

# 健診・保健指導データ分析のてびき

平成 21 年度厚生労働科学研究費補助金

循環器疾患等生活習慣病対策総合研究事業

各種健診データとレセプトデータ等による保健事業の評価に関する研究班

研究協力者： 藤井 仁 国立保健医療科学院人材育成部 主任研究官

研究分担者： 横山 徹爾 国立保健医療科学院人材育成部 部長

研究代表者： 水嶋 春朔 横浜市立大学大学院医学研究科情報システム予防医学  
教授

## 1.はじめに

### (1) 本稿の目的

本稿の目的は、日本の国保データを用い、標準的なレセプトと健診データの処理・分析方法を学ぶことです。具体的には、できるだけ特殊なソフトを使わずに(EXCELを用いて)、厚労省の「標準的な健診・保健指導プログラム(確定版)」に記載されたデータ分析をすることです。

なぜ国保データなのでしょう？

理由は3つあります。

**第一に、各都道府県で手に入れやすいことです。**

健保や共済のデータは都道府県にまたがっており、特定の団体しか持っていないことが非常に多いです。これに対し、国保データは各都道府県の国保連に頼めば、比較的容易に手に入れることができます。

**第二に、生活習慣病対策の対象を多く含むからです。**

言うまでもなく、働き盛りで若い人たちは医療機関にかかることも少なく、生活習慣病も喫緊の問題ではないことが多いと言えます。

**第三に、データのフォーマットの地域差が少ないことです。**

傷病名のデータ(後述します)などの数に差はあるものの、データ項目やデータ形式に大きな差がありません。どこの地方のデータでも、おおよそ同じ方法で解析することが出来ます。

これらの理由から、国保のデータを対象とします。

ただし、他の保険者のデータ形式も、極端に異なっているわけではありませんので、本稿の手法は流用することが可能です。

### (2) 注意

データの分析に取り掛かる前に、あなたが分析しようとしている国保の対象者数と、PC環境を確認してください。

概算ですが、国保の対象者数×10程度、レセプトは生じます。

30万人の国保加入者がいれば、最低年間300万行ほどのレセプトが生じると考えてください。20,000以下のレセプトなら、EXCEL2003で処理可能です。20,000-200,000程度なら、EXCEL2007で処理可能です。200,000-1,000,000程度なら、ACCESSで処理可能です。それを超える場合は、プログラムを組むか、SASなどの統計専用ソフトに頼る必要があります。むしろ、PC環境が何年も前のものである場合、扱えるレセプト数は上記の数よりも格段に減ります(2009年11月現在)。

実際のデータに本稿の内容を適用する場合は、第一にこの制約のことを考えてください。レセプト分析は容量との戦いです。30万程度の国保加入者のレセプトは1年で2GBを超えます。

## 【学習教材】健診・保健指導データ分析の手引き

---

これは新聞縮刷版 5 年分ほどにもなるデータなのです。

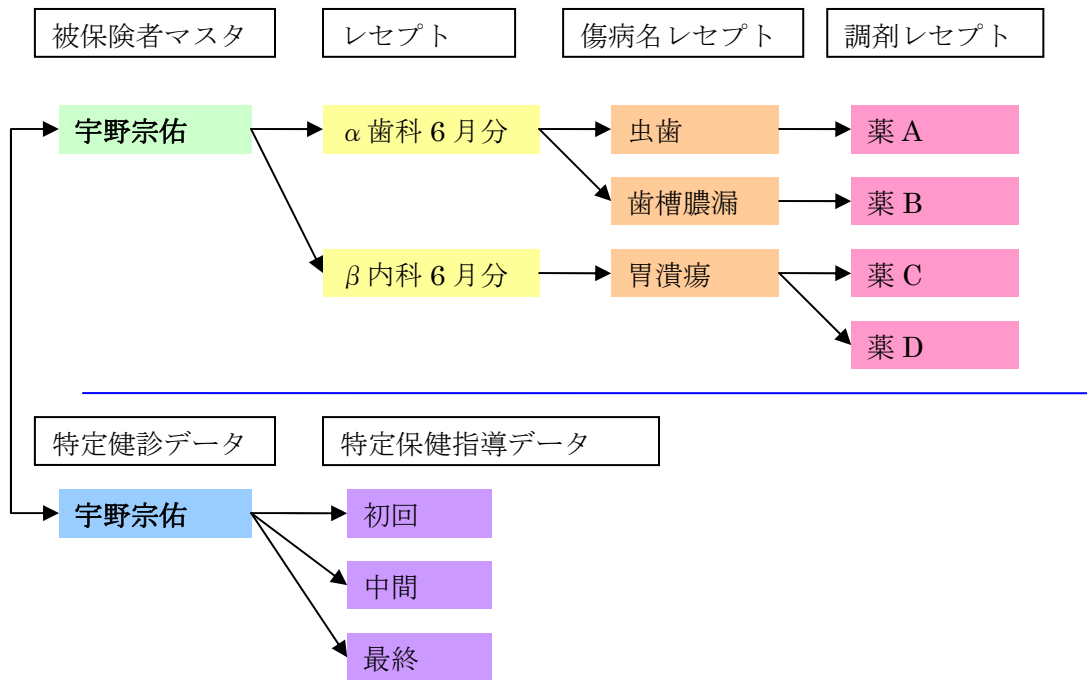
自分の PC 環境では無理だという場合は、データを間引いて分析する方向で考えてください。  
では、データの分析に取り掛かりましょう。

## 2. データセットの構築

### (1) データの性質と形式

国保データは多くの場合、図1のような構成をしています。

図1 国保データ構造



#### ●被保険者マスタ

国保の資格がある人の名簿です。当然ですが、公務員になったり、転出したり、死亡したり、国籍を失えば名簿から削除されます。逆の場合は名簿に加えられます。このデータについては、注意すべき点があります。例えば2009年9月の被保険者マスタを手に入れたとしても、それは何カ月もたってから事後的に変更されることがあります。期限切れの国保を使用した場合などに、このようなことが起こります。また、転出などの場合も、リアルタイムで更新されるわけではありません。当然ながら、上記のようなこと一転入出・転職・死亡などが多いほど信頼性は低くなります。分析する地域の転入出状況などは、分析の前に抑えておいたほうが良いでしょう。

#### ●レセプト

これに関しては多くを説明する必要はないでしょう。年齢、性別から、保険点数や入院日数、保険区分に至るまで、多くの基礎的なデータが含まれています。

レセプトデータの単位は、月にかかった医療機関数です。ある人が4月に歯医者に行き、接骨医に行き、精神科に行ったら、この人のレセプトは5月に3枚生じることになります。

## 【学習教材】健診・保健指導データ分析の手引き

(データはおなじ個人番号のものが3行できます)。

### ●傷病名レセプト

地方自治体の状況によっては、このデータは存在しません。傷病名が一つしか記載されていないような自治体では、多くの場合レセプトに埋め込まれています。一方、東京23区などでは、病名の数に制限がないため、20の病気に同時にかかっている場合、同じレセプト番号で20の病名データが存在することになります(データは同じレセプト番号のものが20行出来ることとなります)。

### ●調剤レセプト

本当はこのデータを活用できればいいのですが、ほとんどの自治体で利用できないのが現状です。利用できる自治体は、傷病名データの代わりにこちらを利用してください。こちらのデータを優先する理由は、傷病名データの信頼性が低いからです。周知のとおり、糖尿病の検査をしたい場合でも、重篤な糖尿病の場合でも、傷病名は「糖尿病」になります。ただし、このデータを利用するには、薬剤名を傷病名に読み替えるー「インシュリン」を「糖尿病」と読み替えるような操作が必要になります。

