

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

平成 15 年度総括・分担研究報告書

主任研究者 小林廉毅

平成 16 (2004) 年 3 月

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

総括・分担研究報告書

平成 15 年度（3 年計画の 3 年目）

主任研究者	小林廉毅	東京大学大学院医学系研究科 教授
分担研究者	岡本悦司	国立保健医療科学院経営科学部 室長
分担研究者	谷原真一	島根大学医学部公衆衛生学 助教授
分担研究者	豊川智之	東京大学大学院医学系研究科 助手

目次

I. 総括研究報告（小林）

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

II. 分担研究報告（岡本）

レセプト傷病分析の原理手法ならびに個人情報保護の技術に関する研究

III. 分担研究報告（谷原）

1. 介護保険と医療保険のレセプトデータの結合および個人単位での受診行動を考慮したレセプト情報の分析例

2. 市町村における国保レセプト情報の活用事例

IV. 分担研究報告（豊川）

レセプトを用いた分析の現状と保険者特性

V. 研究成果の刊行に関する一覧表

厚生労働科学研究費補助金（政策科学推進研究事業）
総括研究報告書

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 東京大学大学院医学系研究科公衆衛生学分野 教授

研究要旨

平成15年5月、個人情報保護法案が国会で可決成立し、2年以内に施行される運びになった。医学関連学会は研究の公益性に鑑み、医療における個人情報の扱いに関して種々の提言を行ってきた。このような背景から、医療・介護に係わる診療報酬明細書（レセプト）の情報についても、個人情報保護とデータの公益利用を両立させるような情報提供のあり方と方策を検討する必要がある。本年度（平成15年度）の研究では、昨年度の調査を元に保険者におけるレセプト情報の利活用について詳細な検討を行った。また分担研究として、レセプトに記載された複数傷病名について、パソコンを用いた自動分類法であるPDM法

（Proportional Disease Magnitude, version 3）の実用的最終版の完成、プライバシー保護の方策、ならびにレセプト情報を用いた地域の健康度指標や介護保険とのリンケージに関する研究を併せて行った。以上の結果、(1)保険者の規模、健康保険と国民健康保険の違い、所在地によってレセプト情報分析の傾向の違いが示唆された。保険者におけるレセプト情報の分析を推進するためには、規模の小さい保険者について分析に関する支援を行うとともに、全国的な関心を高めるようなネットワークの構築などが重要と思われた。(2)PDM法を十分な妥当性をもって一般使用に供するレベルにまで完成させることができた（作成されたプログラムは既にフリーウェアとして提供済みである）。またレセプトの法的性質を各種研究倫理指針、行政通達そして個人情報保護法に照らして検討し、レセプト情報を活用すべき諸条件ならびに現行医療保険関連法規の改善点等を明らかにした。(3)介護保険と医療保険のリンケージはレセプト単独の情報より保健事業の評価や住民の健康度を測定するための指標として有効と考えられた。またレセプト情報を用いて特定の疾患の頻度や継続受診の状況を把握することが可能と考えられた。

（分担研究者）

岡本悦司

国立保健医療科学院
経営科学部 室長

谷原真一

島根大学医学部
公衆衛生学 助教授

豊川智之

東京大学大学院医学系研究科
公衆衛生学 助手

近年の健康保険財政の逼迫は深刻さを増しており、医療・社会保障の効率化や保険者機能強化にむけた政策立案が急務となっている。診療報酬請求明細書（レセプト）情報は、医療費の動向や特定の疾患の流行等について全数かつリアルタイムに利用できるという特色を持ち、医療・社会保障政策へ様々な応用が可能な貴重なデータである。レセプト情報は行政通達により利活用が推奨されており、その調査研究体制の整備が必要と考えられる。

初年度（平成13年度）の報告において、レセプト情報を活用した調査研究の実態と動向について検討した。その結果、医療機関におけるレセプト情報を利用する場合の

A. 研究目的

プライバシー保護対策に関するガイドライン作成の重要性が示唆された。

次年度（平成14年度）の報告では、わが国の保険者におけるレセプト情報の利活用の現状と展望に加え、調査分析等の外部委託に関する調査を行い、その概況を報告した。また分担研究として、レセプトに記載された複数傷病名の客観的かつ自動的 분류法であるPDM（Proportional Disease Magnitude）法の開発、ならびに地域の健康度指標としてのレセプト情報活用に関する研究を併せて行った。

最終年度（平成15年度）は、昨年度の調査を元に保険者におけるレセプト情報の利活用について詳細な検討を行うとともに、分担研究として、レセプトに記載された複数傷病名について、パソコンを用いた自動分類法であるPDM法第3版（Proportional Disease Magnitude, version 3）の開発、ならびにレセプト情報を用いた地域の健康度指標や介護保険とのリンケージの研究を併せて行った。さらに以上の調査結果を総括し、わが国の実状に即した医療・介護に係わるレセプト情報の保護と利活用のあり方と方策について考察を加えた。

B. 研究方法

(1) 保険者特性とレセプト情報を用いた分析の現状（小林・豊川）

調査は一定の調査票を用いて、2003年1月29日より2月28日までの期間に行った。調査対象は国民健康保険（国保）と組合管掌健康保険（健保）の全保険者とした。調査票を送付した5,145保険者のうち2,017保険者より回答があり、回収率は39.2%であった。調査方法の詳細及び調査結果の概要については、昨年度の報告書に記載した。上記の資料を元に保険者特性とレセプトを用

いた調査分析の傾向について詳細な検討を加えた。

(2) レセプト傷病分類の原理と手法（PDM法）及びレセプト情報の利活用におけるプライバシー保護（岡本）

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本研究で開発してきたが、その実用的な完成版（第3版）を作成した。また、レセプト情報の利活用におけるプライバシー上の問題点とその対策についても検討した。

(3) レセプト情報の利活用の研究（谷原）

① 介護保険と医療保険のリンケージによるレセプト情報活用の可能性について

介護保険と医療保険のデータをマッチングさせ、介護保険受給者の医療機関受診状況を分析した。また個人単位の時系列データを用いて特定疾患の医療費の推計を行った。当該研究では、保険者から研究者に対して個人情報を削除した集計データが提供される方法に依った。

