

厚生労働科学研究費補助金

健康安全・危機管理対策総合研究事業

保健医療福祉計画策定のための データウェアハウス構築に関する研究

(H27-健危-一般-001)

平成28年度 総括・分担研究報告書

研究代表者

岡 本 悦 司

平成29(2017)年3月

目 次

I. 総括研究報告

- 保健医療福祉計画データウェアハウスウェブ上公開に関する研究・・・・・・・・・・・・・1
岡本悦司，神谷達夫
（資料）データウェアハウス活用マニュアル

II. 分担研究報告

1. 地域包括ケアシステムのための病床機能報告DWH活用の試み・・・・・・・・・・・・・11
中本稔
2. 医療計画の推進に関する研究・・・・・・・・・・・・・18
大江浩
3. 学校保健統計調査データの活用・・・・・・・・・・・・・31
安藤雄一
4. 母子保健・小児保健におけるハイリスク群（虐待を対象とする）のデータ管理と関連論点
の検討・・・・・・・・・・・・・37
澤口聡子
5. がん検診評価のためのDWH-CUBE活用・・・・・・・・・・・・・48
森永裕美子，松岡宏明
6. 予防接種に関するデータのデータウェアハウス構築と分析・活用の試み・・・・67
松本珠美,永井仁美
7. 健康増進計画策定におけるデータウェアハウスの活用・・・・・・・・・・・・・73
藤井仁
8. データウェアハウスを利用した世帯の主な職業による出産の質の検討・・・・・・・・79
仙田幸子，吉田穂波
9. 乳児死亡原因のRisk factors・・・・・・・・・・・・・90
仙田幸子，吉田穂波

III. 成果資料

保健所長会理事会プレゼン資料(2016年8月17日東京)

第44回北陸公衆衛生学会 講演資料(2016年11月21日富山)

第75回日本公衆衛生学会(2016年11月大阪)演題

保健師ジャーナル2017年3月号

保健医療福祉計画データウェアハウスのウェブ上公開に関する研究

研究代表者 岡本悦司(福知山公立大学)
研究協力者 神谷達夫(福知山公立大学)

研究要旨

市区町村単位の統計データをキューブ化し、ウェブ上で容易にピボットテーブルとして活用できる形態で公開した。ピボットテーブル的な操作を可能にすることはデータの医療圏や保健所管轄区域別集計を容易にするが、クライアント PC にいちいちデータをダウンロードする仕組みのため、大規模データではダウンロードに時間がかかる点等改善の余地を残すものとなった。

A 研究目的

初年度においては、DWH 化した Excel やアクセスファイルを DVD で全国保健所に配布したが、究極的にはウェブ上で自在に集計できる公開が望ましい。そこで、Javascript という言語を用いてウェブ上でピボットテーブルのように自在に扱えるかたちに加工して公開する。

B 研究方法

ウェブ上でリアルタイムでピボットテーブルを作成できる Javascript 言語プログラム pivot.js を用いて HTML ページを作成し、独自サーバーを通じてウェブ上で公開した。URL は <http://www.jmedicine.com>

C 研究結果

ウェブ上で、市区町村単位データを医療圏や保健所管内ごとに、Excel ピボットテーブルのような感覚でドラッグ&ドロップすることで作表できるようになった。しかしながら、Javascript は、対象とする csv ファイルをいったんまるまるダウンロードしなければならず、大容量ファイルではダウンロードに相当な時間がかかることが明らかになった。そのため、たとえば病床機能報告データでは、全国一本の csv ファイルとして公開することは断念し、都道府県ごとに 47 分割されたファイルとして公表せざるをえなかった。

D 考察

ウェブ上でピボットテーブル様にドラッグ&ドロップ操作ができることはデータ活用上きわめて使いやすい形態である。しかしながら、大きな csv ファイルをいちいちクライアントの PC にダウンロードしなければならず、時間がかかる上に、異なる統計調査データを市区町村単位で結合する等ができない、といった限界も残った。今後は、HADOOP 等の並列処理システムを活用して、迅速なデータ処理と転送が可能なシステムへの改善が課題として残った。

初期画面

http://www.jmedicine.comを開くと(3秒後に http://121.101.70.40/に redirect されるが、近く直に開くようにする)下の初期画面が出現する。

左側がメニューであり、右側に使用法が説明されている。

含まれるデータ

28年12月現在下記の統計データが収録されており逐次追加してゆく。元データは csv ファイルであり []はそのファイルサイズである。ファイルサイズがダウンロードに要する時間の目安となる。

