析・SQL Plus・Oracle Rで患者数集計が一定時間内に可能
- 件数・PC台数いずれに対しても、実行所要時間が線形に増加

(注1) 病名抽出条件：傷病名コードに「高血圧」という語を含む

(注2) 実行日時：自由分析 2016/7/11 12:40, SQL Plus 2016/6/1 9:40, Oracle R 2016/7/20 9:17



パフォーマンステストの結論

1. 高負荷のレセプト件数集計、患者数集計は、自由分析、SQL Plus、Oracle Rで一定時間内に可能。

データダウンロードテスト

テスト3

高血圧症性疾患患者の4年分のDPCLレセプトの、患者ID・レセプトID
年齢・性別・診療年月・都道府県・点数（6列）をダウンロード

	SQL * Plus			Oracle R
	データ作成 +10行画面表示	データ作成 +ダウンロード	データ作成 +全行画面表示 +ダウンロード	データ作成 +ダウンロード
実行日時	6月6日18:30	7月7日 18:00	6月2日 20:21	6月7日 14:09
抽出数	10行	26,607,131行	26,607,131行	26,607,131行
ファイルサイズ	N/A	1.36GB	1.36GB	2.6GB
実行時間	5分	約8時間	86時間47分	25分
保存場所	N/A	D（ローカル）	Y（個人ネットワーク）	C（ローカル）

- SQL Plusは、データ作成は短時間で実行可能であるが、データの画面表示、ダウンロードに長時間を要した
- Oracle Rでは一定時間内にデータ作成とダウンロードが可能

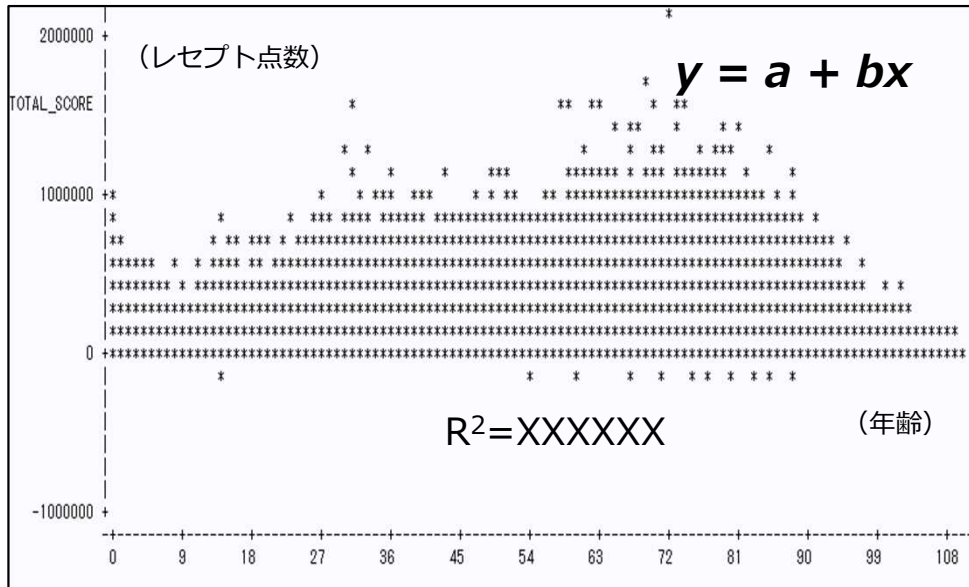


パフォーマンステストの結論

1. 高負荷のレセプト件数集計、患者数集計は、自由分析、SQL Plus、Oracle Rで一定時間内に可能。
2. データのダウンロードは、Rの方が実用性が高い。SQL Plusは件数が増えると長時間かかる可能性がある。



高血圧症性疾患患者の2年分DPCレセプト 12,108,844行・11列 (1.6GB) を解析



	β	β 下限	β 上限	P value
α (切片)	80089.6	79866.1	80313.1	
年齢	-235.6	-238.6	-232.7	***
性別	-3524.5			***
I10 本態性高血圧症	-2871.8			
...				
I15 二次性高血圧症	31481.4			
I150 腎血管性高血圧症	10815.4			***
I151 その他の腎障害による二次性高血圧症	-7783.7	-8586.7	-6980.8	***
I152 内分泌障害による二次性高血圧症	-15724.7	-17665.9	-13783.6	***