② 地域の健康度指標へのレセプト情報活用について

ある自治体のレセプトデータから、疾病コード（中分類）が402（糖尿病）、901（高血圧性疾患）、1901（骨折）のデータを抽出し、当該地域の健康度指標の可能性について検討した。

（倫理面への配慮）

本研究でレセプト情報を扱う場合は個人識別情報を削除した上で取り扱うこととした。

C. 研究結果

(1) 保険者特性とレセプト情報を用いた分析

の現状

レセプト情報を用いた分析の外部委託の有無別に各保険者特性の割合をみると、加入者規模に大きな差はなかった。老人保健制度加入率は、委託した保険者の方が高齢化しており、加入率 30%以上の保険者が 50%近くを占めた。他方、委託しなかった保険者では老人保健制度加入率 10%未満の保険者が 40%を占めた。保険種別にみると、委託した保険者では国民健康保険が 89%占めたが、委託しなかった保険者では国民健康保険(57%)と健康保険(43%)との差は小さかった。所在地については、委託した保険者は東海・甲信越・北陸(24%)が多く、委託しなかった保険者は関東(28%)が多かった。

保険者特性別にみた委託有無のオッズ比から、10万人以上の加入者のいる保険者では委託割合が低く(27%)、それ以外の規模の保険者では委託割合が高かった(40%前後)。また、老人保健の加入率が高いと委託する傾向が見られた。保険の種類では健康保険で委託しない保険者が有意に多く(84%)、所在地別では関東と近畿で委託しない保険者が多かった。

(2)レセプト傷病分類の原理と手法(PDM法)及びレセプト情報の利活用におけるプライバシー保護

PDM法を十分な妥当性をもって一般使用に供するレベルにまで完成させることができた(作成されたプログラムは既にフリーウェアとして提供済みであり、Web上にも詳細を記した)。プライバシー保護の観点からは、レセプトの法的性質を各種研究倫理指針、行政通達そして個人情報保護法に照らして検討し、レセプト情報を活用すべき諸条件ならびに現行医療保険関連法規の改善点等を明らかにした。

(3)レセプト情報の利活用の研究(谷原)

①介護保険と医療保険のリンケージによるレセプト情報活用の可能性について

介護保険受給者に占める医療受診者の割合は全体で 80.2%であり、要介護度が高くなるにつれて低下していった。また、糖尿病などで受診経験を有する者について、一人当たり医療費を求めた結果、60~64歳の年齢階級がもっとも一人当たり医療費が低くなる傾向が認められた。介護保険と医療保険のデータをリンケージすることはレセプト単独の情報を分析するよりも保健事業の評価や住民の健康度を測定するための指標として有効と考えられた。

②地域の健康度指標へのレセプト情報活用について

「骨折」での受診が確認されたレセプトについて分析したところ、男女で年齢分布に大きな格差が存在しており、高齢者の骨粗鬆症の大半が女性であるということと一致する結果であった。また、糖尿病および高血圧性疾患のような慢性疾患と骨折のような急性疾患で、同一医療機関を継続して受診する状況に違いが認められた。レセプトに記載された情報を断面的に分析するだけでなく、特定の情報について時系列分析することによって、重要な知見が得られる可能性を示すことができた。また、基本健康診査などの各種保健事業から得られる情報との統合によりさらに有用な知見が得られると考えられた。

D. 考察

レセプト情報の分析を外部委託している保険者は、委託していない保険者に比べてレセプト情報を分析している割合が総じて

高かった。すなわちレセプト情報の外部委託が、レセプト情報の利活用に結びつく可能性が示唆された。また委託している保険者では、保険者特性に関わらず同様の分析をしていることが示された。このことは外部委託により一定の分析が行われており、これらの保険者では問題分析意識が共通していると推測された。しかし画一的な業務上の分析に終始している可能性も考えられる。他方、委託していない保険者は、規模が大きく、健康保険の割合が高く、関東、近畿に所在している傾向がみられた。

具体的な分析項目をみると、「一人当たりの医療費、一人当たりの件数」、「疾患別医療費の分析」については、規模の大きい保険者で分析している傾向が見られた。加入者数の増加に伴ってこれらの項目の分析が、保険運営上において重要になるためと考えられる。「保健事業の効果分析」は、規模が大きく、老人保健制度加入率が高く、北海道・東北、九州・沖縄の保険者で分析されている傾向が示された。「被保険者指導の参考資料として用いる」については、北海道・東北、九州・沖縄の保険者で利活用されている傾向がみられた。全体としてレセプト情報の利活用については地域差のあることが示唆された。以上から、保険者におけるレセプト情報の分析を推進するためには、規模の小さい保険者について分析に関する支援を行うとともに、全国的な関心を高めるようなネットワークの構築などが重要と思われる。なお、このようなネットワーク構築の一環として、「第6回レセプト情報の活用を考える自由集会」を第62回日本公衆衛生学会総会（京都、2003年10月）において開催した。

また、プライバシー保護の観点と、レセプト情報の利活用に関する法・制度面の分析を通じて、レセプト情報活用と個人情報

保護は十分両立可能であることを、技術的制度的裏付けとともに示すことができた。目下急速に進展しているレセプト情報の電子化と本研究の成果を組み合わせることにより、レセプト情報を医療政策や公衆衛生目的に有効活用することが将来、可能と思われる。

さらに、医療保険のレセプトデータを介護保険など他のデータとリンケージすることによって、医療保険レセプト単独の情報を分析するよりも保健事業の評価や住民の健康度を測定するための指標として有効であることが示唆された。

E. 結論

本研究により、保険者の規模、健康保険と国民健康保険の違い、所在地によってレセプト情報分析の傾向の違いが示唆された。保険者におけるレセプト情報の分析を推進するためには、規模の小さい保険者について分析に関する支援を行うとともに、全国的な関心を高めるようなネットワークの構築などが重要と思われる。また、プライバシー保護の観点と、レセプト情報の利活用に関する法・制度面の分析を通じて、レセプト情報活用と個人情報保護は十分両立可能であることを、技術的制度的裏付けとともに示した。目下急速に進展しているレセプト情報の電子化と本研究の成果を組み合わせることにより、レセプト情報を医療政策や公衆衛生目的に有効活用することが将来、可能になると思われる。さらに、医療保険のレセプトデータを介護保険など他のデータとリンケージすることによって、医療保険レセプト単独の情報を分析するよりも保健事業の評価や住民の健康度を測定するための指標として有効であることが示唆された。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

