

米国ヘルスケア分野における 健康リスク評価のための予測モデルの活用

—その概要と背景、および米国アクチュアリー会 2007 年報告書の解説—

目 次

はじめに

第一部 米国ヘルスケア分野における予測モデル活用の概要と背景

I. ディジーズ・マネジメントにおける予測モデルの活用

II. 健康リスク評価のための予測モデルとはどんなものか

III. 予測モデル活用のその他の具体的事例

IV. 予測モデルの実務活用に向けた取り組み

第二部 予測モデルの比較検討および選択・導入時に考慮すべき諸点

I. 分析対象の予測モデルの概要

IV. 分析結果の概要

II. 比較評価のためのシナリオ設定

V. 今後の課題

III. 予測精度を評価する尺度

付録 感度・特異度・ROC 曲線に関する補足説明

主任研究員 矢倉尚典

要 約

はじめに

米国では、ディジーズ・マネジメントの取り組みを効率的に進めるために、対象集団の健康リスクを評価する予測モデルが盛んに活用されている。本誌第 48 号では、ディジーズ・マネジメントにおいて健康リスクを評価する予測モデルが有効なツールとしてその評価が確立されていることを取り上げたが、ヘルスケア分野の様々な局面においてもこの予測モデルが活用され、様々な予測モデルが開発されている。米国アクチュアリー会は、ヘルスケア分野の様々な予測モデルを分析し、「健康リスク評価のための医療給付金請求データに基づくツールの比較分析」と題する報告書を 2007 年に公表した。本稿ではこの報告書に拠ってヘルスケア分野において、健康リスクを評価する予測モデルがどのような点で有効か、実務に導入・使用するとすればどのような点に留意すべきかについて紹介する。

第一部 米国ヘルスケア分野における予測モデル活用の概要と背景

I. ディジーズ・マネジメントにおける予測モデルの活用

ディジーズ・マネジメントの実務運用における重要なポイントは、働きかけるべき対象者を適切に選び出して階層分けができていくかということである。この課題に対する実用的な解決として、健康リスクを評価するための予測モデルがツールとして使用されている。

II. 健康リスク評価のための予測モデルとはどんなものか

予測モデルは、一般的には、個々人の過去または現在の状態を表す数値を入力し、その個人の健康リスク（将来の状態）を予測した指標を計算する数理モデルである。米国の主要な予測モデルの一つである Impact Pro を例に、予測モデルの基本的な仕組みと 5 つの特徴を解説する。

III. 予測モデル活用のその他の具体的事例

ディジーズ・マネジメント以外の米国ヘルスケア分野における予測モデル活用の具体的事例として、「保険料率設定、保険引受条件判断における活用」、「医療プロバイダーへの医療費支払額決定および医

療の質の維持・向上における活用」、「連邦政府等から健康保険プランへの委託料決定における活用」の3事例について、その背景を含め、説明する。

IV. 予測モデルの実務活用に向けた取り組み

予測モデルの実務活用に当たって個々のユーザーは、各自の活用目的に合わせて、数ある予測モデルの中から最も適当なものを選択することが必要となる。そのためには、予測モデルのもつ制限やその予測精度に影響を与える諸要因を認識することが重要である。予測モデルを実務に導入・使用しようとするればどのような点に留意すべきかについて、米国アクチュアリー会 2007 年報告書を基に説明する。

第二部 予測モデルの比較検討および選択・導入時に考慮すべき諸点

I. 分析対象の予測モデルの概要

米国アクチュアリー会 2007 年報告書で分析対象とされた 12 の予測モデルの概要について述べる。インプットデータの種類で分類すると、その内訳は、診断名に基づくモデルが 4、処方薬剤に基づくモデルが 3、診断名と処方薬剤の両方に基づくモデルが 2、過去の期間のコストデータも用いるものが 3 である。それぞれの予測モデルの開発経緯、特徴を簡単にまとめている。

II. 比較評価のためのシナリオ設定

比較評価に当たっては、10 個のシナリオを設定している。シナリオを特徴付ける要素は次の 5 つである。①「予測モデルを将来予測に使用するのか、当年の現状分析に使用するのか」、②「各予測モデルによるリスク・スコアの計算において、各予測モデルのソフトウェア付属のリスク・ウェイトを用いるのか、今回分析対象とした集団のデータセットに合わせて再調整したリスク・ウェイトを用いるのか」、③「データおよび予測のラグは想定するのか、しないのか」、④「過去の期間のコスト（例えば、前年の医療費コスト）を予測因子として加えてリスク・ウェイトを再調整するか、加えないか」、⑤「給付金請求金額を頭切りする水準の設定」。

III. 予測精度を評価する尺度

各シナリオにおける予測モデルによる計算結果を比較するための尺度として、個々人レベルの予測精度の尺度と集団レベルの予測精度の尺度を設定した。米国アクチュアリー会 2007 年報告書では、前者としては、決定係数と予測誤差の絶対値の平均を、後者としては、予測比率を使用している。

IV. 分析結果の概要

各シナリオにおける 12 の予測モデルによる予測結果の比較に加え、シナリオ間の比較を行うことにより、各予測モデルの特徴を分析している。シナリオ間比較の視点としては、①将来予測に使用する場合は、現状分析に使用する場合に比べてどの程度予測のパフォーマンスが変化するか、②リスク・ウェイトを調整すればパフォーマンスは改善するのか、③データのラグ・予測のラグがあるとパフォーマンスはどの程度低下するのか、④過去の期間のコストを予測因子として追加したらパフォーマンスが改善するか等である。併せて、集団レベルの予測の結果として、主要 6 疾患別の予測比率、金額の大きさによるカテゴリー別の予測比率も比較し、各予測モデルの特徴を分析している。

V. 今後の課題

予測モデルのパフォーマンスに影響を与える制限および要因として、リスク・ウェイトの調整、継続加入条件、データのラグ・予測のラグ、データの入手のしやすさ等を論じるとともに、予測モデルの導入において留意すべき諸点として、事前テストやデータハンドリングの経験、データの妥当性の確認、加入者資格データの鮮度、リスクスコアリングを行う時期および頻度について論じている。最後に、今後行われるべき有益な研究課題として、検査結果数値等をモデルに取り込んだ場合に予測のパフォーマンスがどの程度改善するのか、その可能性を分析すること等を含む 9 つの課題を提示している。

はじめに

損保ジャパン総合研究所では、ディジーズ・マネジメント（以下、「DM」とする）についての調査・研究を継続して行っている¹。DMは「糖尿病、喘息、心臓病等の慢性疾患の患者集団と、彼らを担当する医師その他の医療従事者の双方を主たる対象者として、それぞれの疾病に固有の手法で働きかけを行い、質の高いヘルスケアを低コストで提供できるようにする」²取り組みである。米国ではこの取り組みは研究・試行の段階から、事業化され産業として成立する段階へと既に移行している。この過程で、この取り組みを効率的に進めるための諸課題の実用的な解決が図られてきた。その一つの方法として、米国では、対象集団の健康リスクを評価する予測モデルの活用が盛んであり、その具体的活用事例を本誌第48号（以下、「前稿」とする）で取り上げた。健康リスクを評価する予測モデルはDMのサービス提供を効果的に資源効率的に運用するための有効なツールである。こうした予測モデルの活用はわが国においては未だ一般的ではないが、米国のDMでは多くの利用実績があり、有効なツールとしての評価が確立している。

前稿ではDMにおける予測モデルの活用を取り上げたが、DMに限らず、米国ではヘルスケ

ア分野³の様々な局面で健康リスクを評価する予測モデルが活用されている。また、前稿では、比較的豊富に情報が公開されていることから、米国ジョーンズ・ホプキンス大学が開発した予測モデル（ACG-PM）を取り上げた。米国では、ACG-PMをはじめとして健康リスク評価のための様々な予測モデルが開発されている。こうした環境を受け、米国の保険・年金数理の専門職団体である米国アクチュアリー会は、ヘルスケア分野の様々な予測モデルの分析を継続して数年おきに実施し、公表してきた。本稿では、直近の2007年4月に公表された「健康リスク評価のための医療給付金請求データに基づくツールの比較分析」と題する報告書（“A Comparative Analysis of Claim-Based Tools for Health Risk Assessment”⁴以下、「米国アクチュアリー会2007年報告書」とする）に拠って、ヘルスケア分野において、健康リスクを評価する予測モデルがどのような点で有効か、実務に導入・使用するとすればどのような点に留意すべきかについて紹介する。

本稿の構成は以下の通りである。本稿は、米国ヘルスケア分野における予測モデル活用の概要と背景を取り上げる第一部と、米国アクチュアリー会2007年報告書の詳細説明を行う第二部で構成されている。第一部により、健康リス

¹ これまで、機関誌「損保ジャパン総研クォーターリー」第41号で「米国ヘルスケアにおける新たな潮流－米国における Disease Management の発生と展開－」（2002年10月）を公表したほか、損保ジャパン記念財団主催「欧米のヘルスケアビジネスおよびディジーズ・マネジメント研究会」の事務局を務め、損保ジャパン記念財団叢書で研究成果およびシンポジウム講演討議録を公表してきた。損保ジャパン記念財団叢書第65号「米国におけるディジーズ・マネジメントの発展」（2003年6月）、第67号講演討議録「ディジーズ・マネジメント発展の可能性と課題」（2004年1月）、第72号講演討議録「これからの生活習慣病対策のあり方を探る」（2006年7月）。また、ニュースレター「ディジーズ・マネジメント・レポーター」を2004年3月に創刊し、わが国におけるディジーズ・マネジメントの取組事例を紹介してきた。

² 損保ジャパン記念財団叢書第65号「米国におけるディジーズ・マネジメントの発展」（2003年6月）p.15

³ 「ヘルスケア」という言葉は、狭義には医師などの専門家や病院などの医療施設において患者に供給される医療サービスを指していたが、広義には、より広い供給主体からの商品・サービスや、健康の自己管理までも含むうる広範な概念となっている。当研究所では、過去の調査研究から一貫して、米国の営利保険会社の観点からヘルスケア市場を概観してきた。ただし、米国の場合、保険会社が提供しているのは、保険商品にとどまらず、医療サービス提供者者のネットワークを含めた健康保険プラン（Health Plan）全体を企画設計、運営するサービスであるという特徴がある。本稿でも、「ヘルスケア分野」は、保険会社の観点から捉えたヘルスケア事業分野という意味で使用される。

⁴ Ross Winkelman & Syed Mehmud, “A Comparative Analysis of Claim-Based Tools for Health Risk Assessment”, April, 20, 2007, Society of Actuaries < <http://www.soa.org/research/health/hlth-risk-assement.aspx> > visited June 24, 2008

ク評価のための予測モデルの概要を把握できるように努めた。第一部第I章では、DM実務を例に予測モデルの役割とその重要性を振り返る。第II章で健康リスク評価のための予測モデルの基本的な仕組みと特徴を説明した上で、第III章でDM以外の予測モデルの具体的活用事例をその背景を含めて概観する。第IV章では予測モデルを実務に導入しようとする際の留意点を米国アクチュアリー会 2007 年報告書を基に取りまとめる。第二部では米国アクチュアリー会 2007 年報告書の詳細説明として、第I章で同報告書の分析対象となった 12 の予測モデルの概要を述べ、第II章で比較評価のために設定した 10 種類のシナリオについて、第III章でモデルによる予測の精度を評価するために同報告書が採用した 3 種類の尺度について解説をした上で、第IV章で同報告書に示された分析結果を紹介する。最後に第V章で、分析結果を踏まえた予測モデルに関する諸留意点および同報告書の研究者自らが表明している今後の研究課題について解説を加える。

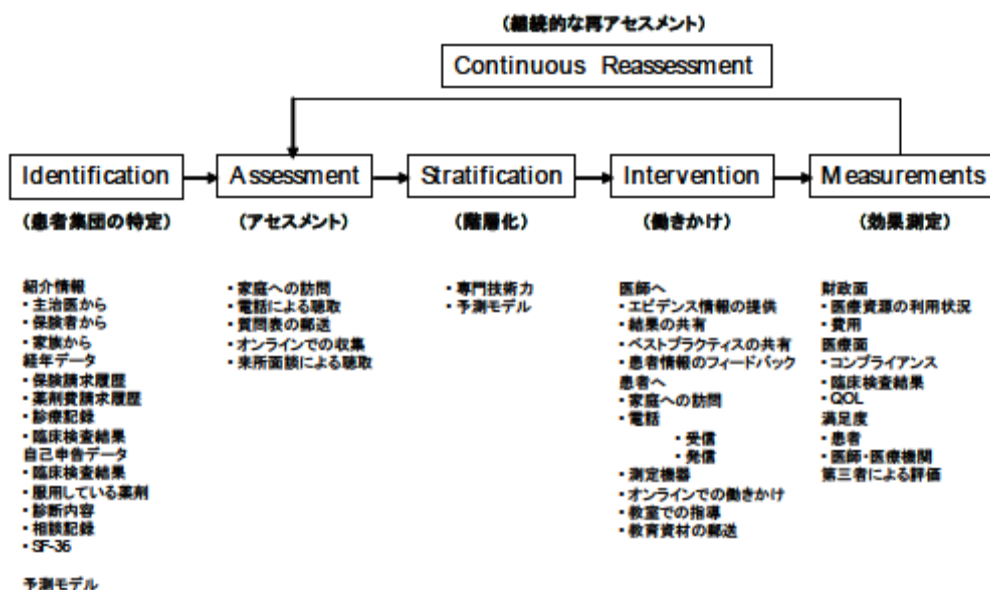
第一部 米国ヘルスケア分野における予測モデル活用の概要と背景

I. デিজーズ・マネジメントにおける予測モデルの活用

ディジーズ・マネジメントに実際に取り組むプログラム（以下、「DMプログラム」とする）の基本プロセスと主な活動は《図表 1》の通りであるが、その特徴は、集団の全員に同じ働きかけを一律に行うのではなく、対象者の特定（Identification）、階層化（Stratification）というプロセスを経て、働きかけ方の強度・頻度を調整していることである。これは、人数的には比較的小さな集団に医療費の大部分が投入される現状を踏まえ、サポートを必要とする対象者を特定して働きかけること、さらに、対象者を階層化して有限で高価なヘルスケア資源の投入を調整することにより効果的・効率的なプログラム運営の実現を図ることを意図したものである。

したがって、DMプログラムの実務運用における重要なポイントは、対象者の特定・階層化、すなわち、働きかけるべき対象者を適切に選び

《図表 1》 デিজーズ・マネジメントの基本プロセスと主な活動



(出典) 前掲注 2 の資料 p.28

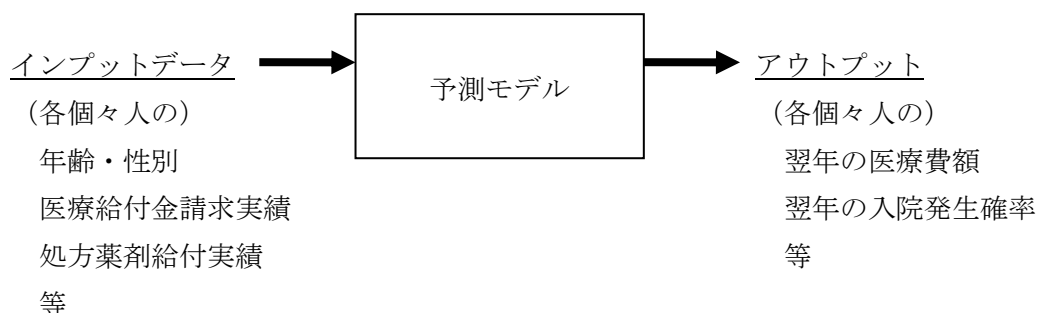
出して階層分けができていないことである。例えば、選び出し方が粗く、働きかけるべき対象者以外の人も含めて選び出してしまうと、働きかけるべき対象者以外の人にも働きかけを実施することになり、資源を非効率に投入することになる。また、逆の意味で選び出し方が粗く、働きかけるべき対象者の選び出し漏れが多いと、働きかけて効果を出すべき機会をみすみす逃していることになる。それでは、この課題に対して、DMプログラムが先行発展した米国では、どのような実用的な解決が図られたのであろうか。

米国の DM プログラム運営において、効果的・効率的な対象者の特定・階層化という課題に対する実用的な解決として使用されているツールが健康リスクを評価するための予測モデルである。仮に利用可能な情報が何もなくて対象者を無作為に選んだとすると、働きかけるべき人々を適切に選び出すことはほとんど期待できない。一方、選び出しの判断材料となる情報がたくさんあれば、より確からしい選び出しの結果が得られる可能性が高くなると期待できる。しかし、例えば、判断材料となる情報をより多く得るために、母集団全員に対して詳細な調査（検査）を実施するとすれば、それには多大な費用の投下が必要となり、実用的ではないかもしれない。そこで、費用対効果の側面も踏まえ

