

## レセプト傷病分析の原理と手法--PDM法について

分担研究者 岡本悦司(国立保健医療科学院技術評価部)  
研究協力者 畑 栄一 (同研修企画部)

### 研究要旨

レセプト電算化によりレセプト情報の利活用は飛躍的に進展すると期待できるが、最大の障害は傷病分析である。従来の人作業に依存した主傷病分類法では、客観的な分類が不可能であるだけでなく、個人情報保護の面でも好ましくない。そこでレセプトに記載された複数傷病名を客観的かつ自動的に分析する原理(PDM法)を考案し、パソコン上で使用できるプログラムを開発した。本稿では、PDM法の原理と概念、ならびにその精度を向上させるための補正法やシミュレーションデータを使った検証法を解説し、その実際の応用例として宮城県名取市におけるインフルエンザワクチン効果評価事業について紹介する。

個人情報保護しつつ、レセプト情報の有効な利活用をはかるためには、含まれる情報が人手を介することなく、全て機械的に処理されることが望ましい。紙レセプト主体の現在では、レセプトのパンチ入力外部委託され、とりわけ傷病分類は人が手と目で判断することがふつうである。レセプトの電子化がすすめば、パンチ入力の委託は必要なくなるが、傷病分類が人手に依存するものであることに変わりはない。

また従来の人判断で複数傷病の中から任意に主傷病を選択するいわゆる主傷病法では、レセプト情報を定量的かつ客観的に処理することは困難である。

こうした問題を解決し、個人情報保護をはかりつつ電子化レセプトを有効活用する目的で、複数傷病の記載されたレセプトの日数、点数といった情報を客観的かつ自動的に分析する原理を考案し、それをWindowsパソコン上で使用できるプログラムを開発した。

#### 【推計法の必要性】

複数傷病の記載されたレセプトにおいて、傷病ごとの日数、点数が求められるとすれば理想である。しかしそれは不可能であり、唯一可能なのは真に近い値を「推計値」するしかない。したがって真の値はわからないが限りなく近い「推計法」を開発することしか方法はない。

その根拠は

方程式の解が求まるのは変数 $\leq$ 方程式でなければならず、変数 $>$ 方程式では求まらない。という数学の公理にある。レセプトには常に1以上の傷病名が記載されている。つまり傷病名数 $\geq$ レセプト件数となる。傷病名数 = レセプト件数(つまり全レセプトに傷病名が一つしかない)ならそもそもPDM法など不必要だから、複数傷病名レセプトを分析するには何らかの方法で推計するしかない。

#### 【説明】

3件のレセプトを1,2,3で、傷病名A,B,Cの一日当点数をa,b,cで、日数をN,点数をPであらわし、レセプト1にはA,B、レセプト2にはA,B,C、レセプト3にはA,Cの傷病名が記載されている。す

ると3つの等式ができる。

$$P1=N1*(a1+b1)$$

$$P2=N2*(a2+b2+c2)$$

$$P3=N3*(a3+c3)$$

これら3つの式で  $a1, a2, a3, b1, b2, c2, c3$  の7つの数値が求められるか？

7つの変数を求めるには式は7つ以上いる。しかるにレセプト件数(式の数)は常に変数の数より少ない。このため7つの数値を求めることは不可である。

もっとも各レセプトの各傷病ごとの点数や日数が同一である場合は事情が異なる。たとえば3つの傷病が同じ点数, 日数で  $a, b, c$  と3つの変数なら

$$P1=N1*(a+b)$$

$$P2=N2*(a+b+c)$$

$$P3=N3*(a+c)$$

となり, 必ず解が求まる(傷病名数が119もあっても一般化逆行列を用いれば瞬時に求まった)。

#### 【重回帰分析】

多変量解析の観点からレセプトをみると, 点数, 日数は目的変数, 多数の傷病名は説明変数となる。重回帰分析を用いない理由について説明する。

日数や点数を目的変数に, 119の傷病分類ごとの傷病名の出現数(無ければゼロ, 1,2,3---)を説明変数に重回帰分析を行うことは可能である。そして各傷病分類の重回帰係数は日数や点数に対する各傷病の寄与度を意味しており「重み」として使うことは考えられる。しかしこの試みは頓挫し断念した。その理由は以下のとおり。

重回帰係数は必ず負値がでる

重回帰分析は可能だが, 119傷病の各係数の相当数が負値になる。点数を比例配分するPDM法では負値はつかえず, このことを数学上では「モデルが破綻している」と呼ぶという。おびただしい負値の理由としては多重共線性の問題や線型性への疑問が考えられる。ただ, 医療機関の種類や患者属性等を厳格に補正すれば全て正の係数になる可能性は残されている。今後厳格な条件をつけて重回帰分析を試みる余地は残されているかもしれない。

重回帰係数は件数が多くなければ信頼できない

重回帰分析は, 説明変数よりもケース数が相当大きくなければならない。100件のレセプトを119の説明変数で分析することは困難である。むろん統計ソフトにデータをつっこむと答えはでるが  $p$  値が0.8にもなったりして「こんな数値が出ましたが当たるも当たらぬも八卦」となる。事実, あるデータセットでは極端に大きな係数がでた傷病が別のデータセットではマイナスになったりと, とても安定した分析に耐えなかった。

原点を通る回帰では正しい推計ができない

通常重回帰分析では

$$Y=B0+B1X1+B2X2+---+BnXn \quad (B0 \text{ は定数項})$$

という回帰式を考え, 119傷病について得られた係数  $B1 \sim B119$  に各傷病の出現数をかけ, 定数  $B0$  を足すと合計点数に一致する。つまり傷病名が全く無いレセプトでも  $B0$  点はあると考える。

しかしレセプトでは傷病名ゼロなら点数もゼロと考えるので(実際には傷病名の無いレセプトは

たまにあり、そんなレセプトでも点数はちゃんとある・・・もっともこれらは記入もれとおもわれる)

$$Y=B1X1+B2X2+---+BnXn$$

と定数項(intercept)の無い回帰を使わなければならない(SPSSでもExcelでもデフォルトは上記のような「定数項有りの回帰」だが「原点を通る回帰」をクリックすれば原点を通る回帰を行える)。

定数項の無い「原点を通る回帰」では、負値の係数は少なくなる(しかし決してなくなる)。ところが定数項有りの通常の回帰式のように、各傷病の係数に出現数をかけて合計しても合計点数Yに一致せず、通常20~30%小さい数字にしかない。

この理由がわからずSPSSに問い合わせたところ、SPSS国際サポート本部に照会していただき、以下のような回答をいただいたので参考まで転載する。

---

#### SPSS国際サポート本部からの回答

n個の点 $\{(x_i, y_i)\}$ ,  $i=1$  to  $n$ , に直線をあてはめた場合、その直線と点との残差の合計は0になります。これはDraperとSmithの「Applied Regression Analysis 第3版」の26ページで証明されています。予測値と残差を保存した場合、従属変数、予測値、残差は定義上次の式に従います。

$$\text{従属変数} = \text{予測値} + \text{残差}$$

これら3つの変数をそれぞれ合計した場合、残差の合計が0になるという事実から、従属変数の値の合計が予測値の合計と一致するという結論に至ります。

しかしながら、定数項を回帰から除外した場合、これはあてはまりません。定数項を除外することは、あてはめの線が原点を通るようにすることです。仮に次の3つのデータポイントがあるとします。

x	y
80	100
90	110
340	130

これに最小二乗法による線をあてはめると、傾きは0.1切片が96あたりになります。この線を原点を通るようにするには、xの平均あたりで回転させ傾斜を急にするので、最初の2点では予測値を下げ、3番目では予測値を上げるという事になります。モデルは最初の二つのxに対するyを過小評価し、最後のxに対するyを過大評価しています。これにより、残差の合計は正の値となり、予測値の合計は観測された従属変数の値の合計よりも少なくなります。

It is true that if you fit a straight line to  $n$  points  $\{(x_i, y_i)\}$ ,  $i = 1$  to  $n$ , then the sum of the residuals about the fitted line is 0. This is proven in Draper and Smith, Applied Regression Analysis, 3rd edition, page 26. If we save the predicted and residual values, the dependent, predicted, and residual variables by definition obey the formula

$$\text{dependent} = \text{predicted} + \text{residual}$$

If we sum the three variables, the fact that the residuals sum to 0 leads to the conclusion that the sum of the dependent values equals the sum of the predicted values.

**THIS IS NOT TRUE, HOWEVER, IF ONE LEAVES OUT THE CONSTANT TERM FROM THE REGRESSION.**

Leaving out the constant term forces the fitted line to pass through the origin. Consider three data points

x	y
80	100
90	110
340	130

If you fit the least squares line to these points, the slope is about 0.1 and the intercept

around 96. Forcing the line to pass through the origin steepens it by rotating it about the mean of the x's, having the effect of decreasing the predicted values at the first two points and increasing it at the third. The model underestimates y for the first two x's and overestimates it for the last, with the net effect that the sum of the residuals is positive and the sum of the predicted values is less than the sum of the observed values.

というわけで、目的変数(従属変数)の値の合計は定数項ありの場合 予測値の合計と一致しますが、なしの場合は原点を通るように強制するのでずれが生じ、残差の合計が 0 にならないために一致はしない、ということになります。

---

### 線型計画法

変数にたとえば5以上という条件をつけて行う線型計画法という手法も検討したので付言する。専門家の意見では、偏回帰係数の解釈まで要求されるような回帰方程式で、100 変数で実行して、解釈可能な結果は出ない、ほぼ必ず多重共線性や抑制効果が生じる、一説によると説明変数の数は2つか3つまでしか無理だという。単に予測値と実績値が近ければよいのであれば、説明変数は 100 でも 200 でも構わない。でも線型計画なら符号条件が付けられるから納得できる結果が得られるのではないかと期待しても結局のところ、分析者が課した範囲条件の境界に解が定まる。つまり、非負と制約すれば 0 となり、10 以上という制約をおけば 10 になる・・・結局のところ分析者の指定をそのまま出力しているのと同じ結果になる。ということでレセプト傷病分析への線型計画法の適用は見送った。

### 重回帰分析とは目的が異なる

重回帰分析も林の数量化Ⅰ類も全て個々のケースの目的変数を推定するための手法です。たとえば傷病 A,B,C がある 15 日の入院レセプトの予想点数は？といった個々のレセプトの点数の予測のための手段が重回帰分析であり数量化理論といえる。しかしここでの目的は、個々のレセプトの点数を予測するための手段ではなく、レセプトの集合(データセット)における傷病別割合を推計することであり、傷病名 A,B,C のレセプトの点数は何点かを予測するための手法ではない。したがって重回帰分析とは目的が異なる。

### 【PDM 法の原理】

PDM(Proportional Disease Magnitude)法はレセプトの点数や日数を、それに記載された全ての傷病名に一定の「重み(magnitude)」を与えて比例配分してゆく分析法である。PDM 法によれば、客観的かつ自動的な傷病分析が可能となり、もし共通の「重み」を用いれば、たとえばA市とB町国保ではどちらが糖尿病の医療費がかかっているか、またC村で糖尿病対策をする前と後とで糖尿病の医療費が増えたか減ったか、等を客観的に比較することも可能となる。裏返せば、では傷病ごとの「重み」をどう決めるか？が課題となる。

重みは基本的にどのように決めてもよく、全ての重みを 1(つまり全ての傷病を等しいとする)としてもかまわない。そうすると単純に傷病名の出現数に比例した配分となる。重みの決め方としては、たとえば DRG も一種の重みといえるし、外国では専門医にアンケートして重みを測定しようという試みもある。PDM 法では、これまで患者調査で得られた「主傷病となる確率」や社会医療診療行為別調査で得られた傷病ごとの「一件当たり点数・日数」を用いたこともある。

### 【PDM 法開発の経緯】

「全ての傷病名を分析し、点数や日数といった医療の資源消費を定量的かつ自動的に傷病分類できないか」そう考えた筆者は 94 年頃「診療報酬明細書による傷病構造の解析」研究に着手(94 年度文部省科学研究費奨励研究)。95 年 6 月米国、シカゴで開催された第 12 回 Association of Health Service Research においてその基本概念を発表し、同年 10 月山形で

開催された日本公衆衛生学会で開催された「第 1 回レセプト情報の活用を考える自由集会」において BASIC で組んだ初歩的なプログラムを公表した。PDM(Proportional Disease Magnitude)法と名付けられた本手法の概念は、厚生統計分野の代表的な学術誌である「厚生指標」96年6月号に掲載された。

PDM法の最大のネックは、記載された多数の傷病名をコード化し入力しなければならない「手作業」にあった。レセプト電算化も期待されたほどには進展せず、実用化へは遠い道のりの状態が続いた。1999年には健康保険組合連合会による「レセプト分析による医療機関評価」研究が行なわれ、そこで PDM 法により傷病構造を補正して医療費の額や在院日数の長短を異なった医療機関間で比較することが試みられた。この時作成された ExcelVBA プログラムはインターネット(resept.com)上で公表され、関係者の関心を集めた。しかし、このプログラムも誰でも気軽に使用できるほどユーザーフレンドリーではなかった。

2001年、岡本は旧国立公衆衛生院に移り、愛媛県、宮城県の国民健康保険団体連合会が実施している全疾病入力レセプトの解析に着手。また本研究に分担研究者として参画、その助成によりそれまでコンピューターに強い人でなければ使用できなかったプログラムが Windows 上で誰でも簡単に使用できるプログラムとして完成をみた。

2002年10月24日埼玉県で開催された「第5回レセプト情報の活用を考える自由集会」で本プログラム(Ver.1)は公表され、その後の研究の進展をふまえて、本年度中には Ver.2 が完成する見通しである。

#### 【PDM法の個人情報保護上のメリット】

PDM法は傷病分類を自動化することにより、手作業の労力と時間を節約するのみならず、個人情報保護の面でも効果がある。データ入力作業はレセプトが完全電算化されれば不必要になるが、傷病分類だけは自動化は不可能で、これまで人間がレセプトを見て判断するしかなかった。熟練したレセプト点検職員は1日に1000件近いレセプトを分類できるが、それでももし年間12億件のレセプト全件を分類するとなると、のべ120万日・人もの人にレセプトを「見せる」必要がある。

どんなに契約書で守秘義務を課してもプライバシー漏洩の危険は、関わる人が多くなると必然的に大きくなる。プライバシー保護を100%確保するには、医療機関から提出された電子化レセプトが、誰からも中身をみられることなく、コンピューター内部で全て処理される、のが理想である。PDM法により、単純集計はもちろん、傷病分類まで、人手を介することなくブラックボックスの中で処理され、結果の表だけが出力されるようになり、プライバシー保護は万全になる。

#### 【日数と点数の扱い】

PDM法ではレセプト総点数を日数と一日当点数に分離して分析している。その理由は日数は1~30(大の月は31)のどの数値でもとりうる自然現象(確率事象)であるが、点数は自然現象ではなく、人為的に決められた数値であることにある。

診察料を例にとると病院の初診料は250点、診療所は270点と決められており、病院のレセプトであれば初診料は必ず250点であって270点である可能性はない。また初診料は一回しか請求できず同一レセプトに2回の請求はない。他の診療行為はもっと複雑だが、とことん細分化してゆくと必ず点数表で決められた点数の和になる。このように日数と点数とは異なった性質の情報であることから分離して分析することがベターといえる。

研究の形態としても、日数は受療状況であって疫学研究に該当し、点数は金額であって経済

研究に当たる。たとえばインフルエンザワクチンの有効性評価は疫学研究であって経済研究ではない。したがってインフルエンザワクチンの有効性評価は日数で行うべきであって点数で行うべきではない。この区分は2002年7月より疫学研究倫理指針が施行され、疫学研究については経済研究とは異なった扱いが求められるようになっただけに重要である。

レセプトが日数という疫学情報と点数という経済情報を合わせもっている点はきわめて豊富な情報源であるといえる。

【平均値を重みに使うことについて】

PDM法プログラムでは「重み」として日数も点数も傷病ごとの単純平均値が用いられている。すなわちある傷病Aの「重み」はその傷病を含む全てのレセプトの日数及び点数の単純平均を算出し、それを重みとして各レセプトごとに比例配分するわけである。いささか乱暴な感じもするが、レセプト全体における特定の傷病の日数点数の合計値だけを知りたいのであれば単純平均でよい。