Okamoto E. Reduction of influenza-related outpatient visits among community dwelling elderly who received influenza vaccination. 薬剤疫学誌8巻2号[2003年12月]55～60頁

岡本悦司、畑栄一. レセプト傷病分析の原理とシミュレーションによる妥当性の検討. 日本公衆衛生雑誌50巻12号[2003年12月]1135～1142頁

岡本悦司、田原康玄. レセプト全傷病分析による町村間ならびに月間変動の分析. 厚生学の指標50巻11号[2003年11月]14～20頁

岡本悦司. 公衆衛生研究における疫学研究倫理指針の適用. 日本公衆衛生雑誌50巻11号[2003年11月]1079～1089頁

2. 学会発表

豊川智之、小林廉毅、岡本悦司、谷原真一. 保険者におけるレセプト情報を用いた調査分析の現状. 第62回日本公衆衛生学会総会、京都、2003年10月.

Toyokawa S, Kobayashi Y, Okamoto E, Tanihara S. Utilization of insurance claims for health services research. The 131th annual meeting of American Public Health Association. San Francisco, November 2003.

岡本悦司、田原康玄. 全傷病分類による傷病別医療費割合の市町村間ならびに月間変動の分析. 第61回日本公衆衛生学会総会、

さいたま、2002年10月. (昨年度追加分)

岡本悦司. 公開個票データを用いたPDM法による虫垂切除術例の傷病別費用推計. 第62回日本公衆衛生学会総会、京都、2003年10月.

岡本悦司. コストデータの日米比較—臨床研究への活用可能性の視点. 第25回臨床研究生物統計研究会. 仙台、2003年11月.

Okamoto E. PDM method for estimating disease specific cost in a population of health insurance claims with multiple diagnoses. The 4th international conference of International Health Economics Association (iHEA). San Francisco, June 2003.

Okamoto E. Effects of add-on copayment of prescription drugs on drug use by disabled elderly. The 130th annual meeting of American Public Health Association. Philadelphia. November 2002. (昨年度分追加)

Okamoto E. Assessment of Effectiveness of FLU Vaccination for the Elderly using Health Insurance Claims. The 131th annual meeting of American Public Health Association. San Francisco. November 2003.

H. 知的所有権の取得状況

なし

レセプト傷病分析の原理手法ならびに個人情報保護の技術に関する研究

分担研究者 岡本悦司(国立保健医療科学院研究情報センター)
研究協力者 畑 栄一(国立保健医療科学院研修企画部)

研究要旨

個人情報保護をはかりつつレセプト情報の有効な利活用を可能にする3つの基盤整備と取り組んだ。ひとつはレセプト情報活用と個人情報保護とのあり方をめぐる法制度的研究、もうひとつは電子化されたレセプト情報を人手を介することなく自動的かつ客観的に分析する原理とコンピュータープログラムの開発(PDM法)、そしてプライバシーを保護しつつデータリンクを行うための技術開発である。

その結果、PDM法を十分な妥当性をもって一般使用に供するレベルにまで完成させることができ、本研究によって作成されたプログラムは既にフリーウェアとして提供済である。法制度面ではレセプトの法的性質を各種研究倫理指針、行政通達そして個人情報保護法に照らして検討し、レセプト情報を活用すべき諸条件ならびに現行医療保険法の改善点等を明らかにした。そしてレセプトと検診カルテと他の個人情報とをリンクすることをプライバシーを保護しつつ行うため簡便な暗号手法を提示し、連結可能匿名化して研究を実施する方法を示した。

本研究を通じて、レセプト情報活用と個人情報保護は十分両立可能であることを、技術的制度的裏付けとともに示すことができた。目下急速に進展しているレセプト情報の電子化と本研究の成果を組み合わせることにより、レセプト情報を医療政策や公衆衛生目的に有効活用する方策を示すことができた。

分担研究者が担当した分野は大きく3の柱からなっているが、法制度的研究については初年度の13年度報告書で詳述したので、最終報告書では2つのテーマについて個別に目的、方法、結果、結論と記載してゆく。

1, PDM法の改善と妥当性の検証

A.研究目的

PDM(Proportional Disease Magnitude)法は、95年に岡本が初めて提案した、複数傷病名が記載されたレセプトデータの集合において、個々のレセプトの合計点数と記載傷病名はわかっているが、個々の傷病ごとの日数や医療費の対応がわからない場合に、傷病別の日数や医療費を推計するための原理である。PDM法は各傷病について共通の「重み(マグニチュード)」を仮定し、各レセプトごとの日数や点数を記載された傷病名ごとの重みに応じて比例配分し(本研究の過程で、本手法がレセプトのみならず、たとえば献立と総カロリーはわかっているが個々の食品ごとのカロリーがわからない場合における栄

養調査のような場合にも適用できる普遍性あるものと判明したため、今後PDMは Proportional Distribution Methodと呼び変える予定である。しかし本報告書では従来通り傷病マグニチュード按分法という名称を用いる、次に傷病ごとの配分日数や点数を合計するものである。しかしながら、どのような数値を重みに使えば妥当か、については本研究開始時点は明確ではなかった。

13年度において、重みに各傷病の記載された全レセプトの日数点数の単純平均を「重み」としてPDM法Ver1を完成させた。妥当性の検証が未着手であったが、ともかくも複数傷病名レセプトを自動的に分析するプログラムの完成は意義深いものであった(初年度報告書にはまにあわず記載されていない)。

14年度においては、乱数で生成した正解既知のデータを用いて妥当性を検証するシミュレーションの手法を導入し、また平均値に一定の補正を加えることによって高額傷病も少額傷病もきわめて良好な妥当性で推計できることを明らかにした(本成果は日本公衆衛生雑誌2003年12月号に掲載)。本成果を踏まえ、PDM法Ver2を作成し、初めてインターネット上(<http://resept.com>)でフリーウェアとして公開した。

最終年度では、前年度の成果をふまえPDM法における傷病別点数の「重み」算定のさらなる改善と今後の方向性、とりわけ14年度報告書で触れた(32頁)「繰り返し法」についてさらなる検討を行う。