た実用的な解決として、母集団構成員の管理データ（年齢、性別等）や医療給付金請求データ、処方薬剤給付データといった、DMプログラムの提供主体あるいはそのスポンサーとなる保険会社等にとって既に収集済みで利用可能な各種情報（情報収集のための新たな費用をほとんど要しない情報）を活用することが考えられた。米国では様々な試行錯誤を経て、また、数理統計理論とコンピュータ科学の進展を受け、このような収集済みで利用可能な情報の大量のデータに対してデータ・マイニングや様々な数理統計解析技法等を行い、こうしたデータを基に各人の健康リスクを評価・予測するための数理モデルが各種開発されている。

こうした予測モデルは、一般的には、各個々人の過去または現在の状態を表す数値（年齢や性別、医療給付金請求の実績、処方薬剤給付の実績等）を入力し、その個人の健康リスク（将来の状態）を予測した指標（例えば、翌年の医療費額や翌年に入院が発生する確率等）を計算する数理モデルである（《図表 2》参照）。DMプログラムにおける対象者の特定・階層化というプロセスにおいては、予測モデルにより計算された予測指標の数値の大きさに着目して、働きかけの対象とするかどうか、どの階層に分類するか等を決定していくという形で活用される。効果的・効率的な対象者の特定と階層化という

《図表 2》 予測モデルの概念図



(出典) 損保ジャパン総合研究所作成

課題に対して、既に収集済みのデータを用いて予測モデルで予測指標を計算し、計算された予測指標に着目して、ある程度対象者を絞り込んだ上で詳細な情報収集を実施するという実用的な解決を図っているのである。

II. 健康リスク評価のための予測モデルとはどんなものか

本章では、具体的な数値例が掲載された資料が公開されていた Impact Pro (米国 Integrated Healthcare Information Services, Inc. が開発した予測モデル) を例に、予測モデルの基本的な仕組みと特徴を解説する。

1. 予測モデルで計算されるもの (Impact Pro の数値例)

Impact Pro により、ある個人について、翌年の医療費の額 (薬剤費を含む。以下本節、次節において同じ。) を予測した算出例を《図表 3》に示す。《図表 3》の合計欄に表示されている 10.12277 は、この個人の翌年の医療費の額が母集団全体における平均的な加入者に要する医療費の額の 10.12277 倍となると予測されることを意味した相対リスクを表示している。

Impact Pro による相対リスク算出のステップは以下の通りである。Impact Pro に対する入力データとして使用されるものは、各個人々の医療給付金請求データ、処方薬剤給付データ、年齢・性別等のデモグラフィックデータである。Impact Pro は、まず、これらのインプットデータを基に、各個人々について、「インシュリン依存性糖尿病で併存疾患があるか」、「直近3か月間に糖尿病を主因とする入院があったか」、「鬱血性心不全で併存疾患があるか」等々のエピソードが存在するかを判断する。Impact Pro では、それぞれのエピソードが存在する場合には医療費の額が母集団全体における平均的な加入者に要する医療費の額の何倍となると予測されるかを示す係数が設定されている。この係数をリスク・ウェイトと呼んでいる。各個人々に存在すると判断されたすべてのエピソードについて、エピソードそれぞれに対して設定されているリスク・ウェイトを合計したものを、その個人に関する合計リスク・ウェイトとする。《図表 3》で例とした個人の場合、入力データにより、①インシュリン依存性糖尿病で併存疾患があり、②直近3か月間に糖尿病を主因とする入院があり、③鬱血性心不全で併存疾患があり、

《図表 3》 Impact Pro による計算例

マーカー	内容	リスクウェイト
02.01.000	Insulin dependent diabetes, with co-morbidity, (base marker) インシュリン依存性糖尿病、併存疾患あり	1.1079
02.01.103	Inpatient Stay, diabetes primary within recent 3 months 直近3か月間に糖尿病を主因とする入院あり	4.5714
08.20.000	CHF, with co-morbidity, (base marker) 鬱血性心不全、併存疾患あり	1.0737
08.20.C3	Significant CHF episode clusters, recent 3 months 直近3か月間に重大なCHFエピソードあり	1.8619
10.05.000	Chronic bronchitis, with co-morbidity, (base marker) 慢性気管支炎、併存疾患あり	0.3978
RX.04.06	Blood, anticoagulants, CHF CHFにより血液抗凝固薬を服用	0.3938
Age-Sex Marker	Males, 55-64 男性、55-64歳	0.7212
合計		10.12277

(出典) Integrated Healthcare Information Services, Inc.のディスクローズ資料 “A Clinical Episode-Based Approach to Predictive Modeling: Impact Pro™” に基づき損保ジャパン総合研究所作成

④直近3か月間に重大な鬱血性心不全の発作があり、⑤慢性気管支炎で併存疾患があり、⑥鬱血性心不全により血液抗凝固剤を服用しており、⑦男性で年齢は55歳～64歳の間という7つのエピソードの存在が把握できた。そのそれぞれのエピソードが存在する場合のリスク・ウェイトは、1.1079、4.5714等々と設定されており、それを合計した10.12277がこの個人に対する相対リスクの予測値となったことが示されている。なお、それぞれのエピソードが存在する場合のリスク・ウェイトは、Impact Proのモデル開発に使用したデータセットに基づき統計的に算出され、Impact Proのソフトウェアに組み込まれている。

2. 予測モデルの特徴

《図表3》は、アウトプットとして、翌年の医療費の額の相対リスクを計算した例であるが、Impact Proでは、(産科以外の)入院が発生する相対リスク計算用のリスク・ウェイトも用意されているとのことである。当然のことながら、医療費の額の相対リスクを計算する場合のリスク・ウェイトと入院発生相対リスクを計算する場合のリスク・ウェイトは異なる数値が設定される。また、インプットデータとして医療給付金請求データと処方薬剤給付データの両方が利用可能な場合もあれば、片方しか利用可能でない場合もあろう。それらによっても、リスク・ウェイトの値は異なる。一般化して予測モデルの特徴を整理すると以下の5つである。

特徴1：インプットデータの種類

《図表3》に示した例は、医療給付金請求データ、処方薬剤給付データ、および年齢・性別等のデモグラフィックデータを使用する場合のものである。しかし、予測モデルを活用しようと

する者の状況によっては、これらのデータがすべて利用可能とは限らない。医療給付金請求データおよび処方薬剤請求データの両方が利用可能な場合もあろうが、いずれか一方しか利用可能でない場合もあろう。利用可能なデータによって、モデルの作りは変わってくる。

特徴2：説明変数の設定

上記のImpact Proの例では、インプットデータを基にエピソードが把握されると説明したが、このエピソードは約600種類設定されており、インプットデータを基にそのそれぞれのエピソードの有無を判断している。ここでは、一つ一つのエピソードの有無を説明変数とし、その値(有=1、無=0)に応じてリスク・ウェイトを合計して、予測指標が計算される。このようにImpact Proではエピソードを説明変数としているが、説明変数をどのように設定するかはモデルにより様々である。医療給付金請求データに記録された診断名コードに着目しその組み合わせを説明変数とするモデルもあれば、処方薬剤給付データに記録された処方薬剤の使われ方・組み合わせを説明変数とするモデルもある。インプットデータを基にどのような説明変数を設定するのかに、予測モデルの開発者は創意工夫、主義主張を込めている。

特徴3：アウトプットの種類

予測モデルのアウトプットとなる指標を何とするのかは予測モデルの使用目的により様々である。健康リスクの評価といっても、様々な指標が考えられる。上記のImpact Proの例では、翌年の医療費の額の相対リスクをアウトプットとするものと、(産科以外の)入院が発生する相対リスクをアウトプットとするものが用意されていると説明したが、他の予測モデルでは、翌

年の医療費の実額をアウトプットとするもの、翌年の医療費の額がその集団における上位 5% 以内に入る各人の確率をアウトプットとするもの等もある。

特徴 4 : アウトプットの時間軸

「予測」モデルという呼び方からすれば、ある年度のデータをインプットデータとし、次年度の医療費を予測するといったイメージが一般的であるが、アウトプットをインプットデータと同一の年度の医療費等とするモデルも開発されている。前者は **prospective** (将来予測) モデル、後者は **concurrent** (または **retrospective** : 現状分析) モデルと呼ばれる。インプットデータと同一の年度の医療費なるものは、既に実績値が判明しているのだから、何故、モデルで推計する必要があるのかと思われるかもしれないが、次のような目的では **concurrent** モデルが必要とされる。

現実の医療費 (実績値) は背景となっているコスト構造に依存して発生するものである。したがって、同じ健康状態であっても背景となるコスト構造により実績値は異なることがある。しかし、**concurrent** モデルでは、医療費の実額そのものは説明変数として用いないでその他の説明変数で当年度の医療費を説明しようとするものであるから、その推計値は、背景となる個別のコスト構造に依存しない (言い換えると、平均的なコスト構造における) 健康リスク (医療費) の評価とすることができる。また、**concurrent** モデルによる推計値は、インプットデータで示されているような健康状態の人々のアウトプット指標はどの程度の値となるのかの平均的な数値を示したものと考えることもできる。したがって、それと実績値を対比することによって、実際に提供された医療サービスの実

態が平均的な姿とどの程度乖離しているのかを分析していくことが可能となる。

特徴 5 : モデル開発のためのデータセット

予測モデルの開発で行われる統計分析は、モデル開発のためのデータセットを用いて、所定の説明変数でアウトプットとする指標がもっともよく説明されるようにモデルの係数 (リスク・ウェイト) を決める分析作業である。データセットに含まれる標本数が多いほど、安定的な結果が期待できる。上記の **Impact Pro** の例では、モデルの開発は、1,700 万人以上のデータに基づいて行われたと、《図表 3》の出典に示したディスクローズ資料で説明されている。データ量の問題だけでなく、用いられるデータセットが異なれば算出される係数の値も異なりうるという特徴もある。人種が異なれば係数の値は異なるだろうし、年齢や性別の分布、背景となる医療制度等々が異なれば係数の値は異なることが考えられる。各予測モデルがリリースされる際には開発時のデータセットに基づいて算出された係数が組み込まれているが、場合によっては、予測モデルの使用者が分析しようとする対象集団における過去のデータを用いて、係数を再算出した上で、モデルを使用するというも行われる。

Ⅲ. 予測モデル活用のその他の具体的事例

第 I 章では DM プログラム運営における予測モデルの活用を振り返ったが、米国ではそれ以外のヘルスケア分野の様々な局面で健康リスクを評価するための予測モデルが活用されている。脚注 3 で示したように、本稿では、「ヘルスケア分野」は、保険会社の観点から捉えたヘルスケア事業分野という意味で使用している。ここでは、その観点から特徴的である、3 つの局面

を背景とともに概観する。

1. 保険料率設定、保険引受条件判断における活用

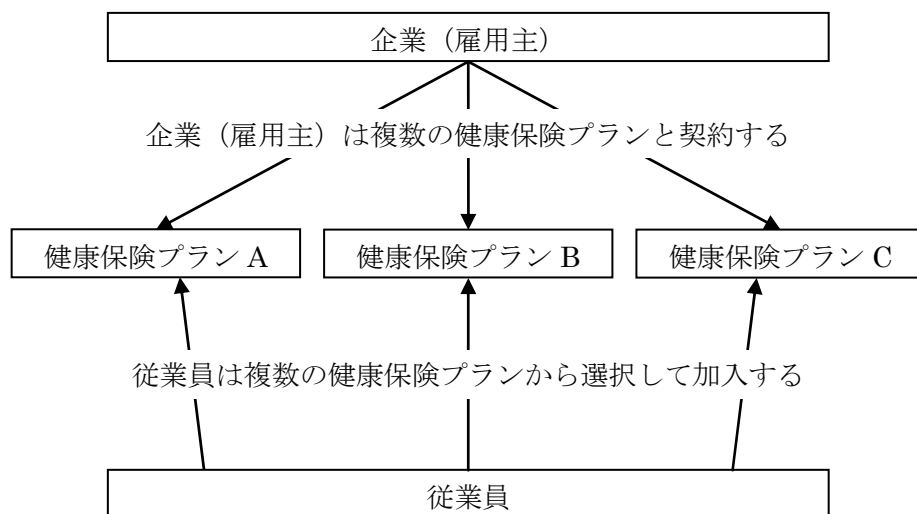
米国では、高齢者等を対象としたメディケア、低所得層を対象としたメディケイドといった公的な医療保障制度はあるものの、一般就業者は職域における団体募集を通じて民間保険制度に加入するのが一般的である⁵（以下、民間保険制度の運営者を「健康保険プラン」とする）。すなわち、企業（雇用主）は健康保険プランと契約してその従業員に医療保険を提供する。企業（雇用主）が従業員に提供する健康保険プランは単一ではなく、通常、複数のプランが提示され、その中から従業員は自分の加入するプランを選択する（《図表4》参照）。

一方、わが国では、健康保険法等の法律の下、国民皆保険の制度であり、政府管掌健康保険、健康保険組合、国民健康保険等運営主体となる保険者は様々であるが、統一の制度で運営されている。企業（雇用主）は、健康保険組合を設

立する（あるいは、既存の加入可能な健康保険組合に加入する）か、政府管掌健康保険に加入するのか、を選択することはできるが、どれか一つに決めなければならない。したがって、企業（雇用主）から従業員に提示される健康保険は一つで、従業員に選択肢はない。健康保険組合の側から見ると、設立母体となっている企業（雇用主）の従業員の増減に従って、加入者が増減するだけで、保険者間の競争は限定的である（《図表5》参照）。

わが国の状況と対比して言えば、米国の健康保険プランは、企業（雇用主）に採用してもらえるか、企業（雇用主）に採用されたとしても、次に、複数提示されている健康保険プランの中から従業員に選択してもらえるかという激しい競争の下で事業を運営しなければならない。さらに、企業（雇用主）が採用する団体保険として運営しなければならないという特徴もある。すなわち、企業（雇用主）が採用する団体保険の場合は、企業（雇用主）が採用を決め、保険の募集を開始したら、健康保険プランは加入申

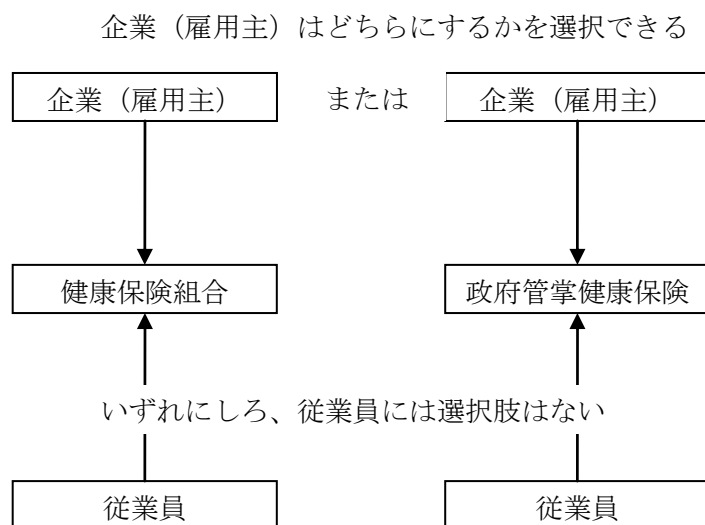
《図表4》米国における健康保険プランの選択



（出典）損保ジャパン総合研究所作成

⁵ 田中健司、森朋也「米国における健康保険市場と保険会社のヘルスケア事業 —2005年を中心とする概況と最近の動き—」（損保ジャパン総研クォーターリーVo.47、2007年）、p.8

《図表 5》わが国における健康保険の選択



込者を全員引き受けることとし、個々の従業員について引き受けるか否かの判断は行わない⁶。また、持病を持った従業員であったとしても、加入後に発生した医療費については支払対象とする。

健康保険プランはこうした環境下で、企業（雇用主）に提示する保険料率の設定を含む保険引受条件を決定しなければならない。したがって、自らが現在保険引受している、あるいは、これから保険引受しようとする各加入者および加入者集団の将来の医療費コストをできる限り正確に予測することが必要である。過大に見積もった場合には高い保険料を提示することになり競争に勝てない。逆に、過小に見積もった場合には収支の悪化を招く。これら为了避免するためには、加入者の健康状態を反映させてリスクを評価し、それに基づき保険料率の設定を含む保険引受条件を適切に調整するプロセスが必要である。このプロセスを実行するための健康状態に基づく

リスクの評価に際して予測モデルが活用されている。

リスクに基づく調整という意味で、米国アクチュアリー会 2007 年報告書では、このプロセスを *risk adjustment* と呼び、そのために使用されるツール（予測モデル）を *risk adjuster* と呼んでいる。また、保険引受条件を判断し決定することを保険業界の用語でアンダーライティング (*underwriting*)、また、その業務を行う人をアンダーライター (*underwriter*) という。

2. 医療プロバイダー⁷への医療費支払額決定および医療の質の維持・向上における活用

米国では、健康保険プランが医療プロバイダーへ支払う医療費の額を決定する際にも、健康リスクを評価するための予測モデルが活用されている。その背景事情は次の通りである。