その理由は、単純平均に個数をかければ、分布がどのようになっていようと同じ合計値になるからである。たとえば人口1万人のA市とB市が一人当たり所得がいずれも100万円であったとしても、所得分布もA市B市同じとは限らない。平均は同じでもA市の方がバラツキ(分布)が大きい、すなわち貧富格差は大きいかもしれない。「所得を比較する時に単純平均で比較するな」は経済学の大原則であり、これが単純平均に対する不信となっているようだ。しかし、A市とB市の所得分布がどうであれ、A市B市の所得総額が100億円であることは間違いのない。したがって両市の所得総額を知りたいのであれば単純平均でよく、むしろ中央値や最頻値は不適當なのである。

傷病名のような名義変数を数値化(数量化)する分析法に林の数量化I類がある。これは、たとえば数学が好き・嫌い、といった生徒の属性で数学の得点を予想する多変量解析法である。予想される数学の得点は外的基準、好き・嫌いという属性はアイテムと呼ばれる。PDM法の重みも、たとえば糖尿病の重みは、あるレセプトで糖尿病という傷病名が有る・無い、で点数がどうなのか予想するものである。119分類の傷病名について重みを算定することは、ある傷病名が有る・無い、をアイテムとして林の数量化I類を119回くりかえすことに相当する。

林の数量化I類は多変量解析の手法であって、たとえば数学が好き・嫌い、国語が好き・嫌いというふうに複数のアイテムで予測を試みるが、アイテムが一つの場合、たとえば糖尿病が有る・無い、でレセプトの点数を予測する場合は、単純に糖尿病の有るレセプト、無いレセプトの平均値をそのまま使うことになる。「常識で考えてもこれは妥当であろう」とされる(藤沢偉作「楽しく学べる多変量解析法」現代数学社63頁)。

ちなみにある市の外来レセプトについて、119の傷病分類ごとに、その傷病の有るレセプト、無いレセプトの点数をプロットしたのが【図1】である。明らかに、ある傷病の有るレセプトの点数はそれぞれ大きく異なっているが、ある傷病の無いレセプトはおしなべて全体平均に一致している。すなわち、ある傷病を持つレセプトが全体に比べて少数である場合は、その平均値は全体平均からの相対的な乖離を示していると考えられる。

もし全傷病が同じと仮定するとその重みは当然ながら全体平均となるが、その逆、つまり「全傷病は同じではない」とすればこれら傷病別平均値を重みとすることは妥当であろう。

PDM法はあくまでレセプトの集合の中のある疾病の総額を求めるものであって、個々のレセプトにおける特定の疾病の額を推計することは、計算の過程でむろん可能だが、その値はあまり信用できないだろう。それは平均所得100万円のA市内で誰か市民をつかまえて所得を調べても

100万円でない可能性の方が大きい。所得は決して正規分布しないからである。

傷病ごとの点数も正規分布するかどうかはまだ十分には解明していない。治療内容が定型的な傷病なら正規分布しPDM法による個々のレセプトにおける推計値もある程度信頼できるであろうが、バラツキの大きい傷病では重みによって比例配分された点数がどれだけ信頼できるかは疑問である。さらに複雑なことはバラツキの小さい傷病も大きな傷病も同じレセプトに「同居」しているということである。極端な場合、レセプトに10の傷病が記載されていても医療費の大半はうち一つの傷病に費やされている、といったことも多々あるだろう。

したがって現時点ではPDM法はレセプト全体における傷病の合計値の推計法であって、個々のレセプトの傷病の点数や日数を推計する上では慎重であるべきと考えられる。

【平均値の算出法】

ある傷病名の記載されたレセプトを全て集めて平均値を出す、傷病名とはよく似た傷病名は同じ傷病分類に一括する。たとえば糖尿病、糖尿病性腎症、糖尿病性網膜症はそれぞれ異なった傷病名だが、119分類ではいずれも糖尿病[402]として一括される。したがって、これら3つの傷病が記載されたレセプトは、402,402,402と同じ分類番号が3つ並ぶ。この場合、PDM法では402というひとつの傷病名にまとめることはせず、402が3回出現するものとして重複カウントする。その根拠を【例題】を用いて説明する。

【例題】診療科別の医師所得を推計したい。一人の医師が単一の診療科を標榜している場合はその医師の所得をそのままその診療科の所得として問題がない。しかしわが国では自由標榜制をとっているので一人の医師が複数の診療科を標榜しているのがふつうである。

そこで、診療科ごとの医師の平均所得を求め、複数標榜の医師の所得はそれに比例して按分する手法をとる。

そのためにはまず診療科ごとの平均所得を算出しなければならない。

以下の例を考える。

表 A	標榜科	所得
Dr.田中	循環器内科	1000 万円
Dr.山本	消化器内科	2000 万円
Dr.佐藤	循環器内科	3000 万円
	消化器内科	

循環器内科と消化器内科の平均所得はそれぞれ $(1000+3000)/2=2000$ 万円、 $(2000+3000)/2=2500$ 万円である。

Dr.佐藤の所得 3000 万円は両科の平均値で比例配分され、以下のようなになる。

表 B	標榜科	所得
Dr.田中	循環器内科	1000 万円
Dr.山本	消化器内科	2000 万円
Dr.佐藤	循環器内科	1333 万円
	消化器内科	1667 万円

よって3人の医師の合計所得 6000 万円の診療科別内訳は、循環器内科 2333 万円、消化器

内科 3667 万円と推計される。

以上が PDM 法プログラムで用いている平均値計算法である。

しかし以下のような計算法も考えられる。

表 C	標榜科	所得
Dr.田中	循環器内科	1000 万円
Dr.山本	消化器内科	2000 万円
Dr.佐藤	循環器内科	1000 万円
	消化器内科	2000 万円

よって3人の医師の合計所得 6000 万円の診療科別内訳は循環器内科 2000 万円，消化器内科 4000 万円と推計される。

つまりそれぞれの診療科の平均値は，それぞれの診療科単独標榜の医師だけでだし，複数標榜科医師の所得は単独の平均値で配分するというものである。この計算法は理想的だが，PDM 法では採用していない。その理由を説明する。

単独のみの診療科（あるいは傷病名）の「純粹」な平均値を計算するのは，医師（あるいはレセプト）の大半は単独標榜で一部のみ複数科標榜という状況では有効であり用いるべきであろう。しかしレセプトでは複数傷病名が多く（だからこそ PDM 法が必要になる）むしろ傷病名一つのレセプトの方が例外的である。

この計算法をとるためには 119 傷病分類の全てについて単独傷病名が相当数（少なくとも 10 件以上）ある必要になる。高血圧や糖尿病のような頻度の高い傷病ならいざしらず，119 全てについて 10 件以上の単独傷病レセプトを確保するには相当規模のレセプト件数が必要になる。小規模な市町村や組合，また単月レセプトやもっと狭い地区のレセプト分析は不可能になる。100 件程度の件数でも比較的安定した分析をしなければならない PDM 法の要請を満たせない。

医師が 10 人おり，循環器内科，消化器内科単独標榜の医師は各一人で残り8人は複数標榜，という状況を考えれば，たった2人の所得の割合で残り8人の所得を配分することの無謀さが理解できる。レセプトの状況はまさにこれに近い。むしろ表3の計算法を否定しているのではなく，レセプト電算化がすすみ都道府県単位の多数のレセプトを分析できるようになればこの計算法の可能性もある。

次に重複カウントの問題に移る。

分類上，循環器内科も消化器内科もともに「内科」というカテゴリーにくられる。すなわち以下のようになる。

表 D	標榜科	所得
Dr.田中	内科	1000 万円
Dr.山本	内科	2000 万円
Dr.佐藤	内科	3000 万円
	内科	

PDM法プログラムでは傷病名を重複カウントしており，これより内科の平均所得は  $(1000+2000+3000+3000)/4=2250$ 万円となる。これは先に出した循環器内科，消化器内科別の

平均2000万円と2500万円のさらに平均となる。

一見するとDr.佐藤について内科3000万円を2回カウントするのはおかしいように見える。「分類がくられて同一分類になったのだから重複カウントはおかしい」とDr.佐藤の内科は1500万円として計算すべきのようにも思われる。

これに対する説明は、分類上3人の医師は同一診療科になったが、実際には3人の医師は異なった診療科であり、Dr.田中とDr.山本は違う診療科だし、Dr.佐藤は他の2人の診療科両方の患者をみていることである。もしDr.佐藤は2つの診療科を標榜しているから、とって診療科当たりの所得1500万円として計算すると内科の平均所得は $(1000+2000+1500+1500)/4=1500$ 万円とおかしなことになる。

【平均値の補正】

上の説明は単純平均を重みとしてPDM法を適用した時のバイアスについても示唆を与えている。3人の医師の所得6000万円の内訳は循環器内科2000万円、消化器内科4000万円がもし真の値とすると、PDM法による推計結果はそれぞれ2333万円、3667万円となった。つまり少ない方の循環器内科は過大評価し、大きい方の消化器内科は過小評価された、すなわち上下の格差が縮小し、均等に近づいた。

これは平均値を用いてPDM法を適用する場合の系統的なバイアスといえ、事実、平均値を重みとしてPDM法で傷病別点数を推計した場合、腎不全のような高額傷病の医療費は過少に、逆に皮膚炎のような低額傷病は過大に評価される傾向にあることが当初から気付かれていた。

そこで平均値を重みとして用いる場合、一定の補正が必要になる。

ここでレセプトの傷病数をNとし、点数をPとすると傷病当たり平均点数は $P/N$ 。そのうちの一つの傷病の点数がPだけ大きく(小さく)なってもレセプトの点数Pは $P/N$ しか増加(減少)しない。

したがって全体平均と同じ傷病については、その傷病の記載されたレセプトの点数の平均をそのまま使っても変化しないが、平均値より大きい傷病は過少に、逆に小さい傷病は過大に評価され、その影響はレセプトの平均傷病数Nによって決まる。

一般にレセプト平均傷病数Nの場合、ある傷病の「真」の平均点数をX、ある傷病の記載された全レセプトの平均点数をYとすると両者は次のような関係になると考えられる。

$$Y = X/N + b \quad (b \text{ は定数})$$

定数bは、全体平均値 $\mu$ で $Y=X$ と交わるように定まる。かりに傷病別点数の全体平均 $\mu$ を500とすると、上式は以下ようになる。

$$N=2 \text{ なら } Y = 0.5 * X + 250$$

$$N=4 \text{ なら } Y = 0.25 * X + 375$$

$$N=8 \text{ なら } Y = 0.125 * X + 437.5$$

このようにレセプトの傷病名数が多くなるほど直線は平べったくなってゆく。ちなみに外来レセプトの平均傷病名数は約4、入院レセプトは約8である。

知りたいのは各傷病の「真」の平均点数Xであるが、実際のレセプト分析ではXは不明で、わかるのはYと $\mu$ だけである。そこでYと $\mu$ だけからXをできるだけ正確に予測するため、補正してできるだけXに近いY'を求める以下の補正式を考案した。

$$Y' = Y * (Y/\mu)^k \quad (k \geq 0)$$

これはYが全体平均 $\mu$ と同じであればそのまま、もしYが全体平均 $\mu$ の2倍であれば4倍に、逆に2分の1であれば4分の1にする補正をk回くりかえす式である。k=0ならY'はYそのまま、つまり無

補正ということになり、 $k=1$ を1次補正、2なら2次補正と呼ぶことにする。その概念をグラフで示した【図2】。そして、この補正を何次( $k$ 回)まで行えばよいかを、レセプト傷病数 $N$ の数に応じてシミュレートした【図2 1~3】。

その結果、 $N=2$ すなわち平均傷病数が2つの時は1次補正( $k=1$ )で十分であるが、 $N=4$ では2次補正まで必要であり、 $N=8$ では5次補正まで必要であることがわかった。

以上をふまえ、PDM法Ver.2には補正式のオプションを加え、外来レセプトでは2次、入院レセプトでは5次補正が望ましいとした。

この補正は、点数についてのみ行い、日数については行なわない(理由は後述)。

#### 【繰り返し法】

PDM法は、まず傷病ごとに平均値を算出し、それを重みとして各レセプトの点数を記載された傷病ごとに配分し、さらにレセプトごとに傷病ごとに配分された点数を傷病ごとに合計する、という演算処理である。したがって演算処理の過程で、レセプトごとに傷病別の平均値で比例配分された新しいレセプトデータがまるまる生成される。

このように処理過程で生成された新しいレセプトデータを再度PDM法にかけ、その過程で生成されたレセプトデータをもういちどPDM法にかける・・・という処理を繰り返すと一定条件の下では最適な値に収斂することが明らかになった。最尤推定法(Maximum Likelihood Estimator)という、推定式を繰り返せば最も尤もらしい値に収斂する、という手法の考えに近いものである。PDM法は、各傷病の点数の平均値を計算し、それを補正式で補正したが、繰り返し法では補正式を使うのではなく、傷病別の平均値を出しては同じ処理を繰り返し行う。

傷病別の平均値は繰り返すごとに微妙に変化してゆく。しかし繰り返しを重ねるごとに変化は小さくなってゆき、やがてはほとんど変化しなくなる。コンピュータープログラムにおいては、たとえば回を重ねるごとの変化がたとえば0.01より小さくなったら「繰り返し打ち切り」とプログラミングしておく。

この方法は、とくにレセプト件数や傷病分類数が多いとコンピューターに大きな負担を強いるのでPDM法にはまだ組み込まれていないが、将来コンピューターの性能が向上すれば可能性の高い手法である。ただし、この手法ではどのような条件でも収斂するとはかぎらず、あくまで加法原理(後述)が成立していることが条件となる。

#### 【シミュレーションデータによる検証】

PDM法開発の上でこれまで最大の障害は、その妥当性の検証ができない、ということであった。すなわち実際のレセプトデータの傷病別の「真」の日数や点数は不明であり、適用して分析結果を出しても、それがどれだけ真実に近いのか検証できなかった。

それが、シミュレーションデータを用いることによって部分的には解決した。シミュレーションデータとはすなわちあらかじめ傷病別の点数等の「正解」を設定し、それに合わせて生成したデータのことである。この「正解」のわかっているデータをPDM法にかけて、推計結果と「正解」とがどれだけ一致しているかで妥当性を検証できる。

社会医療調査や国民健康保険医療給付実態調査等を元にわが国の外来レセプトにできる限り近づけて作製したレセプト1000件分のシミュレーションデータで検証した結果が【図3 1,2】である。ヨコ軸にはシミュレーションデータの傷病別点数(正解,  $x$ )、タテ軸にそれぞれの方法による傷病別点数の推計値( $y$ )を示す。もし完全に一致する $y=x$ になり決定係数( $R^2$ )は1となる。

図からも明らかなように平均値を補正する方法でも繰り返し方法でもきわめて良好な近似が

得られた。しかしながら、日数については平均値をそのまま使う方法では良好な近似であったが、繰り返し法では収斂はしたものの、その近似はきわめて不良であった【図4-1~3】。繰り返し法でなぜ日数でいい結果がえられなかったか、その理由は現時点では断定できないが、ひとつの仮説としては以下に述べる「加法原理が日数ではあてはまらない」ことが考えられる。

これにより日数で評価するインフルエンザワクチンの効果測定は繰り返し法ではなく平均値そのまま重みとして用いる方法で実施した。

(繰り返し法のコンピュータプログラミングでは横山徹爾主任研究官の協力に感謝する)

【加法原理と最大原理】

点数の推計では傷病数で割り、平均値を補正式で補正したが、日数の推計ではそのような処置は行っていない。また繰り返し法では日数については適切な値に収斂しなかった。その理由を考察する。

点数については、たとえば費用1万円を要する傷病Aと2万円の傷病Bが合併するレセプトの医療費は3万円とみなしている(加法原理)。それに対して、月4回通院を要する傷病Aと2回ですむ傷病Bを合併するレセプトの日数は6日ではなく4日とみなしている。明らかに、異なる傷病を合併しても同時に治療が可能であり、たしあわされるのではなくその最大日数が受療日数になる、と考えられる(最大原理)。

したがって傷病別の点数を推計するためには合計点数を傷病数でわる必要があるが、日数についてはそのような処理は行わない。そして傷病数でわる場合は、それによって得られる平均値は傷病数による補正が必要となるが、日数についてはそのような補正は必要とされない。