病床機能報告

- ▷レセプト件数[2.8MB]
- ▷回答項目別[8.3MB]
- ▷数値データ[都道府県別]

健康増進事業報告

- ▷胃がん[108MB]
- ▷肺がん[291MB]
- ▷大腸がん[139MB]
- ▷子宮頸がん[106MB]
- ▷子宮体がん[8MB]
- ▷乳がん[107MB]
- ▷肝炎[16MB]

地域保健事業報告

- ▷予防接種[59MB]

人口動態統計

- ▷人口動態総覧[47MB]
- ▷母年齢階級別出生数[60MB]
- ▷性・年齢階級別死亡数[117MB]
- ▷住民基本台帳人口[79MB]

介護保険関係

- ▷在宅ケア見える化[11MB]

国民健康保険関係

- ▷被保険者数[24MB]

▷所得[3MB]

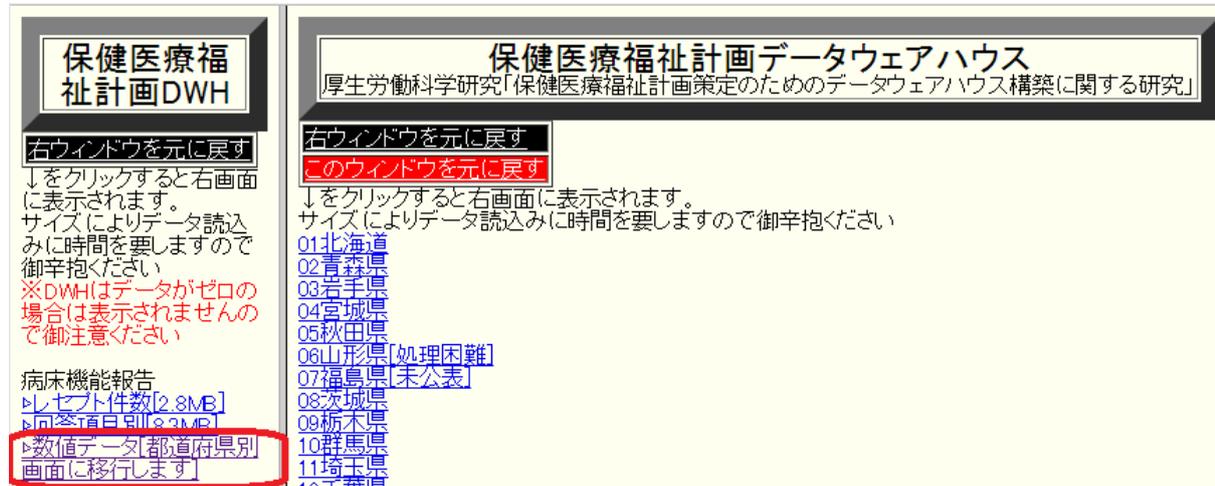
その他の統計

▷市区町村の指標

▷警察自殺統計[46MB]

操作法 1:病床機能報告を例に

病床機能報告は一般病院及び有床診療所(約2万)からの報告を都道府県がサイト上に公表したものである。最新の2015年7月現在報告をDWH化した。福島、奈良、兵庫県は未公表となっている。病床機能報告は膨大であるため、3部分に分かれている。レセプト件数、回答項目(たとえば救急指定の有無等)そして病床数等の数値データである。数値データは都道府県別のファイルとなっている。数値データをクリックすると右画面に都道府県のリストが表示される。



都道府県をクリックするとデータのダウンロードが始まり、ダウンロード状況が%で表示される。ダウンロードの時間はPCの性能、回線速度による。ダウンロードが100%完了するとデフォルト画面が出現する。都道府県単位の場合、デフォルトでは市町村が行、大分類が列に表示される。

画面の説明

下図の赤い枠がそれぞれ行、列の見出しが入る部分。デフォルトでは、全国データなら行に都道府県、列に年を入れてあるが、自在に入れ換えることができる。

データダウンロード状況 100.0%

病床機能報告2015数値【京都府】

DATAは触らない → DATA - 中分類 - 小分類 - 病院 - 医療圏 - 保健所 変数名のリスト

合計(総数) DATA 大分類 - 列見出し

市町村 - 行見出し

市町村	大分類	リハビリテーションの実施状況	入院患者の状況(年別)	入院者の退院月間	全管理状況	医療機器の台数	急性期又は在宅療養後の療養の状況	手術状況	救急医療状況	病床状況	看護を行った患者数	職員の状況	退院を要する者	退院を要する者のうち在宅療養者	重症患者への対応状況	重症患者の入院	長期療養者状況	(件)
26101 京都市北区		1,008	41,180	1,644	502	11	32	632	13,358	2,810	1,150	761	16	403		218	132	
26102 京都市上京区		1,976	700,649	7,646	3,379	44	716	5,230	75,675	4,518	52	2,250	3,242	49	124		1,628	
26103 京都市左京区		6,615	94,149	7,117	2,380	53	640	3,051	47,639	5,243	3,232	2,869	79	259	17	274	1,172	