わが国の健康保険の場合、医療機関への支払

⁶ 個人で加入するタイプの保険の場合は、個々の加入者の状況（健康状態等）を見て、民間保険会社は保険を引き受けるか、辞退するか、割り増し保険料を課すかといった引受判断を行うことが一般的である。

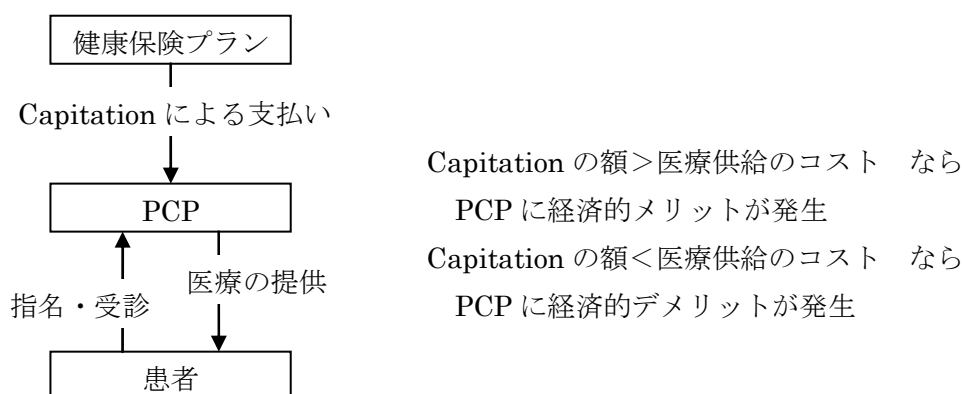
⁷ 米国では、医師・医療機関等のヘルスケア・サービスの供給者を指して医療プロバイダーという用語を使うことが多い。本節タイトルの医療プロバイダーはこの意味である。なお、米国アクチュアリー会 2007 年報告書等では、文脈によっては、単に“*provider*”と表現されていることもある。

は、診療報酬体系として一制度で運用されている。支払の基本は、診療・治療行為の種類等により定められた単価と実施された診療・治療行為の量によって決まる出来高払い制である。一部には、包括払い点数となっているものもある。

一方、米国の健康保険プランの形態は多岐にわたる。伝統的な形態はインデムニティプランと呼ばれるもので、加入者は医療機関を自由に選択でき、医療機関に対する支払は出来高払いで行われるものである。一方、今日主流となっている形態は、HMO (Health Maintenance Organization) 等のマネジドケア型と呼ばれるプランで、ヘルスケアのコスト、質、およびアクセスを管理しようとして、医療サービスの提供に一定の制約をかけることが行われる。マネジドケア型のプランでは、例えば、医療費の補償を健康保険プランが指定した医療機関ネットワーク内の医療機関で受診した場合に限ることや、加入者がプライマリーケア医 (primary care physician、以下「PCP」とする) を指定すること等が行われる。PCPは、プラン加入者から指名を受け、加入者に対して医療サービスを提供することにより、健康保険プランから支払いを

受ける。健康保険プランから PCP への支払は提供された医療サービスの量に基づく出来高払いではなく、**Capitation** と呼ばれる人頭払いの場合が多い⁸。人頭払いはその名の通り、自分が PCP に指名された加入者の人数に応じて、健康保険プランから支払いを受けることである。提供した医療サービスの量に係らず人数に応じて支払われるので、きちんとした患者対応等により重症化や合併症が回避できれば、投入する医療サービスの量が少なくて済み、結果として PCP には経済的メリットが生じる。逆に、想定以上に医療サービスの投入を要するようになった場合は、PCP に経済的デメリットが発生する (《図表 6》参照)。つまり、人頭払いは、出来高払いでは健康保険プランの側にのみ存在していた医療費変動リスクを医療プロバイダー側にも分担させる機能がある。しかし、人頭払いだからといって、担当する加入者の健康状態に係りなく医療プロバイダーに支払われる金額が一律だと、医療プロバイダー側はなるべく健康で医療費のかかりそうにない加入者のみを担当したがる傾向を招き、制度運営に支障をきたすことが懸念される。したがって、人頭払いでの支

《図表 6》 Capitation の構造



(出典) 損保ジャパン総合研究所作成

⁸ 西村徹、江頭達政、中村岳「米国ヘルスケア市場と保険会社のヘルスケア事業 —沿革、現状および最近の動向—」(安田総研クォーターリーVol.37、2001年)、p.21

払金額等、健康保険プランから医療プロバイダーへ支払う額の決定においても、PCP が担当する加入者の健康リスクに応じて金額を適切に設定することが重要となり、ここでも、健康リスクを評価する予測モデルが活用されている。

また、米国の場合、保険会社が提供しているのは、保険商品にとどまらず、医療プロバイダーのネットワークを含めた健康保険プラン全体を企画設計、運営するサービスであるという特徴がある⁹。その意味では、自社のネットワークに参加している医療プロバイダーとの連携を密にして、医療プロバイダー・ネットワークの提供する医療の質を維持することも保険会社の重要な戦略となる。例えば、**Capitation** が上記のような構造であるので、医療費変動リスクを医療プロバイダー側にも分担させる機能もある一方、経済合理性だけで言えば、医療供給を抑えようとするインセンティブが PCP に働くことも懸念される。もし医療の供給が過小であると、将来的に重症化や合併症の発症等に繋がり、それは医療の質の劣化を意味する。こうした事態に対応するためにも予測モデルは活用されている。医療費を推計する予測モデルは、同じような健康状態にある患者に要する平均的な医療費の額と考えることができる。保険会社は予測モデルによる推計値と実際に提供された医療サービスを比較する。医療サービス供給が推計値を大きく下回る場合には、過小供給 (*underutilization*) が生じているかもしれない。その場合、保険会社は当該医療プロバイダーとの連携を密にして、状況を確認し、医療の質に対する懸念の解消に努める。予測モデルは、医療プロバイダー・ネットワークの提供する医療の質の維持・向上に向けた取り組みにも活用されている。

3. 連邦政府等から健康保険プランへの委託料決定における活用

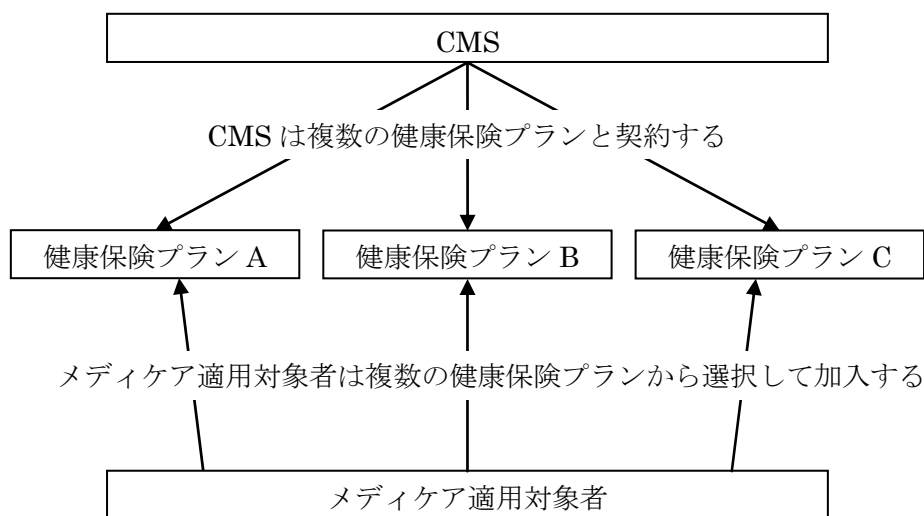
米国においては、連邦政府や州が運営するメディケアやメディケイドのような公的医療保障制度の適用対象者に対しても民間保険会社が健康保険プランを提供している¹⁰。65歳以上の者、65歳未満の障害者、および末期腎臓病患者を対象とするメディケアは、連邦政府の機関である **Centers for Medicare & Medicaid Services** (以下、「CMS」とする) が運営しており、連邦政府の予算と被保険者、企業(雇用主)の負担金等によって運営されている。給付内容は、CMSの運営するプラン(パートA、パートB:従来型プログラムで出来高払い方式)と、民間の健康保険プラン(メディケア・アドバンテージ。パートCと呼ばれることもある。主としてマネジドケア型)があり、メディケア適用対象者が任意に選択する。また、2006年1月から処方箋薬剤給付(パートDとも呼ばれる)の運用も開始された。このパートCやパートDはCMSが競争入札により選定し、契約した複数の民間健康保険プランにより提供される(《図表7》参照)。

また、低所得層を対象とした医療給付の公的制度であるメディケイドは、連邦政府の定めるガイドラインを満たした上で、各州で独自の運営ルールを作成し、州政府が運営主体となっている。メディケイドにおいても、州政府自らが給付を行うケースと、州政府が民間健康保険プランと契約して間接的に給付を行うケースがある。1967年に複数の州が従来型の出来高払いの健康保険に加え、マネジドケア型の健康保険プランの提供を開始して以降、対象者をマネジドケア型健康保険プランに加入させる取り組みは任意のものであったが、「1981年包括財政調整

⁹ 脚注3参照

¹⁰ 前掲注5、田中健司他、p.3

《図表 7》メディケア・アドバンテージにおける健康保険プランの選択



(出典) 損保ジャパン総合研究所作成

法」(Omnibus Budget Reconciliation Act of 1981)の下で、「社会保障法」に条項が追加され、州政府が適用対象者をマネジドケア型プランに強制加入させることも可能とされた。

こうしたメディケアやメディケイドにおけるマネジドケア型のプログラムの運営については、CMS や州政府が競争入札により選定した複数の民間健康保険プランに委託して実施されている。したがって、ここでも、CMS や州政府から、プログラム運営を委託した民間健康保険プランに支払う委託料(医療費)の決定は重要である。すなわち、CMS や州政府は、民間健康保険プランがマネジドケア型で運営することで、自らが出来高払い方式で運営しているコストよりも安いコストで運営されることを期待している。しかし、健康保険プラン側からすれば、契約した委託料(医療費)で加入者の1年間の医療費を賄わなければならないという医療費変動リスクを負担することになる。複数の健康保険プランが参画し、適用対象者が健康保険プランを選択する形で運営されるが、適用対象者の健康状態は様々である。したがって、CMS や州

政府から民間健康保険プランに支払われる委託料(医療費)が、各民間健康保険プランが引き受けている適用対象者の健康状態(リスク)に拘らず、一律であると、医療費変動リスクを全面的に負担することになる民間健康保険プランの収支は、引き受けた適用対象者の分布に大きく影響を受けることになり、公平な運営が図れない。場合によっては、健全な収支の確保が困難として、多くの民間健康保険プランがこのスキームから撤退する事態も生じ得る。そこで、制度を安定的に持続可能な運営とするためには、CMS や州政府から民間健康保険プランに支払われる委託料(医療費)を、各民間健康保険プランが引き受けている適用対象者の健康状態(リスク)を調整して決定することが重要となり、このリスク調整にも予測モデルが活用されている。

IV. 予測モデルの実務活用に向けた取り組み

上記の通り、健康リスクを評価するための予測モデルは米国のヘルスケア分野で様々な課題解決のために使用されている。また、その需要

に合わせて各種の予測モデルが開発されており、開発された様々な予測モデルがソフトウェアとして商用ベースで販売もされている。したがって、予測モデルの活用にあたって個々のユーザーは、各自の活用目的に合わせて、数ある予測モデルの中から最も適当なものを選択することが必要になる。そのためには、予測モデルのもつ制限やその予測精度に影響を与える諸要因を明らかにし、予測モデルを実務に導入・使用するとすればどのような点に留意すべきかを理解することが重要である。

米国アクチュアリー会 2007 年報告書は、米国の主要な 12 の予測モデル¹¹について、実務に即して設定した 10 個のシナリオ¹²における予測結果¹³を詳細に比較検討し、予測モデルのもつ制限やその予測精度に影響を与える諸要因、および実務への導入における留意点を報告している。米国アクチュアリー会 2007 年報告書の詳細な説明は第二部に譲り、本節では、米国アクチュアリー会 2007 年報告書に述べられている実務への導入における留意点について、その概要を説明する。

1. 予測モデルの実務への導入における留意点

(1) 予測モデルの特異性と分析対象集団への適用可能性

予測モデルにより、その開発意図や開発に使用したデータセットが異なる。その特徴が、実務で分析対象とする集団についても当てはまるのか慎重に検討することが重要である。第Ⅱ章第 1 節において、予測モデルのソフトウェアには、そのモデル開発に使用したデータセットに基づいて統計的に算出されたリスク・ウェイト

が組み込まれていると述べた。米国アクチュアリー会 2007 年報告書では、この報告書で分析対象とした集団のデータセットに合わせて各予測モデルのリスク・ウェイトの再調整を行い、組み込みリスク・ウェイトを使用したシナリオにおける予測結果と再調整リスク・ウェイトを使用したシナリオにおける予測結果の比較を行っている。その結果、リスク・ウェイトを再調整して用いれば、よいパフォーマンスが得られることがあることが示された。しかし、説明変数として設定される種々のカテゴリー分けが、実務で分析対象とする集団に対しても適当なものであるかどうか留意する必要があると指摘している。例えば、65 歳以上の集団を想定して開発された予測モデルは、妊娠関連の疾病や幼児期の疾病に対する取り扱いが十分考慮されていないかもしれないからである。

(2) データと予測のタイムラグ

予測モデルを実務に適用する局面においては、インプットデータとする医療給付金請求データや処方薬剤給付データの一部が支払い完了となっていない場合（データのラグ）や、インプットデータの収集期間末と予測対象期間の開始との間に間隔が空いている場合（予測のラグ）が一般的である。米国アクチュアリー会 2007 年報告書では、ラグなしのシナリオにおける各予測モデルの予測結果とラグありのシナリオにおける予測結果を比較している。ラグの存在が実務局面では一般的であることから、ラグが予測モデルのパフォーマンスに与える影響の程度を理解しておくことが重要であると指摘している。

¹¹ 12 の予測モデルの概要については、第二部第Ⅰ章参照。

¹² シナリオ設定の要素とシナリオの内容については、第二部第Ⅱ章参照。

¹³ 米国アクチュアリー会 2007 年報告書では、各予測モデルによる計算結果を比較評価する指標尺度として、決定係数、予測誤差の絶対値の平均、予測比率を用いている。これらの指標尺度については、第二部第Ⅲ章参照。

(3) 予測モデルのアップデートの必要性

米国アクチュアリー会 2007 年報告書で分析対象とした予測モデルを各モデルで使用されるインプットデータの種類で分けると、医療給付金請求データに記録された診断名に関するコードを使用するもの、処方薬剤に関するコードを使用するもの、その両方を使用するもの、過去の期間のコストデータも使用するものがある。記録された診断名のコードや処方薬剤のコードは時間の経過とともにそのパターンが変化することがあり、予測モデルは定期的にアップデートすべきことが指摘されている。

診断名に関して、コーディングのパターンが時間の経過に伴い変化することが想定される。コーディングに慣れるに従い、新規患者の主傷病や重症度のコーディングの正確性が向上したり、治療に影響を与えるすべての関連する状態を記録すべく記載されるコードの数が増えたりする。こうしたコーディング実務の変化により、予測モデル開発時の集団に比べて、リスクの高い集団と見えるようになることもあり得る。

処方薬剤に関しても、同様に、時間の経過に伴うパターンの変化に留意する必要がある。処方薬剤に基づいた予測モデルは一般的に、処方に関して任意性（自由裁量性）がないと考えられている薬剤を基にしているといわれている。しかし、短期間で新薬の導入が進むことや、承認されている適応症以外での処方が進むこと等も考えられる。したがって、処方薬剤に基づく予測モデルについても定期的にアップデートし、リスク・ウェイトの調整を行うことが重要であると指摘されている。

(4) 予測モデル選択の基本的視点

医療プロバイダーまたは健康保険プランに対して支払う医療費等の額の決定に予測モデルを

活用する場合には、その副次的インセンティブの懸念を考慮し、診断名に基づいた予測モデルが選択されることが多いと述べられている。これは、診断名に基づいた予測モデルの方が、処方薬剤に基づいた予測モデルに比べ、支払を受ける側に都合のいいようにデータに記録されるコードを使い分けるインセンティブが少ないと考えられるからとしている。

また、集団全体に対する予測ではなく、特定の疾病等、ターゲットを絞った集団に対する予測を行う場合には、重症度やターゲットとする疾病に関連する合併症を最も適切に捕捉しているのはどのモデルであるかを詳細に検討の上、モデル選択を行うべきことを指摘している。

(5) 予測モデル導入前の準備

予測モデルの導入を計画した段階で、事前テストや実際のデータハンドリングの経験を一定程度積むべきであると助言している。予測モデルに用いる医療給付金請求データや処方薬剤給付データは莫大なデータ量となる。莫大な量のデータのハンドリングやシミュレーションには習熟が必要であり、事前の経験も積まざりにいきなり導入することは極めてチャレンジング（困難）であるとしている。