もっともこのような最大原理と加法原理が厳密になりつつわけではなく、月4回の通院を要する傷病Aに別の傷病が合併すると通院日数は平均して4回よりも多くなる可能性はあるし、また1万円の傷病Aと2万円傷病Bとが合併しても3万円になるとはかぎらず、たとえば診察料のように傷病数とは無関係に一定額しか請求できない診療報酬もある。投薬や検査の中には複数の傷病を適応症にしているものもある。これまで漠然と点数とは総点数を指してきたが、点数をさらに診察料、投薬料、検査料と細分化してPDM法で分析する場合、最大原理と加法原理とは適宜使いわける必要がある。

ちなみにPDM法Ver.2では、傷病数でわらず、一件当点数をそのまま使うオプションも加えた。これは手術料等単一の傷病に月一回しか行なわれない診療行為の報酬分析を意図している。

現時点では、上に述べたような「わりきり」を行ったが、日数と点数が傷病との関連で最大原理、加法原理がどれだけ通用するかは、今後の研究課題である。

【PDM法プログラムの概要】

本研究の最大の成果物は、以上のPDM法の原理をウィンドウズパソコン上で行えるコンピュータプログラムである。そのVer.1が2002年10月埼玉で開催された日本公衆衛生学会総会における第5回レセプト情報の活用を考える自由集会で発表された【マニュアルを参考資料に添付】。その後、シミュレーションによる補正式の開発等ととりくみ、近くVer.2が完成する見通しである。

Ver.2の概要を【表1】に示すとともに、それを実際にある市の外来レセプト15771件に適用し、点数および日数の傷病別割合を出した結果を示す【図5】。

【PDM法の応用例】

PDM法を具体的に応用した例として、インフルエンザワクチン接種の有効性評価をあげる。

本研究は宮城県名取市の委託によるもので報告書が2003年2月名取市に提出された【資料2】。

ワクチンの有効性評価で最も適切なものは無作為割付比較試験であるが、高齢者へのインフルエンザワクチンの有効性は無作為割付比較試験で既に確立されているため、評価目的に一般人口に行くことは倫理的に許容されない。そこでインフルエンザワクチン接種の有効性評価をレセプト情報を用いて行ったところ、無作為割付比較試験と同様の結果が得られた。

名取市の65歳以上高齢者のレセプトをワクチン接種者台帳と個人情報でリンケージし接種、非接種者にわけた。しかしそれだけでは、分母がわからないため発生率を比較することはできない。そこで接種、非接種者両群の総受療日数に占めるインフルエンザ及び関連呼吸器疾患の「割合」を測定することによって比較する。【日数と点数の扱い】でも述べたようにワクチンの有効性は疫学情報である日数で評価すべきであって、経済情報である点数で評価すべきではない。そして日数については補正式を用いることなくPDM法で傷病別日数を推計すべきである。

接種、非接種者の疾病リスクは受療日数に反映されるとみなし、もしワクチンの効果がなければ、総受療日数に占めるインフルエンザの割合は両群で同じであると仮定する。もしワクチンの効果があれば、それはインフルエンザによる受療日数の減少として現れ、結果として総受療日数に占めるインフルエンザの割合は減少する。こうすることによって、発生率が不明でも、接種、非接種者両群のリスクを補正しつつ、ワクチンによるインフルエンザ受療日数の減少を客観的に測定できる。そして得られた結果は既に確立された無作為割付比較試験の結果ときわめて近似していた。

インフルエンザによる受療日数を客観的にきりわけるPDM法により、たとえ発生率がわからなくてもレセプトだけで簡便にワクチン効果を測定できる手法が確立された。これにより、今後全国の自治体はインフルエンザワクチン接種事業の有効性をレセプトだけから行えることになり、行政評価の観点からも、疫学研究の観点からもきわめて有意義な成果であった。