B, 方法

PDM法Ver2では、たとえば平均傷病数が3.7なら2次補正が妥当、というふうに全ての傷病名の記載レセプトが同一の傷病数であることを前提に主動で補正次数を決定していたが、それを自動化し、傷病ごとにその傷病が記載されたレセプトの平均傷病数を算出し、それに応じて補正次数を自動的に決定する補正式の考案を試みた。

【重み推計法1】平均値補正法の自動化

平均値とその補正法については昨年度報告書で詳説したが、今年度で開発した補正法を説明するため必要部分を以下に引用する。

-----<14年度報告書31頁からの引用>-----

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ΔPだけ大きく(小さく)なってもレセプトの点数PはΔP/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μ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μを500とすると、上式は以下のようなになる。

$$N=2 \text{ なら } Y = 0.5 * X + 250$$

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

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

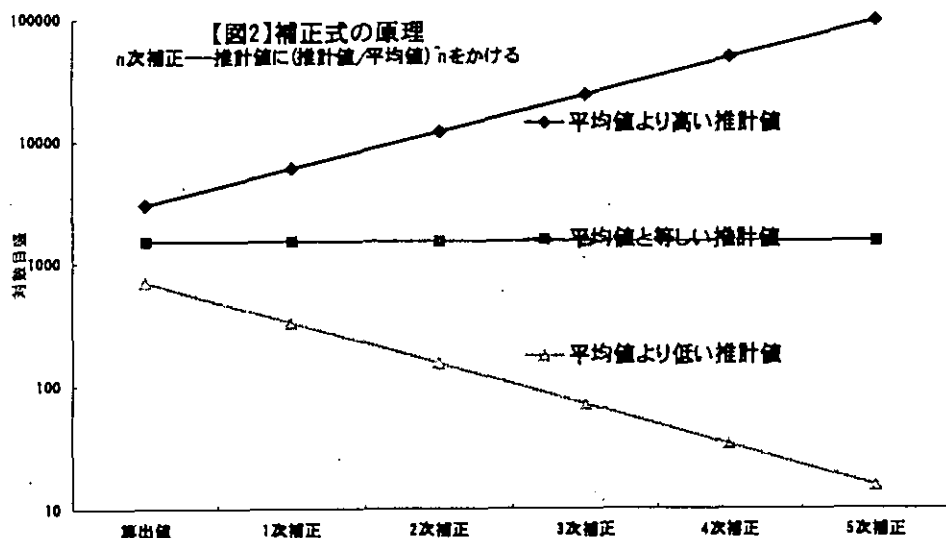
このようにレセプトの傷病名数が多くなるほど直線は平べったくなってゆく。ちなみに外来レセプトの平均傷病名数は約4, 入院レセプトは約8である。

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μだけである。そこでYとμ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μと同じであればそのまま, もしYが全体平均μ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すと【図2】のようになる。



-----<引用終わり>-----

14年度にはシミュレーションの手法を用いて経験的に, たとえば傷病数4の場合は2次補正が適当, とした。しかしながら傷病によってはたとえばインフルエンザのように単独でしか出現しない傷病や糖尿病のように多数の傷病を伴うものもある(常に単独でしか出現しない傷病のレセプトの平均値は補正せずそのまま用いなければならない)。それらがいっしょくたに同一の補正を受けることになる。また外来レセプトは平均傷病数が少ないから二次補正, 傷病数の多い入院レセプトは3次補正, というふうに手動で選択しなければならない。

Ver3では, 傷病ごとにその傷病の記載されたレセプトの傷病数(N)をコンピューターが自動的に計算し, 適切な補正次数kを決定する方法と取組んだ。そしてその傷病の記載されたレセプトの平均傷病数をNとするとその傷病の平均値を以下のように補正する式を考案した。

$$1) \text{ 傷病平均} > \text{全体平均}: \text{傷病平均} \times \left(\frac{\text{傷病平均}}{\text{全体平均}} \right)^{LN(N)}$$

2) 傷病平均 < 全体平均: 全体平均 $\times \left(\frac{\text{傷病平均}}{\text{全体平均}} \right)^N$

上のようの場合わけしたのは、全体平均より小さい傷病については補正をきつく、大きい傷病では緩めにするためである。

すなわち昨年度の研究では平均傷病数 2 の場合は一次補正、4 の場合は 2 次補正そして 8 の場合は 3 次補正とおおまかな傾向を明らかにしたが、上記の補正式に従えば、平均値が全体平均より大きい傷病については、平均傷病数 2.7 で 1 次補正、7.3 で 2 次補正とゆるやかな補正となっており、逆に平均値が全体平均より小さい傷病については平均傷病数 2 なら 2 次補正、3 なら 3 次補正ときつい補正が行なわれる。

その理由は、全体平均より小さい傷病についてはどんなにきつい補正をしても決してゼロにはならないが、全体平均より大きい傷病についてはきつい補正を行うと極端に大きな数値になってしまうからである。それゆえ全体平均より大きい傷病については「控えめ」な補正式を選択した。逆に全体平均より小さい傷病については、後述の Excel ソルバーを用いた方法でも触れるように、レセプト傷病名には重みが限りなくゼロに近い傷病名が含まれることからむしろゼロに近づける補正を行う方が推計の妥当性を高める、と判断したからである。

【重み推計法 2】繰返しによる最適化法

PDM法は傷病ごとに共通の重みを設定し、その重みにしたがって各レセプトの日数や点数を比例配分する。平均値補正法では、まず傷病ごとの平均値を求めそれに一定の補正式を適用して重みとした。

ここで視点を変えて、初めに重みを決めてしまうのではなく、PDM法で推計した各レセプトの日数や点数が実際の日数や点数に最も近くなるような数値を探し、逆にそれを重みとする、方法を考える。このような作業は手作業では不可能だが、コンピューターを用いれば100以上の傷病ごとに数字をいろいろに変えてPDM法を適用し、その結果がどれくらい個々のレセプトの日数や点数に近いかを分析することができる。いろいろと変える作業をおびたしい回数くり返せば、少なくとも最も近似する数値に到達することが可能であり、その数値は「最適解」と呼ばれる。14年度報告書で述べた「繰返し法」はこのことを指す。