### ●特定健診データ

このデータに関しては、国が定めた形式(XML形式)に則っているため、全国共通でデータ形式が整っています。取り扱いしやすいデータです。

### ●特定保健指導データ

このデータも国が定めた形式(XML形式)に則っているため、全国共通でデータ形式が整っています。特定健診データと異なるのは、介入ごとにデータが増えていくことです。たとえば、積極的支援で初回、中間、最終と三回介入したとすれば、同じ人のデータが3回生じることとなります。データはほとんど共通しているので、多くの場合は最終データのみを用いれば事足ります。

ここまでで説明したとおり、それぞれのデータは単位が異なります。組み合わせて使用する際は、単位と組み合わせ方に気をつけて下さい。

たとえば、レセプトと特定健診データを組み合わせて分析するとします。その場合以下のようなパターンがあります。

## 【学習教材】健診・保健指導データ分析の手引き

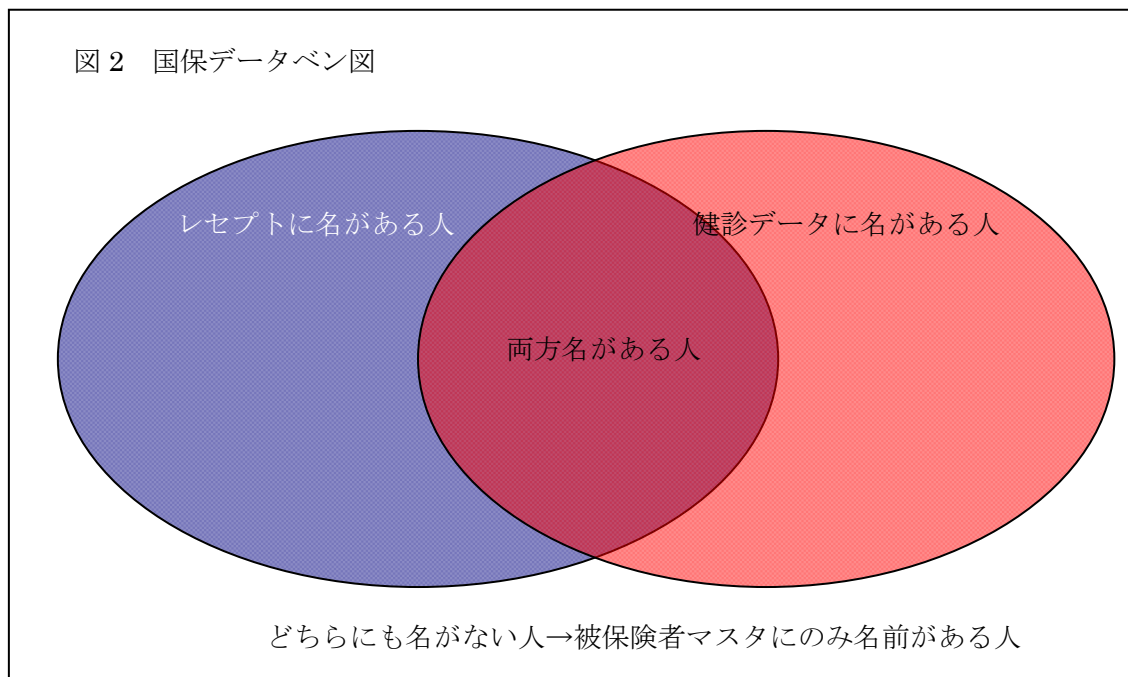
表1 レセプト単位の集計

No	氏名	診療科	決定点数	特定健診実施日	体重
1	A	歯科	200	2008/10/3	55
2	A	精神科	450	2008/10/3	55
3	A	耳鼻科	300	2008/10/3	55

表2 人別の集計

氏名	特定健診実施日	体重	診療科	決定点数
A	2008/10/3	55	歯科,精神科,耳鼻科	950

さらに、多人数になると、ベン図をイメージする必要があります。



レセプトに名がある人を基準にデータをそろえるのか、健診に名がある人を基準にデータをそろえるのか、両方名がある人だけを抜き出すのか、あらかじめ決めておく必要があります。以下のようなデータを突合するにもいろいろなやり方があります。

健診	実施日	体重
A	2008/10/3	56
C	2008/10/4	70
D	2008/10/5	82

レセプト	診療科	決定点数
A	歯科	522
B	耳鼻科	241
C	内科	695

## 【学習教材】健診・保健指導データ分析の手引き

表3 レセプトに名がある人基準

氏名	診療科	決定点数		特定健診実施日	体重
A	歯科	522		39724	56
B	耳鼻科	241			
C	内科	695		39726	70

表4 健診に名がある人基準

氏名	特定健診実施日	体重	診療科	決定点数
A	2008/10/3	56	歯科	522
C	2008/10/4	70	内科	695
D	2008/10/5	82		

表5 両方に名がある人基準（レセプト∩健診）

氏名	特定健診実施日	体重		診療科	決定点数
A	2008/10/3	56		歯科	522
C	2008/10/4	70		内科	695

表6 どちらかに名がある人基準（レセプト∪健診）

氏名	特定健診実施日	体重		診療科	決定点数
A	2008/10/3	56		歯科	522
B				耳鼻科	241
C	2008/10/4	70		内科	695
D	2008/10/5	82			

本稿では、私が一番使いやすいと考える方法でデータを組み合わせます。ただし、この方法は融通がききますが、目的によっては違う組み合わせ方のほうが適している場合もあるでしょう。いくつか例を挙げますので、皆さんも目的に合わせてデータ集計の最適な方法を考えてください。

## 【学習教材】健診・保健指導データ分析の手引き

### (2) EXCEL の操作

最低限、相対参照と絶対参照だけは理解してから、以後の文章を読み進めてください（関数の中に \$ マークをつける意味を理解してください）。

### (3) データの読み込み

取り扱うデータには、およそ三つの形式があります。

- ・ CSV データーレセプトなど
- ・ 固定長データー傷病名など
- ・ XML データー健診・保健指導データなど

こう書くとややこしいですが、いずれのデータもただのテキストデーターメモ帳で開くことができるデータです。ちなみに、メモ帳で開くと「景=¥・縉= m 二」&- &!&&KM 鶴郎 RM ケヲヲ;L:L 𠂔」こういう意味不明な文字列になるデータをバイナリデータといいます。話は少々脱線しますが、ひとつぐらいはテキストエディタ（高性能なメモ帳のようなもの）を用意しておくと、データのチェックの際に非常に役立ちます。お勧めは EMeditor という製品です。秀丸など他に有名なエディタはいくつもあります。EMeditor のみが 1000 万行を超えるデータを読み込むことが出来ます。レセプトは時に 1000 万行を超えるので、事実上このソフトだけが行数を気にせず操作することが可能です。今のうちにダウンロードされることをお勧めします（もうすぐフリー版の公開が停止されるといううわさがあります）。

#### ● CSV ファイル

CSV ファイルというのは、いわゆるカンマ区切りテキストです。

(CSV 例)

名前,目,科,属

イヌ,ネコ目,イヌ科,イヌ属

カモノハシ,単孔目,カモノハシ科,カモノハシ属

普通に EXCEL で読み込むことが出来ますので、特に問題はありません。

#### ● XML ファイル

XML データというのは、タグと呼ばれる記号でデータを区切ったテキストファイルです。

(XML 例)

<名前>イヌ</名前>

<目>ネコ</目>

<科>イヌ</科>

<族>イヌ</族>

<>で囲まれた部分をタグといいます。自分で定義することも可能ですが、健診・保健指導ファイルの場合は事前に定められたタグが用意されています。この形式のファイルは、普通に EXCEL で読み込むことが出来ますが、容量が大きくなる傾向があるので、しばしば EXCEL