【添付資料】

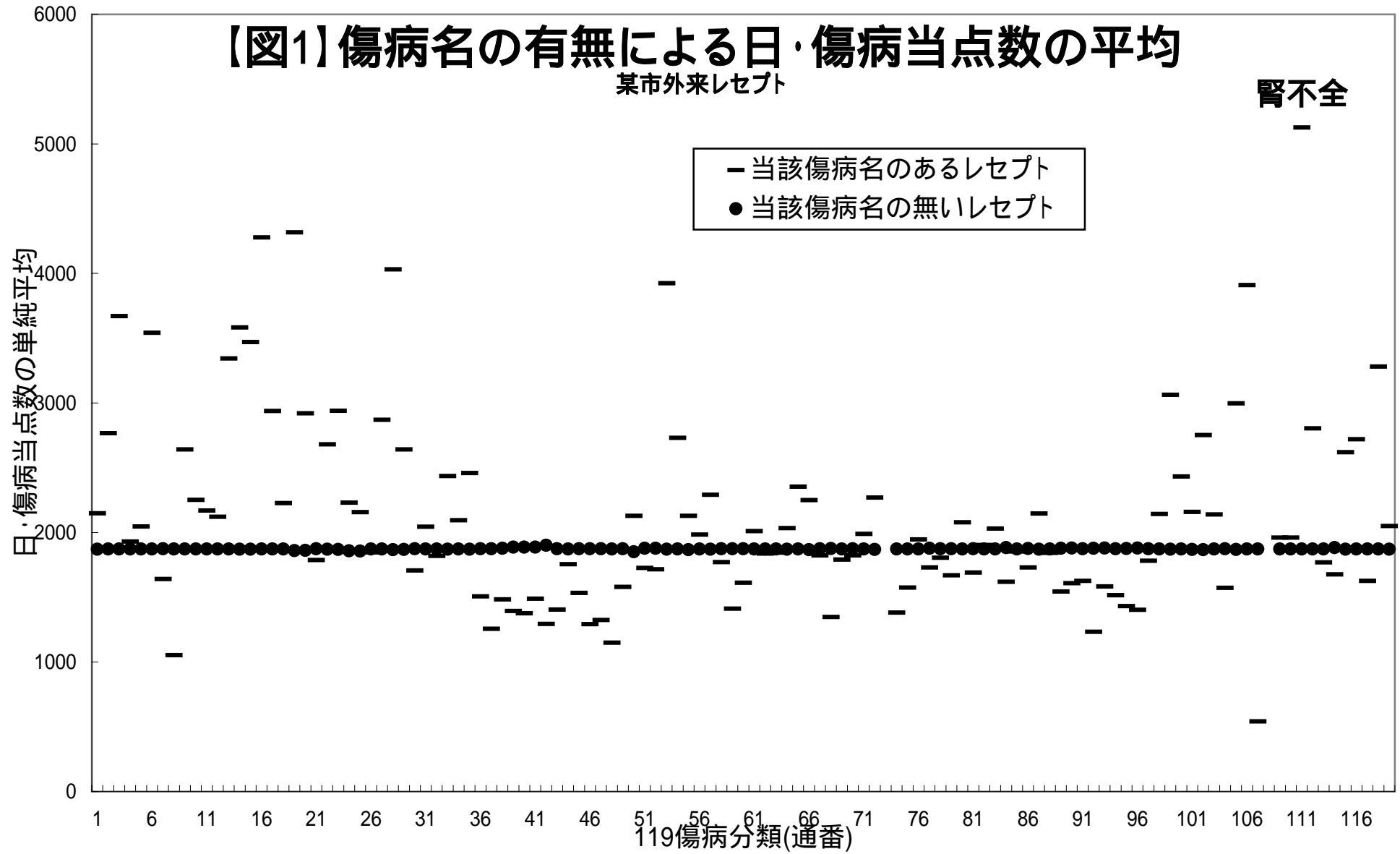
【資料1】PDM法使用マニュアル2002年10月24日公表

【資料2】名取市「高齢者インフルエンザ予防接種の効果分析」報告書2003年2月

# 【図1】傷病名の有無による日・傷病当点数の平均

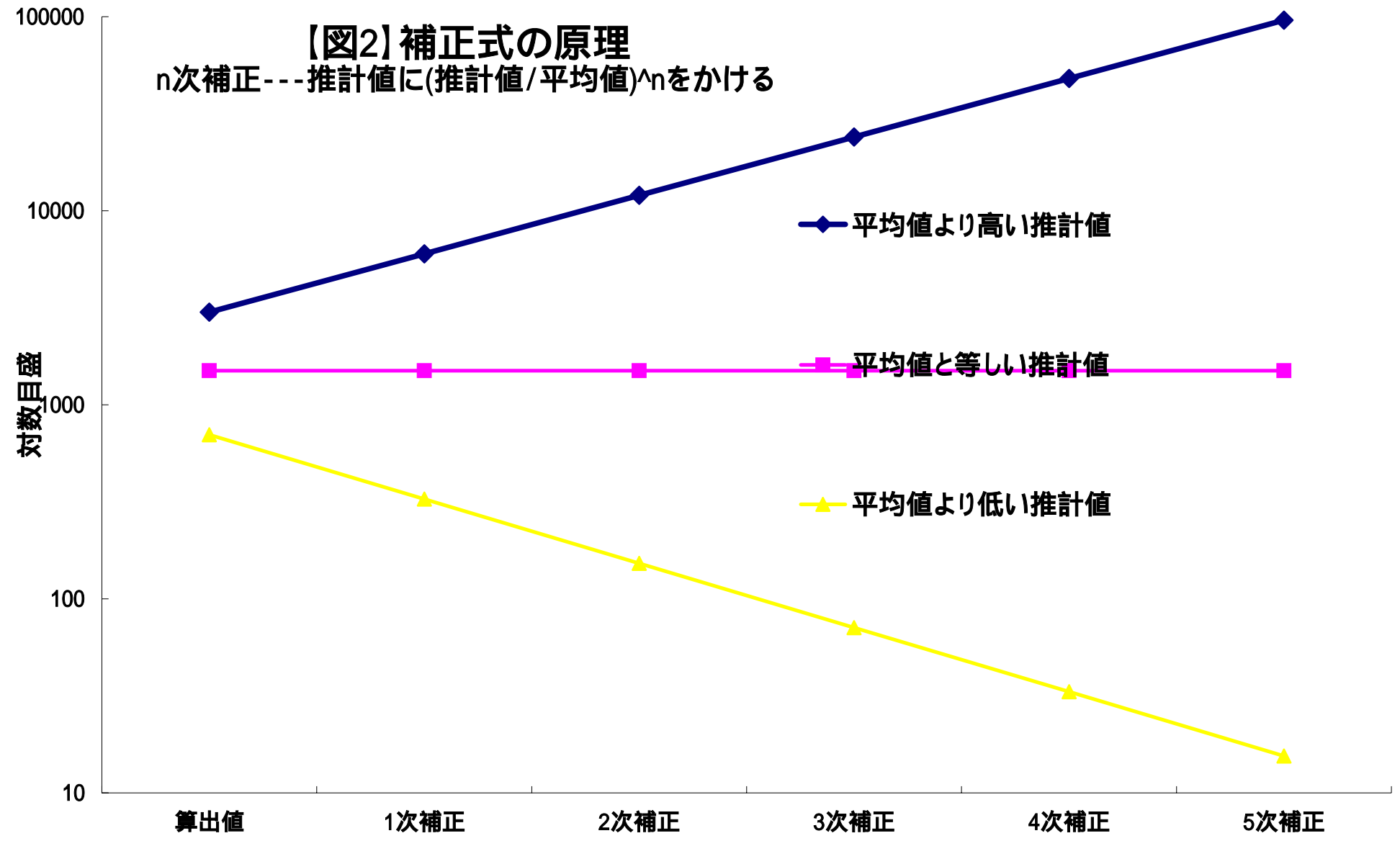
某市外来レセプト

腎不全

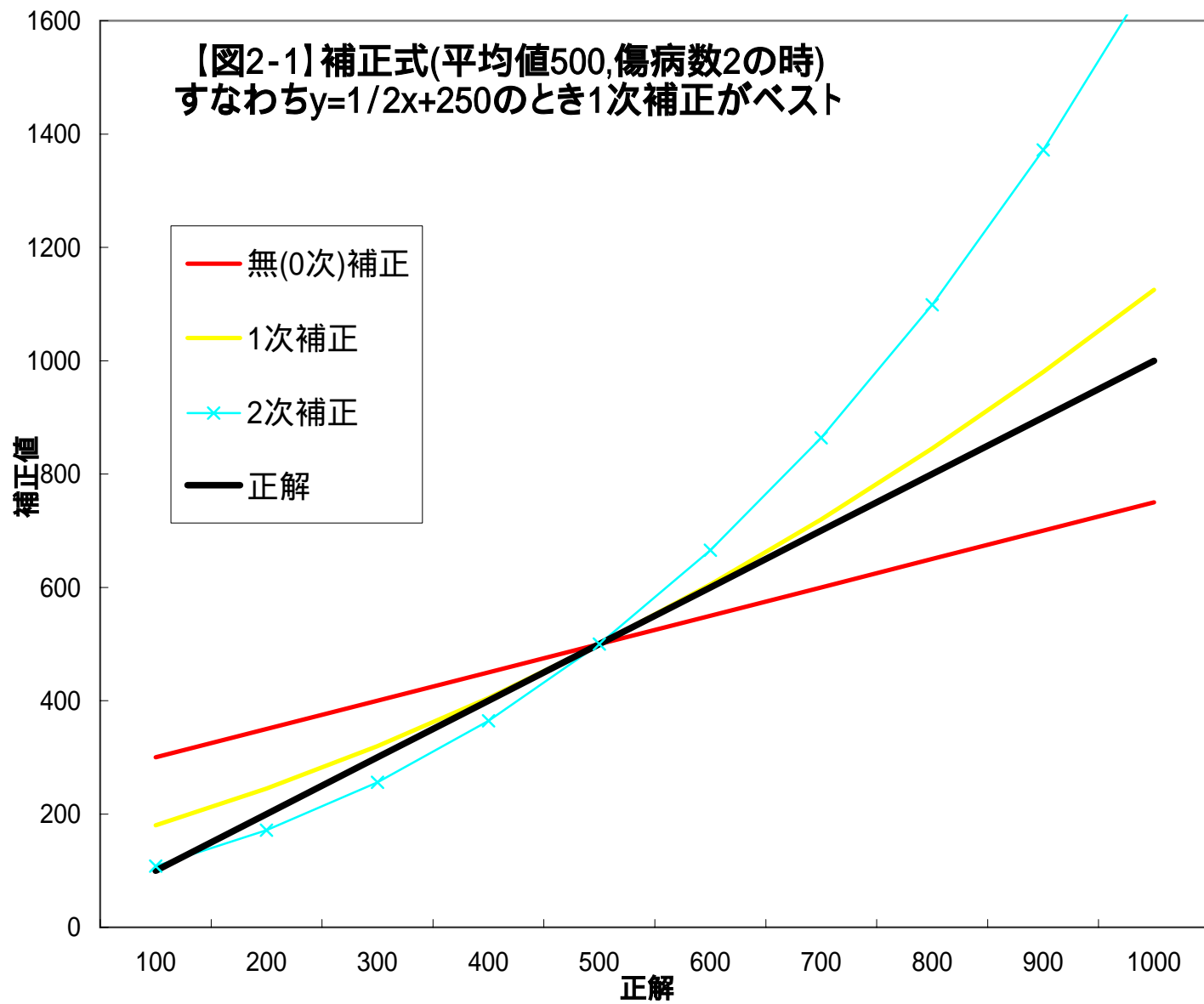


# 【図2】補正式の原理

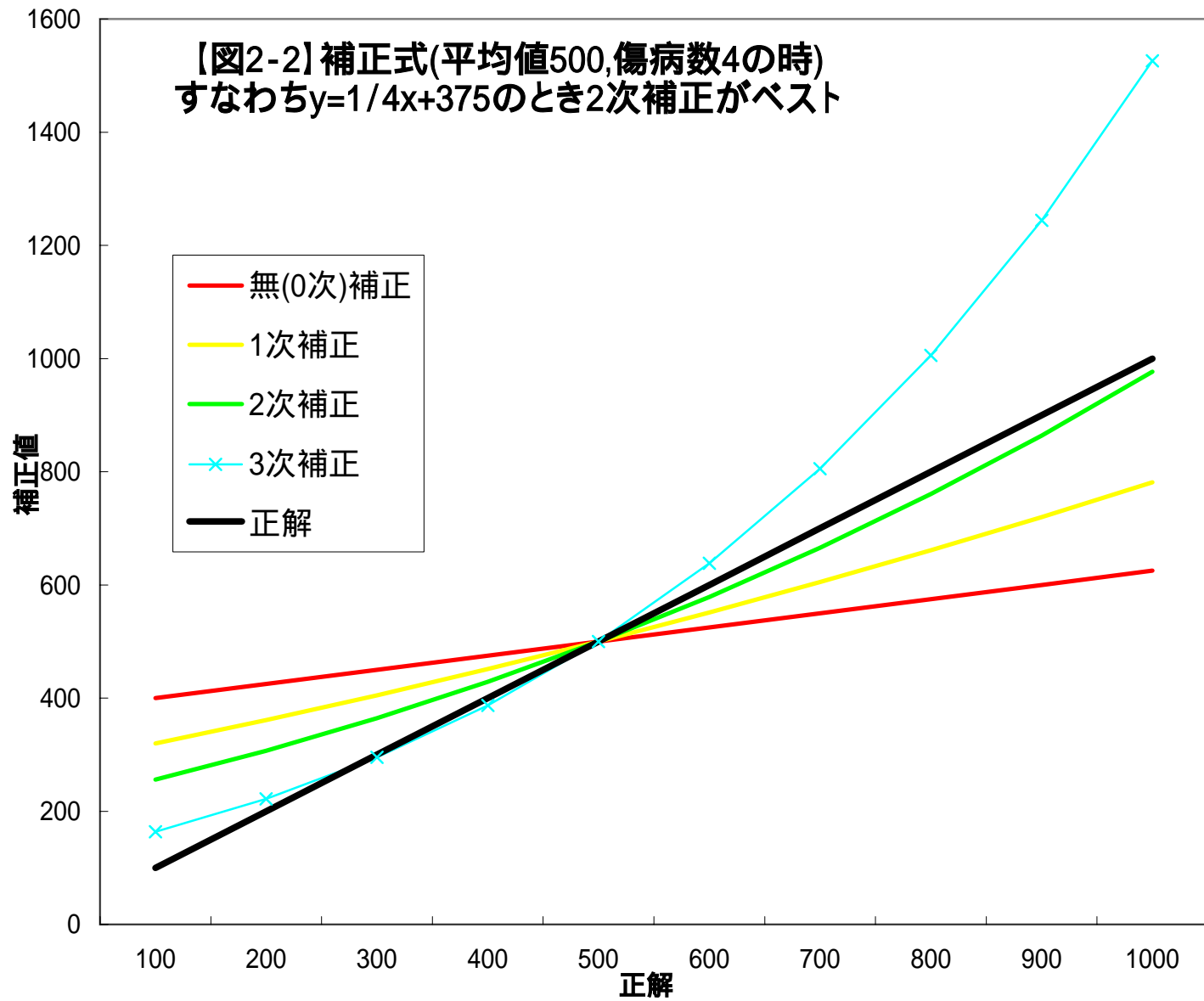
n次補正 --- 推計値に  $(推計値 / 平均値)^n$  をかける



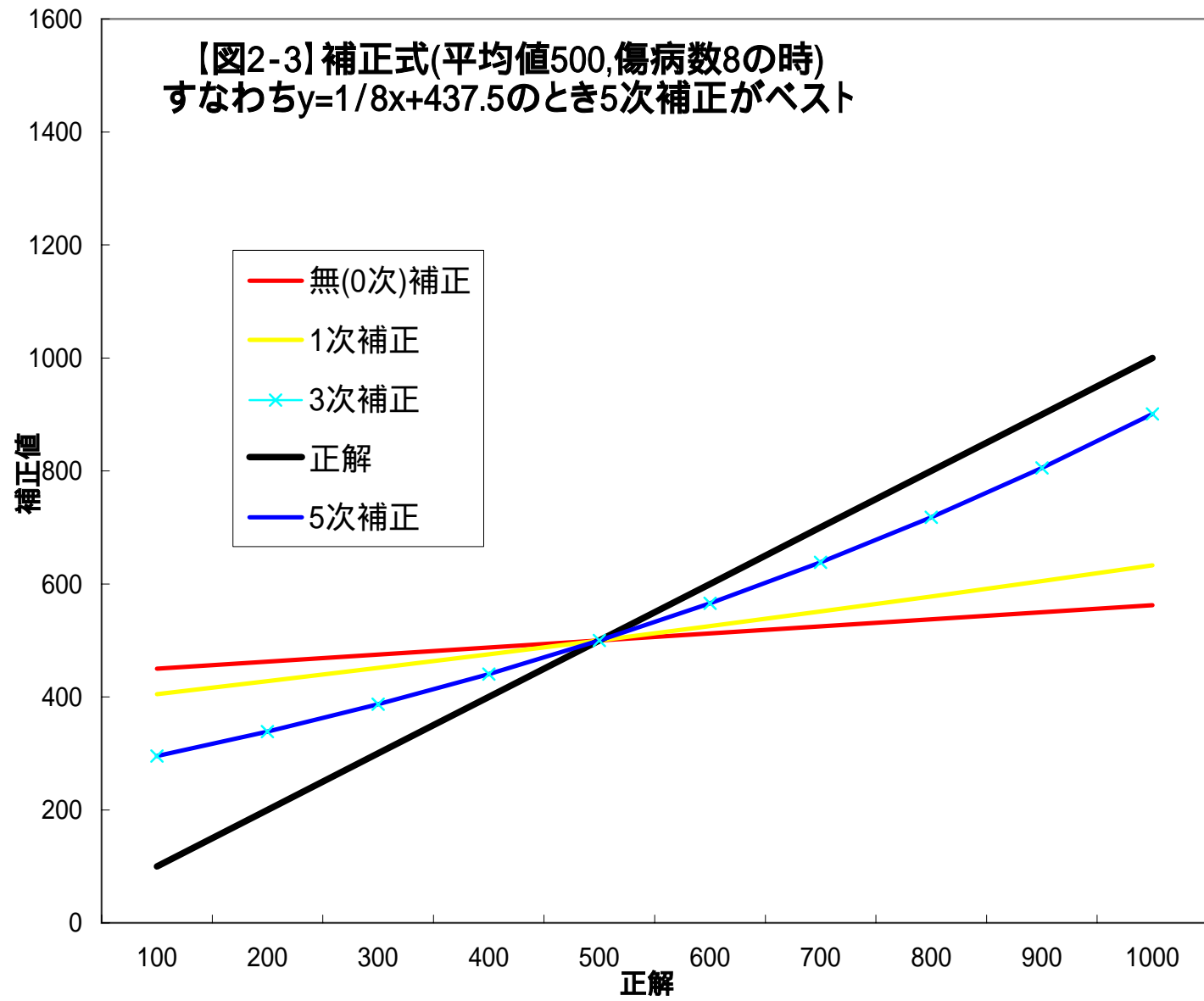
【図2-1】補正式(平均値500,傷病数2の時)  
すなわち $y=1/2x+250$ のとき1次補正がベスト



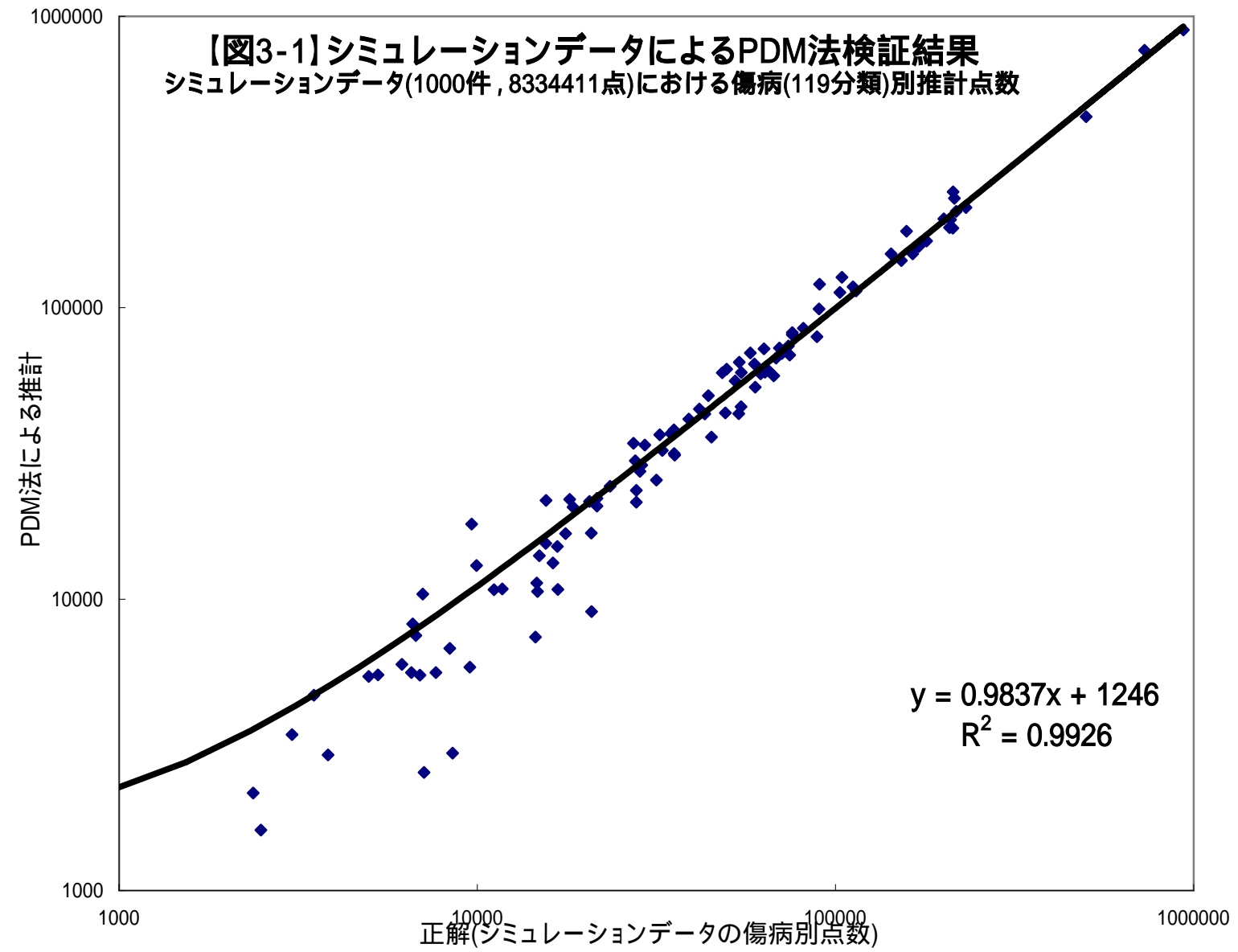
【図2-2】補正式(平均値500,傷病数4の時)  
すなわち $y=1/4x+375$ のとき2次補正がベスト