理論に基づくのではなくコンピューターによる力づくの方法ではあるが、数理計画法等で用いられ、Excelにもソルバーという機能がアドインソフト(必要に応じて追加するソフト)として備えられている。そこでExcelソルバーを用いてこの手法を試みた。

i を個々のレセプト、 j を傷病分類を指すとする。 i 番目のレセプトの点数を P_i 、 j 番目の傷病分類の重み P_j 、 i 番目のレセプトの j 番目の傷病の数を N_{ij} とする。かりに1000件のレセプトを100の傷病分類で分析とすると i 番目のレセプトの点数推計値 \hat{P}_i は以下のようなになる。

$$\hat{P}_i = \sum_{j=1}^{100} N_{ij} * P_j \dots\dots [1]$$

P_j の最適解とは実測値である P_i と推計値である \hat{P}_i の差、 $P_i - \hat{P}_i$ を最小にする P_j のことである。そのためには100ある P_j の数値をいろいろに変化させて以下の値を最小にする100

のPjを定めればよい。

$$\sum_{i=1}^{1000} (P_i - P_j)^2 \dots [2]$$

実際のExcelソルバーでこの処理を行うイメージを以下に示す(実際はもっと複雑であるが簡略化して表示した)。

i はレセプト, j は傷病を指す (1=<i=<1000, 1=<j=<100)
 Nij, Pij はそれぞれ i 番目のレセプトの傷病 j の傷病数, 点数である
 レセプトの総点数 P = Σ Pi = Σ Σ Pij, 総傷病数 N = Σ Σ Nij
 i 番目レセプトの点数は Pi, j 番目の傷病の「平均」点は Pj

		傷病分類 (j)		使用する Excel 関数
		Pj	Σ (Pi - Σ (Nij * Pj)) ²	← SUMXMY2(Pi, Σ (Nij * Pj))
レセプト (i)	Pi	Nij	Σ (Nij * Pj)	← SUMPRODUCT(Nij, Pj)

ソルバーによる傷病平均 Pj の求め方

- は既知のデータ。これより Pj の最適解を求めるには
- を最小化するように
- のセルを変動させる (119 分類なら 119 セル) 非負条件を入れる

[2]の式はExcel関数ではSUMXMY2(Piの列, Pjの列)で表される[SUMXMY2とはSum of (X minus Y)²の意味]。この関数の入ったセルをExcelソルバーの「目的セル」に指定し、その値を最小化するように「変化させるセル」に100のPjの入ったセル(下の図ではE2:E101に入っている)を指定して変化させる。この後で実行をクリックすればExcelソルバーは[2]式を最小化する100のPjを与える。

	A	B	C	D	E
1	=SUMXMY2(B2:B1001,C2:C1001)	実測値Pi	推計値Pi		Pj
2	レセプト1	886	219		
3	レセプト2	811	883		
4	レセプト3	103	160		

ソルバー パラメータ設定

目的セル(E): [fx]

目標値: 最大値(M) 最小値(N) 値(V)

変化するセル(B): [fx]

制約条件(U)

実行(S) 閉じる オプション(O) 追加(A) 変更(C) 削除(D) リセット(R) ヘルプ(H)

【シミュレーションデータ】

上の2つの重み推計法は14年度と同様、シミュレーションデータに適用し、その妥当性を検証した。シミュレーションデータの諸元は以下の通り。

- レセプト件数1000件 ($1 \leq i \leq 1000$)
- 傷病分類100分類 ($1 \leq j \leq 100$)
- 傷病分類jの総傷病数 $N_j \cdots 1$ から100までの数をランダムに発生
- 傷病分類jの仮平均 \cdots 平均100, 分散100の正規分布により発生(非負)
- i番目のレセプトのj番目の傷病分類の点数 $P_{ij} \cdots$ 傷病分類jの仮平均として生成された数を平均, その30%を分散とする正規分布により発生(非負)

- i番目のレセプトの点数 P_i (実測値) $\cdots \sum_{j=1}^{100} P_{ij}$

こうして生成されたシミュレーションデータに2つの重み推計法を適用し、得られた重みを用いてPDM法を用い傷病別医療費を推計する。得られた傷病別重みと傷病別医療費推計値のそれぞれについて妥当性を、既知の重みと実測値との回帰直線で評価する(ここでは「正解」にあたる実測値をX, 推計値をYとして単回帰をとる)。回帰直線の傾きで正確度(accuracy)を、決定係数(R²)で精度(precision)を評価する。いずれも1に近いほど正確度でも精度でも妥当性が高いといえる。

【モンテカルロ法】

上記のシミュレーションデータを100セット作成し、100の回帰直線の傾きと決定係数を、重みと傷病別医療費のそれぞれについて得る。100の数値の平均値を持って最終的な評価とする。

C, 結果

結果は以下の通り。正確度, 精度ともに最適化法は平均値補正法に勝っている。精度の

点では両者に大きな差は無いが、正確度の面では平均値補正法は明らかに劣っていることが回帰直線の傾きが1より小さく、回帰直線が平らになっていることよりわかる。すなわち平均値補正法では、高額傷病の医療費は低く推計され、低額傷病の医療費は高く推計される傾向がある。

14年度の報告書でも指摘したように、これはPDM法の本質的な性向であり、本来ならゼロに近い傷病でも一定の点数を配分するためこのような結果になる。さらに今回用いた補正法では、全体平均より高い傷病については補正次数を「甘く」したため、高額傷病の推計値が正解より小さめに出たことも原因といえる。

重み推計法		回帰直線の傾きと決定係数	平均	S.D.
最適化法	傷病別医療費(Pj)	傾き(accuracy)	0.995	0.014
		R ² (precision)	0.984	0.004
	傷病別重み(Pj)	傾き(accuracy)	0.974	0.104
		R ² (precision)	0.908	0.092
平均値補正法	傷病別医療費(Pj)	傾き(accuracy)	0.865	0.021
		R ² (precision)	0.958	0.008
	傷病別重み(Pj)	傾き(accuracy)	0.700	0.081
		R ² (precision)	0.799	0.075