## 【学習教材】健診・保健指導データ分析の手引き

の限界を超えることがあります。

### ●固定長ファイル

固定長ファイルというのは、テキストファイルですが、何バイト目に何のデータがあるということが決まっています。

(固定長例)

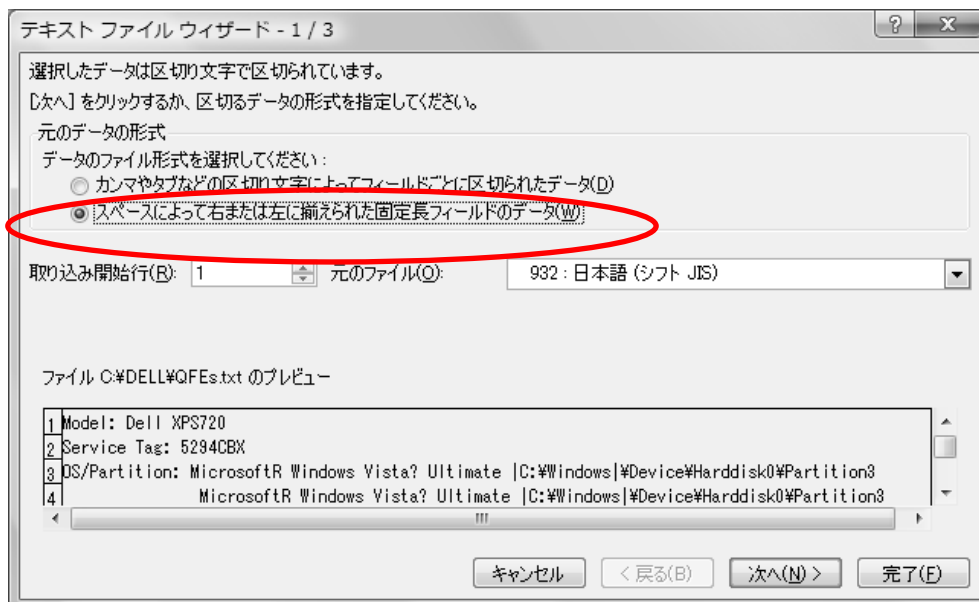
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

名前            目            科            属

イヌ            ネコ目    イヌ科            イヌ属

カモノハシ    単孔目    カモノハシ科    カモノハシ属

このファイルを EXCEL で読み込むには、多少の操作が必要です。普通に「ファイルを開く」から、「すべてのデータ (\*.\*)」を選択し、テキストデータをクリックすると以下のようなウィンドウが開きます。



ここで、固定長を選択し、「次へ」をクリックします。

## 【学習教材】健診・保健指導データ分析の手引き



ここで、データの区切り目の位置にマウスで線を引いてください。データ項目があまりに多く、何度も同じ操作をする場合は ACCESS のほうが便利ですが、たいていの固定長データは EXCEL でも処理できます。

身近にあるテキストファイルを、試しに EXCEL で分割して読み込んでみましょう。

### (4) データの突合

(3)のような操作で EXCEL データができれば、簡単な突合から始めてみたいと思います。

はじめに、レセプトデータに健診データを突合してみましょう。レセプト番号別(表1)に、レセプトに名がある人基準(表3)で突合します。

ユニークな(同じ番号が二つとない)個人番号があれば、それをキーにしてレセプトと健診データを突合します。なければ個人番号を作るところから始めます。番号と言っても数字である必要はなく、ふたつと同じものができなければ問題ありません。たとえば被保険者番号+氏名で十分です。

EXCEL では、

=被保険者番号の入ったセル&氏名が入ったセル

で、文字や数字を連結できます。

個人番号ができれば、vlookup 関数を用いてデータを突合します。

**=VLOOKUP(R2,\$A\$1:\$AV\$151,3,FALSE)**

レセプトの  
個人番号

健診データの位置  
(先頭は個人番号)

健診データの何列  
目を突合するか

近似値=true  
完全一致=False

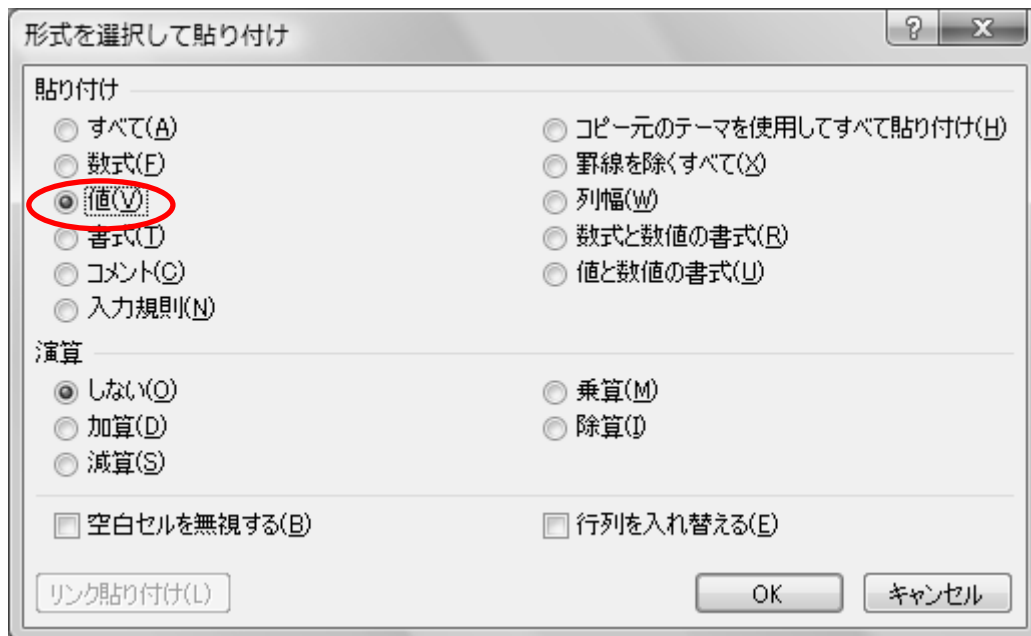
## 【学習教材】健診・保健指導データ分析の手引き

レセプトに名はあるけれど、健診には名がない人は「#N/A」と表示されます。レセプトにも健診にも名がある人は、指示された列のデータが表示されているはずですが、

問題なくできたでしょうか？

実際の国保データを利用する際は、データの型に気を付けてください。片方が数値、片方が文字列であると、一見同じに見えても突合できません。EXCEL の場合、桁数が 15 を超えると数値では突合できなくなりますので、両方文字列でそろえたほうがよいでしょう。

Vlookup 関数が残っていると、のちの計算速度が遅くなるので、Vlookup を入力した全セルをコピーし、「形式を選択して貼り付け」から値のみを張り付けておきましょう。貼り付けた範囲を指定したら、右クリックでメニューを呼び出し、形式を選択して貼り付けを選びます。ラジオボタンの「値」の所を選択しましょう。