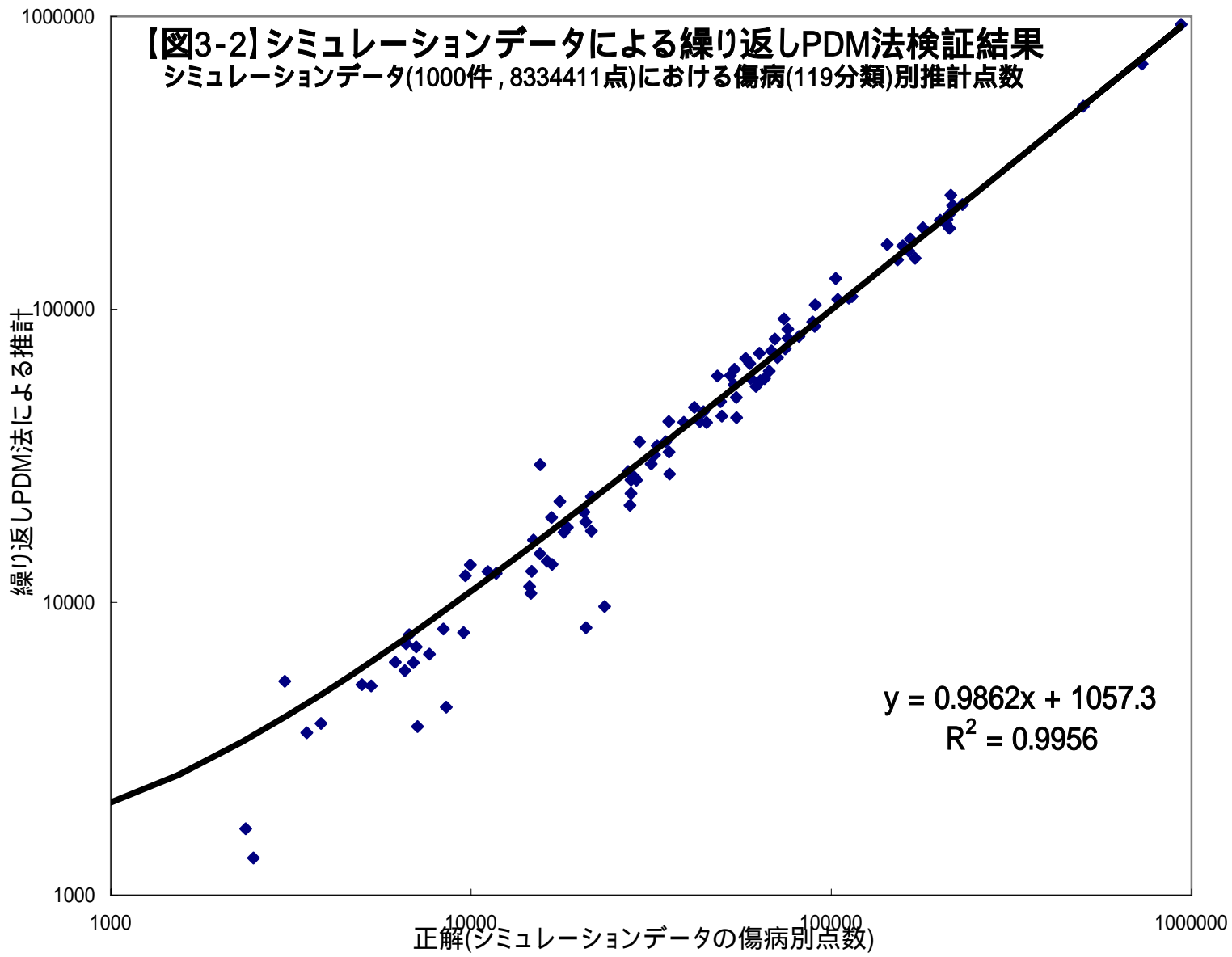


【図2-3】補正式(平均値500,傷病数8の時)  
すなわち $y=1/8x+437.5$ のとき5次補正がベスト

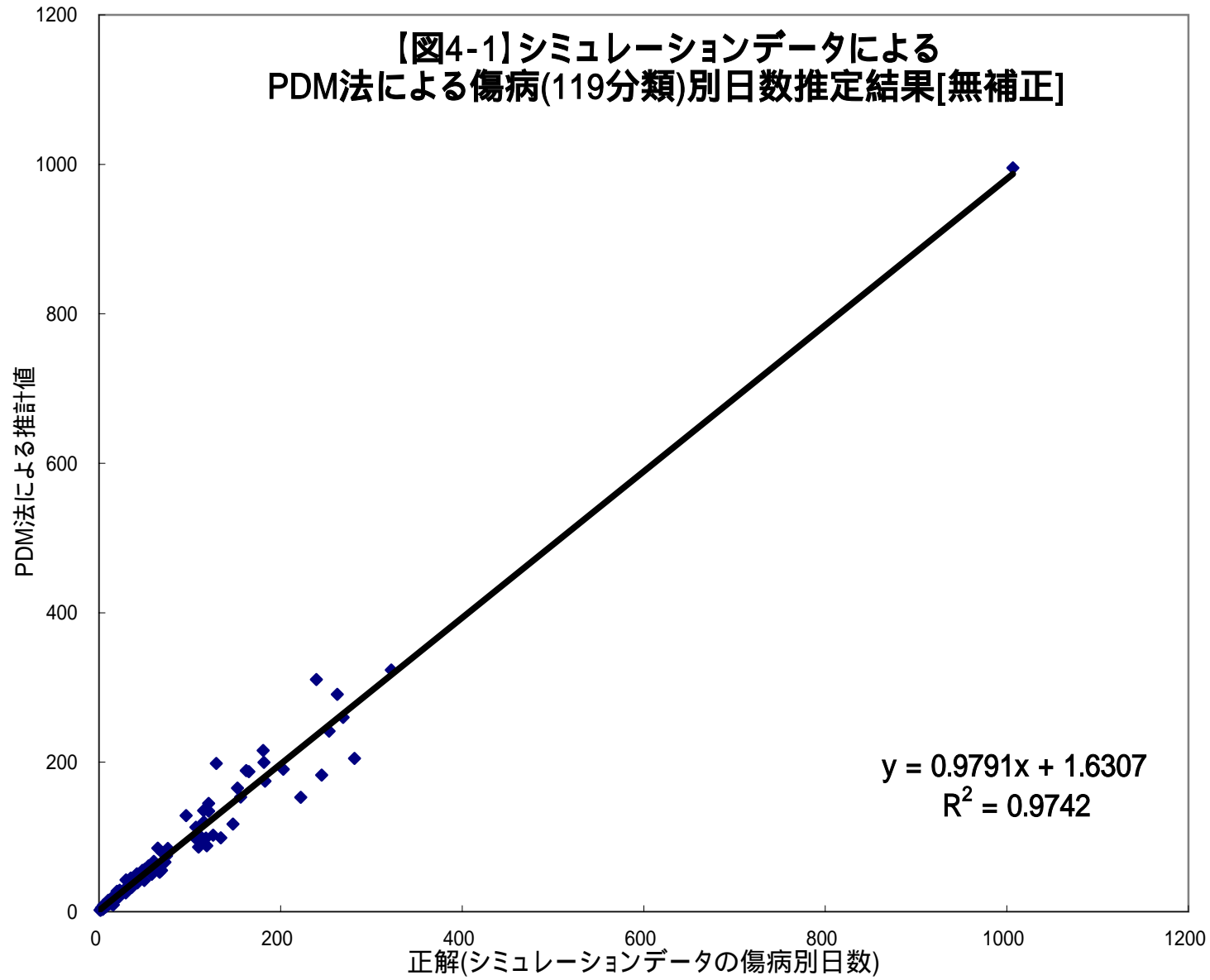


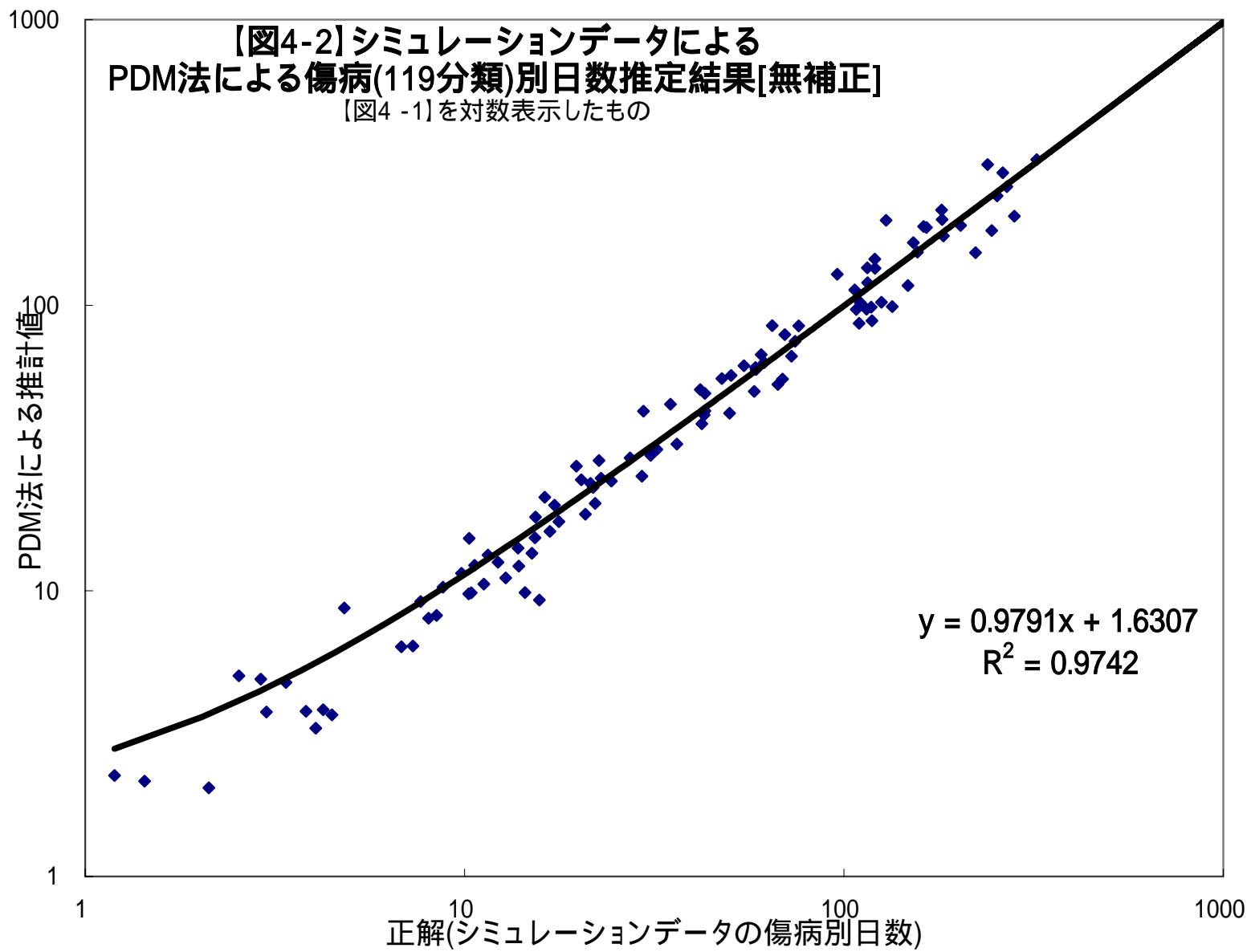
**【図3-1】シミュレーションデータによるPDM法検証結果**  
シミュレーションデータ(1000件, 8334411点)における傷病(119分類)別推計点数



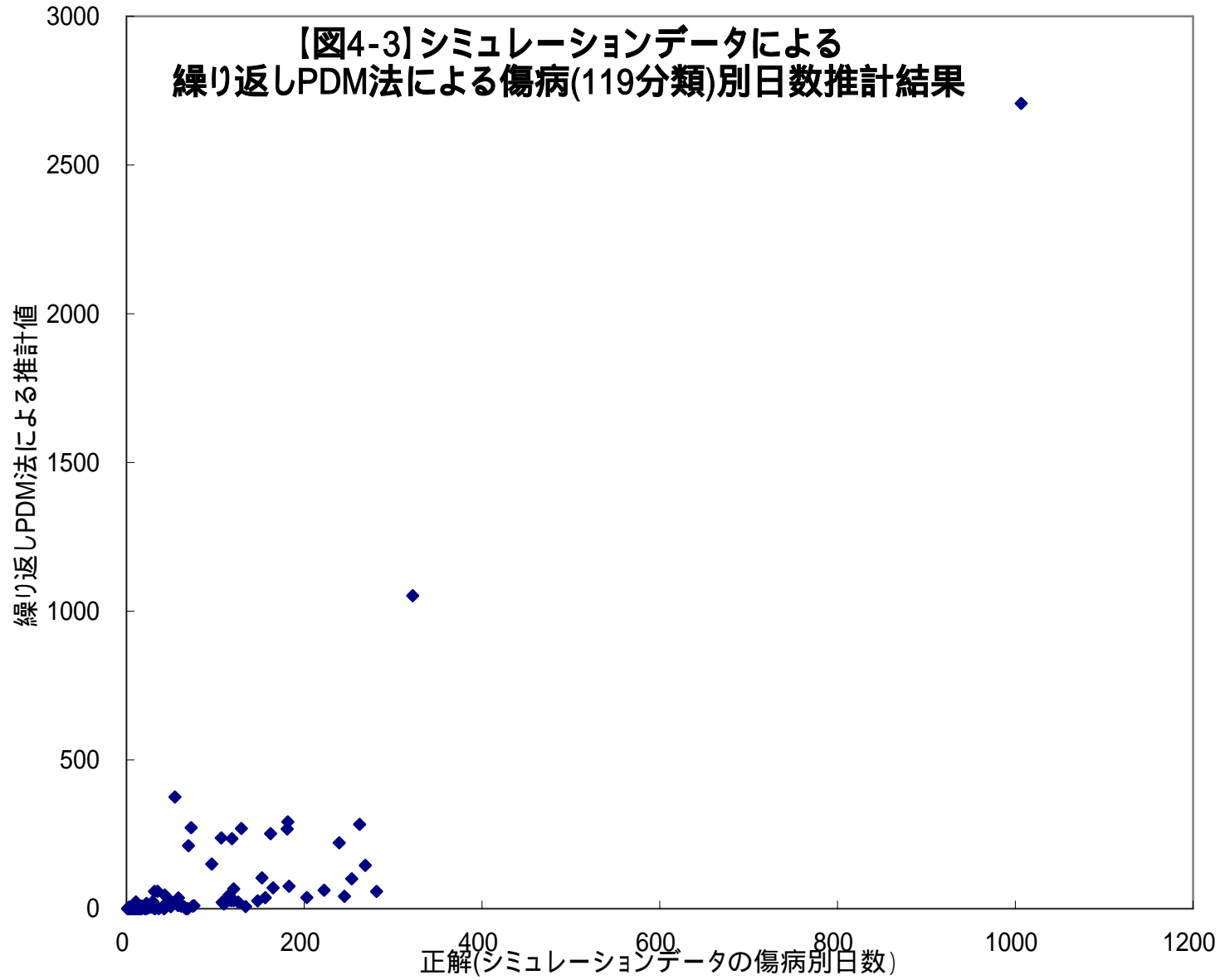


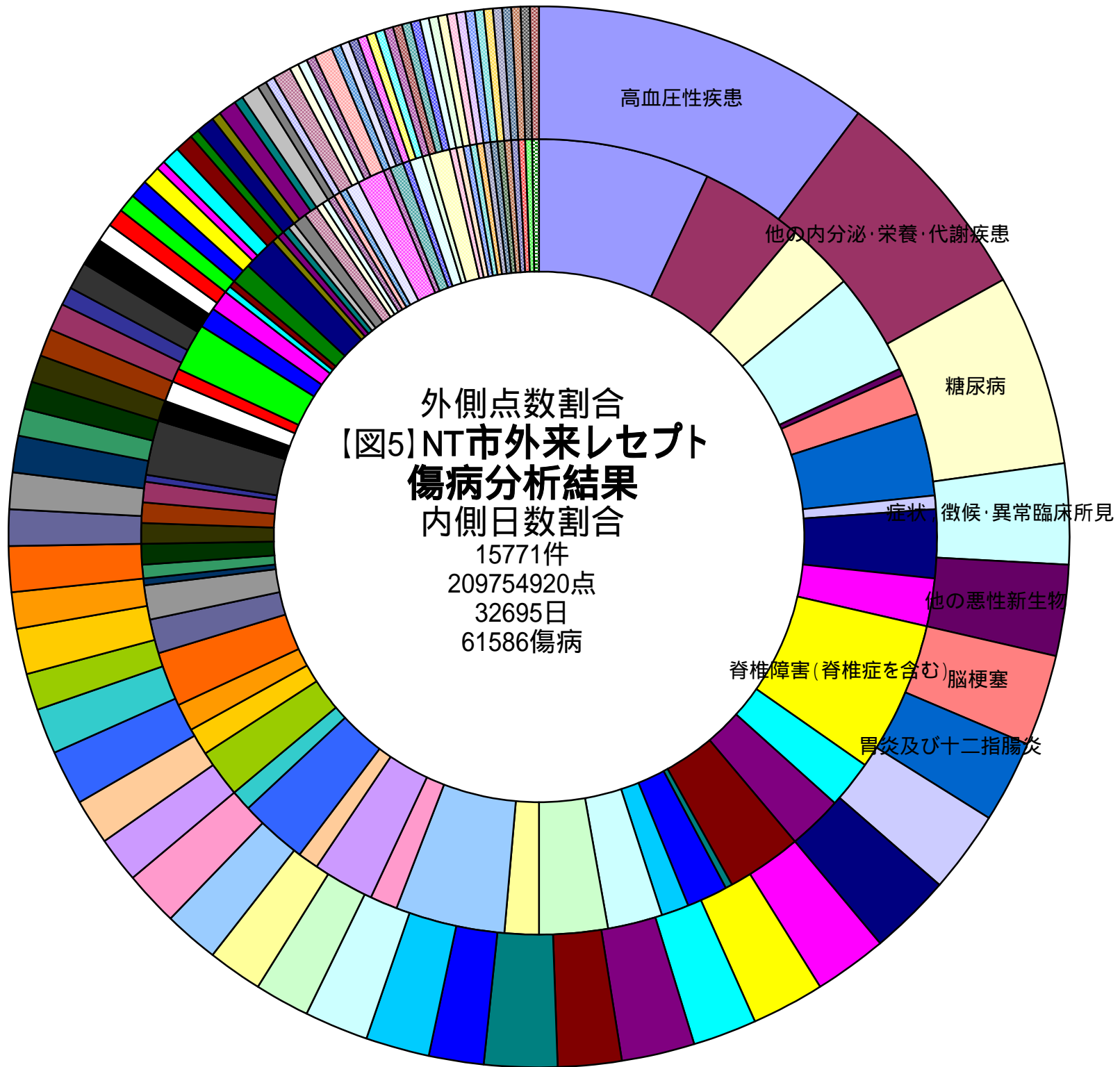
【図4-1】シミュレーションデータによる  
PDM法による傷病(119分類)別日数推定結果[無補正]





【図4-3】シミュレーションデータによる  
繰り返しPDM法による傷病(119分類)別日数推計結果





**【表1】PDM法プログラムの概要**

【プログラムの内容】csv形式(またはExcelファイル)化されたレセプトデータ(通番, 日数, 点数, 傷病名コード)を機械的に「重み」によって配分集計し, 結果をレポートおよびcsvファイルとして出力したり, グラフ化して表示するプログラム

【動作環境】Windows/パソコン。OSはWin2000またはXp。CPU,Pentium4(1GHz), メモリ1GBに対応。

**プログラム仕様(Ver.1,2を対比させつつ)**

Ver.1	Ver.2			
各レセプトから日数D, 点数P, 傷病コードを読み取る				
csvファイル。傷病数は常に15であり, 15未満のレセプトについては,,,を必要数つけないとエラー出る。	傷病数を最大30まで任意に設定			
重みの計算法は一つのみ	日数	点数配分		
		重みに件当点数を用いる[手術料等一件につきしかない点数分析に適用]	重みに日当点数を用いる [Ver.1]	重みに日・傷病当点数を用いる[デフォルト]
各レセプトについてP/Dを算出	Dをそのまま用いる (Ver1と変更無し)		各レセプトについてP/Dを算出	各レセプトの傷病数Nを計算 各レセプトについてP/D/NとD*Nを算出
[参考]Ver1の詳細表の出現数, 日数, 点数(AVG)の合計はそれぞれ (N), (D*N), (P)に相当する。		件当点数平均 (P)/件数を算出	日当点数平均 M= (P)/ (D) を算出	日・傷病当点数 M= (P)/(D*N)を算出
各傷病の出現するレセプトについてP/Dの平均値[P/D]を求める	各傷病の出現するレセプトについてDの平均値[D]を求める	各傷病の出現するレセプトについてPの平均値[P]を求める	各傷病の出現するレセプトについてP/Dの平均値[P/D]を求める	各傷病の出現するレセプトについてP/D/Nの平均値 [P/D/N]を求める
		補正次数の設定n=0,1,2,3,4,5 [デフォルト=2]		
	[P]を補正する。 $[P]^{(n+1)}/M^n$	[P/D]を補正する。 $[P/D]^{(n+1)}/M^n$	[P/D/N]を補正する。 $[P/D/N]^{(n+1)}/M^n$	
各レセプト内の日数(D), 点数(P)の配分				
全レセプトの傷病別日数, 点数の集計				

レセプト自動傷病分類ソフト

**P D M**

(Proportional Disease Magnitude)

操作説明書

関連サイト <http://resept.com>

2002年10月24日

発案・設計 岡本悦司 (国立保健医療科学院)  
製 作 バイオコミュニケーションズ株式会社

本プログラムは下記の助成により制作された  
厚生労働科学研究費補助金政策科学推進研究事業  
「レセプト情報の利活用と個人情報保護のあり方に関する研究」  
(主任研究者小林廉毅)