また上の結果においては、傷病別重みの推計結果と最終目的である傷病別医療費の推計結果も比較すべきである。いずれの推計法においても、元となる重みの推計の妥当性は正確度でも精度でもあまり高くないが、その重みをPDM法にかけると、いずれの推計法でも最終的な傷病別医療費の推計での妥当性が向上している。とりわけ平均値補正法においては重み推計の妥当性は正確度でも精度でもきわめて低いが、最終的な傷病別医療費推計の妥当性では、最適化法に匹敵するほどの妥当性になっている。

これぞ、どのような重み推計法を使うかにかかわらず、PDM法の最大の強みといえる。つまり元となる重みの推計の妥当性は低くてもPDM法にかけると最終的な妥当性は必ず向上する。

D, 考察

傷病別医療費推計の妥当性の面では最適化法は優れている。にもかかわらず、最終的なPDM法プログラムVer.3には平均値補正法を取り入れられており、最適化法はとり入れなかった。その理由として2点考察を加える。

ひとつはExcel Solverそのもののツールの限界である。Excelのアドオンソフトであるから当然ながらデータはExcelのワークシートでなければならない。そのため6万5000件というExcelの処理能力の限界があり、またSolverの作動には膨大なメモリが必要となる。

今回は扱ったシミュレーションデータは1000件だったので数分で最適解が得られたが、もし数万件ものデータであれば相当な長時間とコンピューターのメモリが必要となる。レセプト分析では将来的には億に近い件数を扱わなければならないことを考えるとこれらは大きな制約要因である。それに対して平均値補正法はきわめてシンプルかつ簡便な手法であり、現在フリーウェアとして公開しているプログラムでも数百万件の処理が迅速に行える。

正確度が若干劣ったとしてもきわめて実用的な分析法である。

もうひとつは医療関係者の許容性である。レセプト傷病分析に重回帰分析が用いられないのは係数に必ずといっていいほど負値がでる、すなわちマイナスの医療費になる傷病が出現することであった(このような場合、その病気については医療機関が保険者に金を払うことになる)。理論的には正しくてもこれは医療関係者の通念に反し到底受け入れられる数字ではない。

そのことを実例で示す。PDM法はレセプトの傷病別医療費の推計だけでなく、たとえば栄養調査等、総計と項目だけがわかっているが、項目ごとの数値が不明なデータの解析にも活用できる。以下はある人の10回の食事のカロリーと献立である。このように、食事ごとの総カロリーと何を食べたか、は比較的調査しやすいがそれぞれの献立から何カロリー摂取したか推計することは難しい。栄養士は、そこで、それぞれの献立をどれだけ食べたか聞き取ったり、食事の写真をとる等してせいっぱいの推計を試みる。しかしこうした聞き取りや写真撮影ができず、単に以下のデータだけしかない場合はPDM法が役立つ。

カロリー	献立				
1600	リンゴ	味噌汁	ハンバーグ	コロッケ	
1800	ハンバーグ	ギョーザ	おでん	味噌汁	
1100	チキン	ぶたまん			
900	コロッケ	味噌汁			
600	ぶたまん	おでん			
800	味噌汁	ハンバーグ			
700	おでん	味噌汁			
1000	やき魚	ぶたまん	ギョーザ		
1300	ギョーザ	ぶたまん	ハンバーグ	コロッケ	
1650	ラーメン	味噌汁	おでん	ギョーザ	コロッケ
11450					

以上のデータを、平均値補正法を用いたPDM法Ver3と、最適化法を用いたPDM法(プログラム化にいたらなかったの以下PDM法Xと呼ぶ)とで、献立別カロリーを推計した結果は以下のようなになる。

	PDM-Ver3	PDM-X
リンゴ	551	870
ハンバーグ	1702	1542
チキン	651	1100
コロッケ	1551	1457
ぶたまん	1063	
味噌汁	2054	1020
おでん	1239	1686
やき魚	207	
ギョーザ	1907	3351
ラーメン	524	423

合計	11450	11450
----	-------	-------

上の例ではぶたまんとやき魚のカロリーがゼロとなっている。ぶたまんは4回も食べているが、それからの摂取カロリーがゼロということに納得する栄養士はいるだろうか？ このように最適化法では負値こそ出ないがやはり少なからぬ重みゼロの傷病がでる。しかしながらレセプトは医療費の請求書であるからそこに記載された傷病に対しては何らかの医療費がかかっていると考えるのがたてまえである。その点平均値補正法なら医療費ゼロの傷病は決して出現せず、医療従事者にとって受け入れやすい。

最終年度に開発したPDM法Ver3プログラムでは、最適化法による重み推計法は採用しなかった。その理由は第一にExcelアドインソフトである技術的な制約であったが、第二はこうした受け入れ可能性という理論的問題があった。

しかしながら、最適化法が傷病別医療費の推計としてはより正確である点は事実であり、たとえば全体平均より大きい傷病については最適化法、全体平均より小さい傷病については平均値補正法、というように両方の重み推計法を併用することも検討に値する。

また平均値補正法も、単純な方法でありながら複雑かつ膨大な計算を要する最適化法にかなり近い正確度と精度を有しており、膨大な件数をおおざっぱに分析するのに十分であるといえる。

E, 結論——PDM法進化の過程

複数傷病のレセプトの日数点数を客観的、自動的かつ全ての傷病を加味して配分する原理としてPDM法を1995年に考案、同年10月山形公衆衛生学会での第一回集会において原始的なBASICプログラムを公表した。

PDM法は各傷病について共通の日数、点数の「重み」を定め、全レセプトをこの値に比例配分し、傷病ごとに合計する、という原理であり現在にいたるまで一貫してPDM法の基本原理をなす。重みはどのように定めてもよいが、数学的に妥当な方法の開発には長年いたらなかった。2001年度よりスタートした本研究において共通の日数、点数の計算法を考案しWindowsプログラム化し一般公開した。