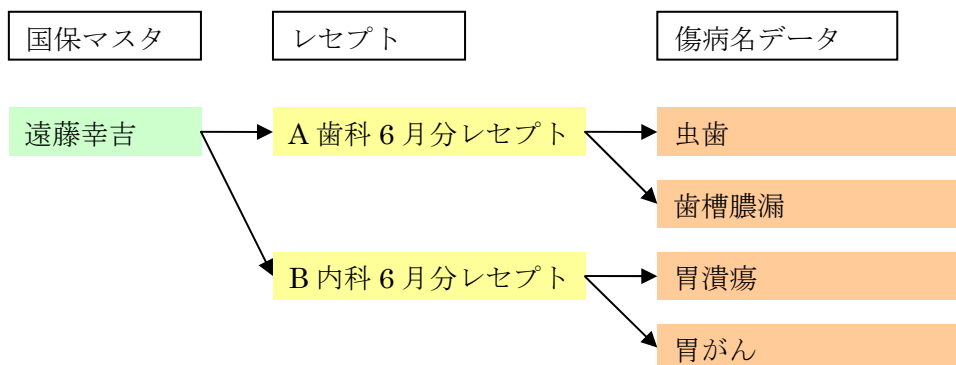
OK で関数なしの、値だけのデータにすることができます。

## 【学習教材】健診・保健指導データ分析の手引き

### (5) 傷病名データの処理

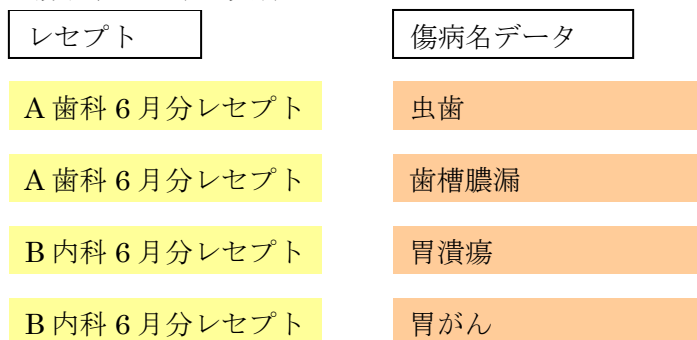
すでに説明したように、レセプトデータ 1 つにつき、いくつかの傷病名が付きます。

図 3 国保データ構造



これらのデータについては、いろいろなまとめ方が考えられます。傷病名データを基準にすると、以下のようなまとめ方になるでしょう。しかし、このまとめ方をするとデータ量が莫大になってしまいます。

図 4 傷病名を基準に突合



レセプトを基準にすると、以下のようなまとめ方になります。一見よさそうに見えますが、傷病名は多い場合で 40 前後付きます。と、いうことは、すべてのレセプトデータに「傷病名 1」～「傷病名 40」というデータ列を作る必要性が生じます。やはり、データ量は甚大になってしまいます。

図 5 レセプトを基準に突合



## 【学習教材】健診・保健指導データ分析の手引き

レセプト基準でまとめつつ、データ量を抑えるために、本稿では以下のようにまとめることにします。

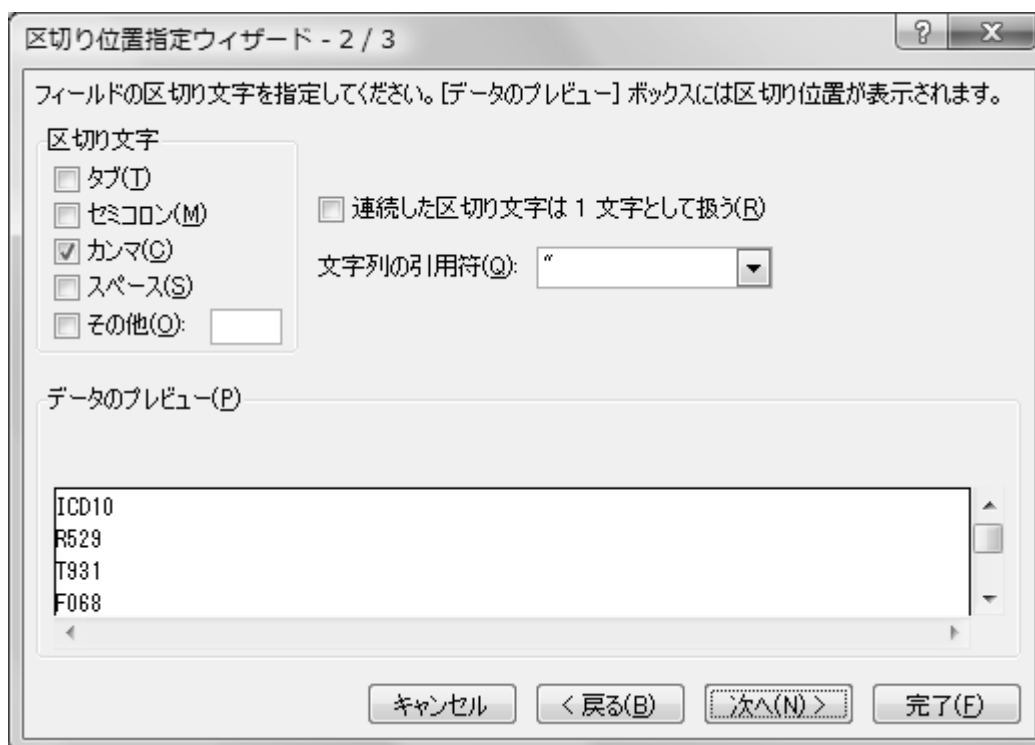
図 6 レセプトを基準に突合-傷病名は連結

レセプト	傷病名
A 歯科 6 月分レセプト	虫歯, 歯槽膿漏,
B 内科 6 月分レセプト	胃潰瘍, 胃がん,

このようにすれば、たくさんの傷病名列を作らなくても済みますし、いざとなれば EXCEL の「区切り位置」機能で図 5 の形に戻すことができます。「区切り位置」機能はツールバー「データ」→区切り位置で選択できます。



## 【学習教材】 健診・保健指導データ分析の手引き



上図のように選択していけば、図 6 のようにカンマで区切られたデータは、図 5 のようなセルごとに傷病名が入った状態に変形することが出来ます。

では、実際に図 6 のようなデータを作ってみましょう。

傷病名データは以下のようにしているとしましょう

レセ No	性別	年齢	病名	ICD10
11234	1	54	疼痛	R529
11234	1	54	糖尿病	E14
11234	1	54	習慣性便秘	K590
11234	1	54	慢性胃炎	K295
25351	2	37	統合失調症	F209
25351	2	37	便秘症	K590
25351	2	37	頭痛	R51

これを以下のような形に直します。

レセ No	性別	年齢	病名	ICD10
11234	1	54	疼痛, 糖尿病, 習慣性便秘	R529
25351	2	37	統合失調症, 便秘症, 頭痛	E14

## 【学習教材】 健診・保健指導データ分析の手引き

最初にデータをレセプト番号順並べ替えておきましょう。それができたら病名が最大でいくつあるかを調べます。

**=COUNTIF(A1:A81,A41)**

検索する範囲

検索する条件。この場合は A41 と等しいものを探せという命令になる。

この COUNTIF 関数は、「その行のレセプト番号と同じものが、前後 40 行の中にいくつあるか」を調べたものです。全体を検索範囲にしないのは、処理速度の問題です。この計算をした列の最大値が、病名の最大数になります。かりにそれが 40 だったとします。そうしたら以下の関数で病名に 1-40 までの番号を振ります。

**=COUNTIF(A41:A81,A41)**

次に、縦に並んでいる傷病名データを横に並べなおす作業をします。

**=IF(G2=1,E2,"")**

条件

条件を満たす場合の値。

条件を満たさない場合の値。  
""は空白を意味する。

1-40 までの番号を振った列が G 列だとします。その場合、このような IF 関数を、ICD コードの横に並べていきます

**=IF(G2=2,E3,"") =IF(G2=3,E4,"") =IF(G2=4,E5,"")**

すると、縦並びだった傷病名が横並びになります。

最後に&で傷病名データを結合します。

**=H2&","&I2&"," . . .**

&で文字列同士をつなぐことができます。あとで区切りを入れる時のために、間にカンマを入れておきましょう。

ここまでの作業をすべて終えたら、G 列の値が 1 の行には、ICD コードが連結した形で並んでいるはずですが、これで一応完成なのですが、ICD10 コードだけではなかなか病名が分かりづらいので、病名も入れておきましょう。