$$y = a + b_1x_1 + b_2x_2$$

実行日時：2016年7月11日 14時30分

単回帰分析が1分以内に完了

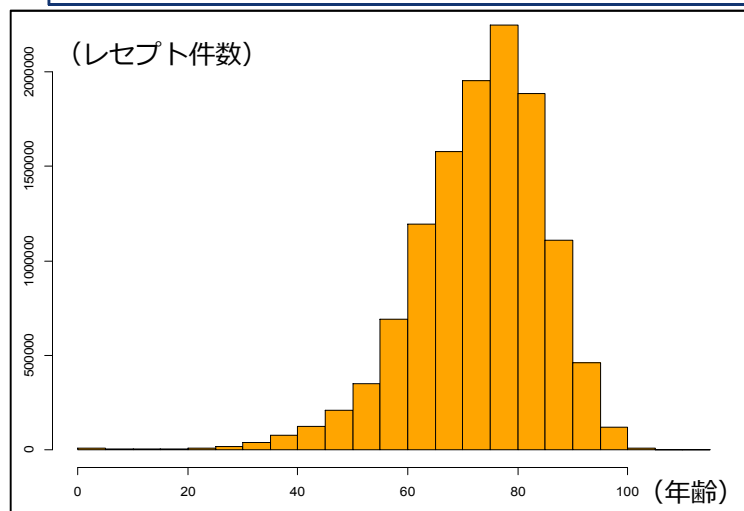
重回帰分析が1分以内に完了

ローカルでSASを用いる解析は、一定時間内に可能

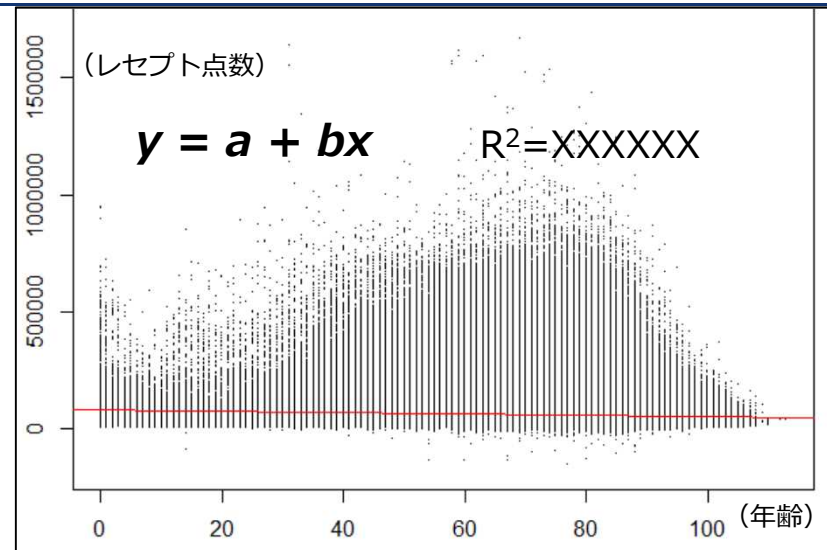
ローカルでの統計解析テスト (R)

テスト4

高血圧症患者の2年分DPCレセプト 12,108,844行・11列 (1.6GB) を解析



① 年齢層別レセプト数の作図
14:30



② 単回帰分析 14:31

	β	β 下限	β 上限	P value
α (切片)	80089.6	79866.1	80313.1	
年齢	-235.6	-238.6	-232.7	***
性別	-3524.5	-3595.8	-3453.3	***
I10 本態性高血圧症	-2871.8			
⋮				
I15 二次性高血圧症	31481.4			
I150 腎血管性高血圧症	10815.4			
I151 その他の腎障害による二次性高血圧症	-7783.7			

$y = a + b1x1 + b2x2 + \dots$

③ 重回帰分析 14:32

Warning messages;
Reached total allocation of
16267MB:
see help(memory.size)