3か年にわたるVer1, Ver2, Ver3の進化を総括する。

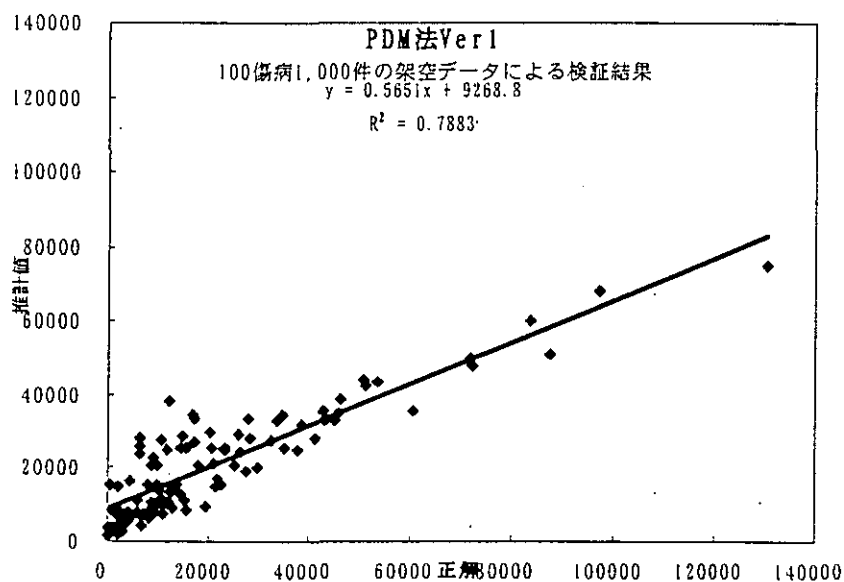
【日数の重み】日数については各傷病の記載されたレセプトの単純平均値を重みとして用いる。これはVer1,2,3を通じて不変。日数は診療報酬改定に影響されない最も安定した指標であり、インフルエンザワクチン事業の効果測定等には日数を用いた【添付資料2】。

【点数の重みと補正】Ver1,2,3を通じて最も進化が顕著だったのは点数の重みの推計である。以下に進化の過程を14年度報告書で用いたシミュレーションデータを分析した結果を使ってその正確度や精度の改善のプロセスを概観する。

●Ver.1(2001年度成果物。2002年10月23日埼玉公衆衛生学会における第5回レセプト情報活用を考える集会で公表)

日数と同様、各傷病の記載されたレセプト点数(一日当たり点数)の単純平均値を用いた。以下のグラフは1,000件の架空レセプトデータを用いて、100傷病の点数を推計した結果。単純平均そのままでは、併記される傷病数の多い傷病では、その傷病固有の点

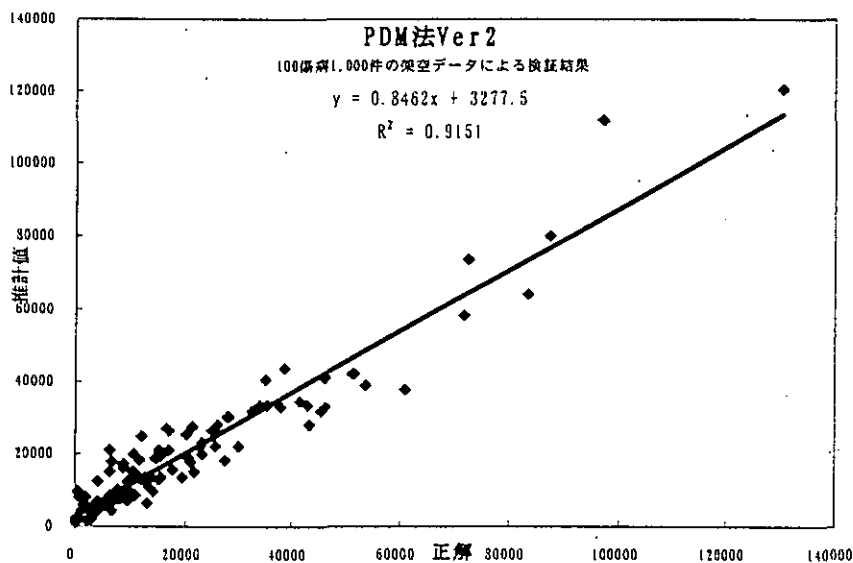
数が「薄め」られるため、高額傷病の推計値は少なく、低額傷病の推計値は高めにでる、という矛盾があった。



● Ver.2(2002年度成果物。2003年4月23日インターネット上でフリーウェアとして公表)
日数の重み推計はその Ver1 と同じだが、点数については傷病平均値そのままでは正確な推計値にならないため、補正式を導入した。

$$\text{傷病平均} \times \left(\frac{\text{傷病平均}}{\text{全体平均}} \right)^n$$

もし $n=2$ (2次補正) なら全体平均が 10 で傷病Aの記載されたレセプトの平均値が 15 であったなら、傷病Aの値は $15 * (15/10)^2 = 33.75$ と推計される。何次補正がいいかは傷病Aの記載されたレセプトの傷病数に依存し、わが国外来レセプトの平均傷病数 3.87 では 2次補正が最もよいことが架空データで明らかになった。【添付資料3】



● Ver3(2003年度成果物。2003年10月23日第6回集会で公表)

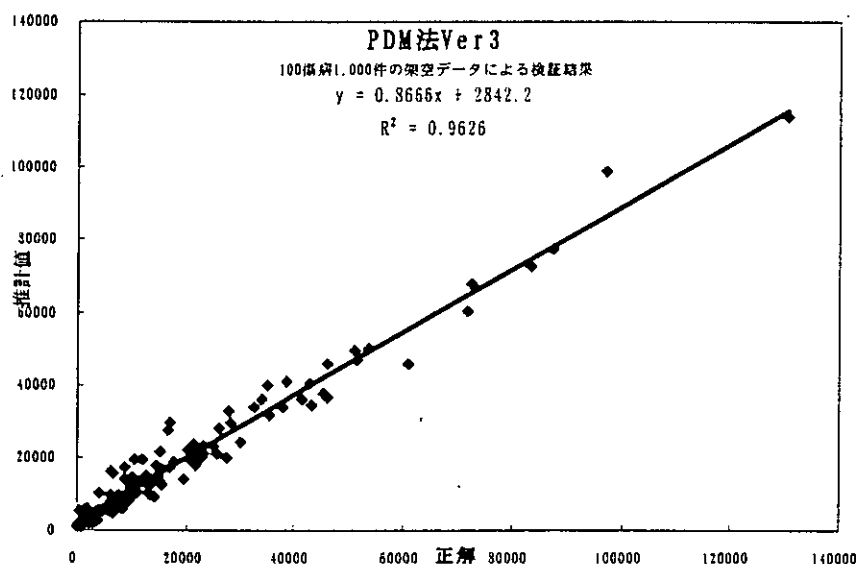
Ver2 では、全ての傷病についてたとえば 2 次補正を行っていた。しかしながら傷病によってはたとえばインフルエンザのように単独でしか出現しない傷病や糖尿病のように多数の傷病を伴うものもある(常に単独でしか出現しない傷病のレセプトの平均値は補正せずそのまま用いなければならない)。それらがいっしょくたに同一の補正を受けることになり、また外来では二次補正、傷病数の多い入院レセプトは 3 次補正というふうにユーザーが手動で選択しなければならなかった。