## PDM 法とは

皆保険制をとるわが国では、レセプトは地域や集団の傷病構造を把握する有力な医療情報です。レセプト統計において傷病分類は、複数の傷病が記載されたレセプトについては分類する人の主観的判断で主傷病を一つ選択する、いわゆる「主傷病選択法」がとられています。しかしながら、この方法は分類する人によって判断が異なったりして客観性や再現性に乏しく、また、たとえば糖尿病と高血圧が記載されたレセプトがもし高血圧に分類されるとインスリンも高血圧の治療費に分類される、といった矛盾があります。何より主傷病分類法は人間による作業であるため、時間と手間がかかり、何億もの膨大なレセプトを短期間に分類することは不可能です。

PDM 法はレセプトの点数や日数を、それに記載された全ての傷病名に一定の「重み」を与えて比例配分してゆく分類法です。PDM 法によれば、客観的かつ自動的な傷病分類が可能となり、もし共通の「重み」を用いれば、たとえばA市とB町国保ではどちらが糖尿病の医療費がかかっているか、またC村で糖尿病対策をする前と後とで糖尿病の医療費が増えたか減ったか、等を客観的に比較することも可能となります。裏返せば、では傷病ごとの「重み」をどう決めるか？が最大の課題となります。

重みの決め方としては、たとえば DRG も一種の重みといえますし、外国では専門医にアンケートして重みを測定しようという試みもあります。PDM 法では、これまで患者調査で得られた「主傷病となる確率」や社会医療診療行為別調査で得られた傷病ごとの「一件当たり点数・日数」を用いたこともあります。本プログラムではデフォルトでは、各傷病名を含むレセプトの一日当たり点数、一件当たり日数を単純平均する方法がとられます(むろんそれ以外の独自の重みファイルを使うことも可能です)。単純平均とは一見単純すぎるようにもみえますが、その重みによる分類結果は、社会医療診療行為別調査のような全国規模のデータを重みに使った場合ときわめてよく一致することが確認されています。PDM 法は、対象となるレセプトデータがある程度異なっても結果が大きくは左右されない「頑健(robust)」な手法といえるでしょう。

## PDM 法のあゆみ

「全ての傷病名を分析し、点数や日数といった医療の資源消費を定量的かつ自動的に傷病分類できないか」そう考えた岡本(当時近畿大学講師)は94年頃「診療報酬明細書による傷病構造の解析」研究に着手(94年度文部省科学研究費奨励研究)。95年6月米国、シカゴで開催された第12回 Association of Health Service Research においてその基本概念を発表し、同年10月山形で開催された日本公衆衛生学会で開催された「第1回レセプト情報の活用を考える自由集会」において BASIC で組んだ初歩的なプログラムを公表しました。PDM(Proportional Disease Magnitude)法と名付けられた本手法の概念は、厚生統計分野の代表的な学術誌である「厚生指標」96年6月号に掲載されました。

PDM 法の最大のネックは、記載された多数の傷病名をコード化し入力しなければならない「手作業」にあります。レセプト電算化も期待されたほどには進展せず、実用化へは遠い道のりの状態が続きました。1999年には健康保険組合連合会による「レセプト分析による医療機関評価」研究が行なわれ、そこで PDM 法により傷病構造を補正して医療費の額や在院日数の長短を異なった医療機関間で比較することが試みられました。この時作成された Excel の VBA プログラムはインターネット(resept.com)上で公表され、関係者の関心を集めました。しかし、このプログラムも誰でも気軽に使用できるほどユーザーフレンドリーではありませんでした。

2001年、岡本は旧国立公衆衛生院に移り、愛媛県、宮城県の国民健康保険団体連合会が実施している全疾病入力レセプトの解析に着手。また「レセプト情報の利活用と個人情報保護のあり方に関する研究(主任研究者:小林廉毅)」に分担研究者として参画、その助成によりそれまでコンピューターに強い人でなければ使用できなかったプログラムが Windows 上で誰でも簡単に使用できるプログラムとして完成しました。

## PDM 法の個人情報保護上のメリット

PDM 法は傷病分類を自動化することにより、手作業の労力と時間を節約しますが、個人情報保護の面でも効果があります。データ入力作業はレセプトが完全電算化されれば不必要になりますが、傷病分類だけは自動化は不可能で、人間がレセプトを見て判断するしかない、というのが「定説」でした。熟練したレセプト点検職員は1日に1000件近いレセプトを分類できるといいますが、それでももし年間12億件のレセプト全件を分類すると、のべ120万日・人もの人にレセプトを「見せる」必要があります。

どんなに契約書で守秘義務を課してもプライバシー漏洩の危険は、関わる人が多くなると必然的に大きくなります。プライバシー保護を100%確保するには、医療機関から提出された電子化レセプトが、誰からも中身をみられることなく、コンピューター内部で全て処理される、のが理想です。PDM 法により、単純集計はもちろん、傷病分類まで、人手を介することなくブラックボックスの中で処理され、結果の表だけが出力されるようになり、プライバシー保護は万全になります。

## PDM 法の今後

PDM 法の実用化とレセプト電算化が進展すれば、全国民の受療状況や疾病の流行状況を月単位で把握することが可能になり、はれて医療「費」情報を医療情報として活用する可能性が開けます。レセプト情報については以前より MEDIAS という月報情報システムがインターネット上で提供されており医療機関種別医療費も近年では提供されていますが、傷病別医療費はまだ提供されていません。将来は MEDIAS にも PDM 法による傷病別医療費情報が提供され、たとえばインフルエンザの流行状況をレセプトの件数や医療費で把握することや、個人情報保護法下で運用が困難になりつつあるがん登録に代わるレセプトがん登録も可能となるでしょう。

PDM 法は医学疫学研究のためだけでなく行政上も役立つと考えられます。たとえば毎年公表される医療費マップと高医療費市町村の指定も、現在では一人当たり医療費のみで評価され傷病別には評価されません。PDM 法を年齢階級別に適用すれば、たとえばC型肝炎の多い地域ではC型肝炎の医療費を除外して地域差指数を算出することも可能になります。あるいは健康日本21の事業として生活習慣病の医療費マップを作ることも検討に値しましょう。レセプト情報から得られる傷病分類は、地域保健医療計画の策定から保健事業の経済評価にいたるまで幅広い応用可能性を秘めています。

本プログラムが国保、社保をとわず、医療保険や保健事業関係者によって積極的に活用されることを祈ります。

### レセプトデータの準備

**[重要]** 何より重要なのはレセプトデータの整理です。傷病名無記載, 点数ゼロ点のものは除いて下さい。傷病名無記載は「分類不能」に別に計上して下さい。点数ゼロのものは処理に無関係であるのみならず, 分母をゼロにしてエラーの原因になります。日数ゼロ(電話再診のみ等)は日数を1以上にして下さい。これを怠ると後々エラーで悩まされます。

レセプトデータは整理, 日数, 点数, 傷病コード ~ 傷病コード の18項目のデータをカンマ区切りのCSV形式でご用意ください。Excel形式のレセプトデータがある場合には、上記の18項目について、カンマ区切りのCSV形式で保存したものをご用意ください。**傷病名が無い場合も必ずカンマだけは入れて下さい**【後述のよくあるエラー参照】。CSVで保存する前に全セルを右クリックし「セルの書式設定」を「数値」「小数点以下桁数ゼロ」にして下さい。さもないと101が0101と保存されPDMを走らせると「傷病名無し」エラーです。

#### レセプトデータフォーマット

	項目	摘要
1	整理	修正のため便宜的に(処理とは無関係)
2	日数	レセプトの日数(ゼロ, 空白はエラー出ます)
3	点数	レセプトの点数
4	傷病コード	通常は社会保険表章用の119分類コードを使用しますが、 <b>独自の分類コード使用も可</b> 。その場合必ず傷病マスターファイルと揃えてください。異なった番号があるとエラーになります。
5	傷病コード	
	…<略>…	
17	傷病コード	
18	傷病コード	

傷病マスターファイル(通常は119分類ですが、独自の傷病分類も可能です)

	項目	摘要
1	傷病コード	レセプトデータと合致すればどんな番号でもよい(119分類の他, 18大分類でも160の分類で何でもよい)
2	傷病名称	出力レポートにはこの名称が出力される

重みづけファイル(自動算出される平均値とは違う重みで分類し比較したい時に使用)

	項目	摘要
1	傷病コード	<b>必ず傷病マスターファイルと一致させる</b>
2	一件当たり日数	レセプトの日数(ゼロ, 空白はエラー出ます)
3	<b>一日当たり点数</b>	他のレセプト統計を用いる時は必ず一日当点数に直して下さい

\* 傷病マスターファイルは、社会保険表章用の119分類のものがデフォルトで用意されています。

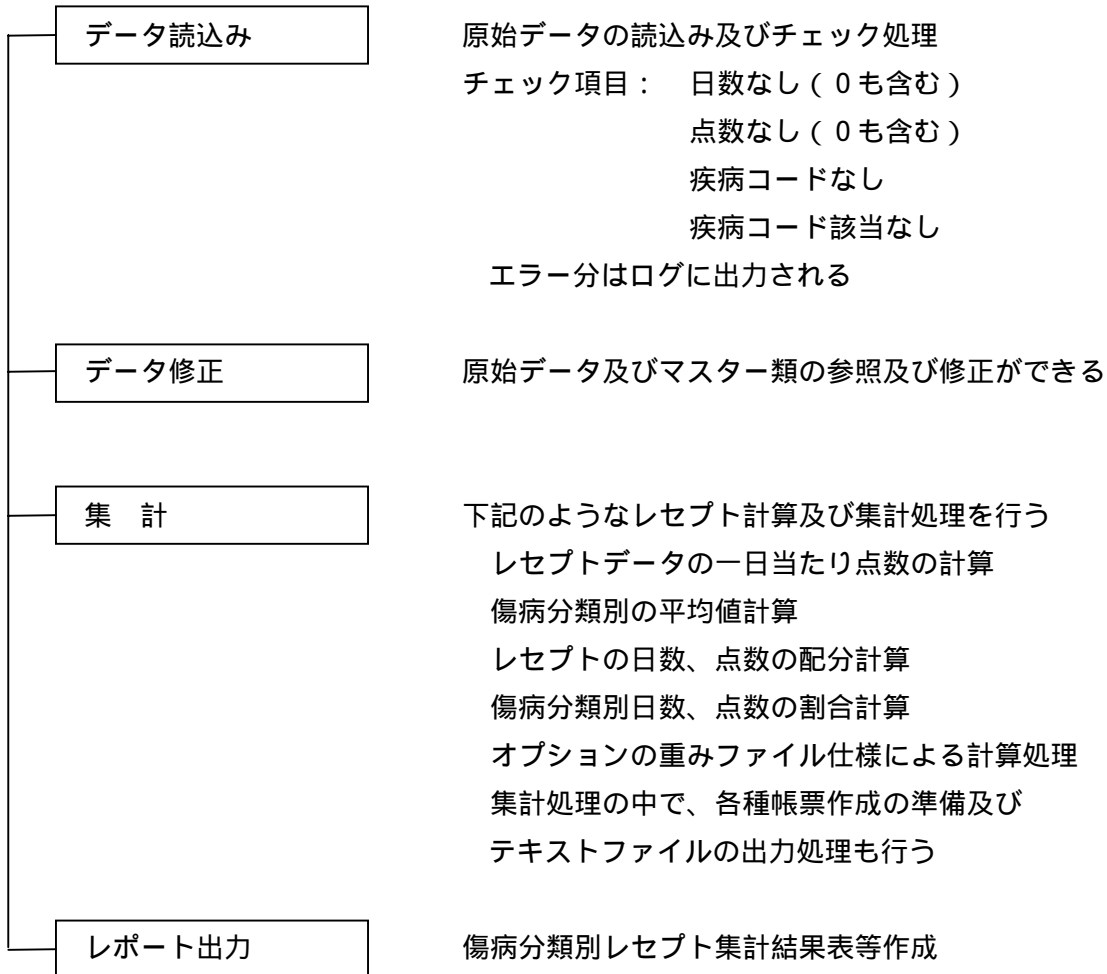
\* 重み付けファイルのフォーマットは上記参照。

\* 参考までに、下記の重み付けファイルが用意されています。

- ・1999SICS-ip.csv : 1999年社会医療調査による傷病別1件当り日数, 1日当り点数[入院]
- ・1999SICS-op.csv : 1999年社会医療調査による傷病別1件当り日数, 1日当り点数[外来]
- ・1999PtSurvey-ip.csv : 1999年患者調査による傷病別の主傷病となる割合(日数, 点数共通)[入院]
- ・1999PtSurvey-op.csv : 1999年患者調査による傷病別の主傷病となる割合(日数, 点数共通)[外来]
- ・all-one.csv : 全ての傷病の重みを等しいとしたもの