④ エラーで実行中断
14:33



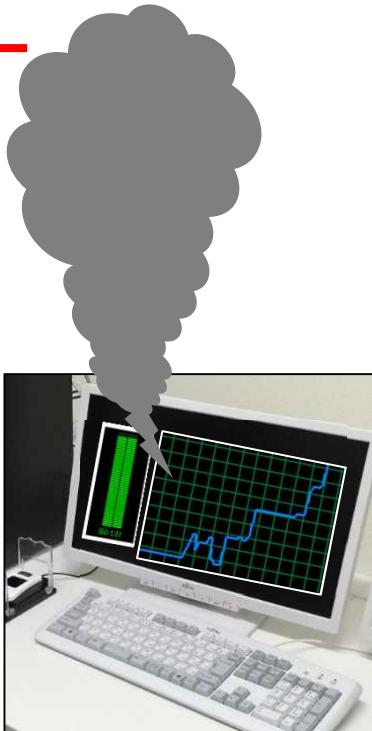
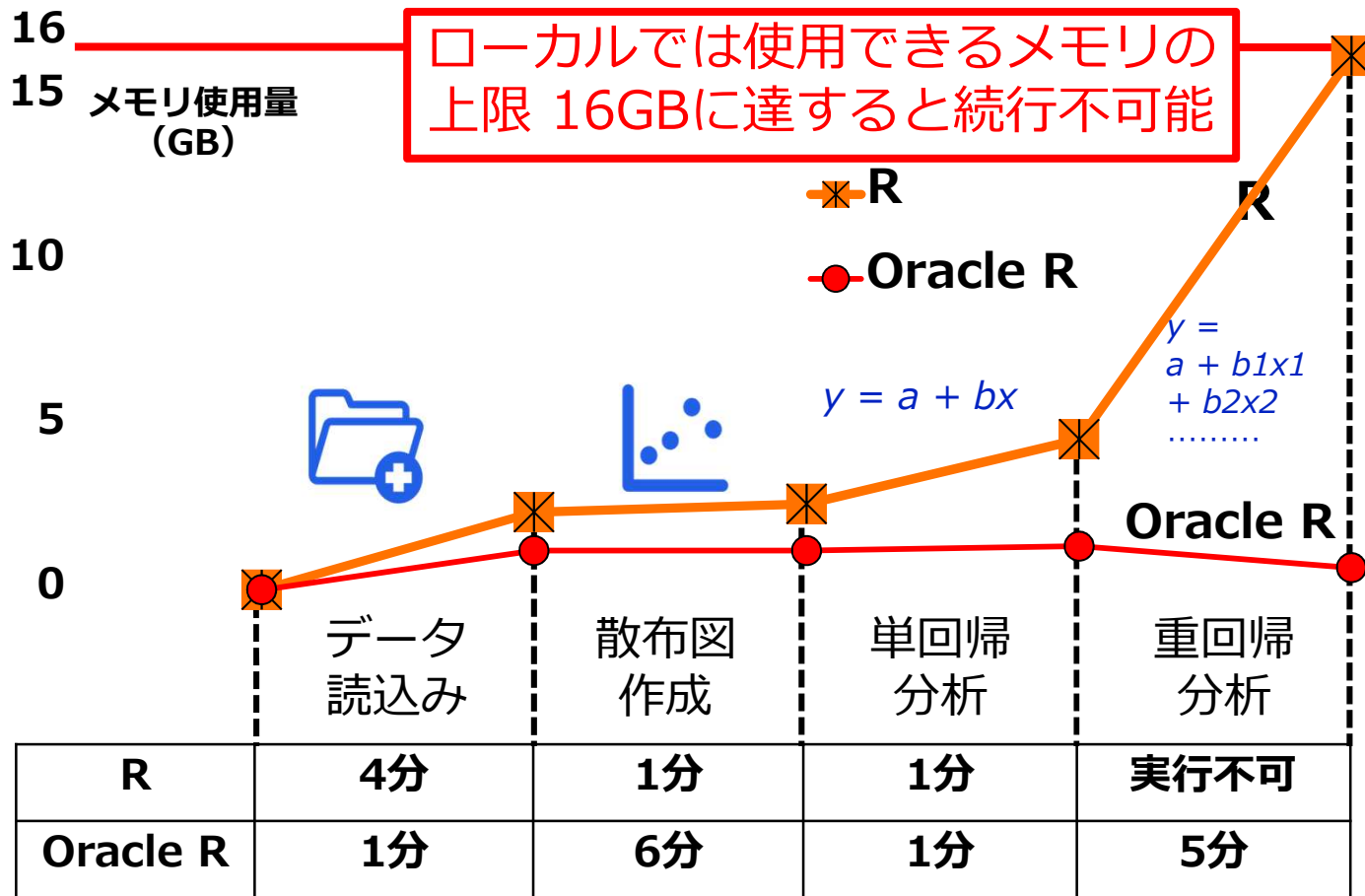


ローカル VS サーバー内の 統計解析パフォーマンスの比較

テスト4

テスト5

高血圧症患者の2年分DPCLレセプト 12,108,844行・11列 (1.6GB) を解析



実行日時
2016年7月11日 16:40

- ローカルでRを用いる解析は、メモリによる制約が大きい
- サーバー上では、Oracle Rで一定時間内にデータ解析が可能



パフォーマンステストの結論

1. 高負荷のレセプト件数集計、患者数集計は、自由分析、SQL Plus、Oracle Rで一定時間内に可能。
2. データのダウンロードは、Rの方が実用性が高い。SQL Plusは件数が増えると長時間かかる可能性がある。
3. ローカルで統計解析を行う場合、Rはメモリの制約が大きい。
4. 現状、大規模データ解析は、SASによるローカルでの解析、またはOracle Rによるサーバー上での解析の実用性が高い。



大規模データ解析のパフォーマンス

		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	出来る		
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		



結語

- オンサイトリサーチセンター（京都）のアプリケーションを試行的に利用し、データ抽出、件数集計、統計解析機能を検証した。
- 今後、模擬申出の課題に即して、更なる検証を行う。