(6) 分析に使用するデータの妥当性の確認

予測モデルへのインプットに使用するデータの妥当性を十分確認すべきことが指摘されている。米国では、多様な給付条件の健康保険プランが運営されており、医療プロバイダーに対する支払方式も多様である。人頭払いで医療サービスが提供されている場合で受診内容に関するデータに欠損がないか確認する必要がある。また、メンタル関連の給付は別の健康保険プランで提供される場合もあるが、メンタル関連の

サービス提供者からのデータがそっくり欠損しているといったこともあり、注意を喚起している。

さらに、予測モデルへのインプットに使用するデータを生成する管理システムにも留意が必要である。例えば、管理システム上、1レコードあたりに記録される診断名コードの数に制限を設けている場合、診断名コードの一部が欠落する可能性があり、リスクの評価が過小となることがあるかもしれない。

また、加入者資格データの更新頻度にも留意が必要である。米国では、健康保険プランの加入者の加入資格状況の変更に関するデータを企業（雇用主）から受け取るのに数か月かかる場合もある。例えば、加入者の死亡が加入資格データに反映されるのに数か月かかるとすれば、死亡後のその数か月間、その死亡した加入者は医療費が発生していないように見える。このことは、予測モデルを現状分析に使用する場合に最も大きな影響を与えると注意を喚起している。

(7) リスク評価を行う時期及び頻度

予測モデルによるリスク評価を行う時期、頻度、およびどの程度速やかに行うかも、予測モデルの実務活用上の留意点である。給付金請求が支払い完了となるまでに要する期間も考え合わせれば、通常、この期間の幅は6か月から24か月の間となろうと助言している。

2. 米国アクチュアリー会 2007 年報告書の背景

前節の通り、米国アクチュアリー会 2007 年報告書は極めて実務的な内容である。このような実務的な報告書が作成された背景を簡単に説明し、第一部の締めくくりとする。

米国アクチュアリー会 2007 年報告書における分析を行ったのは、アクチュアリー

(actuary) と呼ばれる保険・年金分野の数理統計の専門職である。一般の商品であれば、販売する時点で原価（製造コスト）はほぼ確定しているが、保険・年金の分野では、販売時点では原価（保険金等の支払がいくらになるか）は確定しておらず、掛け金（保険料）を設定するにも、将来の保険金支払を推計する等の確率や統計の技術を要する。また、一般の商品であれば、販売した時点で売上げを収益と認識し、原価を費用と認識し、その差額で損益が確定する。しかし、保険・年金の分野では、掛け金（保険料）は通常前払いで契約時点で収入し、保険金等を支払う責任はそれから長期にわたって負うものである。したがって、会計期間に合わせて期間損益を認識するためには、将来の保険金支払に備えた負債の評価（支払備金、責任準備金の算出）が必要となる。これらも決算期末では確定していないものなので、確率や統計の技術を用いて推計・評価することが必要である。これら以外にも、資産運用等々、保険・年金の分野では確率や統計の技術を要する業務が各種存在する。アクチュアリーはこれらの業務に要する数理統計技術を担う専門職である。なお、アクチュアリー会という組織は、アクチュアリーを会員とした専門職能団体で、アクチュアリー資格試験の運営、会員の継続教育、関連調査・研究等を行っている。各国にアクチュアリー会が組織されており、米国アクチュアリー会 2007 年報告書のスポンサーとなった米国アクチュアリー会とは、米国の Society of Actuaries と呼ばれている組織である。

米国アクチュアリー会は、医療給付金請求データに基づいた健康リスク評価のための各種手法・ツールの比較分析を継続して数年おきに実施し、公表してきた。米国アクチュアリー会によるこの研究プロジェクトは、1996年、2002

年にも実施されており、2007年4月に公表された米国アクチュアリー会2007年報告書は第3回目の報告書である。予測モデルのユーザーとして保険会社等がその導入を検討する際、モデルの選択、導入方法の検討を担う一員がアクチュアリーである。各アクチュアリーがこうした実務を担うために必要な知識と技術に関する調査研究を行い、その結果を広く会員に知らしめることも米国アクチュアリー会という専門職団体の使命の一つであり、その目的でこの報告書が作成されたと考えられる。したがって、この報告書は、予測モデルの開発者サイドが、自社のモデルの優秀さを主張するためのレポートではなく、ユーザーサイドが理解すべき予測モデルの精度比較方法の特徴、限界、導入に際しての留意点を詳細に検討している点が興味深い。詳しくは第二部を参照されたい。

第二部 予測モデルの比較検討および選択・導入時に考慮すべき諸点

第二部では、米国アクチュアリー会 2007 年報告書に基づいて、米国アクチュアリー会のスポンサーシップにより実施された 12 の予測モデルの比較検討結果、およびそれを踏まえた予測モデルの選択・導入時に考慮すべき諸点について取り上げる。第 I 章で米国アクチュアリー会 2007 年報告書の分析対象とした 12 の予測モデルの概要を述べ、第 II 章で比較評価のために設定した 10 種類のシナリオについて、第 III 章でモデルによる予測の精度を評価するために同報告書が採用した 3 種類の尺度について解説をした上で、第 IV 章で同報告書に示された分析結果の概要を紹介する。最後に第 V 章で、分析結果を踏まえた予測モデルに関する諸留意点および同報告書の研究者自らが表明している今後の研究課題について解説を加える。

I. 分析対象の予測モデルの概要

2002 年の研究では 7 つのモデルが取り上げられたが、米国アクチュアリー会 2007 年報告

書では 12 のモデルが分析対象とされた。米国アクチュアリー会 2007 年報告書の原タイトルに **claim-based** とあるが、**claim** とは給付金請求の意味で、**claim-based** とは健康保険プランが給付金請求受付・支払のために保持する管理データをインプットデータとすることを意味する。米国アクチュアリー会 2007 年報告書では、給付金請求データにおける診断名に関するコードと処方薬剤に関するコードのいずれか一方または両方を使用するモデルに焦点を当てている。分析対象とした 12 のモデルの内訳は、診断名に基づくモデルが 4、処方薬剤に基づくモデルが 3、診断名と処方薬剤の両方に基づくモデルが 2、過去の期間のコストデータも用いるものが 3 である。分析対象としたモデルの略名、開発者、インプットに使われている情報項目の一覧表は《図表 8》の通りである。また、〈インプット〉欄の表記に関する簡単な説明は《図表 9》の通りである。

各ソフトウェア・ベンダーから提供された説明資料からの要約として、米国アクチュアリー会 2007 年報告書では、各モデルの概要を次の

《図表 8》米国アクチュアリー会 2007 年報告書における分析対象モデル一覧表

<モデル>	<開発者>	<インプット>
ACG	Johns Hopkins	Diag
CDPS	Kronick/UCSD	Diag
Clinical Risk Groups	3M	Diag
DxCG DCG	DxCG	Diag
DxCG RxGroups	DxCG	Rx
Ingenix PRG	Ingenix	Rx
MedicalRx	Gilmer/UCSD	Rx
Impact Pro	Ingenix	Med+Rx+Use
Ingenix ERG	Ingenix	Med+Rx
ACG w/ Prior Cost	Johns Hopkins	Diag+\$Rx
DxCG UW model	DxCG	Diag+\$Total
<サービス・ベンダー>		<インプット>
MEDai	MEDai	All

(出典) 前掲注 4 の資料、p.1 の Table I.1 より損保ジャパン総合研究所作成

《図表 9》 インプット欄の表記の説明

表記	説明
Diag	ICD-9(*1) の診断名コード
Med	ICD-9 の診断名コードおよび処置の情報
Rx	処方薬剤の NDC(*2)コード
\$Rx	過去の期間の処方薬剤のコスト
\$Total	過去の期間の総コスト
Use	過去の期間の資源利用（コスト以外）の指標
All	上記のすべて

*1 the International Classification of Diseases – 9th Revision

*2 National Drug Codes

(出典) 前掲注 4 の資料、p.17 の表より損保ジャパン総合研究所作成

ように説明している¹⁴。

1. Adjusted Clinical Groups (Vender: Johns Hopkins University, School of Public Health)

Adjusted Clinical Groups (ACGs)は、ジョンズ・ホプキンス大学の Jonathan Weiner およびその他の研究者によって開発された診断名に基づくリスク評価モデルである。ACG システムにはハイ・コスト・ケースを特定する目的で開発された予測モデルのパッケージソフトも含まれている。米国アクチュアリー会 2007 年報告書では ACG Case-Mix System 7.1 を用いた。このモデルでは、ACG 分類、拡大された診断名クラスター (Expanded Diagnosis Clusters: EDCs)、および将来の入院または医学的に虚弱な状態が発生する可能性を示す診断上の指標が組み込まれている。ACG 分類とは、このモデルの基本となっている病的状態 (morbidity) に基づいた分類のことである。また、EDCs とは、一部の、インパクトの大きな、疾病特異的な診断名の区分であると説明されている。

米国アクチュアリー会 2007 年報告書で用い

た現状分析用モデル¹⁵は、回帰分析に基づいたモデルではなく、actuarial cell アプローチに基づいたモデルである。ACG でいう actuarial cells とは、同様のリスクのレベルにある患者に関する分類のことであり、それは臨床的に定義され、かつ、相互に排他的に取り扱われる。他の条件がすべて同等であれば、通常、このアプローチでは予測精度は低くなる。しかしながら、actuarial cells は、導入が容易で、解釈もしやすく、また安定性もあるという性格を持つので、このモデルは支払金額の推計に適用することが望ましいとの ACG チームの見解が紹介されている。

2. Chronic Illness and Disability Payment System

Chronic Illness and Disability Payment System (CDPS)は、カリフォルニア大学サンディエゴ校の Richard Kronick およびその他の研究者により開発された診断名に基づくリスク評価モデルである。米国アクチュアリー会 2007 年報告書では CDPS Version 2.5 を用いた。こ

¹⁴ 米国アクチュアリー会 2007 年報告書で分析対象とされたモデルが 12 種類であるが、ACG および ACG w/Prior Cost は合わせて ACG として概要説明しているので、モデル概要での標題は 11 種類となっている。

¹⁵ 予測モデルには、将来の状態を予測するためのモデルと現状を分析するためのモデルがある。詳細は後記第 II 章第 1 節参照。

のモデルは、そもそもは、障害者および Temporary Aid for Needy Families (TANF) を含むメディケイド母集団向けに開発された。CDPS モデルは、Disability Payment System (DPS: Kronick がメディケイドの障害者集団を対象に開発し、1996 年に公表した先行モデル) を更新・拡張したものである。

CDPS では、各人に対して、診断名コードに基づき、67 の医的な状態を示す分類の一つもしくは複数を割り当てる。各人には、同時に、年齢と性別からなる 16 の分類の一つが割り当てられる。割り当てられた医的な状態分類と年齢性別分類に基づいて、各人の総医療費が予測される。このモデルは、2 種類のリスク・ウェイトを提供している。一つは、TANF 集団を対象に算出されたものであり、もう一つは障害者集団を対象に算出されたものである。TANF 集団の方が米国アクチュアリー会 2007 年報告書で用いた民間保険の加入者集団に類似しているため、米国アクチュアリー会 2007 年報告書では、TANF 集団用のリスク・ウェイトを用いた。また、このモデルでは、成人用、小児用の 2 種類の異なるリスク・ウェイトが用意されており、米国アクチュアリー会 2007 年報告書ではその両方を用いた。

3. Clinical Risk Groups (Vender: 3M)

Clinical Risk Groups (CRGs) は 3M により開発された診断名に基づくリスク評価モデルである。米国アクチュアリー会 2007 年報告書では、2006 年にリリースされた CRG Version 1.4 を用いた。CRG は、人頭払い制度におけるリスク調整やマネジドケア型のプラン運営組織の管理用ツールとして使用されている。このモデルのデザインと開発は、メディケアの入院における包括前払い制度の影響を受けている。すべての

加入者には、臨床的な基準に基づいて、約 1,100 種類設定されているリスク分類のどれか一つが割り振られる。CRGs は臨床面に着目しているため、ケアのパスを支援する管理システムや商品ライン管理、ケース管理の基礎となるものとして設計されている。

4. Diagnosis Cost Groups (Vender: DxCG)

Diagnosis Cost Groups (DCG) は、DxCG の製品である RiskSmart Models の構成要素である。DCG の研究は、ボストン大学において 1984 年に開始され、その後の 20 年間で、ボストン大学の Arlene Ash と Randall Ellis の指揮の下、多くの改良と拡張が実施されてきた。DCG は診断名に基づいたモデルで、分析対象の集団のタイプ（民間保険加入者、メディケイド対象者、あるいは、メディケア対象者）やデータのソース（入院データのみ、あるいは、入通院すべての受診データを含む）、モデルの目的（支払額の決定に使用する、あるいは、現状の説明に使用する）に合わせて様々な変数が用意されている。

米国アクチュアリー会 2007 年報告書では、RiskSmart Version 2.1.1 を用いた。DCG モデルは、民間保険の加入者集団におけるすべての受診を対象としたモデルで、総支払額（医療費と処方薬剤費の合計）の将来予測および現状分析の両方に使用される。先行研究が実施された時点においては、総支払額の現状を分析するモデルはリリースされていなかった。

DxGroups が DCG モデルの主要な構成要素である。すべての診断名コードは 781 の臨床的に均質なグループ (DxGroups) に分類される。これらのグループはさらに 184 の階層化された状態カテゴリー (hierarchical condition categories: HCC) にマッピングされる。併せて、

各患者には 32 の年齢・性別カテゴリーが割り振られる。このモデルは、HCC と年齢・性別カテゴリーに基づいて、各患者の総コストを予測する。

5. Episode Risk Groups (Vendor: Ingenix)

Episode Risk Groups (ERGs)は、Ingenix Inc.の子会社である Symmetry Health Data Systems が開発したリスク評価モデルである。ERGs は、同じく Symmetry が開発した Episode Treatment Groups (ETGs)モデルに基づいている。ETGs は医療サービスをケアのエピソードに分類するものである。ERGs は 2001 年に開発され、リリースされている。米国アクチュアリー会 2007 年報告書で用いた ERGs は、ETGs の Version 5.3 に基づいたものである。

ERG モデルは、医療費請求データおよび処方薬剤費請求データにより把握可能な診断名および治療行為の情報に基づき、各加入者を episode risk groups と呼ばれる 120 の医的な状態を示すカテゴリーの一つまたは複数に割り付ける。各加入者の ERG によるプロファイルは年齢、性別、および割り付けられた ERG を考慮して決定される。このプロファイルを用いて、将来および過去のリスク・スコアが推計される。

6. Impact Pro (Vendor: Ingenix)

Impact Pro は、Ingenix Inc.の子会社である Integrated Healthcare Information Services, Inc. (IHCIS) により開発された。このシステムは、報告システムとリスク調整アルゴリズムを一体化したシステムで、加入者データ、医療給付金請求データおよび処方薬剤費請求データを組み込んでいる。このシステムでは、個々の給

付金請求を、episodes of care およびそれ以外の診断名に基づくカテゴリー (Impact Clinical Categories: ICCs) のいずれか一つに分類する。これらのカテゴリーは、加入者に観察された疾病、医的な状態、および共存症や合併症の組み合わせを表している。ICCs はさらに、“base-markers”と呼ばれる均質なリスク・カテゴリーにグループ化される。各加入者は 1 つまたはそれ以上の base-markers および 1 つの demographic marker に分類される。これらのマーカーに基づき、各加入者のリスク・ウェイトがアウトプットされるが、いくつかの異なる条件設定 (i.e. 頭切り水準¹⁶) が用意されている。

7. MEDai (Vendor: MEDai, Inc.)

Risk Navigator ClinicalTM は MEDai, Inc.により開発された予測モデルのソリューションでありレポーティング・ツールである。Risk Navigator ClinicalTM は、医療給付金請求データ、処方薬剤費請求データ、年齢・性別等の属性、検査結果および健康リスク評価結果 (health risk assessment: HRAs) を用いて、コスト、入院期間、緊急治療室の利用、処方薬剤コストおよび節減額を予想する。疾病に関するエピソード (Symmetry ETGs)、薬剤のカテゴリー、年齢、性別、保険の種類およびその他のリスク・マーカー (治療または診断のタイミングや頻度等)等の臨床上の要因の組み合わせに基づいて、加入者ごとの個別の予測が行われる。

Risk Navigator ClinicalTM では、モデルの構築、改善・精緻化、および検証に 2 年間分のデータを用いる。収集され、有効性が確認されたデータは MEDai の予測エンジン (Multiple Intel-

¹⁶ 分析にあたって、所定の額を超える金額の給付金請求については頭切りを適用して、当該所定の金額であったものとして取り扱うことをいう。詳細は第 II 章第 6 節参照。

ligent Tasking Computer Heuristics: MITCH) で処理される。このエンジンは線形の分析技術と非線形の分析技術が組み合わされている。