元が非常に長い固定長のファイルのばあい、余計な空白が多く全体が表示されません。処理上の問題はありませんが、気になる人は TRIM 関数で余計な空白を除去できます

**=TRIM(AQ2)**

これで傷病名の処理は完了です。すでに述べた Vlookup 関数を使い、レセプトデータに、傷

## 【学習教材】健診・保健指導データ分析の手引き

---

病名データを突合せましょう。この場合のキーはレセプト番号になります。

日本語表記の傷病名は表記にゆらぎがあるので、ICD10 コードを突合せさせて分析に使いましょう。N/A が非常に多く出ることがありますが、多くの場合それは歯科レセプトです。歯科以外のデータが突合できていれば成功です。

### (6) 保健指導データの処理

すでに説明したとおり、保健指導データは介入があるたびにデータが増えます。ただ、分析に途中の介入のデータを使うことはあまりありませんので、最終データだけを使うようにしましょう。手を抜いて、保健指導実施年月日（シート「保健指導」BJ 列）で降順に並べ替えておけば十分です。

いずれ分析の必要に応じて、Vlookup 関数でデータを引っ張ってくることになります。そこで、引っ張ってくる番号が複数ある場合（一対多の関係になっている場合）、Vlookup 関数は一番上の番号を引っ張ってくる性質があります。ですので、ここで日付の多い順に並べておけば、日付の最も大きいデータを参照します。

これで準備は完了しましたので、いよいよ分析にうつります。



# 【学習教材】健診・保健指導データ分析の手引き

## 3.標準的な健診・保健指導プログラム(確定版)のデータ分析

まず、分析の前に標準的な健診・保健指導プログラム（確定版）をダウンロードしてください。

<http://www.niph.go.jp/soshiki/jinzai/koroshoshiryo/index.html>

まずは、ここに挙げられている表を作ります。

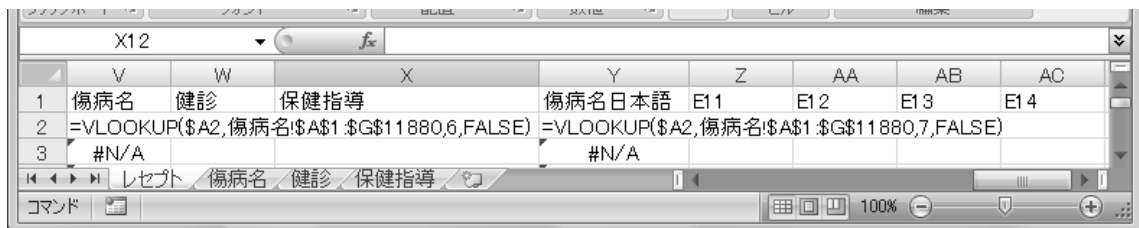
### (1) 様式 1-1 200 万円以上となった個別レセプト一覧

様式 1-1 200万円以上となった個別レセプト一覧(医療費の高い順)

番号	被保険者番号	年齢	性別	入院・入院外	費用額	基礎疾患					循環器疾患				その他	傷病名1	傷病名2	傷病名3	傷病名4	傷病名5
						高血圧症	糖尿病	高脂血症	高尿酸血症	虚血性心疾患	(再)バイパス・ステント手術	大動脈疾患	脳血管疾患	動脈閉塞						
1			男	入院		●				●	●									
2			男	入院		●	●			●	●									
3			男	入院		●				●	●									
4			女	入院			●						●	●						
5																				
6																				
7																				
8																				
9																				
10																				
11																				
12																				
13																				
14																				
15																				
16																				
17																				
18																				
19																				
20																				
21																				
22																				
23																				
24																				
25																				
26																				

準備として、ICD10 コードを突合できているでしょうか？ICD10 コードだけでは何の病気か分かりにくいという人は、ほかの列に日本語の傷病名を貼り付けておきましょう（重くなるので、値だけコピー&ペーストして関数を消すのを忘れずに）。

まず、レセプトと傷病名・健診結果の突合をしたシートをクリックします。そこで入外コードを探しましょう。保健区分等で番号が振られているはずですが、Ctrl+H で置換ウインドウを呼びだし、入院を表す数字を「1」、その他の数字を「0」に置き換えます。

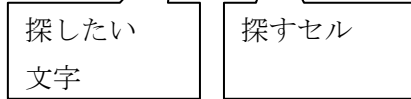


次に、あいた列に E11、E12、E13、E14 と打ち込みます。これらは、標準的な健診・保健指導プログラム（確定版） p146 で分析の対象とされている、糖尿病の ICD10 コードです。

次に、以下のコマンドを入力します。

## 【学習教材】健診・保健指導データ分析の手引き

=FIND("E11",V2)/FIND("E11",V2)



Find 関数は、指定した文字がどの位置にあるかを示します。

たとえば、探したい文字に「花」、探す対象のセルに「まんさくの花」と入力されていたら、Find 関数は「6」を返します。ここでは、糖尿病の ICD10 コード E11-E14 が、傷病名の中に含まれているかどうかを確認しています。割り算の形式にしたのは、含まれている場合を「1」で統一するためです。この操作が E12,E13,E14 でも終了したら、もう一度セルをコピーし、「形式を選択して貼り付け」で値のみを張り付けます。Ctrl+H で置換ウインドウを呼びだし、#VALUE!、#N/A のエラーを空白に置換しましょう。

次に Ctrl+A で全データを選択します。選択できたら、「挿入」→「ピボットテーブル」を選択します (EXCEL2003 の場合は後述します)。

個人番号	年齢	入外	決定点数	E11
6752	73	1.5	907400	
10032	66	1.235294118	712811	
15413	70	1.52173913	624626	
16211	68	1.854166667	560907	
22481	36	1.895833333	554877	
18848	71	1.942857143	536612	
7551	41	2.315789474	509156	
5911	74	1.945454545	507913	
22773	54	1.142857143	496420	
9609	57	1.466666667	491930	
7373	62	1	463267	
9517	72	1.789473684	450822	
6796	69	1.925	420484	
9944	73	2	412875	
21751	63	5	392663	
10069	57	1.625	378992	
8396	66	1.966666667	344902	
10372	75	1.916666667	325528	
18357	49	2	319371	
7875	73	1.842105263	313315	
20331	50	2	295353	
6274	60	1.111111111	286002	
23335	74	1.222222222	279685	
10200	71	1.951219512	251728	
8587	66	5.566037736	247549	
21282	32	1	230877	
11339	54	2	230014	
7522	70	1.777777778	227551	
8138	75	1.920634921	226700	
7939	68	1.925	224159	
9623	75	1.964285714	222179	

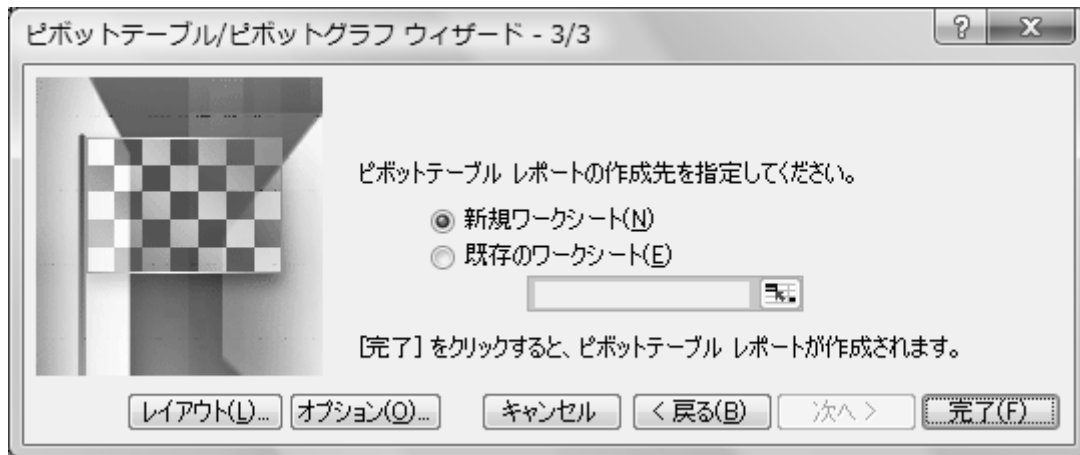
行ラベルに「個人番号」、値に「年齢」、「性別」、「入外」、「決定点数」、「E11」～「E14」をドラッグして入れます。「決定点数」は合計、それ以外は平均で集計します。値の欄に一旦入れたデータをクリックして、「値フィールドの設定」で集計方法を選択できます。