変数名の説明

変数名のうち以下のものは共通である。

DATA…表に表示される数値である。DATAは通常は、ドラッグドロップ等の操作は行なわない。

医療圏…医療計画にいう二次医療圏を指す[2014年時点]。

保健所…保健所管轄区域を指す[2012年時点]。

変数の操作

変数名リストより表示したい変数を行または列にドラッグ&ドロップする。この場合、大分類→中分類→小分類、都道府県→保健所[又は医療圏]→市町村→病院という階層構造に合わせてドラッグすることが原則である。

たとえば大分類の内訳を知りたいければ「中分類」を変数名リストより列見出しにドラッグ&ドロップする。すると「リハビリテーションの実施状況」の下に中分類が以下のように表示される。

保健医療福祉計画DWH

右ウィンドウを元に戻す
↓をクリックすると右画面に表示されます。
サイズによりデータ読み込みに時間を要しますので御辛抱ください
※DWHはデータがゼロの場合は表示されませんので御注意ください

病床機能報告
レセプト件数[2.8MB]
回答項目別[8.3MB]
数値データ[都道府県別画面に移ります]

健康増進事業報告
胃がん[108MB]
肺がん[291MB]
大腸がん[139MB]
子宮頸がん[106MB]
子宮体がん[8MB]
乳がん[107MB]
肝臓[16MB]

地域保健事業報告
予防接種[59MB]

人口動態統計
人口動態総覧[47MB]
母体年齢別出生数

データダウンロード状況 100.0%

病床機能報告2015数値【京都府】

表 ▼ DATA ▼ 小分類 ▼ 病院 ▼ 医療圏 ▼ 保健所 ▼

合計(整数) ▼ DATA ▼ 大分類 ▼ 中分類 ▼

市町村 ▼

大分類	リハビリテーションの実施状況														入院患者の状況(年間)		入院患者の状況(月間)	
	がん患者リハビリ	高齢者リハビリ	脳卒中リハビリ	呼吸器リハビリ	心大血管系リハビリ	認知症	早期リハビリ	脳血管系リハビリ	認知症リハビリ	運動器リハビリ	過去1年間の退院患者数	障害者リハビリ	在宅患者数	新規入院患者数	退院患者数	新規入院患者数	退院患者数	
25101 京都市北区	19	51		87	19		11	173	235	201	242		393,192	9,362	9,126	921	893	

同様に各市町村内の病院を表示させたいければ「病院」を行見出しにドラッグ&ドロップする。すると以下のようになり病院ごとに数値が表示される(なお、病床機能報告では10未満の数値は非表示となっているので留意)。

データの抽出

病床機能報告2015数値【京都府】

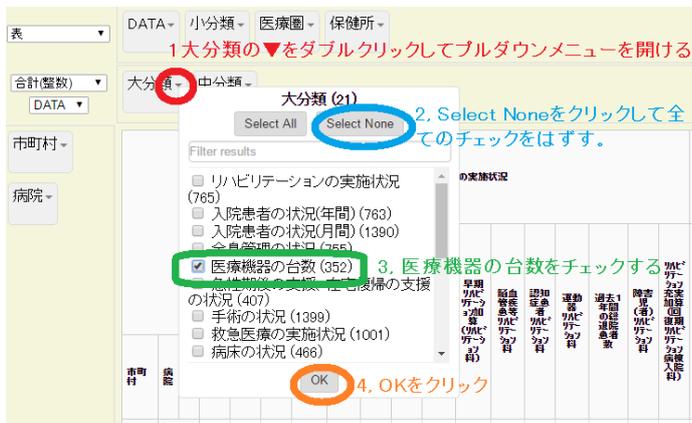
表 ▼ DATA ▼ 小分類 ▼ 医療圏 ▼ 保健所 ▼

合計(整数) ▼ DATA ▼ 大分類 ▼ 中分類 ▼

市町村 ▼ 病院 ▼

大分類	リハビリテーションの実施状況															
	がん患者リハビリ	高齢者リハビリ	脳卒中リハビリ	呼吸器リハビリ	心大血管系リハビリ	認知症	早期リハビリ	脳血管系リハビリ	認知症リハビリ	運動器リハビリ	過去1年間の退院患者数	障害者リハビリ	在宅患者数	新規入院患者数	退院患者数	
京都市左京区								28			45					
富田病院								17	61		13					
北山武田病院								10			39					