8. MedicaidRx

MedicaidRx は、カリフォルニア大学サン・ディエゴ校の Todd Gilmer およびその他の研究者により開発された処方薬剤に基づいたリスク評価モデルである。このモデルは、もともとはメディケイド集団を想定してデザインされたもので、Group Health Cooperative of Puget Sound の研究者により開発された Chronic Disease Score model を更新・拡張したものである。

MedicaidRx モデルでは、各加入者に対して、各加入者が使用した処方薬剤に基づき 45 の医的条件カテゴリーのうちの一つまたはいくつかを割り振るとともに、11 の年齢・性別によるカテゴリーの一つを割り振り、各メンバーの医療費の総額を予測する。このモデルのリスク・ウェイトは成人用と小児用に分けて設定されている。

9. Pharmacy Risk Groups (Vendor: Ingenix, Inc.)

Pharmacy Risk Groups (PRGs) は、Ingenix Inc. の子会社である Symmetry Health Data Systems が開発した処方薬剤に基づくリスク評価モデルである。米国アクチュアリー会 2007 年報告書では PRGs の Version 5.3 を用いた。PRGs の構成要素は、各人の処方薬の組み合わせと処方されたそれぞれの薬が他の薬とどのように関係するかである。処方された薬剤は National Drug Codes (以下、「NDC」とする) というコード体系で記録されるが、記録された各 NDC は 107 の PRGs の一つに分類される。各加入者に対する PRG プロファイルは、年齢、

性別および各加入者に割り振られた PRGs によって作成され、各加入者の将来および過去のリスクスコアが算出される。

10. RxGroups (Vendor: DxCG, Inc.)

RxGroups は、DxCG の製品である RiskSmart Models の一部を構成するツールである。米国アクチュアリー会 2007 年報告書では、RiskSmart Models 2.1.1 を用いた。RxGroups は処方薬剤に基づいたリスク評価モデルで、Kaiser Permanente、CareGroup of Boston および Harvard Medical School の研究者および臨床家により開発され、2001 年にリリースされた。このモデルでは、NDC は、各薬剤のもつ治療上の意味合いに基づき、164 の相互に排他的なカテゴリー (RxGroups と呼ばれる) に分類される。各患者には 32 の年齢・性別カテゴリーの 1 つが割り振られる。このモデルでは、RxGroups と年齢・性別カテゴリーに基づき、各患者の総医療費を予測する。

11. Underwriting Model: RiskSmart (Vendor: DxCG, Inc.)

RiskSmart underwriting model は、DxCG の製品である RiskSmart Models に新たに追加されたモデルで、2006 年にリリースされた。米国アクチュアリー会 2007 年報告書では、RiskSmart Version 2.1.1 を用いた。Underwriting Model は、給付金請求がすべて完了する前の時点で、アンダーライターが医療保障の保険契約更改および保険料設定の目的で職域集団を評価する際に使用されている。このモデルでは、予測の元となるデータ期間の末日と予測対象期間との間に 6 か月間のラグがあることを前提として、給付金請求のラグを予測に組み込んでいる。このモデルでは、将来の医療

コストを予測するために、HCC、疾病の相互作用、年齢・性別カテゴリーおよび過去の期間のコスト（prior cost）を使用している。このモデルは、予測に役立てるために過去の期間のコストを変数として含んでいる。また、複数の頭切りオプション（\$25,000、\$100,000、\$250,000）を用意している。

II. 比較評価のためのシナリオ設定

米国アクチュアリー会 2007 年報告書では、各モデルの予測精度の比較評価のために、10 個のシナリオを設定している。シナリオを分ける要素は、①将来予測か現状分析か、②計算に使用するリスク・ウェイトはソフトウェア付属のものか、今回の分析対象集団のデータに合わせて再調整されたものか、③データおよび予測のラグを設定しているかいないか、④過去の期間のコストを変数として含めるか含めないかの 4 要素である。以下、この 4 要素およびその他の取り扱いについて補足説明する。

1. Prospective vs. Concurrent 将来予測か現状分析か

予測モデルを将来予測（prospective）に使用するとは、過去の給付金請求データを用いて将来の期間の医療費請求の金額を予測するということである。現状分析（concurrent or retrospective）に使用するとは、ある期間の給付金請求データを用いて当該期間の医療費請求の金額を予測することである。現状分析に使用する場合は、現実の背景となっているコスト構造に係らず、各人の健康状態を推定することができる。

米国アクチュアリー会 2007 年報告書の将来予測シナリオでは、2003 年の診断名データおよび処方薬剤データを用いて、2004 年の各人の医

療費請求の金額を予測している。また、現状分析シナリオでは、2003 年の診断名データおよび処方薬剤データを用いて、2003 年の各人の医療費請求の金額を予測している。米国アクチュアリー会 2007 年報告書における現状分析シナリオの設定は米国アクチュアリー会の先行研究とは少し異なっている。米国アクチュアリー会 2007 年報告書では研究のためのデータ期間の初年度を対象としているのに対し、1998 年と 1999 年のデータが利用可能であった先行研究では、1999 年（研究のためのデータ期間の第 2 年度）を対象としていた。

2. Offered vs. Calibrated Risk Weights ソフトウェア付属のリスク・ウェイトか再調整したリスク・ウェイトか

各モデルとも、医的な状態（medical condition）のカテゴリー毎にリスク・ウェイトが設定されている。リスク・ウェイトは、特段の医的な状態を持たない個人のベースとなるコストに対する、それぞれの医的な状態における相対的な限界費用の推計値である。リスク・ウェイトを例示すると《図表 10》（《図表 3》の再掲）の通りである。

前章で簡単に説明したように、Impact Pro の場合、エピソードに基づき、各種のマーカーが設定されるが、それぞれのマーカーの内容に対して、リスク・ウェイトが設定されている。例えば、インシュリン依存性糖尿病で併存疾患ありの場合のリスク・ウェイトは《図表 10》の例では 1.1079 である。この人には他にも該当するマーカーが存在し、リスク・ウェイトの合計は 10.12277 となっている。これは、この人のコストが特段の医的な状態を持たない個人のベースとなるコストに対して、相対的に 10.12277 倍となることを意味している。

《図表 10》リスク・ウェイトの例 (Impact Pro)

マーカー	内容	リスクウェイト
02.01.000	Insulin dependent diabetes, with co-morbidity, (base marker) インシュリン依存性糖尿病、併存疾患あり	1.1079
02.01.103	Inpatient Stay, diabetes primary within recent 3 months 直近3か月間に糖尿病を主因とする入院あり	4.5714
08.20.000	CHF, with co-morbidity, (base marker) 鬱血性心不全、併存疾患あり	1.0737
08.20.C3	Significant CHF episode clusters, recent 3 months 直近3か月間に重大なCHFエピソードあり	1.8619
10.05.000	Chronic bronchitis, with co-morbidity, (base marker) 慢性気管支炎、併存疾患あり	0.3978
RX.04.06	Blood, anticoagulants, CHF CHFにより血液抗凝固薬を服用	0.3938
Age-Sex Marker	Males, 55-64 男性、55-64歳	0.7212
合計		10.12277

(出典) IHCIS 社のディスクローズ資料 “A Clinical Episode-Based Approach to Predictive Modeling: Impact Pro™” に基づき損保ジャパン総合研究所作成

マーカーの設定や分類の仕方はモデルにより様々であるが、各モデルとも、それぞれのモデルで前提としているマーカーや分類に対して設定したリスク・ウェイトを基に、各人の医療費を予測する。各モデルがソフトウェアとして市場に出る際には、そのモデルの開発・検証に使用したデータセットに基づいて算出されたリスク・ウェイトを、標準的なリスク・ウェイトとしてソフトウェアに組み込んで提供されている。米国アクチュアリー会 2007 年報告書でいう *offered risk weights* とは、このソフトウェアに付属しているリスク・ウェイトのことである。米国アクチュアリー会 2007 年報告書に掲載されている図表等ではこれを略して、単に “Offered” と表記されていることがある。

上記のように、各モデルの *offered* リスク・ウェイトは、それぞれのモデルの開発に使用したデータセットを統計的に解析し、算出されたものである。分析の対象とする集団(データセット)が異なれば、同じ手法で解析しても結果として算出されるリスク・ウェイトが異なる可能性もある。ベンダーによっては、一定人数規模以上の集団に対してモデルを使用する顧客に対

しては、その集団のデータを用いて解析し、その集団独自のリスク・ウェイトを推計するサービスを提供しているところもある。米国アクチュアリー会 2007 年報告書では、分析対象として、約 62 万名の健康保険プラン加入者集団のデータを用いているが、各モデルの *offered* リスク・ウェイトによる分析に加え、この分析対象データを用いてリスク・ウェイトの調整を行い、調整したリスク・ウェイトによる分析も実施している。米国アクチュアリー会 2007 年報告書では、この調整したリスク・ウェイトを *calibrated risk weights* と表現しているが、ページによっては、*recalibrated* と表現されていることもある。

米国アクチュアリー会 2007 年報告書で採られたリスク・ウェイト調整の具体的方法、計算手順は次の通りである。

ステップ 1：データセットの分割

米国アクチュアリー会 2007 年報告書では、米国の商用データベース MedStat MarketScan のデータを用いている。このデータには、ICD-9 コード、CPT-4 コード (Current Procedural

Terminology, 4th Edition)、NDC コードが記録されている。このデータベースのデータのうち、comprehensive major medical と呼ばれる種類の健康保険プランに2年間連続して加入していた被保険者約 62 万名、年間の給付金請求金額にして約 30 億ドルの集団のデータを米国アクチュアリー会 2007 年報告書では用いている。

調整リスク・ウェイトの導出と検証のために、米国アクチュアリー会 2007 年報告書では 50/50 split design を採用した。具体的には、上記の 62 万人分のデータの個人を、(1) 調整リスク・ウェイト算出用のサブセットおよび(2) 検証 (validation) 用のサブセットのどちらか一方に無作為に割り付け、データセットを2分割した。

ステップ 2: Offered リスク・ウェイトによるリスクスコアの算出

各モデルは、それぞれのモデルで設定している医的な状態のカテゴリー、処方薬剤のカテゴリーおよび年齢・性別のカテゴリーに対して、該当するかしないかを表す変数 (indicator variable (0 or 1)、以下「指標変数」という) を当てはめ¹⁷、該当するカテゴリーのリスク・ウェイトを合計して各人のリスクスコアを計算している。調整リスク・ウェイト算出用のサブセットに対して、各モデルで offered リスク・ウェイトを用いて、各人のリスクスコアを算出した。

ステップ 3: 回帰分析 (調整リスク・ウェイトの算出)

リスク・ウェイトを調整する方法は各種考えられる。米国アクチュアリー会の先行研究では、個人属性や状態等の指標変数を給付金請求合計

額の実績値に回帰させる手法が取られたが、米国アクチュアリー会 2007 年報告書では、個人属性や状態等の指標変数を、給付金請求合計額の実績値と offered リスク・ウェイトによる予測値との差額に回帰させることで、リスク・ウェイトの調整を算出している。概念的に示すと、次式の通りである。

$$Y_{Actual} - Y_{Prediction} = \sum_{i=1}^A \alpha_i \times AgeBin_i + \sum_{i=1}^B \beta_i \times ConditionBin_i$$

ここで、

Y_{Actual} : 給付金請求合計金額の実績値 (医療費、処方薬剤費を含む)

$Y_{Prediction}$: offered リスク・ウェイトを用いた各モデルによる給付請求合計金額の予測値 (医療費、処方薬剤費を含む)

$AgeBin_i$: 年齢・性別等のデモグラフィックベースのカテゴリーの指標変数

α_i : 年齢・性別等のデモグラフィックベースのカテゴリーのリスク・ウェイトを調整するための回帰係数

$ConditionBin_i$: 医的な状態、処方薬剤等のカテゴリーの指標変数

β_i : 医的な状態、処方薬剤等のカテゴリーのリスク・ウェイトを調整するための回帰係数

さらに、調整の実施に当たっては、上記の回帰分析で得られた回帰係数の p 値に基づいた信頼度を考慮している。回帰分析を行う際には、同時に、得られた回帰係数について、「回帰係数 = 0」を帰無仮説とする検定を行う。ここでいう p 値は、この帰無仮説が成立するならばという

¹⁷ 細かく言うと、CRG と MEDai の取り扱いは少し異なる。CRG では、各人を約 1,000 個設定されているリスク・カテゴリーの一つまたは複数に当てはめる。MEDai では、医的な状態、処方薬剤、年齢・性別といったカテゴリーに対する指標変数に加え、過去の期間のコストのような連続変数も用いている。

仮定の下で、回帰係数が得られた値となる確率を示したもので、 p 値が小さいほど、回帰係数は 0 ではないことが確からしいと考えることができる。そこで、米国アクチュアリー会 2007 年報告書では、上記の回帰分析で得られた回帰係数に信頼係数を乗じて、最終的な調整を求めている。なお、信頼係数としては、 $(1-p)$ 値^{5,95}を使用したと説明されている。

また、現状分析のシナリオに対するリスク・ウェイトの調整では、デモグラフィック関連の変数 ($AgeBin_i$) を説明変数として用いないという取り扱いとしたことが説明されている。現状分析のシナリオでは、給付金請求の発生していない人に対してはリスクが計算されないことが望ましい。しかし、上式のように、デモグラフィック関連の変数を説明変数に含めて回帰分析を行うと、デモグラフィック関連の変数にも回帰係数が計算されることになり、結果として、給付金請求の発生していない人に対しても「0」でないリスクを計算してしまうことになる。これを避けるためこの取り扱いとしている。

3. With and Without Data and Prediction Lag (“Lagged” and “Nonlagged”) データおよび予測のラグの有無（「ラグあり」および「ラグなし」）

米国の健康保険プランでは、給付金請求は支払われるまでに平均的に数か月、場合によっては、もっと長い期間（数年に及ぶこともある）を要することもある。データのラグとは、予測モデルを使った分析を実施する時点では情報の把握に漏れがあり、給付金の請求・支払データに欠損が存在する状況をいう。また、多くの実務局面においては、支払データの期間末と予測期間の始まりとの間にラグが存在する。これを予測ラグという。米国アクチュアリー会 2007 年報告書では、「ラグあり」シナリオ、「ラグな

し」シナリオの両方の分析を行っているが、「ラグあり」シナリオとは、データのラグおよび予測のラグの組み合わせが存在するシナリオのことをいう。

「ラグなし」シナリオでは、2003 年発生の請求事案に、2005 年 8 月までの支払を反映させたデータを使用した。すなわち、「ラグなし」の現状分析においては、2003 年 1 月 1 日から 2003 年 12 月 31 日までに発生した給付金請求を対象として分類を行い、同期間に対応する予測値を算出している。「ラグなし」の将来予測においては、同じく、2003 年 1 月 1 日から 2003 年 12 月 31 日までに発生した給付金請求を対象として分類を行い、予測期間としては、2004 年 1 月 1 日から 2004 年 12 月 31 日までに発生した給付金請求を対象としている。ただし、いずれも 2005 年 8 月 31 日までの支払が反映されている。

「ラグあり」シナリオでは、2003 年 1 月から 8 月発生の請求事案に 2003 年 8 月までの支払を反映させたデータを使用した。すなわち、「ラグあり」の現状分析においては、分類期間・予測期間とも、2003 年 1 月 1 日から 2003 年 8 月 31 日までに発生した給付金請求を対象とし、2003 年 8 月 31 日までの支払が反映されている。また、「ラグあり」の将来予測においては、分類期間としては 2003 年 1 月 1 日から 2003 年 8 月 31 日までに発生した給付金請求（ただし、2003 年 8 月 31 日までの支払を反映）を対象として分類を行い、予測期間としては、2004 年 1 月 1 日から 2004 年 12 月 31 日までに発生した給付金請求（ただし、2005 年 8 月 31 日までの支払が反映）を対象としている。

不完全なデータでは、一般的に予測の精度は落ちることとなるが、「ラグあり」シナリオは、アクチュアリーやアンダーライターが仕事をし

なければならぬ環境を適切に反映したものであると米国アクチュアリー会 2007 年報告書では説明している。また、処方薬剤に基づいたモデルは、診断名（および診断名プラス処方薬剤）に基づくモデルよりもデータのラグの悪影響を受けにくいことにも触れている。なぜなら、処方薬剤のデータの方がより早く支払いがなされるからである。このことは情報のラグの軽減に寄与するが、予測のラグの軽減にはつながらない。

4. Including Prior Costs as a Predictor 過去の期間のコストを予測因子として加えること

過去の期間のコストの集計値を明示的な予測因子(説明変数)としてモデルに用いることは、医療プロバイダーまたは健康保険プランに対する支払を目的とするケースでは適切ではない。しかしながら、保険数理分析や保険引受条件判断が目的の場合は、過去の期間のコストを組み込むことで、モデルのパフォーマンスが大幅に改善する。事実、いくつかのモデル (DxCG underwriting model、本研究で使用された MEDai model および ACG prior cost model) では、過去の期間のコストが組み込まれている。米国アクチュアリー会 2007 年報告書では、その他のモデルについても、過去の期間のコストを説明変数として追加するという修正も、分析に加えている。