EXCEL2003 の場合は、以下のようになります。

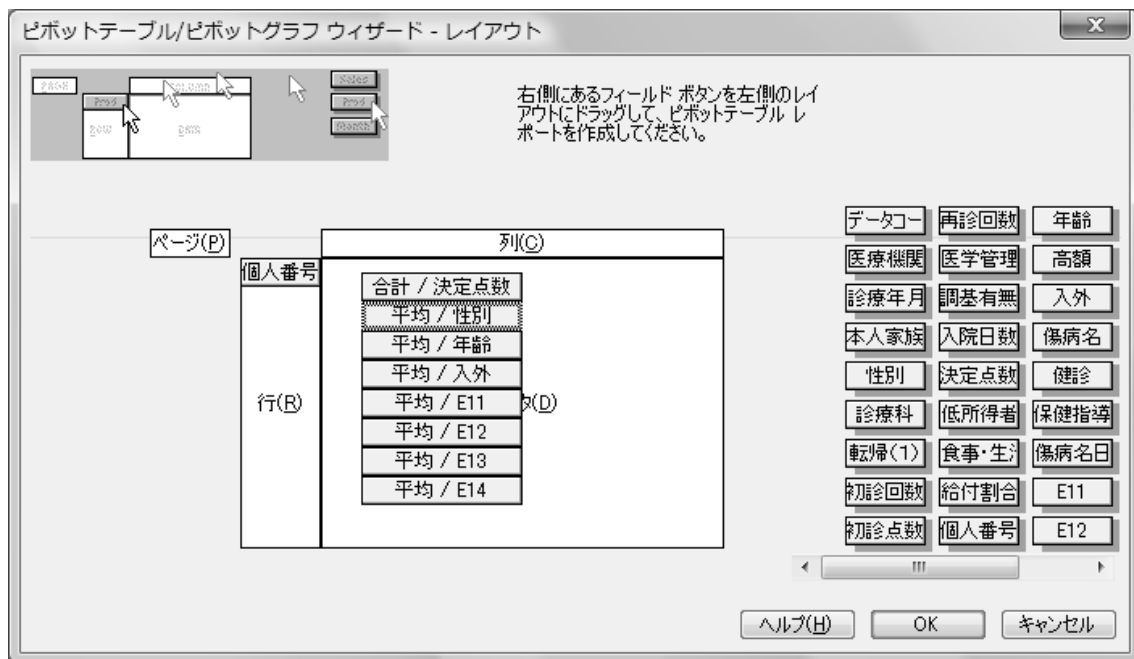
データの入ったセルをクリックしてから Ctrl+A で全データを選択します。選択できたら、

## 【学習教材】健診・保健指導データ分析の手引き

ツールバーから「データ」→「ピボットテーブルとピボットグラフレポート」をクリックします。



3/3 まで進んだら、レイアウトをクリックし、下図のように変数をドラッグしてください。合計・平均などは変数をダブルクリックすると、切り替えできます。

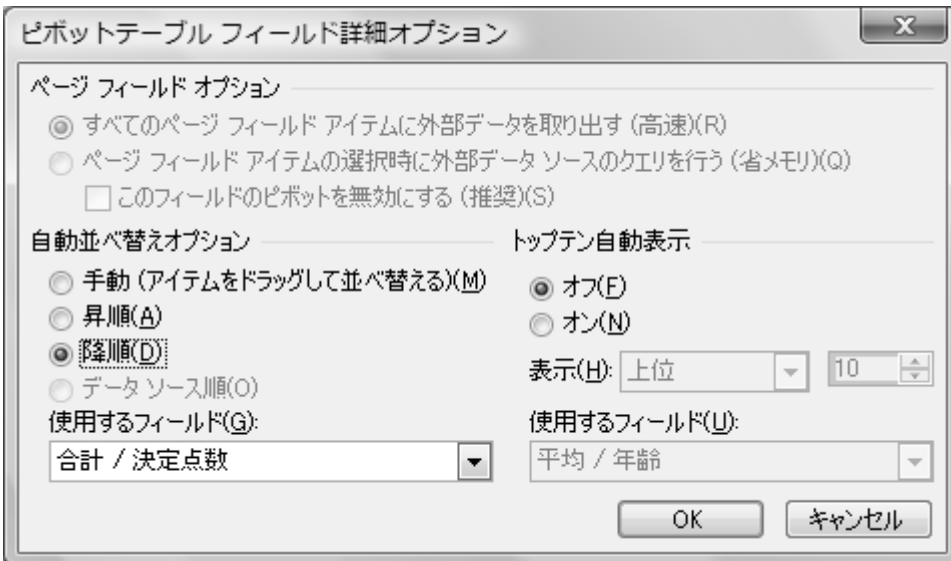


テーブルが完成したら、「データ」を下図のようにドラッグします。

2				
3	個人番号	データ	集計	
4	4841	合計 / 決定点数	101938	
5		平均 / 性別	2	
6		平均 / 年齢	69	
7		平均 / 入外	0.176470588	
8		平均 / E11		
9		平均 / E12		
10		平均 / E13		
11		平均 / E14		
12	4851	合計 / 決定点数	43131	
13		平均 / 性別	2	

## 【学習教材】健診・保健指導データ分析の手引き

「個人コード」を右クリックし、「フィールドの設定」を選びます。「詳細」をクリックし、以下のように選択します。



下図のように表ができたならばほぼ完成です。

	個人番号	合計 / 決定点数	平均 / 性別	平均 / 年齢	平均 / 入外	平均 / E11	平均 / E12	平均 / E13	平均 / E14
5	6752	907400	1	73	0.5				
6	10032	712811	1	66	0.764705882				
7	15413	624626	1	70	0.47826087				1
8	16211	560907	1	68	0.145833333				
9	22481	554877	2	36	0.104166667				1
10	18848	536612	2	71	0.057142857				1
11	7551	509156	1	41	0				
12	5911	507913	2	74	0.054545455	1			1
13	22773	496420	1	54	0.857142857				1

様式 1-1 と比較してみましょう。病名が多少少ないですが、同じものが作成できました（入外は、0 でないデータは何らかの形で入院していると判断してください）。

ここまで完成したら、個人番号をクリックし、Ctrl+A で全体を指定し、Ctrl+C でコピーしましょう。ピボットテーブルを右クリックしても「コピー」のメニューは表示されないので、ショートカットキーを用います。コピーした内容は、新しいシートに「形式を選択して貼り付け」で値のみを張り付けておきましょう。