\* 上記の重みづけファイルは119分類以外では使えません。

## システムの機能構成



## 操作方法

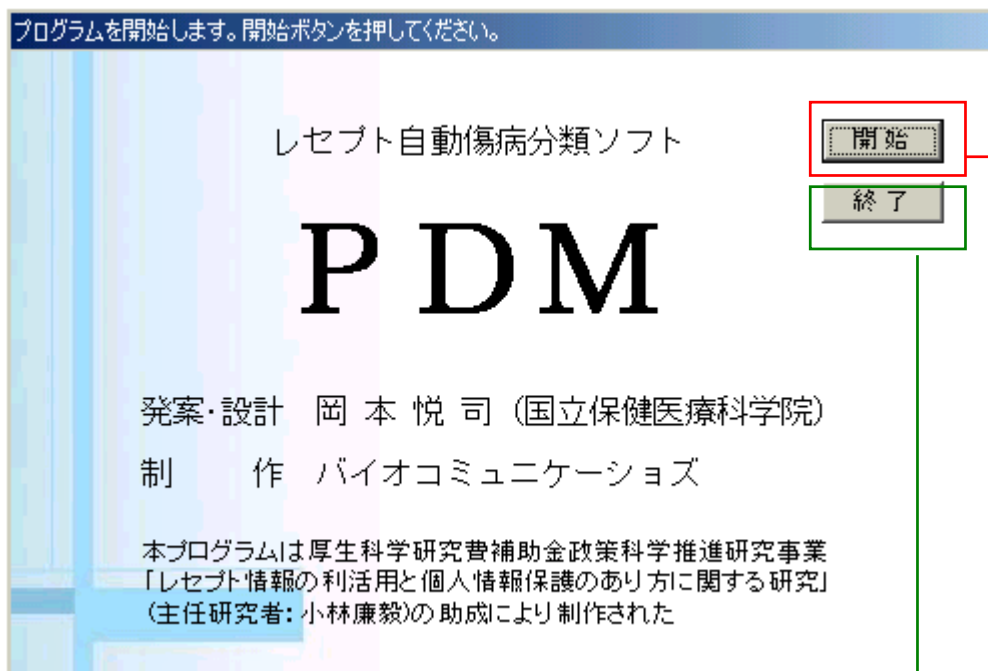
### 1. 起動方法

#### 1-1. 起動方法

Windowsを立ち上げた状態で操作を始めて下さい。

Windowsのスタートメニューから、プログラム - レセプト自動傷病分類ソフトをダブルクリックして下さい。

下図の画面が立ち上がります。



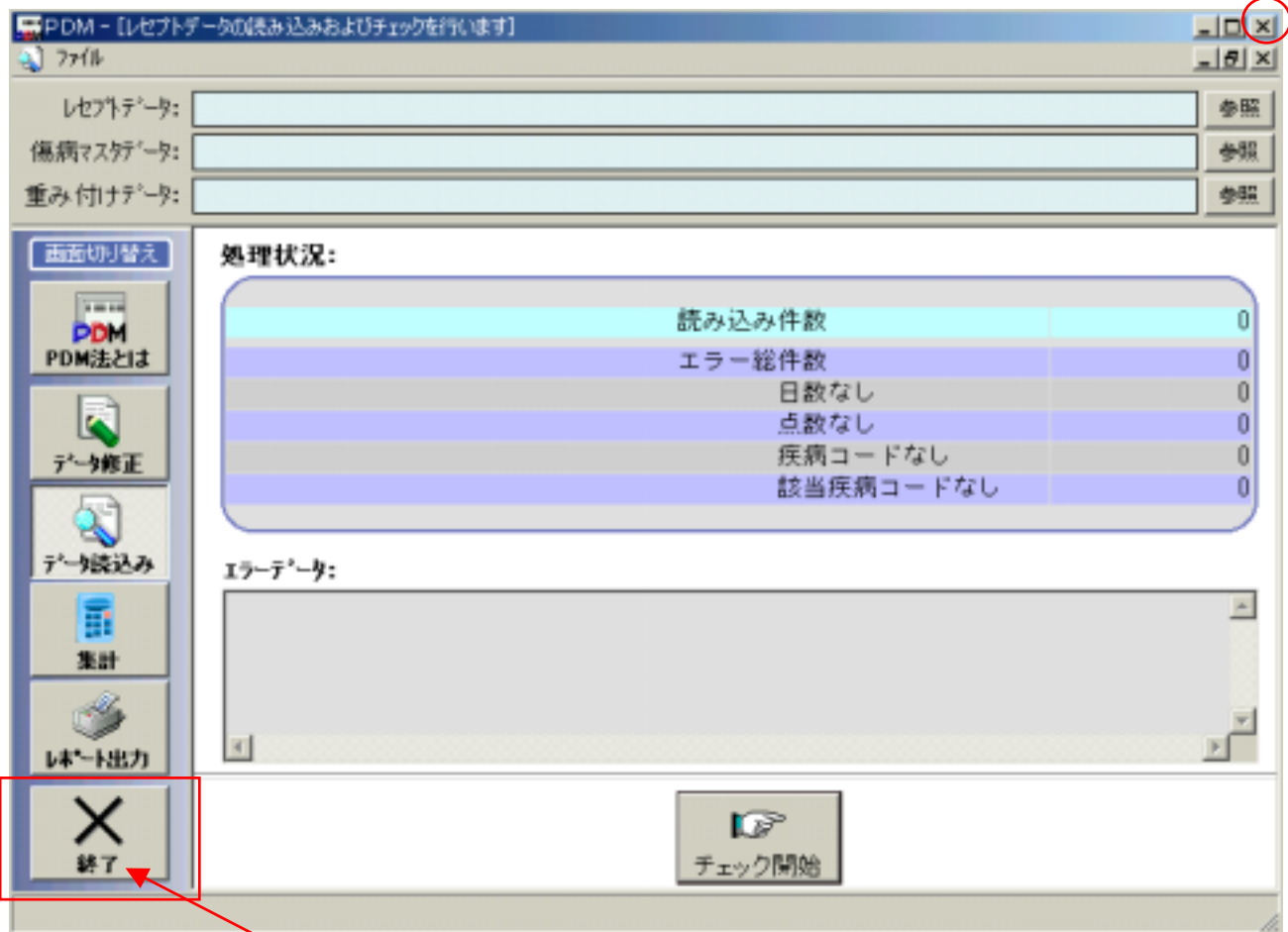
“開始”をクリックするとPDMシステム画面が表示されます。

当システムの立ち上げを中止する場合は“終了”を押して下さい。

## 2. PDMシステム画面

### 2 - 1 . PDM システム画面

PDMシステム画面が表示されます。



システムを終了させる場合は、画面切り替えメニューの“終了”ボタンか右上の×をクリックして下さい。

通常の処理手順は下記ようになります。

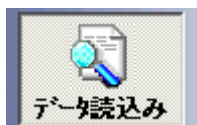
- ・ レセプトデータ及び傷病マスタデータ、重み付けデータの指定
- ・ データ読み込み
- ・ データ修正
- ・ 集計
- ・ レポート出力

上記の順に沿って、以下説明をします。

画面切り替えメニューのうち、処理したいメニュー項目ボタンをクリックして下さい。

### 3. データ読み込み

#### 3 - 1 . データ読み込み画面



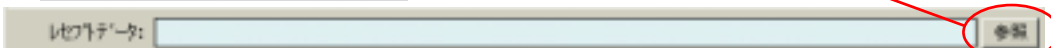
画面切り替えメニューで“データ読み込み”を選択すると、次の画面を表示します。



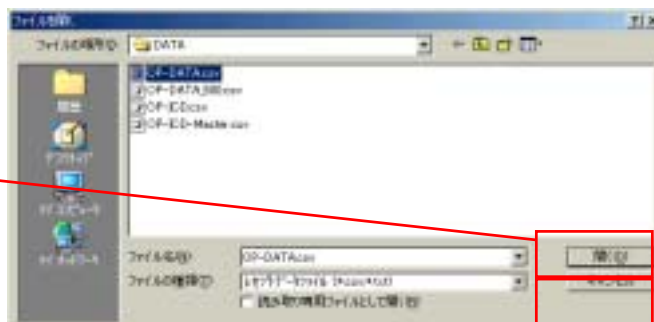
#### 3 - 2 . ファイルの指定

読み込むデータファイルを指定します。参照ボタンをクリックしてください。下記の画面を表示します。

レセプトデータを指定します。



該当する読み込み用ファイルを選択して下さい。  
その後“開く”ボタンをクリックしてください。  
元の画面に戻ります。



ファイル指定の処理を取り止める場合は、“キャンセル”ボタンが右上の×をクリックして下さい。



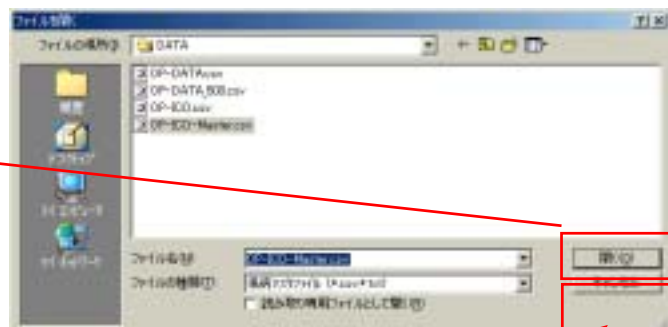
読み込むレセプトデータが指定されました。

読み込むデータファイルを指定します。参照ボタンをクリックしてください。下記の画面を表示します。

傷病マスタデータを指定します。

傷病マスタデータ:

該当する読み込み用ファイルを選択して下さい。  
その後“開く”ボタンをクリックしてください。  
元の画面に戻ります。



ファイル指定の処理を取り止める場合は、“キャンセル”ボタンが右上の×をクリックして下さい。

傷病マスタデータ: C:\PDM\DATA\OP-ICD-Master.csv

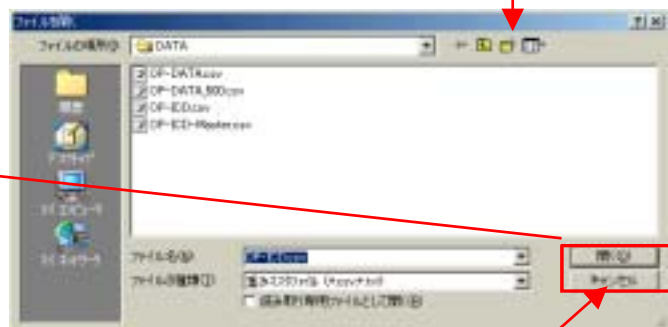
読み込む傷病マスタデータが指定されました。

読み込むデータファイルを指定します。参照ボタンをクリックしてください。下記の画面を表示します。

重み付けデータを指定します。

重み付けデータ:

該当する読み込み用ファイルを選択して下さい。  
その後“開く”ボタンをクリックしてください。  
元の画面に戻ります。



ファイル指定の処理を取り止める場合は、“キャンセル”ボタンが右上の×をクリックして下さい。

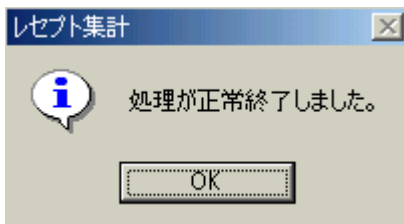
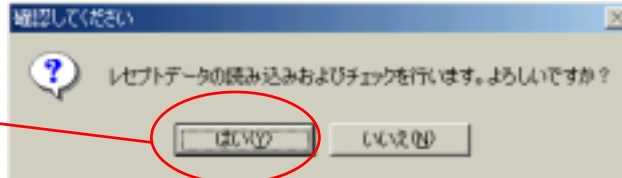
重み付けデータ: C:\PDM\DATA\OP-ICD.csv

読み込む重み付けデータが指定されました。

### 3 - 2 . データ読み込み処理

読み込みデータファイルを指定後、 をクリックすると、下記画面を表示します。

読み込みファイルが正しいか確認し、はい(Y)をクリックしてください。  
データを読み込み、エラーチェックを開始します。



処理終了のメッセージが出力されます。OKをクリックしてください。

処理中、処理状況が表示されます。エラーデータについて確認し、必要であれば、データを修正してください。修正方法については後述します。

処理状況:

読み込み件数	15754
エラー総件数	10
日数なし	0
点数なし	0
疾病コードなし	0
該当疾病コードなし	10

エラーデータ:

```

<00036> RCEP-04: 17197,1,19708,281,1112,509,1908,903,861,74,0,0,0,0,0,0
<00162> RCEP-04: 17697,1,4470,603,74,0,0,0,0,0,0,0,0,0,0,0
<00375> RCEP-04: 16886,2,5240,801,75,1886,0,0,0,0,0,0,0,0,0,0
<00795> RCEP-04: 1978,2,16290,1185,1306,1880,1309,74,1305,1319,982,1003,1806,0,0,0
<00783> RCEP-04: 12916,1,6720,1281,75,0,0,0,0,0,0,0,0,0,0,0
<00168> RCEP-04: 3388,8,16810,901,75,483,0,0,0,0,0,0,0,0,0,0
    
```

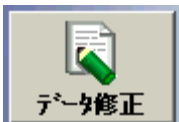
何件めのデータがエラーなのか示している。

<エラーコード>

- RCEP-01: レセプトデータに日数が入っていない
- RCEP-11: レセプトデータの日数に不正な値が入っている
- RCEP-02: レセプトデータに点数が入っていない
- RCEP-21: レセプトデータの点数に不正な値が入っている
- RCEP-03: レセプトデータに傷病コードが1つも入っていない  
もしくは正しい傷病コードが1つも入っていない
- RCEP-04: レセプトデータの傷病コードが傷病マスタに存在しない

## 4. データ修正

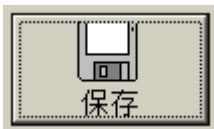
### 4-1. データ修正



画面切り替えメニューで“データ修正”を選択すると、次の画面を表示します。タグをクリックすると、レセプトデータ、傷病マスタデータ、重みマスタデータそれぞれの画面が表示されます。

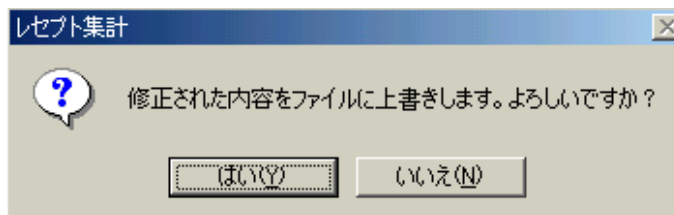


修正したい箇所をクリックし、修正します。



をクリックし、修正したデータを上書き保存します。次のメッセージが表示されるので、「はい」をクリックしてください。

上書き保存を取り止めるときは、「いいえ」をクリックしてください。

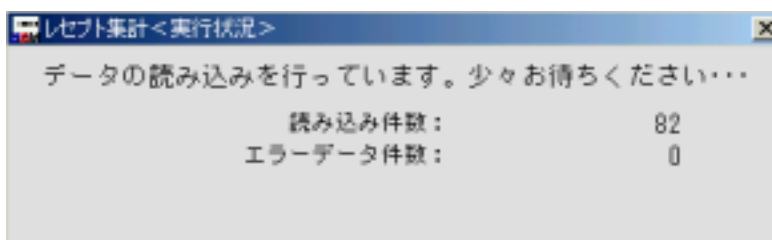


<メモ>



データを読み込む前にデータ修正画面を表示すると、左のようなメッセージが表示されます。

はいをクリックすると、読み込みを始めます。ただし、エラー情報などは表示されませんので、前述の **データ読み込み処理** 手順に従うことを推奨いたします。



## 5. 集計

### 5 - 1 . 集計



画面切り替えメニューで“集計”を選択すると、次の画面を表示します。

集計処理時のオプション:

重み付けデータを使用する  
 集計結果をテキストファイルに出力する  
ファイルパス: C:\PDM\RESULT\_2002XXXX\XX時XX分.csv

傷病分類別 日数、1日当り点数の集計:

読み込み件数	0
集計済み件数	0

各レセプト内の傷病分類別 配分日数、配分点数の集計:

読み込み件数	0
集計済み件数	0

集計開始

“集計処理時のオプション”を確認します。

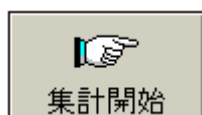
デフォルトでは次のような設定になっています。適宜、変更してください。

#### 集計処理時のオプション:

- 重み付けデータを使用する  
 集計結果をテキストファイルに出力する  
ファイルパス: C:\PDM\RESULT\_20021015\_13時54分.csv

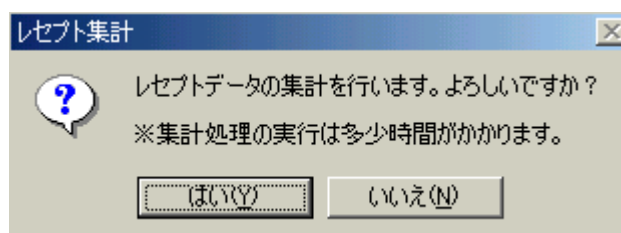
<デフォルトで設定されている内容>

- 重み付けデータを使用しない  
集計結果をテキストファイルに出力する  
出力するファイル名



集計処理を開始します。次のメッセージが表示されます。

もう一度集計処理時のオプションを確認の上、「はい」をクリックしてください。  
集計開始の指示を取り止める場合は「いいえ」をクリックしてください。



集計処理を行います。

傷病分類別 日数、1日当たり点数の集計、各レセプト内の傷病分類別 配分日数、配分点数の集計件数がカウントアップ表示されます。

傷病分類別 日数、1日当たり点数の集計:	
読み込み件数	500
集計済み件数	488

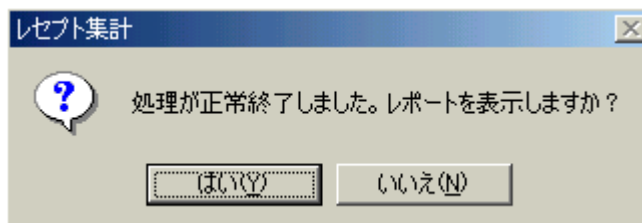
各レセプト内の傷病分類別 配分日数、配分点数の集計:	
読み込み件数	500
集計済み件数	488

集計処理が終わると下記のメッセージが表示されます。

すぐにレポートを表示したいときは「はい」を、そうでなければ「いいえ」をクリックしてください。

「いいえ」を選んでも、あとでレポート表示することは可能です。

レポート表示には多少時間がかかります。



### 【しばしば遭遇するエラーメッセージ】

エラーが出た場合はレセプトデータファイルのどの行でストップしたか、整理番号を控えてチェックして下さい。

#### 実行時エラー'6' オーバーフローしました。

これはゼロでわろうとしたときに出来ます。レセプトデータの日数にゼロや空白はありませんか？重みづけファイルの一件当たり日数にゼロや空白はありませんか？ PDM では一日当たり点数を重みにしますので、点数を自動的に日数で割ります。もし日数にゼロや空白があるとこのエラーが出ます。

#### 実行時エラー'9' インデックスが有効範囲にありません

csv ファイルを必ず、ワードパッド又はメモ帳といったテキストエディターで観察して下さい。このエラーはレセプトデータや重みづけファイルに必要なカンマがたりない時にでます。たとえば8番目の4日、2000点、傷病名が901、402の2傷病だけのレセプトの場合