過去の期間のコストを説明変数として追加するという修正は、第 2 節の調整リスク・ウェイトの算出のステップの一部として行われた。Offered の状態では過去の期間のコストを説明変数として使用していないモデルでは、リスク・ウェイトの調整を行わないまま、過去の期間のコストを変数追加することは適切でないの、各モデルのリスク・ウェイトの調整を算出

する際に同時に、過去の期間のコストを説明変数として追加して、回帰分析を実施した。概念的に示すと、次式の通りである。

$$Y_{Actual} - Y_{Prediction} = \sum_{i=1}^A \alpha_i \times AgeBin_i + \sum_{i=1}^B \beta_i \times ConditionBin_i + \gamma \times PriorCost$$

ただし、第 2 節の式における記号の定義に加えて、

γ : 過去の期間のコストに対するリスク・ウェイトを調整するための回帰係数

PriorCost : 過去の期間のコスト

5. 米国アクチュアリー会 2007 年報告書で設定したシナリオ

上記の 4 要素の区分けに従えば、シナリオは、計算上 $2 \times 2 \times 2 \times 2 = 16$ 通り考えられるが、過去の期間のコストの変数追加に関しては、将来予測で調整リスク・ウェイトを用いる場合(「ラグあり」「ラグなし」の 2 シナリオ)にしか実施しないので、米国アクチュアリー会 2007 年報告書で設定したシナリオは $2 \times 2 \times 2 + 2 = 10$ 通りである。これを一覧にまとめたものが、《図表 11》である。

6. Claim Truncation 給付金請求金額の頭切り

米国アクチュアリー会 2007 年報告書の分析では、Claim Truncation (給付金請求金額の頭切り)も考慮されている。給付金請求金額の頭切りとは、分析に際して、一定金額以上の高額給付金請求事案については、その定めた一定金額を上限とする取り扱いのことである。米国アクチュアリー会 2007 年報告書では、頭切りの取り扱いとして、頭切りを適用しない、

《図表 11》米国アクチュアリー会 2007 年報告書で設定したシナリオ一覧表

将来予測	Offered RW	Lagged		○
将来予測	Offered RW	Lagged	Prior Cost 付加	△
将来予測	Offered RW	Nonlagged		○
将来予測	Offered RW	Nonlagged	Prior Cost 付加	△
将来予測	調整 RW	Lagged		○
将来予測	調整 RW	Lagged	Prior Cost 付加	○
将来予測	調整 RW	Nonlagged		○
将来予測	調整 RW	Nonlagged	Prior Cost 付加	○
現状分析	Offered RW	Lagged		○
現状分析	Offered RW	Lagged	Prior Cost 付加	
現状分析	Offered RW	Nonlagged		○
現状分析	Offered RW	Nonlagged	Prior Cost 付加	
現状分析	調整 RW	Lagged		○
現状分析	調整 RW	Lagged	Prior Cost 付加	
現状分析	調整 RW	Nonlagged		○
現状分析	調整 RW	Nonlagged	Prior Cost 付加	

○：前掲注 4 の資料、p.9 に列挙されている 10 シナリオ

△：上記 10 シナリオ以外で、前掲注 4 の資料の APPENDIX (pp.40-63) に掲載されている 2 シナリオ

無印：分析対象とはされなかったシナリオ

(出典) 前掲注 4 の資料より、損保ジャパン総合研究所作成

\$ 100,000 で頭切りする、\$ 250,000 で頭切りする、の 3 種類を検討している。頭切りは、2004 年度の給付金総金額（現状分析の推計においては 2003 年度の給付金総金額）に対して適用した。

調整リスク・ウェイトの算出は、給付金請求金額頭切りの各水準（頭切りなし、\$250,000、\$100,000）それぞれについて行われた。将来予測／現状分析、ラグあり／ラグなし、過去の期間のコストを追加／非追加のそれぞれのシナリオにおいて、3 種類の頭切りの取り扱い方法のそれぞれに対する調整リスク・ウェイトが算出された。

高額給付金請求に対する頭切りは、次のような様々な理由により、予測モデルの予測精度を分析する際には一般的に使われる手法であると

米国アクチュアリー会 2007 年報告書では説明されている。

■ 頭切りを行うことにより、外れ値 (outliers) による影響を制限する。こうすることによって、モデルのリスク・ウェイト調整の結果や予測精度を分析する際の結果がより安定的に得られる。

■ 特定の個人に関する高額な給付金請求は一般的に予測不能である。従って、分析に際しては外れ値を除外するか、何らかの制限を加えるべきであると主張する研究者もいる。

■ 頭切りを実施することによって、再保険、すなわち、頭切りの水準をエクセスポイントとする stop loss 再保険の影響をシミュレートできる。

■ 予測精度に関するいくつかの尺度は、高額の

給付金請求に過度に反応する。

Ⅲ. 予測精度を評価する尺度

米国アクチュアリー会 2007 年報告書では、各予測モデルの予測精度を比較するために、個人レベルの 2 尺度および集団レベルの 1 尺度、計 3 つの尺度を使用している。これらの尺度はいずれも給付金請求金額の実績値と各予測モデルによる予測値を比較するものである。

1. 個人レベルの予測精度の尺度

米国アクチュアリー会 2007 年報告書で検討している予測精度に関する個人レベルの尺度は、R-squares (以下、「決定係数」とする) と mean absolute prediction error (予測誤差の絶対値の平均。以下、「MAPE」とする) の 2 つである。これらの尺度について簡単に説明する。

(1) 決定係数

決定係数の定義式は以下の通りである。

$$\text{決定係数} = 1 - \frac{\sum(\text{Actual} - \text{Predicted})^2}{\sum(\text{Actual} - \text{Average_of_Actual})^2}$$

ここで、

Actual : 実績値

Predicted : 各モデルによる予測値

Average_of_Actual : 実績値の平均値

右辺第 2 項の分母は、各実績値と、実績値の平均値との差の二乗をすべての標本について加えたものであるから、実績値がその平均値に対してどの程度ばらついているかを示している。一方、右辺第 2 項の分子は、各実績値と予測値との差の二乗をすべての標本について加えたものであるから、実績値が予測値に対してどの程度ばらつきが残っているかを示している。その意味で、決定係数は、実績値のばらつきのうち、

モデルによる予測で説明できる割合 (パーセンテージ) を表している。

上記の定義式は統計の教科書に出てくる一般的な表現であるが、米国アクチュアリー会 2007 年報告書では次の解説を付け加えている。すなわち、上記の定義式については、基本的な決定係数の算式から導出されていること、および、予測が最小二乗法に基づいている場合にこの導出が成り立つことに留意することが重要である。米国アクチュアリー会 2007 年報告書の場合、予測は、grouping algorithms や臨床的な有意性等に基づいているので、この導出は必ずしも成り立たない。したがって、ここで示されているものは、決定係数の統計的なエッセンスを持ってはいるが、厳密な意味での決定係数の計算にはなっていない。

決定係数はモデルによる予測結果を評価するための標準的な統計尺度として、一般的に使用されているものである。決定係数は、0 から 1 までの値をとるように標準化された尺度で、予測結果を集約して単一の尺度で表現している。0 は、モデルが各個人のコストのばらつきの 0%しか説明していないことを意味し、1 は、モデルが各個人のコストのばらつきを 100%説明する(予測の精度が 100%である)ことを意味する。0-1 で標準化されているので、異なる研究間での比較も可能である。しかしながら、ある研究における決定係数と他の研究の決定係数を比較する場合には、データセットの差、研究デザインの差、データの質の差等、その比較を適切または妥当でなくす可能性のある多くの潜在的な課題に留意する必要がある。

決定係数には若干の難点もある。予測誤差(各実績値と予測値との差)が二乗されるので、給付金請求金額の大きな個人に対する予測誤差に過度に反応する傾向がある。米国アクチュア

リー会の先行研究である1996年の報告書には、「決定係数は予測誤差を二乗しているため、相対的に少ない件数だが非常に大きな額の予測誤差をもつ標本点の影響を大きく受ける可能性がある。少数の個人が相対的に大きな支出額となっているという医療費支出の典型的な分布を前提とすると、この影響が懸念される」と述べられている。予測精度の尺度として決定係数を用いる場合には、大きな額の給付金請求については頭切りを行う理由のひとつがここにある。

(2) MAPE

MAPEの本来の定義式は以下の通りである。

$$MAPE = \frac{\sum |Actual - Predicted|}{Sample_Size}$$

すなわち、実績値と予測値との差額で算出した予測誤差の絶対値をすべての標本について加えたものを、標本の数で割ったものである。

MAPEも予測結果を集約して単一の尺度で表現するものである。MAPEの良い点は、MAPEでは予測誤差を二乗していないので、大きな金額の給付金請求に過剰に反応することがない点である。しかし、標準化された尺度とはなっていないので、異なる研究同士の比較は困難である。そこで、米国アクチュアリー会2007年報告書では、上記定義式で算出された値を、加入者1人当たりの年間のコストの平均値で除して得た値（パーセンテージ）を用いることとし、それをMAPEと表示している。

2. 集団レベルの予測精度の尺度

予測精度に関する集団レベルの尺度として、米国アクチュアリー会2007年報告書では、predictive ratio（以下、「PR」とする）が取り上げられている。PRの定義式は以下の通りである。

$$PR = \frac{\sum Predicted}{\sum Actual}$$

分子は、集団の各個々人の給付金請求の予測値を足し上げたもので、分母は、同集団の実績値を足し上げたものである。PRの値が1.0に近いほど、予測の当てはまりがよいことを意味する。

IV. 分析結果の概要

本章では、分析対象としたデータの特徴について説明した後、各シナリオに基づいた分析結果について説明する。

1. 分析対象としたデータに関する記述

前述のように、米国アクチュアリー会2007年報告書では、MedStat Marketscanのデータ、約62万人分を用いて分析を実施した。《図表12》（米国アクチュアリー会2007年報告書におけるTable III.1）は、分析対象とした集団におけるデモグラフィック（年齢・性別等）の分布と典型的な被保険者集団（表上「Reference」（参照集団）と表示）における分布を示している。米国アクチュアリー会2007年報告書では参照集団としてthe Milliman Health Cost Guidelines, 2006 editionにおける被保険者集団を取り上げたことが明記されている。《図表12》をみれば明らかなように、分析対象集団においては高齢者（50歳以上）の割合が参照集団に比べて大きい。加えて、子供の割合は、分析対象集団の方が小さい。この属性の違いは、米国アクチュアリー会2007年報告書においては、他の分布の集団の場合に比べ、慢性疾患の予測可能性をより大きく強調し易いことを示唆している。このことは、後に示す誤差計算においても見られるだろうし、こうした分析対象とする集団のデモグラフィックの差異が、各ソフトウェア・

《図表 12》分析対象集団の年齢・性別等の分布

TABLE III.1 Demographic Characteristics of Study Population Compared to Reference Population				
Demographic Category	% of Total		% of Category	
	Study	Reference	Study	Reference
Male, To 25	0%	2%	1%	7%
Male, 25-29	1%	3%	4%	11%
Male, 30-34	2%	4%	5%	13%
Male, 35-39	2%	5%	6%	15%
Male, 40-44	3%	5%	10%	16%
Male, 45-49	5%	5%	15%	15%
Male, 50-54	7%	4%	20%	13%
Male, 55-59	8%	2%	24%	7%
Male, 60-64	5%	1%	15%	4%
Demographic Category	% of Total		% of Category	
	Study	Reference	Study	Reference
Female, To 25	0%	2%	1%	6%
Female, 25-29	1%	3%	4%	10%
Female, 30-34	2%	4%	5%	13%
Female, 35-39	2%	5%	6%	14%
Female, 40-44	4%	5%	10%	16%
Female, 45-49	6%	5%	15%	15%
Female, 50-54	8%	4%	20%	13%
Female, 55-59	10%	3%	24%	8%
Female, 60-64	7%	2%	16%	5%
Demographic Category	% of Total		% of Category	
	Study	Reference	Study	Reference
Child, 00-01	1%	3%	5%	7%
Child, 02-06	4%	7%	14%	20%
Child, 07-18	16%	21%	61%	59%
Child, 19-22	5%	5%	21%	13%

(出典) 前掲注 4 の資料、p.11 の Table III.1 を引用。

パッケージの予測能力の差を強調することになることが考えられる。

分析対象集団の地理的な分布は《図表 13》(米国アクチュアリー会 2007 年報告書における Table III.2) の通りである。また、疾病分類による分布は《図表 14》(米国アクチュアリー会 2007 年報告書における Table III.3) の通りである。疾病分類の各群に属する人の数は、各人についてその診断がいつ下されたかによって変動するが、《図表 14》は、「ラグなし」の将来予測の分析において使用した 2003 年度の疾病分類ごとの人数である。

《図表 13》分析対象集団の地理的な分布

TABLE III.2 Geographical Characteristics	
Region	Members
Northeast	43,330
North Central	392,743
South	128,436
West	52,301
Unknown	873
Total	617,683

(出典) 前掲注 4 の資料、p.12 の Table III.2 を引用。

《図表 14》分析対象集団の疾病分類ごとの分布

Condition Category	Unique Members
Asthma	6,806
Breast Cancer	2,299
Diabetes	19,690
Heart Disease	19,270
HIV	170
Mental Illness	22,421
Total	70,656

(出典) 前掲注 4 の資料、p.12 の Table III.3 を引用。

なお、疾病で分類する際に使用された各診断名の定義は《図表 15》(米国アクチュアリー会 2007 年報告書における Table III.4) の通りである。疾病による分類は、この診断名のコードおよび偽陽性に関する調整(入院に関しては当該コードが 1 回記録されていれば十分とし、外来については 2 回以上記録されていることを要する)に基づいて決定したと説明されている。

2. 個人レベルの予測の結果

(1) 「ラグなし」将来予測の結果

《図表 16》(米国アクチュアリー会 2007 年報告書における Table IV.1) は、offered リスク・ウェイトを用いた各モデルの決定係数の結果および最適化したリスク・ウェイトを用いた各モデルの決定係数の結果を示している。Offered リスク・ウェイトとは、前述のように、各モデルのソフトウェア付属のリスク・ウェイトで、この分析対象集団およびデータに合わせてカスタマイズは行っていないものをいい、また、最適化したリスク・ウェイトとは、この分析対象集団およびデータに合わせてリスク・ウェイトの調整を行い、併せて、過去の期間のコストを予測のための変数として加えることも実施したものである。決定係数が大きな値であるほど、よく当てはまったモデルであることを意味する。

MEDai のプロセスが、決定係数(および後記の MAPE でも)で見た場合、最も当てはまりがよかった。ただし、MEDai については個別事情のあることが明記されている。MEDai 以外のモデルについては、リスク・ウェイトの調整等の作業を、すべて米国アクチュアリー会

《図表 15》疾病による分類の定義

Condition	ICD-9
Breast Cancer	174-174.9
Heart Disease	390-398, 402, 404-429
Asthma	493-493.9
Mental Illness	290-298.9, 300-312.9
HIV	042
Diabetes	250.1, 250.10, 250.11, 250.12, 250.13, 648.0, 648.00, 648.01, 648.02, 648.03, 648.04, 648.8, 648.80, 648.81, 648.82, 648.83, 648.84, 250.0, 250.00, 250.01, 250.02, 250.03, 250.2, 250.20, 250.21, 250.22, 250.23, 250.3, 250.30, 250.31, 250.32, 250.33, 250.40, 250.41, 250.42, 250.43, 250.5, 250.50, 250.51, 250.52, 250.53, 250.60, 250.61, 250.62, 250.63, 250.70, 250.71, 250.72, 250.73, 250.8, 250.80, 250.81, 250.82, 250.83, 250.9, 250.90, 250.91, 250.92, 250.93, 362.0, 362.01, 362.02, 362.1, 775.1, 790.2, 790.21, 790.22, 790.29, 253.5

(出典) 前掲注 4 の資料、p.16 の Table III.4 を引用。

《図表 16》「ラグなし」将来予測の決定係数

TABLE IV.1		R-Squared for Prospective Nonlagged (Offered vs. Optimized) by Truncation Level (Offered Compared to Recalibrated, with Prior Costs)					
Risk Adjuster Tool	Inputs	Offered Models			Optimized Models (Include Prior Costs)		
		100K	250K	None	100K	250K	None
ACG	Diag	20.8%	19.2%	16.2%	24.2%	23.0%	20.2%
CDPS	Diag	17.6%	14.9%	12.4%	27.4%	24.6%	21.2%
Clinical Risk Groups	Diag	19.3%	17.5%	14.9%	21.5%	20.5%	18.4%
DxCG DCG	Diag	22.3%	20.6%	17.4%	29.7%	26.5%	22.9%
DxCG RxGroups	Rx	23.8%	20.4%	16.8%	30.6%	27.1%	23.4%
Ingenix PRG	Rx	25.0%	20.5%	17.2%	30.9%	27.4%	23.7%
MedicaidRx	Rx	19.3%	15.8%	12.9%	29.7%	26.3%	22.7%
Impact Pro	Med+Rx+Use	26.3%	24.4%	21.3%	29.3%	27.2%	24.0%
Ingenix ERG	Med+Rx	23.7%	19.7%	16.2%	30.0%	26.5%	22.8%
ACG w/ Prior Cost	Diag+\$Rx	25.6%	22.4%	18.7%	27.7%	25.4%	22.1%
DxCG UW Model	Diag+\$Total	31.3%	27.4%	23.6%	33.1%	29.1%	25.2%
Service Vendor	Inputs	100K	100K	250K	None	100K	250K
MEDai*	All	N/A	N/A	N/A	35.7%	32.1%	27.6%

* The offered MEDai model was not tested in the study.