## 【学習教材】健診・保健指導データ分析の手引き

### (2) 様式1-2 200万円以上となったレセプト基礎疾患（費用額別・疾患別）

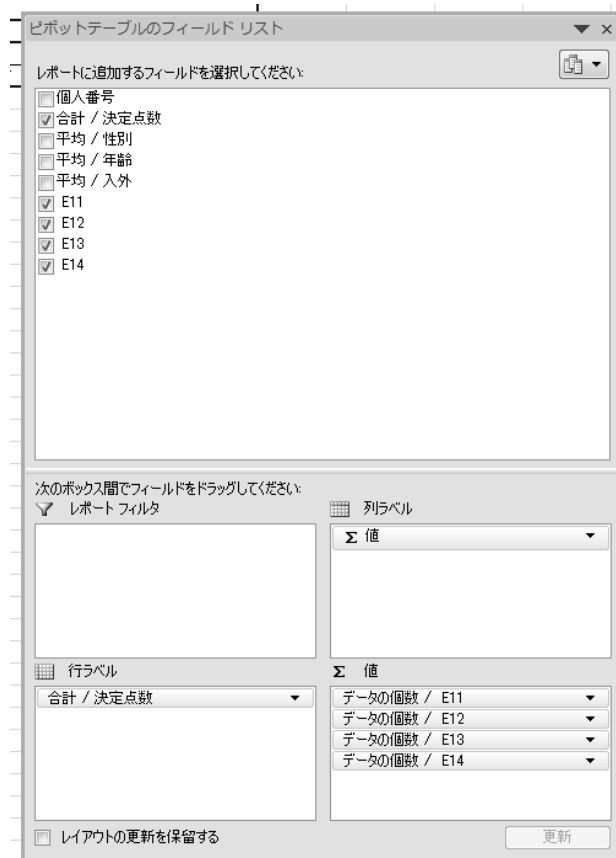
様式1-2 200万円以上となったレセプト基礎疾患（費用額別・疾患別）

	循環器疾患								その他		総合計	
	虚血性心疾患		大動脈疾患		脳血管疾患		動脈閉塞					
	件数	割合	件数	割合	件数	割合	件数	割合	件数	割合	件数	割合
1000万円以上												
900万円台												
800万円台												
700万円台												
600万円台												
500万円台												
400万円台												
420万円以上小計	血管に関する疾患											
	件数				割合							
300万円台												
200万円台												
合計	(再)血管病に関する疾患											
	件数				割合							

審査基準、高額レセ ←

より高額なものが血管病であることを知るため ←

様式1-1で作ったピボットテーブルの値をコピー&ペーストできたら、データの末尾一総計の行を削除します。データの左上「個人番号」をクリックしたら、Ctrl+Aで全体を指定し、もう一度ピボットテーブルを作ります。下図のように変数をドラッグしてください。行ラベル



は合計／決定点数です。

下図のようなテーブルができたなら、表側の決定点数を右クリックします。

## 【学習教材】 健診・保健指導データ分析の手引き

	A	B	C	D	E
1					
2					
3		データ			
4	合計 / 決定点数	データの個数 / E11	データの個数 / E12	データの個数 / E13	データの個数 / E14
5	42				1
6	57				
7	136				
8	152				
9	166				
10	214				
11	217				
12	240				

メニューから「グループ化」を選択し、以下のように数値を入力します。

すると、以下のような表が完成します。

	データ				
3		データ			
4	合計 / 決定点数	データの個数 / E11	データの個数 / E12	データの個数 / E13	データの個数 / E14
5	0-49999	14			227
6	50000-99999	2		2	53
7	100000-149999	3		1	17
8	150000-199999				2
9	200000-249999				4
10	250000-299999	1			2
11	300000-349999				4
12	350000-399999				
13	400000-449999				2
14	450000-499999				4
15	500000-549999	1			2
16	550000-599999				1
17	600000-649999				1
18	700000-749999				
19	800000-849999				
20	850000-899999	1		1	1
21	総計	22		4	320
22					

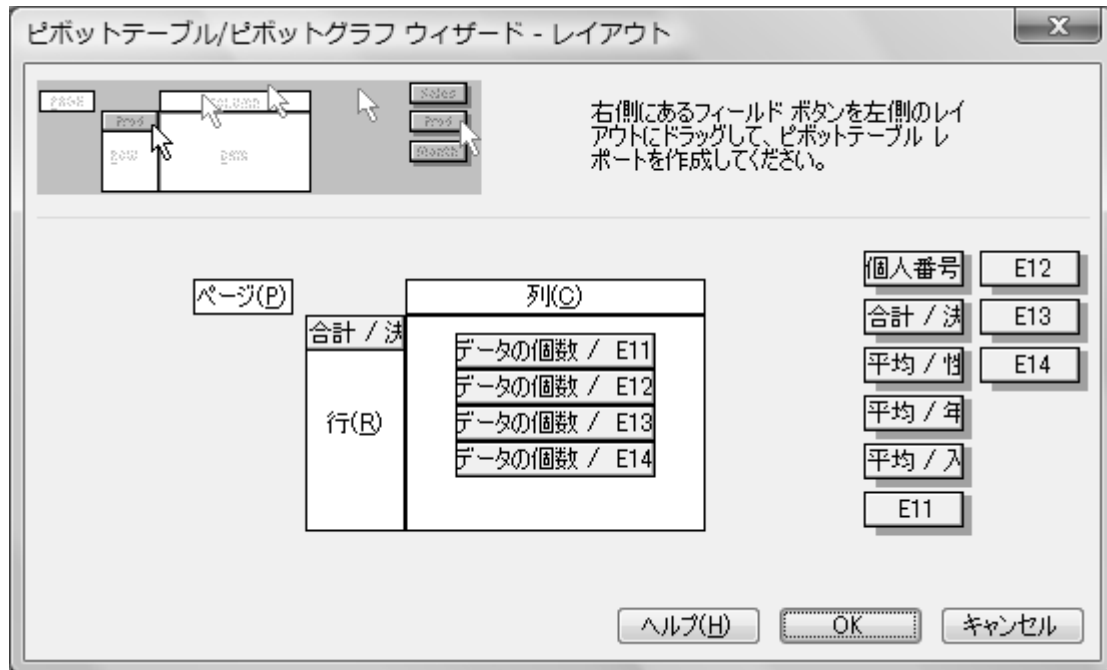
病名は糖尿関連だけですが、これで、ほぼ様式 1-2 と同じものができました。様式 1-2 に付加されている割合 (%) は、これらのデータから容易に計算が可能です (決定点数のデータ個数を出し、それで各病気の個数を割れば算出できます)

EXCEL2003 の場合は以下ようになります。

様式 1-1 で作ったピボットテーブルの値をコピー&ペーストできたら、データの末尾ー総計の行を削除します。データの左上「個人コード」をクリックしたら、Ctrl+A で全体を指定し、

## 【学習教材】健診・保健指導データ分析の手引き

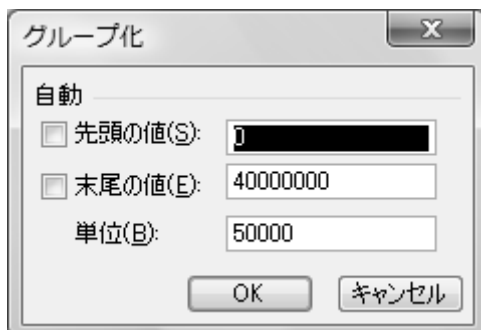
もう一度ピボットテーブルを作ります。下図のように変数をドラッグしてください。行は合計／決定点数です。



下図のようなテーブルができたなら、表側の決定点数を右クリックします。

	A	B	C	D	E
1					
2					
3	合計 / 決定点数	データ	集計		
4	42	データの個数 / E11			
5		データの個数 / E12			
6		データの個数 / E13			
7		データの個数 / E14	1		
8	57	データの個数 / E11			
9		データの個数 / E12			