Ver3 では、傷病ごとにその傷病の記載されたレセプトの傷病数(N)をコンピューターが自動的に計算し、適切な補正次数を決定するようにした。その傷病の記載されたレセプトの平均傷病数を N 、とするとその傷病の平均値を以下のように補正する式を考案した。

1) 傷病平均 > 全体平均: $\text{傷病平均} \times \left(\frac{\text{傷病平均}}{\text{全体平均}} \right)^{\text{LN}(N)}$

2) 傷病平均 < 全体平均: $\text{全体平均} \times \left(\frac{\text{傷病平均}}{\text{全体平均}} \right)^N$

上のようの場合わけしたのは、全体平均より小さい傷病については補正をきつく、大きい傷病では緩めにするためである。これによりさらに精度は向上した。各傷病の平均傷病数の格差が小さい架空データでは Ver2 との違いは大きくないようだが、傷病ごとの記載傷病数のバラツキの大きい現実データでの精度向上はきわめて大きいと考えられる。

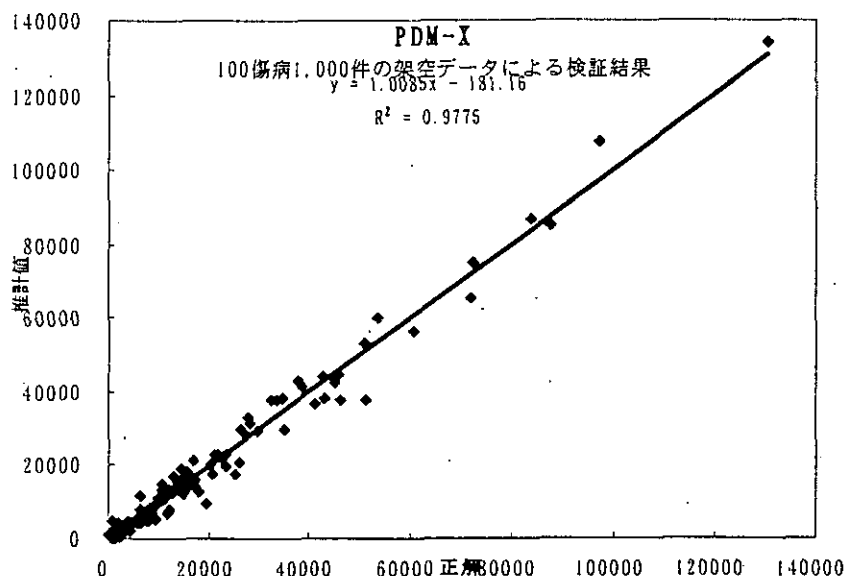


●今後の発展の方向(最適化法を用いたPDM-X)

Ver1,2,3 と日数、点数ともに平均値を求め、それをそのまま、あるいは補正して使っていた。Ver3において十分実用に耐えるものとなったが、ほぼ完全な推計値が得られる全く異なった方法として Excel ソルバーという最適化ツールを用いた手法も開発した。ただ、Excel のアドオンソフトであること、コンピューターのメモリを大量と必要とすること、扱える件数に制限があること、時間がかかること、また考察でも述べたようにゼロと評価される傷病が出て医療関係者にとっては受け入れにくいといった問題も残されている等の事情から本研究期間中には実用化にいたらなかった。

PDM 法の技術的開発はとりあえず本研究期間の満了をもって区切りとし、今後は PDM

法の電子レセプトへの普及を専らはかってゆく。一方、新しい手法による仮称 PDM-X についても、平均値補正法との併用やコンピューターの進歩をにらみつつ実用化の方向をさぐってゆきたい。



【本研究の意義】レセプト情報が医療費情報としてのみならず、医療情報として活用されるためには、レセプトの日数、点数の傷病別内訳が客観的かつ定量的に推計されなければならない。複数傷病の記載された医療費請求書の科学的な分析法は国際的にも先行研究が無く、外国でもICDコーディングによる主傷病やDRG, MDCによって「一人の患者は一つのカテゴリー」に「分類」しているのが現状である。わが国保険診療に導入されたDPCも発想は同じである。

9年前に着想されたPDM法は、「一人の患者は必ず一つのカテゴリーに分類される」というこれまでの常識を根本から覆すものであった。一人の患者が複数の傷病を有することは当然であり、もし集団の傷病別医療費や日数を推計する場合は、一人の患者にかかった費用や日数を全てひとつの傷病カテゴリーに「分類」するのではなく、それぞれの傷病の割合を量的に「分析」すべき、がPDM法の発想である。PDM法が傷病「分析」法であって「分類」法ではないゆえんである。

本研究を通じてPDM法は、妥当性の面でも技術的な面でも完成の域に達し、今後年間12億件のレセプトが完全電子化されても、瞬時に傷病分析ができ、市町村間比較、経時変動そして、たとえばワクチン接種非接種者間といった集団間の傷病比較が可能となる。また研究の過程でPDM法はレセプト傷病分析のみならず、たとえば栄養調査のように、複数食品の品目と合計カロリーのみがわかっている場合の品目ごとのカロリー推定といった目的にも活用できることが明らかとなり、本手法はいわば重回帰分析の開発に匹敵する新しい統計分析手法であると自負している。

今後はPDM法の活用をレセプト分析のみに限定せず、たとえばExcelのアドインソフトとして一般化させる、ことも検討してゆく。名称も今回の研究後は従来のProportional Disease Magnitude法からより普遍的でわかりやすいなProportional Distribution Method(比例配分法)と変更する。