(出典) 前掲注 4 の資料、p.17 の Table IV.1 を引用。

2007 年報告書の研究者サイドで行ったのに対して、MEDai は、研究者の手でテスト可能な製品を提供せず、逆に、調整リスク・ウェイト算出用のデータおよびテストのためのインプット情報を MEDai に対して提供し、リスク・ウェイトの調整は MEDai が行った。MEDai 以外のモデルについても分析対象集団およびデータに合わせてリスク・ウェイトを自ら調整する機会が各社に与えられていたら、その他のモデルも、もっとよいパフォーマンスを示したかもしれない。このため、米国アクチュアリー会 2007 年報告書においては、MEDai を、ソフトウェア・ベンダーではなく、サービス・ベンダーと性格付け、他のベンダーとの公平性を保つため、分けて表示されている。MEDai は《図表 16》の最適化モデル以外のモデルも提供したが、ベンダーとしての関与度が異なり、公平な比較にならない可能性があるため、その他の比較表において MEDai は比較対象に含めず、「N/A」と表

示したと解説されている。

概して言えば、処方薬剤のみに基づいたモデルは、offered リスク・ウェイトを用いたモデルでも、調整リスク・ウェイトを用いたモデルでも、パフォーマンスがよかった。MedicaidRx モデルは、offered リスク・ウェイトを用いた場合、決定係数は相対的に低いですが、これは、このモデルがメディケイドの被保険者集団向けを意図して開発されたものであり、米国アクチュアリー会 2007 年報告書では民間健康保険プランの被保険者集団を分析対象として用いたことを考え合わせれば、予想されたことである。調整リスク・ウェイトを用いたモデルは、offered リスク・ウェイトを用いた場合に比べ、パフォーマンスが大幅に改善したことが示されているが、それは、主として、過去の期間のコストを独立変数として追加したことによると解説されている。その理由として、もともと過去の期間のコストを組み込んでいたモデルの方が、調整リス

ク・ウェイトを用いることによる改善幅が小さいことが挙げられている。また、《図表 16》における「100K」「250K」「None」はそれぞれ、「\$ 100,000 で頭切りした場合」「\$ 250,000 で頭切りした場合」「頭切りを適用しない場合」の結果を示しているが、予想通り、実績金額の頭切りを適用した場合には、決定係数は大幅に改善する。改善の度合いはモデルによって異なり、いくつかのモデルにおいては他のモデルより改善幅が大きかった。

《図表 17》(米国アクチュアリー会 2007 年報告書における Table IV.2) は、決定係数の結果に代えて、MAPE の結果が表示されていることを除けば、《図表 16》と同様である。決定係数とは異なり、MAPE は値が小さいほど望ましい。調整リスク・ウェイトを用いたモデルでは、MEDai と DxCG underwriting models の MAPE は最小であり(すなわち、パフォーマンスがよい)、一方、offered リスク・ウェイトを用いた場合では、CRGs モデルおよび CDPS モ

デルの MAPE が最大であった。

前述のように、決定係数に比べると、MAPE では、外れ値に対する予測誤差の影響が小さくなるという特徴がある。米国アクチュアリー会 2007 年報告書では、MAPE による結果が有用な場合として、小規模集団に対する契約更改の保険引受条件判断を挙げている。米国では、健康状態に応じて保険料率を平均値からどの程度乖離させてよいかの幅を大きく制限している州がいくつもある。こうした州政府の規制により保険料率設定の自由度が制限されているため、外れ値の重要性がより小さい状況にあり、小規模集団における外れ値を予測する精度を上げても余り役には立たないことが考えられるとしている。

(2) 米国アクチュアリー会の先行研究(2002年)との結果の比較

《図表 18》(米国アクチュアリー会 2007 年報告書における Table IV.3) は、米国アクチュ

《図表 17》「ラグなし」将来予測の MAPE

TABLE IV.2		MAPE for Prospective Nonlagged (Offered vs. Optimized) by Truncation Level (Offered Compared to Recalibrated, with Prior Costs)					
		Offered Models			Optimized Models (Include Prior Costs)		
Risk Adjuster Tool	Inputs	100K	250K	None	100K	250K	None
ACG	Diag	87.7%	89.9%	90.4%	84.6%	86.2%	86.6%
CDPS	Diag	93.4%	95.3%	95.8%	83.7%	85.6%	86.3%
Clinical Risk Groups	Diag	88.7%	90.9%	91.4%	85.2%	86.6%	87.0%
DxCG DCG	Diag	85.3%	87.5%	88.0%	80.5%	82.5%	83.2%
DxCG RxGroups	Rx	82.9%	85.3%	85.9%	78.7%	80.7%	81.4%
Ingenix PRG	Rx	83.4%	85.8%	86.4%	78.9%	80.9%	81.5%
MedicaidRx	Rx	87.3%	89.6%	90.2%	79.9%	81.9%	82.6%
Impact Pro	Med+Rx+Use	79.3%	81.8%	82.4%	78.7%	80.6%	81.2%
Ingenix ERG	Med+Rx	84.1%	86.4%	87.0%	79.1%	81.2%	81.9%
ACG w/ Prior Cost	Diag+\$Rx	85.1%	85.6%	85.6%	80.3%	82.1%	82.6%
DxCG UW Model	Diag+\$Total	80.1%	80.4%	80.4%	76.1%	78.3%	78.9%
Service Vendor	Inputs	100K	100K	250K	None	100K	250K
MEDai*	All	N/A	N/A	N/A	73.0%	75.2%	75.6%

* The offered MEDai model was not tested in the study.

(出典) 前掲注 4 の資料、p.18 の Table IV.2 を引用。

《図表 18》2002 年研究との比較（「ラグなし」将来予測、offered リスク・ウェイトの決定係数）

TABLE IV.3	Comparison to 2002 Study of Offered Weight R-Squared Prospective Nonlagged by Claims Truncation Level					
	2002 Study			Current Study		
Risk Adjuster Tool	50K	100K	None	100K	250K	None
ACG	N/A	N/A	N/A	20.8%	19.2%	16.2%
CDPS	13.4%	12.5%	10.3%	17.6%	14.9%	12.4%
DCG	19.5%	18.0%	14.3%	22.3%	20.6%	17.4%
MedicaidRx	11.6%	9.8%	7.1%	19.3%	15.8%	12.9%
RxGroups	20.6%	18.1%	13.4%	23.8%	20.4%	16.8%
RxRisk	17.5%	14.8%	11.1%	N/A	N/A	N/A
ERG	21.8%	19.3%	14.6%	23.7%	19.7%	16.2%

（出典）前掲注 4 の資料、p.19 の Table IV.3 を引用。

アリー会 2007 年報告書における決定係数の結果と 2002 年度研究における決定係数の結果を比較したものである。

概して言えば、各モデルとも先行研究の時点よりはパフォーマンスが向上している。これは、モデル自身の改善およびデータのコーディングの改善によるものと思われる。米国アクチュアリー会 2007 年報告書では解説している。厳密に言えば、頭切りの水準は 2 つの研究で異なっているが、全般的な医療費コストのトレンドおよび標本を均質で比較の容易な comprehensive benefit-type の保険加入者に限定したこと等から、相対的には比較可能であると米国アクチュアリー会 2007 年報告書では述べている。

なお、2002 年研究では RxRisk と呼ばれる処方薬剤に基づく予測モデルも分析対象に含まれていたが、米国アクチュアリー会 2007 年報告書では分析対象とされなかった。理由として、米国アクチュアリー会 2007 年報告書では、最近アップデートが実行され、かつ、広く使われている予測モデルに限られた資源を集中するこ

ととしたが、RxRisk について入手したコピーは、2002 年 3 月以降アップデートされていないものだったことを挙げている。

（3）「ラグなし」将来予測、offered リスク・ウェイト（過去の期間のコストなし）の結果

《図表 19》（米国アクチュアリー会 2007 年報告書における Table IV.4）は、もともと過去の期間のコストを組み込んでいないモデルについて、「ラグなし」将来予測シナリオで、offered リスク・ウェイトを用いた場合の決定係数と MAPE の結果が示されている。数値的には、《図表 16》と《図表 17》の Offered Models 欄の再掲であるが、ACG w/ Prior Cost および DxCG UW Model は、もともと過去の期間のコストを組み込んでいないモデルなので、《図表 19》上は「N/A」と表示されている。

《図表 19》によれば、Impact Pro が、MAPE および決定係数の両方で最もよいパフォーマンスを示した。Ingenix PRG も、100K で頭切りした場合の決定係数でよいパフォーマンスを示した。

《図表 19》「ラグなし」将来予測、offered リスク・ウェイト、
（過去の期間のコストなし）の決定係数および MAPE

TABLE IV.4		R-Squared and MAPE Prospective Nonlagged Offered (Without Prior Cost) by Claims Truncation Level					
Risk Adjuster Tool	Inputs	R-Squared			MAPE%		
		100K	250K	None	100K	250K	None
ACG	Diag	20.8%	19.2%	16.2%	87.7%	89.9%	90.4%
CDPS	Diag	17.6%	14.9%	12.4%	93.4%	95.3%	95.8%
Clinical Risk Groups	Diag	19.3%	17.5%	14.9%	88.7%	90.9%	91.4%
DxCG DCG	Diag	22.3%	20.6%	17.4%	85.3%	87.5%	88.0%
DxCG RxGroups	Rx	23.8%	20.4%	16.8%	82.9%	85.3%	85.9%
Ingenix PRG	Rx	25.0%	20.5%	17.2%	83.4%	85.8%	86.4%
MedicaidRx	Rx	19.3%	15.8%	12.9%	87.3%	89.6%	90.2%
Impact Pro	Med+Rx+Use	26.3%	24.4%	21.3%	79.3%	81.8%	82.4%
Ingenix ERG	Med+Rx	23.7%	19.7%	16.2%	84.1%	86.4%	87.0%
ACG w/ Prior Cost*	Diag+\$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model*	Diag+\$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Inputs	100K	250K	None	100K	250K	None
MEDai*	All	N/A	N/A	N/A	N/A	N/A	N/A

* These models include prior cost as input.

(出典) 前掲注 4 の資料、p.19 の Table IV.4 を引用。

Impact Pro モデルの MAPE の結果は、《図表 17》で示されているように、調整リスク・ウェイトを用いた場合（リスク・ウェイトの調整が実施され、過去の期間のコストも組み込まれた場合）でも offered リスク・ウェイトを使用した場合に比べ改善は左程大きくない。Impact Pro は保険引受条件判断に使うことを意図されたものであるから、この結果は少々驚きであるが、これは、Impact Pro が、直接的に過去の期間のコストを組み込んでいるわけではないが、過去の期間の医療資源消費（prior use）を変数として組み込んでいるからと思われると米国アクチュアリー会 2007 年報告書の研究者は述べている。また、MedicaidRx および CDPS はメディケイドの集団向けに開発されたもので、民間健康保険プランの加入者集団を対象に使用することは意図されていなかったため、offered リスク・ウェイトを用いた予測の尺度はこの点が反映されていると解説している。

（４） offered リスク・ウェイトを用いたモデルと調整リスク・ウェイトを用いたモデルの比較

《図表 20》（米国アクチュアリー会 2007 年報告書における Table IV.5）は、もともと過去の期間のコストを組み込んでいないモデルについて、「ラグなし」将来予測シナリオ、\$ 250,000 で頭切り、offered リスク・ウェイトを用いた場合と調整リスク・ウェイトを用いた（ただし、過去の期間のコストの変数追加は行わない）場合の決定係数と MAPE の変化を示している。

リスク・ウェイトの調整の結果、最も大きな改善を示したのは CDPS および MedicaidRx である。先に言及したように、これらのモデルは、もともとはメディケイドの集団向けに開発されたものである。したがって、offered リスク・ウェイトもメディケイド集団のデータを基に算出されたものである。今回の分析対象の集団は民間健康保険プランの加入者集団であるので、この対象集団に合わせてリスク・ウェイトを調整す

《図表 20》「ラグなし」将来予測シナリオ、\$250,000 で頭切り、
 offered リスク・ウェイトを用いた場合と調整リスク・ウェイトを用いた
 (ただし、過去の期間のコストの変数追加は行わない) 場合の決定係数と MAPE の変化

TABLE IV.5		R-Squared and MAPE Prospective Nonlagged Offered vs. Recalibrated (Without Prior Cost, 250K Truncation)					
Risk Adjuster Tool	Inputs	R-Squared			MAPE%		
		Offered	Recalibrated	Change	Offered	Recalibrated	Change
ACG	Diag	19.2%	19.6%	0.4%	89.9%	88.8%	-1.1%
CDPS	Diag	14.9%	17.7%	2.8%	95.3%	91.9%	-3.4%
Clinical Risk Groups*	Diag	17.5%	N/A	N/A	90.9%	N/A	N/A
DxCG DCG	Diag	20.6%	21.3%	0.7%	87.5%	87.0%	-0.5%
DxCG RxGroups	Rx	20.4%	20.5%	0.1%	85.3%	85.3%	0.0%
Ingenix PRG	Rx	20.5%	21.2%	0.7%	85.8%	85.6%	-0.2%
MedicaidRx	Rx	15.8%	17.7%	1.9%	89.6%	88.4%	-1.2%
Impact Pro	Med+Rx+Use	24.4%	25.6%	1.2%	81.8%	81.6%	-0.2%
Ingenix ERG	Med+Rx	19.7%	20.0%	0.3%	86.4%	86.1%	-0.3%
ACG w/ Prior Cost**	Diag+\$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+\$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Inputs	Offered	Recalibrated	Change	Offered	Recalibrated	Change
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input.

(出典) 前掲注 4 の資料、p.20 の Table IV.5 を引用。

ることによって予測精度の改善が示されたと考えることができる。また逆に、いくつかのモデルにおいては、リスク・ウェイト調整による改善は比較的小さなものであった。これらのモデルは、とても堅牢に設計されたモデルであるか、または、米国アクチュアリー会 2007 年報告書で分析対象としたものと類似のデータセットで offered リスク・ウェイトが算出されているモデルのどちらかであったと考えられると米国アクチュアリー会 2007 年報告書では解説されている。

米国アクチュアリー会 2007 年報告書におけるリスク・ウェイトの調整は前記の通り、極めて単純な方法で行われた。そのアプローチは先行研究で使用されたものとは少し異なる。米国アクチュアリー会 2007 年報告書では、offered リスク・ウェイトと同じ構成のものを一から作

り直すのではなく、offered リスク・ウェイトの値に対する調整係数を算出するアプローチを取った。このアプローチは、先行研究で採ったアプローチより計算手順としては単純である。いくつかのモデルでは、offered リスク・ウェイトと同じ構成のものを一から作り直すのは容易ではない。また、米国アクチュアリー会 2007 年報告書のアプローチでは、統計的検定の p 値に基く信頼係数を用いて調整係数を算出している点が特徴である。

(5) 「ラグあり」と「ラグなし」との比較

前項までの比較は、「ラグなし」のシナリオにおける比較であったが、《図表 21》(米国アクチュアリー会 2007 年報告書における Table IV.6) は、「ラグあり」のシナリオとした場合に、結果がどの程度変わるかを示している。

《図表 21》 将来予測シナリオ、\$250,000 で頭切り、offered リスク・ウェイト
「ラグあり」と「ラグなし」の決定係数と MAPE の変化

TABLE IV.6		R-Squared and MAPE Offered Prospective Lagged vs. Nonlagged (Without Prior Cost) 250K Truncation					
Risk Adjuster Tool	Inputs	R-Squared			MAPE%		
		Lagged	Nonlagged	Change	Lagged	Nonlagged	Change
ACG	Diag	14.5%	19.2%	4.7%	93.7%	89.9%	-3.8%
CDPS	Diag	11.9%	14.9%	3.0%	98.8%	95.3%	-3.5%
Clinical Risk Groups	Diag	14.1%	17.5%	3.4%	93.9%	90.9%	-3.0%
DxCG DCG	Diag	15.1%	20.6%	5.5%	91.6%	87.5%	-4.1%
DxCG RxGroups	Rx	18.0%	20.4%	2.4%	87.4%	85.3%	-2.1%
Ingenix PRG	Rx	18.0%	20.5%	2.5%	87.8%	85.8%	-2.0%
MedicaidRx	Rx	13.6%	15.8%	2.2%	91.7%	89.6%	-2.1%
Impact Pro	Med+Rx+Use	21.4%	24.4%	3.0%	85.5%	81.8%	-3.7%
Ingenix ERG	Med+Rx	16.9%	19.7%	2.8%	88.7%	86.4%	-2.3%
ACG w/ Prior Cost*	Diag+\$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model*	Diag+\$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Inputs	Lagged	Nonlagged	Change	Lagged	Nonlagged	Change
MEDai*	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model includes prior cost as input