8,4,2000,901,402,,,,,,,,,

とカンマが12続かなければなりません(傷病名は必ず15あるとみなされるから)が、もし

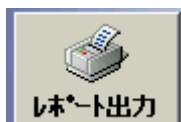
8,4,2000,901,402,,,

と十分な数のカンマがないとこのエラーになります。15傷病あるべきものが途中でなくなっているのですから…。

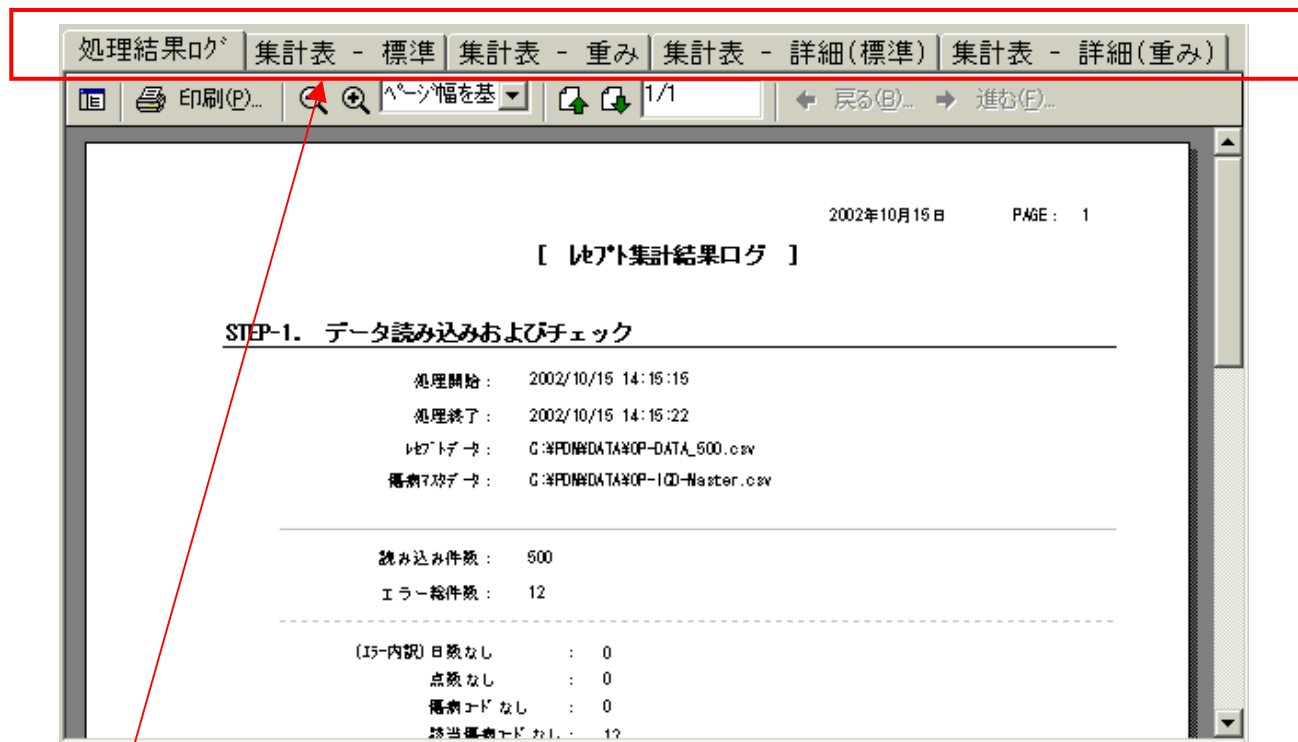
レセプトデータファイルは Excel の「ファイル」名前つけて保存 カンマつき(csv)で保存、の作業で行なわれることが多いと思いますが、Excel がレセプトデータの途中の行から12つけるべきカンマを8つかつかなかったり、といった異常がでるようです。これは本プログラムのバグではなく、Excel の欠陥です。原因は今のところ不明で、少なくともこのエラーが出たらテキストエディターで中身をチェックすることを忘れないで下さい。

## 6. レポート出力

### 6-1. レポート表示



画面切り替えメニューで“レポート出力”を選択すると、次の画面を表示します。  
レポート表示には多少時間がかかります。



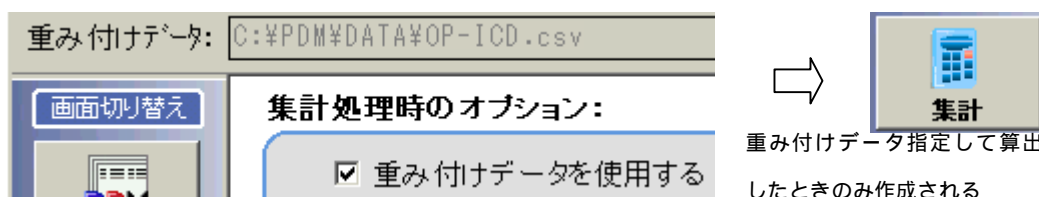
表示したいレポートのタグをクリックしてください。

#### <レポートの種類>

集計処理結果ログ …………… 集計処理において使用したデータ明細，エラー内訳等の処理結果情報

集計表（標準） …………… レセプトデータの平均日数，1日平均あたり点数等をもとに当システムで傷病分類毎の重み付けを決定して算出したもの

集計表（重み） …………… 集計処理時に重み付けデータを別途指定して算出したもの  
重み付けデータを指定して算出したときのみ作成される

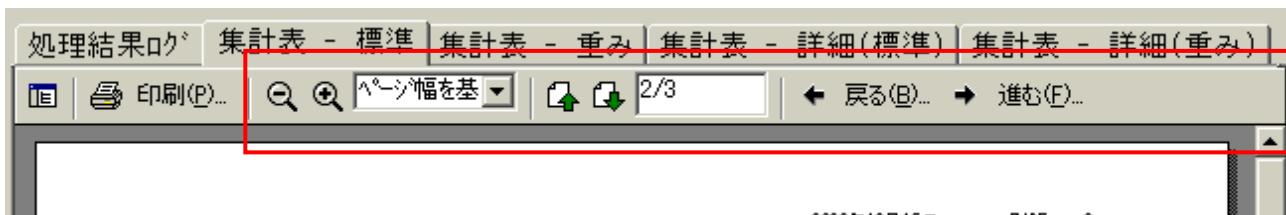


重み付けデータ指定して算出したときのみ作成される


集計表 - 詳細（標準） …… 集計表（標準）の詳細


集計表 - 詳細（重み） …… 集計表（重み）の詳細

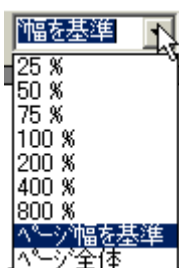
## 各種ボタンの説明



### 縮小 / 拡大ボタン


 レポートを縮小表示したいとき、このボタンをクリックすると縮小されます。


 レポートを拡大表示したいとき、このボタンをクリックすると拡大されます。




表示拡大率を指定すると、その拡大率で表示されます。


### 表示ページ

 レポートが複数ページある場合、現在表示しているページの前ページを表示します。

 レポートが複数ページある場合、現在表示しているページの次ページを表示します。

現在表示しているページ数が表示されます。  
また、表示したいページ数を入力すると、そのページが表示されます。

 戻る(B)... 表示した履歴に従って、前の状態に戻ります。

 進む(F)... 表示した履歴に従って、次の状態へ進みます。

## 6 - 2 . レポート印刷



印刷したいとき、このボタンをクリックしてください。  
プリンタ設定画面が表示されます。

# 出力帳票

< 出力帳票 例 1 >

2002年10月10日

PAGE : 1

## [ 傷病分類別 レポート集計結果表 ]

傷病名	ICD	出現数	日数(AVG)	点数(AVG)	日数(%)	点数(%)
1 腸管感染症	101	121	1.5	7334	9.75	6.849
2 結核	102	27	1.7	15956	1.21	2.666
3 主として性的伝播様式を	103	16	1.8	19014	1.301	2.319
4 皮膚及び粘膜の病変を伴	104	58	2.9	4815	9.411	4.232
5 ウイルス肝炎	105	10	2.2	6235	0.464	0.298
6 その他のウイルス疾患	106	13	2.8	18318	1.629	1.775
7 真菌症	107	92	1.7	5357	6.854	3.636
8 感染症及び寄生虫症の続	108	0	0	0	0	0
9 その他の感染症及び寄生	109	13	1.4	6771	0.701	0.449
10 胃の悪性新生物	201	71	1.9	10606	3.631	4.628
11 結腸の悪性新生物	202	50	1.6	8196	2.24	2.493
12 直腸S状結腸移行部及び直	203	24	1.9	8128	1.516	1.328
13 肝及び肝内胆管の悪性新	204	11	2.1	14926	0.701	1.445
14 気管、気管支及び肺の悪	205	44	2.5	14593	3.744	6.082
15 乳房の悪性新生物	206	22	1.5	8936	0.973	1.114
16 子宮の悪性新生物	207	1	1	19490	0.023	0.092
17 悪性リンパ腫	208	0	0	0	0	0
18 白血病	209	0	0	0	0	0
19 その他の悪性新生物	210	11	1.8	16614	0.238	0.777
20 良性新生物及びその他の	211	58	1.6	10556	1.81	2.933
21 貧血	301	35	2	11319	1.357	1.452
22 その他の血液・造血器の疾	302	7	1.9	23613	0.238	1.226
23 甲状腺障害	401	7	1.6	12770	0.17	0.325
24 糖尿病	402	41	2.3	12792	1.957	2.94
25 その他の内分泌、栄養及	403	65	1.9	8985	2.669	2.233
26 血管性及び詳細不明の痲	501	1	1	23380	0.011	0.064
27 精神作用物質使用による	502	0	0	0	0	0
28 精神分裂病、分裂病型障	503	0	0	0	0	0
29 気分[感情]障害(躁うつ病	504	1	3	3877	0.045	0.006
30 神経症性障害、ストレス関連障	505	12	4.1	10449	0.905	0.951
31 精神遅滞	506	0	0	0	0	0
32 その他の精神及び行動の	507	1	2	5340	0.057	0.023
33 パーキンソン病	601	0	0	0	0	0
34 アルツハイマー病	602	0	0	0	0	0
35 てんかん	603	3	1	11767	0.057	0.045
36 脳性麻痺及びその他の麻	604	2	4.5	13212	0.226	0.234
37 自律神経系の障害	605	0	0	0	0	0
38 その他の神経系の疾患	606	4	2.5	10706	0.362	0.643
39 結膜炎	701	7	4.3	10020	0.826	0.352
40 白内障	702	10	1.7	6845	0.17	0.11
41 屈折及び調節の障害	703	7	1.4	6751	0.102	0.122
42 その他の眼及び付属器の	704	8	1.9	10070	0.17	0.204
43 外耳炎	801	0	0	0	0	0
44 その他の外耳疾患	802	0	0	0	0	0
45 中耳炎	803	0	0	0	0	0
46 その他の中耳及び乳様突	804	0	0	0	0	0
47 メニエール病	805	0	0	0	0	0
48 その他の内耳疾患	806	0	0	0	0	0
49 その他の耳疾患	807	0	0	0	0	0
50 高血圧性疾患	901	56	1.8	12584	1.561	2.195
51 虚血性心疾患	902	49	2.2	11316	1.685	2.051
52 その他の心疾患	903	41	2.1	12521	1.21	1.495
53 くも膜下出血	904	1	1	700	0.011	0
54 脳内出血	905	1	2	30650	0.011	0.167
55 脳梗塞	906	15	1.6	10363	0.339	0.304
56 脳動脈硬化(症)	907	0	0	0	0	0
合計		1796	1907	941728	99.999	99.999

[ 傷病分類別 レポート集計結果詳細表 ]

傷病名	ICD	出現数	Σ日数	Σ点数	日数(AVG)	点数(AVG)	総配分日数	総配分点数	日数(%)	点数(%)
1 膣管感染症	101	121	178	887471	1.5	7334	86.2	480051	9.75	6.849
2 結核	102	27	45	430812	1.7	15956	10.7	186868	1.21	2.666
3 主として性的伝播様式を	103	16	29	304225	1.8	19014	11.5	162522	1.301	2.319
4 皮膚及び粘膜の病変を伴	104	58	169	279297	2.9	4815	83.2	296618	9.411	4.232
5 ウイルス肝炎	105	10	22	62350	2.2	6235	4.1	20835	0.464	0.298
6 その他のウイルス疾患	106	13	37	238140	2.8	18318	14.4	124406	1.629	1.775
7 真菌症	107	92	160	492855	1.7	5357	60.6	254844	6.854	3.636
8 感染症及び寄生虫症の統	108	0	0	0	0	0	0	0	0	0
9 その他の感染症及び寄生	109	13	18	88022	1.4	6771	6.2	31495	0.701	0.449
10 胃の悪性新生物	201	71	134	753016	1.9	10606	32.1	324381	3.631	4.628
11 結腸の悪性新生物	202	50	81	409776	1.6	8196	19.8	174724	2.24	2.493
12 直腸S状結腸移行部及び直	203	24	45	195083	1.9	8128	13.4	93068	1.516	1.328
13 肝及び肝内胆管の悪性新	204	11	23	164181	2.1	14926	6.2	101282	0.701	1.445
14 気管、気管支及び肺の悪	205	44	109	642076	2.5	14593	33.1	426259	3.744	6.082
15 乳房の悪性新生物	206	22	32	196597	1.5	8936	8.6	78086	0.973	1.114
16 子宮の悪性新生物	207	1	1	19490	1	19490	0.2	6483	0.023	0.092
17 悪性リンパ腫	208	0	0	0	0	0	0	0	0	0
18 白血病	209	0	0	0	0	0	0	0	0	0
19 その他の悪性新生物	210	11	20	182755	1.8	16614	2.1	54460	0.238	0.777
20 良性新生物及びその他の	211	58	92	612252	1.6	10556	16	205598	1.81	2.933
21 貧血	301	35	71	396165	2	11319	12	101790	1.357	1.452
22 その他の血液・造血器の疾	302	7	13	165288	1.9	23613	2.1	85954	0.238	1.226
23 甲状腺障害	401	7	11	89388	1.6	12770	1.5	22756	0.17	0.325
24 糖尿病	402	41	95	524453	2.3	12792	17.3	206040	1.957	2.94
25 その他の内分泌、栄養及	403	65	121	583999	1.9	8985	23.6	156494	2.669	2.233
26 血管性及び詳細不明の病	501	1	1	23380	1	23380	0.1	4495	0.011	0.064
27 精神作用物質使用による	502	0	0	0	0	0	0	0	0	0
28 精神分裂病、分裂病型障	503	0	0	0	0	0	0	0	0	0
29 気分[感情]障害 [味うつ病	504	1	3	3877	3	3877	0.4	426	0.045	0.006
30 神経症性障害、ストレス関連障	505	12	49	125383	4.1	10449	8	66641	0.905	0.951
31 精神遅滞	506	0	0	0	0	0	0	0	0	0
32 その他の精神及び行動の	507	1	2	5340	2	5340	0.5	1629	0.057	0.023
33 パーキンソン病	601	0	0	0	0	0	0	0	0	0
34 アルツハイマー病	602	0	0	0	0	0	0	0	0	0
35 てんかん	603	3	3	35300	1	11767	0.5	3153	0.057	0.045
36 慢性麻痺及びその他の麻	604	2	9	26424	4.5	13212	2	16386	0.226	0.234
合計		1798			190.7	941728	984.1	7009082	99.999	99.999

< 出力帳票 項目説明 >

出現数 : 該当傷病名のあるレセプト件数

日数 : 該当傷病名のあるレセプトの日数を集計したもの

点数 : 該当傷病名のあるレセプトの1日当り点数を集計したもの

日数 (AVG) :  $\text{日数} / \text{出現数}$  で算出

点数 (AVG) :  $\text{点数} / \text{出現数}$  で算出

総配分日数 : レセプト単位の日数を傷病毎の平均日数で配分し、傷病コード毎に集計したもの

総配分点数 : レセプト単位の点数を傷病毎の1日当り平均点数で配分し、傷病コード毎に集計したもの

日数 (%) : 総配分日数合計のうち、各傷病分類別の総配分日数が占める割合

点数 (%) : 総配分点数合計のうち、各傷病分類別の総配分点数が占める割合

注) 重みデータ使用の場合には、レセプト単位の日数及び点数配分の際にレセプトより計算した傷病毎の平均日数, 平均1日当り点数を使用せずに、重みデータとして指定したもので計算を行います。

## 出力データの形式

下記のレイアウトで、カンマ区切り CSV 形式のデータ出力を行います。

NO	項目	備考
1.	傷病コード	x x x x
2.	傷病名称	
3.	件数	
4.	日数	
5.	1日当り点数	
6.	A V日数	x x . x 日
7.	A V 1日当り点数	
8.	配分日数	
9.	配分点数	
10.	日数割合	x x . x ( % )
11.	点数割合	x x . x ( % )
12.	A V日数	オプションテーブルより引用
13.	A V 1日当り点数	オプションテーブルより引用
14.	配分日数	
15.	配分点数	
16.	日数割合	オプション使用の場合
17.	点数割合	オプション使用の場合

「(資料) 高齢者インフルエンザ予防接種の効果研究」は省略