京都市左京区の病院の「医療機器の台数」を知りたい場合は以下のようにする。「大分類」の▼をダブルクリックして「医療機器の台数」を選択する。



「市町村」についても同様の操作を行い京都市左京区を選択する。すると左京区内病院の医療機関の台数が表示される。

市町村		大分類	中分類	医療機器の台数			Totals	
市町村		大分類	中分類	CT	MRI	その他	Totals	
26108京都市左京区	京都大学医学部附属病院			8	6	15	29	
	公益社団法人信和会 京都民医連第二中央病院			2	2		4	
	医療法人 寿尚会 洛陽病院			1	1		2	
	医療法人一仁会脳神経リハビリ北大路病院			1	1		2	
	医療法人社団円珠会 京都大原記念病院			1			1	
	医療法人社団食原会 吉川病院			2	2		4	
	医療法人社団厚和会 京都下鴨病院			1	1		2	
	医療法人社団連徳会比较病院			2			2	
	医療法人種門会いわくら病院			1			1	
	医療法人高済会 高折病院			2			2	
	総合病院日本ハブレスト病院			2	2		4	
	Totals				23	15	15	53

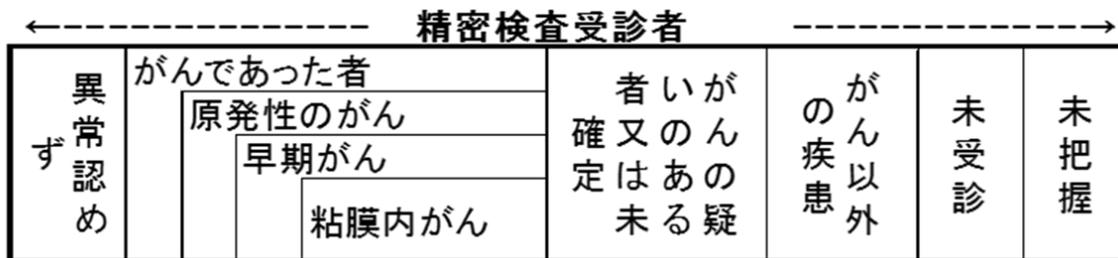
操作法 2: 健康増進事業報告がん検診を例に

2009年度より健康増進事業報告の内容が詳しくなり、がん検診については前年度受診者の追跡調査も含まれるようになったのがん発見率等の精度管理が可能になった。2009～14年度の6年間分の市町村データを収録した。

がん検診データには以下の変数名が含まれる。

個別集団…個別検診か集団検診かの2つの値をとる
 初回非初回…初回(過去3年以内に受診していなかったという意味で生まれて初めてではない)、非初回そして不詳(=総数-初回-非初回)の3つの値をとる(特に初期では初回非初回を把握していない、不詳が多い点に留意)。
 データ型…偶発症の有無別人数、前年度受診者数そして要精密検査者数の3つをとる。
 データ項目…データ型に含まれる項目。データ型>データ項目という関係であり、データ項目は**必ず**データ型の下に配置しなければならない。

データ型=要精密検査者数のデータ項目は複雑である。以下のように原発性のがん、早期がん、粘膜内がんは(再掲)と表示され「がんであった者」に含まれる。よって合計は(再掲)とあるデータ項目のチェックを外す必要がある(そのまま合計すると実数より大きくなる)。



大腸がんを例に説明する。初期画面は他と同様、行に都道府県、列に年となっている。

健康増進事業報告【大腸がん】

表: DATA 年齢階級: 性: 初回非初回: 個別集団: データ型: 市町村: 保健所: 医療圏: データ項目

合計(整数): DATA 年:

都道府県:

都道府県	年	2009	2010	2011	2012	2013	2014	Totals
01北海道		17,409	18,026	19,038	20,273	25,332	23,617	124,295
02青森県		5,861	7,200	7,888	5,808	9,415	9,720	45,992
03岩手県		8,270	7,804	8,206	7,710	12,187	9,539	53,716
04宮城県		11,953	4,544	14,600	13,513	20,323	14,968	80,001
05秋田県		6,244	6,459	6,116	6,331	8,376	7,385	40,912