メニューから「グループと詳細の表示」→「グループ化」を選択し、以下のように数値を入力します。



すると、以下のような表が完成します。

## 【学習教材】健診・保健指導データ分析の手引き

1					
2					
3		データ			
4	合計 / 決定点数	データの個数 / E1	データの個数 / E2	データの個数 / E3	データの個数 / E4
5	0-49999	14			227
6	50000-99999	2		2	53
7	100000-149999	3		1	17
8	150000-199999				2
9	200000-249999				4
10	250000-299999	1			2
11	300000-349999				4
12	350000-399999				
13	400000-449999				2
14	450000-499999				4
15	500000-549999	1			2
16	550000-599999				1
17	600000-649999				1
18	700000-749999				
19	900000-949999				
20	39150000-39199999	1		1	1
21	総計	22		4	320
22					
23					

病名は糖尿関係だけですが、これで、ほぼ様式 1-2 と同じものができました。様式 1-2 に付加されている割合 (%) は、これらのデータから容易に計算が可能です (決定点数のデータ個数を出し、それで各病気の個数を割れば算出できます)



## 【学習教材】健診・保健指導データ分析の手引き

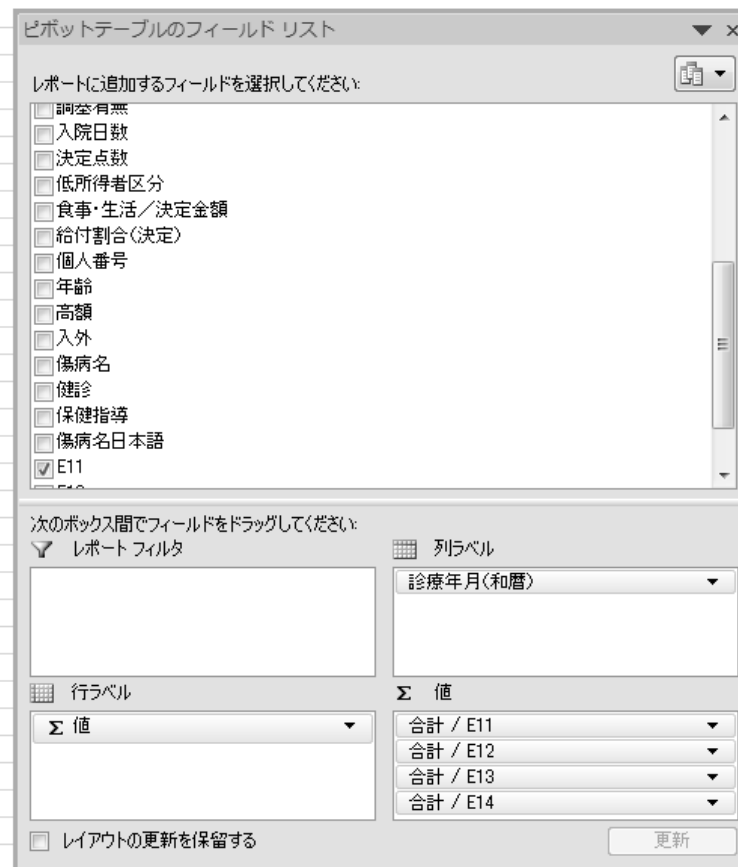
(3) 様式 5-7 一昨年・昨年・今年生活習慣病と診断された患者の重症化

様式5-7 一昨年・昨年・今年生活習慣病と診断された患者の重症化

	一昨年	昨年	今年	増減数(昨年と今年の比較)	増減率(昨年と今年の比較)
生活習慣病患者数(総数)					
糖尿病					
高血圧症					
高脂血症					
高尿酸血症					

この様式は、様式 1-1 を改造するだけでできます。

様式 1-1 のピボットテーブルをもう一度呼び出してください。フィールドリストが表示されていない時は、データを右クリックして「フィールドリストを表示する」を選択します。



上図のように変数をドラッグすると、以下のようなテーブルができます。

## 【学習教材】健診・保健指導データ分析の手引き

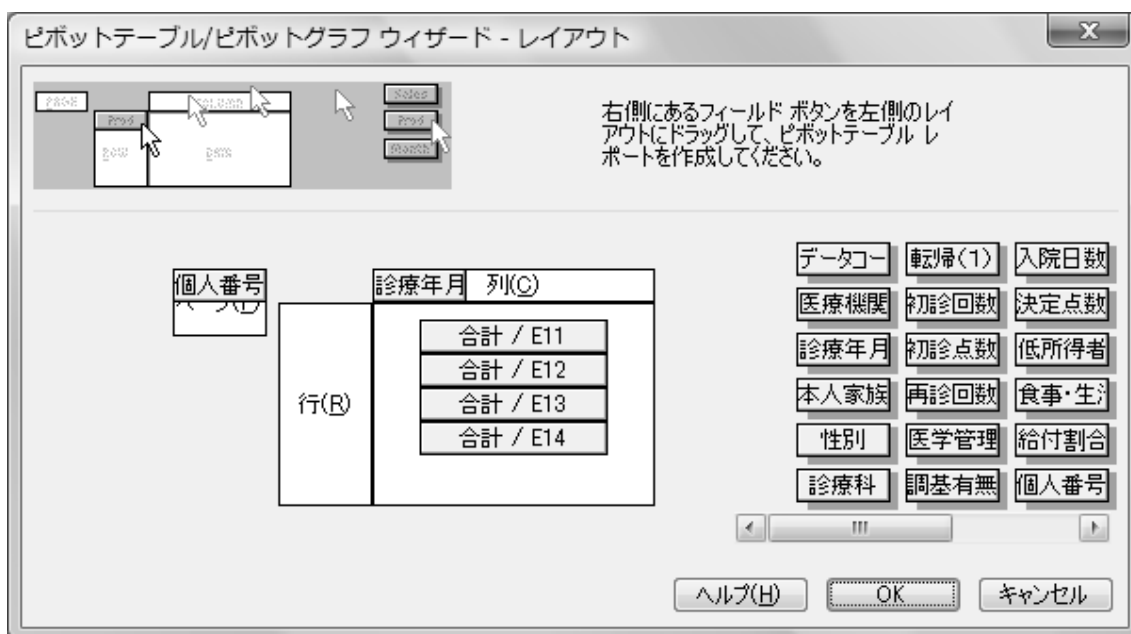
102	42003	42004	42005	42006	42007	42008	42009	42010	42011	42012	42101	42102	42103	総計
		3	6	5	6	6	6	8	6	11	9	13	10	89
	1	2	2	2	2	2	2	2	2	2	3	3	2	27
	6	158	151	173	171	161	175	180	159	169	150	163	149	1965

様式 5-7 では年単位でデータを見ていますが、この表は月単位でデータを見ています（上図の和暦データの 4 は平成、その次の 2 桁は年、最後の 2 桁は月を表しています：参考 23 区データ）。その差はあれど、操作は基本的に同じです。年単位にしたい場合は月をグループ化するといいでしょう。

EXCEL2003 の場合は以下のとおりです。この様式は、様式 1-1 を改造するだけでできます。

様式 1-1 のピボットテーブルをもう一度呼び出してください。

データを右クリックしてピボットテーブルウィザードを選びます。



上図のように変数をドラッグすると、以下のようなテーブルができます。

	42001	42002	42003	42004	42005	42006	42007	42008	42009	42010	42011
				3	6	5	6	6	6	8	6
			1	2	2	2	2	2	2	2	2
			6	158	151	173	171	161	175	180	159

様式 5-7 では年単位でデータを見ていますが、この表は月単位でデータを見ています（上図の和暦データの 4 は平成、その次の 2 桁は年、最後の 2 桁は月を表しています：参考 23 区データ）。その差はあれど、操作は基本的に同じです。年単位にしたい場合は月をグループ化するといいでしょう。このように、ほとんどの様式はピボットテーブルを利用することで作成できます。挑戦してみてください。