【具体例】 個別検診か、集団検診か、初回か非初回かでがん発見率に違いはあるか？
 この問題に答えるためには、以下のような表を6年間分全国のデータで出せばよい。

		がんであった者(A)	前年度受診者数(B)	がん発見率(A/B)
個別	不詳			
	初回			
	非初回			
集団	不詳			
	初回			
	非初回			

全国データを6年分まるまる用いるので、まず行、列より都道府県と年は不要だから外す(上の変数リストにドラッグ&ドロップして移動させる)。

データ型は「要精密検査者数」と「前年度受診者数」の2つのみチェックを入れる。また下に来る「データ項目」は「がんであった者」をチェックするとともに「空白」も必ずチェックをいれる。これはデータ型が「要精密検査者数」の場合はデータ項目に「がんであった者」、「未受診」…といろいろなデータ項目があるのに対してデータ型「前年度受診者数」はデータ項目が無く空白だからである。

表: DATA 年齢階級: 都道府県: 性: 年: 市町村: 保健所: 医療圏:

合計(整数): DATA

個別集団: 初回非初回:

データ型: データ項目:

データ項目 (12)

Select All Select None

前年度受診者数にはデータ項目が空白である。だから前年度受診者数と「がんであった者」の両方に表示するには空白もチェックをいれる!

偶発症の前年度受診者数 (449,279)

がんであった者 (55,250)

がんの疑いのある者または未確定 (13,285)

がん以外の疾患であった者 (225,013)

未受診 (133,169)

未把握 (90,686)

異常認めず (173,672)

精密検査中または精密検査後に重篤な偶発症を確認 (151)

OK

その結果、以下のように、6年間でのべ4515万5020人が大腸がん検診を受け、10万3525人ががんと診断されている。

表	DATA	年齢階級	都道府県	性	年	市町村	保健所	医療圏
合計(整数)	データ型	データ項目						
DATA								
個別集団		データ型	前年度受診者数	要精密検査者数				
		データ項目		がんであった者				
初回非初回	個別集団	初回非初回						
	個別	不詳	5,920,158	18,219	5,938,377			
		初回	4,992,879	20,949	5,013,828			
		非初回	12,694,752	20,802	12,715,554			
	集団	不詳	3,565,788	9,581	3,575,369			
		初回	3,790,025	13,102	3,803,127			
		非初回	14,191,418	20,872	14,212,290			
		Totals	45,155,020	103,525	45,258,545			

発見率を出すには、がんであった者数÷前年度受診者数を計算すればよいが、残念ながらウェブ版データウェアハウスでは割り算等の演算ができないので、データを Excel にはりつけてがん発見率を算出する。発見率は初回は非初回より高く、個別検診は集団より発見率がやや高い傾向がある(下図は Excel の「条件付書式」で棒グラフ表示してある)。

(ハイライトして Excel へのコピペがうまくゆかない時は、一旦メモ帳にはりつけて、その後にはりつけるとうまくゆことあり)

		がんであった者(A)	前年度受診者数(B)	がん発見率(A/B)
個別	不詳	5,920,158	18,219	0.3%
	初回	4,992,879	20,949	0.4%
	非初回	12,694,752	20,802	0.2%
集団	不詳	3,565,788	9,581	0.3%
	初回	3,790,025	13,102	0.3%
	非初回	14,191,418	20,872	0.1%

棒グラフとヒートマップ表示

ウェブ DWH は計算機能は無いが、棒グラフやヒートマップ表示機能はある。表示させるには左上のプルダウンメニューをデフォルトは「表」となっているのを「バーチャート(棒グラフ)」「ヒートマップ(数値を赤色の濃淡で示したもき。全体、行、列の3メニューあり)」にそれぞれ変更すればよい。

ヒートマップ(列)	DATA	年齢階級	都道府県	性	年	市町村	保健所	医療圏
合計(整数)	データ型	データ項目						
DATA								
個別集団		データ型	前年度受診者数	要精密検査者数				
		データ項目		がんであった者				
初回非初回	個別集団	初回非初回						
	個別	不詳	5,920,158	18,219	5,938,377			
		初回	4,992,879	20,949	5,013,828			
		非初回	12,694,752	20,802	12,715,554			
	集団	不詳	3,565,788	9,581	3,575,369			
		初回	3,790,025	13,102	3,803,127			
		非初回	14,191,418	20,872	14,212,290			
		Totals	45,155,020	103,525	45,258,545			