(出典) 前掲注 4 の資料、p.21 の Table IV.6 を引用。

《図表 21》に示されるように、完全に「ラグなし」の場合のパフォーマンスの改善は顕著である。「ラグあり」では、処方薬剤に基づいたモデルの方が診断名のみに基づいたモデルよりパフォーマンスがよい。DxCG DCG および ACG が「ラグあり」の影響を最も大きく受けている。また、いくつかのベンダーでは、彼らの提供する製品パッケージソフトとして、ラグを考慮したモデルを提供している（例えば、DxCG underwriting models および Impact Pro）。

米国アクチュアリー会 2007 年報告書の Appendix には、調整リスク・ウェイトを用いた場合のラグありの結果も掲載されている。

(6) 現状分析シナリオと将来予測シナリオとの比較

《図表 22》（米国アクチュアリー会 2007 年報告書における Table IV.7）では、offered リ

スク・ウェイトを用いた現状分析シナリオの結果が示されている。

DCG モデルが、決定係数においても MAPE においても最もよいパフォーマンスを示した。現状分析シナリオでは、過去の期間（この場合は、当該期間）のコストをインプット変数として組み込むことは適切でないと米国アクチュアリー会 2007 年報告書の研究者は考えたので、過去の期間のコストを変数として使用しているモデルについては、《図表 22》上、“N/A”と表示されている。また、モデルの設計上、現状分析シナリオのリスク・スコアをアウトプットしないモデルについても同様に“N/A”と表示されている。

《図表 23》（米国アクチュアリー会 2007 年報告書における Table IV.8）では、将来予測シナリオと現状分析シナリオにおける決定係数と

《図表 22》「ラグなし」現状分析、offered リスク・ウェイトの決定係数と MAPE

TABLE IV.7		R-Squared and MAPE Offered Concurrent Nonlagged by Claims Truncation Level					
Risk Adjuster Tool	Inputs	R-Squared			MAPE%		
		100K	250K	None	100K	250K	None
ACG	Diag	29.4%	29.7%	27.4%	73.0%	75.0%	75.4%
CDPS	Diag	35.5%	32.9%	31.0%	79.0%	80.6%	81.0%
Clinical Risk Groups	Diag	47.1%	43.3%	39.9%	68.6%	70.5%	70.9%
DxCG DCG	Diag	57.2%	51.8%	49.8%	61.6%	65.0%	65.4%
DxCG RxGroups*	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG*	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	32.1%	28.1%	24.6%	77.2%	79.1%	79.6%
Impact Pro*	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	46.5%	42.4%	38.6%	65.8%	67.7%	68.2%
ACG w/ Prior Cost**	Diag+\$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+\$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Inputs	100K	250K	None	100K	250K	None
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* These models do not include a concurrent option.

** These models include prior cost as input.

(出典) 前掲注 4 の資料、p.21 の Table IV.7 を引用。

《図表 23》「ラグなし」offered リスク・ウェイト、\$250,000 頭切り、将来予測シナリオと現状分析シナリオにおける決定係数と MAPE の比較

TABLE IV.8		R-Squared Offered Nonlagged (Without Prior Cost & 250K truncation) – Prospective vs. Concurrent					
Risk Adjuster Tool	Inputs	R-Squared			MAPE%		
		Prospective	Concurrent	Change	Prospective	Concurrent	Change
ACG	Diag	19.2%	29.7%	10.5%	89.9%	75.0%	-14.9%
CDPS	Diag	14.9%	32.9%	18.0%	95.3%	80.6%	-14.7%
Clinical Risk Groups	Diag	17.5%	43.3%	25.8%	90.9%	70.5%	-20.4%
DxCG DCG	Diag	20.6%	51.8%	31.2%	87.5%	65.0%	-22.5%
DxCG RxGroups*	Rx	20.4%	N/A	N/A	85.3%	N/A	N/A
Ingenix PRG*	Rx	20.5%	N/A	N/A	85.8%	N/A	N/A
MedicaidRx	Rx	15.8%	28.1%	12.3%	89.6%	79.1%	-10.5%
Impact Pro*	Med+Rx+Use	24.4%	N/A	N/A	81.8%	N/A	N/A
Ingenix ERG	Med+Rx	19.7%	42.4%	22.7%	86.4%	67.7%	-18.7%
ACG w/ Prior Cost**	Diag+\$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+\$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Inputs	Prospective	Concurrent	Change	Prospective	Concurrent	Change
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* These models do not include a concurrent option.

** These models include prior cost as input.

(出典) 前掲注 4 の資料、p.22 の Table IV.8 を引用。

MAPE の比較結果を示している。

現状分析シナリオにおけるパフォーマンスは、モデル作成に使われたデータのレベル（診断名のみか、処方薬剤のみか、両方か）に関係するようだと米国アクチュアリー会 2007 年報告書ではコメントしている。また、将来予測シナリオのパフォーマンスもデータのレベルの影響を受けるが、その影響は現状分析シナリオの方が大きい。医学的な診断名と服用している薬剤の両方の情報に基づいて現期間の支出総額（医療費＋薬剤費）を予測する方が、診断名あるいは処方薬剤どちらか一方のタイプの情報のみに基づいて医療費＋薬剤費の合計を予測するより容易と考えられるから、この結果は直観的に理解できるものである。将来予測シナリオではどのモデルも正確性が低いので、すべてのタイプのデータを用いることの有用性は小さい。

（7）過去の期間のコストをリスク・ウェイトの調整時に組み込むことのインパクト

過去の期間のコストを予測のための独立変数として付け加えることで、大半のモデルの精度は格段に改善する。特に、もともと過去の期間のコストを変数として組み込んでいなかったモデルは、その改善が顕著である。健康保険プランが予測モデルを契約更改時の保険引受条件判断に使用する場合は、職域集団に対する更改契約の保険料率決定のために、個々人に対するリスク調整の予測の集計値とともに、当該集団の過去の期間のコストを使用するのが一般的である。過去の研究で、集団の被保険者数が多いほど、信頼度すなわち過去の期間のコストに割り振られるウェイトが大きくなることが示唆されている。したがって、もしそのリスク調整のためのソフトウェアが個々人に対する予測において過去の期間のコストの尺度を組み込んでい

ならば、過去の期間のコストの集計値に適用されるべきウェイトにどのように影響を与えているかを考慮することは重要である。職域集団を対象とした異なるモデルの精度をモデル化することは本研究の範囲外（ただし、推奨される将来研究の範囲としてリストアップされている）である。一般的に言って、集団の被保険者数規模が大きいほど、モデル間の精度に関する相対的な差異は小さくなると考えられる。

《図表 24》（米国アクチュアリー会 2007 年報告書における Table IV.9）は、もともとは過去の期間のコストを組み込んでいなかったモデルに過去の期間のコストを組み込んでリスク・ウェイトの調整を行った場合の影響を示している。

《図表 24》の R-Squared, with Prior の欄は、《図表 16》の Optimized Model (Include Prior Costs), 250K の欄と同じものである。また、《図表 24》の MAPE%, with Prior の欄は、《図表 17》の Optimized Model (Include Prior Costs), 250K の欄と同じものである。先に述べたように、MEDai は、他のモデルに比して圧倒的にパフォーマンスが優れている。しかし、《図表 24》により、処方薬剤に基づいたモデルは過去の期間のコストを組み込むことにより大きなメリットを受けることが見て取れる。事実、過去の期間のコストを組み込むことにより、MedicaidRx モデルは、決定係数においては商用販売されている 3 つのモデルより、MAPE においては商用販売されている 4 つのモデルより優れたパフォーマンスを示している。ここで商用販売されているとは、使用にあたってライセンスフィーを支払う必要のあるものという意味である。

《図表 24》「ラグなし」将来予測、調整リスク・ウェイト、
\$250,000 頭切り、過去の期間のコストを組み込んだ場合の影響

TABLE IV.9		R-Squared and MAPE Prospective Recalibrated Nonlagged (Without Prior Cost vs. With Prior Cost) 250K truncation					
Risk Adjuster Tool	Inputs	R-Squared			MAPE%		
		w/out Prior	with Prior	Change	w/out Prior	with Prior	Change
ACG	Diag	19.6%	23.0%	3.4%	88.8%	86.2%	-2.6%
CDPS	Diag	17.7%	24.6%	6.9%	91.9%	85.6%	-6.3%
Clinical Risk Groups*	Diag	N/A	20.5%	N/A	N/A	86.6%	N/A
DxCG DCG	Diag	21.3%	26.5%	5.2%	87.0%	82.5%	-4.5%
DxCG RxGroups	Rx	20.5%	27.1%	6.6%	85.3%	80.7%	-4.6%
Ingenix PRG	Rx	21.2%	27.4%	6.2%	85.6%	80.9%	-4.7%
MedicaidRx	Rx	17.7%	26.3%	8.6%	88.4%	81.9%	-6.5%
Impact Pro	Med+Rx+Use	25.6%	27.2%	1.6%	81.6%	80.6%	-1.0%
Ingenix ERG	Med+Rx	20.0%	26.5%	6.5%	86.1%	81.2%	-4.9%
ACG w/ Prior Cost**	Diag+\$Rx	N/A	25.4%	N/A	N/A	82.1%	N/A
DxCG UW Model**	Diag+\$Total	N/A	29.1%	N/A	N/A	78.3%	N/A
Service Vendor	Inputs	w/out Prior	with Prior	Change	w/out Prior	with Prior	Change
MEDai**	All	N/A	32.1%	N/A	N/A	75.2%	N/A

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input.

(出典) 前掲注 4 の資料、p.22 の Table IV.9 を引用。

3. 医的な条件でグループ化された集団に対する結果

米国アクチュアリー会 2007 年報告書では、医的な条件により疾病でグループ化された集団に対する結果を PR¹⁸ で示している。PR が 100%に近いほど当てはまりがよいことを意味する。以下に掲げる各図表に示されているように、PR は通常 100%より小さな値となる¹⁹。これは、多少なりとも予想されることである。なぜなら、予測モデルは、通常、高額なコストとなる個人に対するコストを過小に予測する

からである。この傾向を理解しておくことは、実務適用に際して重要である。

医的な条件による疾病でのグループ化を 2003 年または 2004 年のいずれの年度における所定の状態の有無により行ったかは、シナリオによって異なる。将来予測シナリオでは、各個人 2003 年における所定の状態の有無により分類した場合と、2004 年における所定の状態の有無により分類した場合の両方について、2004 年の PR (2004 年の予測値と 2004 年の実績値の割合) が示されている。また、現状分析シナ

¹⁸ PR : predictive ratio、特定の集団に対する予測コストの平均値と実績コストの平均値の比率。第三章第 2 節参照。

¹⁹ 米国アクチュアリー会 2007 年報告書ではこの点に関し、William Gilmore, ASA, MAAA, of Blue Cross Blue Shield of Mississippi が提起した興味深い疑問を脚注で引用している。Gilmore 氏の主張は次の通りである。彼が予測モデルを実際に使用したところ、加入者個人に対する予測の平均値と実績コストの平均値は非常に近い値となった。しかしながら、これを男女別に見ると、それぞれの性別の予測の平均値は実績コストの平均値と等しくはならなかった。差異は相対的には小さなものであったが、それなりの大きさであった。原因は何なのか。状態カテゴリーのウェイトは通常、デモグラフィック・カテゴリー (性別または年齢) 毎には算出されず、集団全体に対して最適化されているので、この結果は理に適っている。信頼性と資源の節約を理由にこの方式が採られているのである。デモグラフィック毎ではない調整方法に起因する予測尺度の変化を検定する機会は与えられていない。状態カテゴリーのウェイトをデモグラフィック・カテゴリー毎に算出することによって予測尺度が少しは改善することは期待できる。そうすることによって得られる結果で、おそらくもっと重要なことは、その結果は年齢・性別カテゴリーに渡って妥当なもので、組織内の他の人に説明するときに役に立つということであろう。

リオでは、各個々人の 2003 年における所定の状態の有無により分類し、2003 年の PR (2003 年の値 [現状なので、厳密に言えば予測ではない] と 2003 年の実績値の割合) が示されている。

(1) 将来予測シナリオ

《図表 25》(米国アクチュアリー会 2007 年報告書における Table V.1) は、各個々人の 2003 年における所定の状態の有無により分類した場合の 2004 年の PR を示したものであり、《図表 26》(米国アクチュアリー会 2007 年報

《図表 25》将来予測の PR (2003 年における医的状态で分類した場合)

TABLE V.1		Predictive Ratios by Medical Condition in 2003 (Recalibrated Nonlagged Prospective without Prior Costs, 250K Truncation)					
Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	98.3%	90.9%	96.2%	100.8%	99.1%	98.0%
CDPS	Diag	97.1%	81.3%	97.7%	93.5%	94.9%	91.1%
Clinical Risk Groups*	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	93.5%	91.1%	97.5%	96.0%	92.5%	98.5%
DxCG RxGroups	Rx	95.2%	72.4%	95.7%	86.7%	84.0%	89.2%
Ingenix PRG	Rx	93.0%	73.2%	96.0%	86.3%	85.6%	87.4%
MedicaidRx	Rx	91.9%	74.0%	95.2%	78.8%	84.7%	88.1%
Impact Pro	Med+Rx+Use	99.3%	97.5%	98.3%	97.0%	101.6%	97.8%
Ingenix ERG	Med+Rx	97.3%	92.6%	99.4%	94.5%	81.5%	92.3%
ACG w/ Prior Cost**	Diag+\$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+\$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input.

(出典) 前掲注 4 の資料、p.24 の Table V.1 を引用。

《図表 26》将来予測の PR (2004 年における医的状态で分類した場合)

TABLE V.2		Predictive Ratios by Medical Condition in 2004 (Recalibrated Nonlagged Prospective without Prior Costs, 250K Truncation)					
Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	71.6%	63.8%	83.7%	60.1%	71.9%	70.8%
CDPS	Diag	69.2%	57.5%	84.1%	55.1%	63.3%	65.7%
Clinical Risk Groups*	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	68.2%	64.6%	84.4%	57.7%	66.0%	70.5%
DxCG RxGroups	Rx	68.2%	64.6%	84.4%	57.7%	66.0%	70.5%
Ingenix PRG	Rx	74.1%	52.9%	86.8%	58.3%	60.8%	69.5%
MedicaidRx	Rx	72.6%	53.6%	87.1%	57.9%	63.0%	68.2%
Impact Pro	Med+Rx+Use	73.9%	65.2%	88.6%	58.8%	57.7%	69.2%
Ingenix ERG	Med+Rx	73.9%	65.2%	88.6%	58.8%	57.7%	69.2%
ACG w/ Prior Cost**	Diag+\$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+\$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models included prior cost as input.

(出典) 前掲注 4 の資料、p.25 の Table V.2 を引用。

告書における Table V.2) は、各個々人の 2004 年における所定の状態の有無により分類した場合の 2004 年の PR を示したものである。

この 2 つの図表を見比べると、2004 年における状態で分類した場合の方が、PR が悪いことが見てとれる。これは、2004 年においてこれらの医的な状態を持っている人は、2003 年においてこれらの医的な状態を持っていた人よりも、コストの平均値が高くかつコストのばらつき（分散）も大きかったことによると、米国アクチュアリー会 2007 年報告書では説明している。当然のことながら、翌年のコストの平均値が高くなり、かつばらつきが大きくなると PR は悪くなる。

モデル別のコメントとして、米国アクチュアリー会 2007 年報告書は、Impact Pro、Ingenix ERG および ACG が、PR の尺度では他のモデルより相対的にパフォーマンスがよかったと述べている。また、興味深い結果として、処方薬剤のみに基づく予測モデルの PR は①疾病分類により著しく変動していること、②概して言えば、診断名に基づいたモデルほどには 100% に近い値となっていないこと、③この傾向は、状態による分類を 2003 年データで行った場合の方が目立つことを指摘している。この分析では、疾病による分類は診断名に基づく基準で行われ、服用されている処方薬剤（NDC コード）に基づいて行われたものではないので、この結果は意外なものではない。しかし、この例は、予測モデルの実務適用に当たって、適切なツールを使用することの重要性を示唆している。すなわち、集団内を階層化して使用する前提で予測モデルの選択を考慮する場合には、行おうとしている集団内の階層化のための定義の仕方が、その予測モデルにおける特徴識別（グループ化）

の仕組みに反映されているかどうかに基づいて選択すべきであると米国アクチュアリー会 2007 年報告書は述べている。

（２）現状分析シナリオ — 2003 年の医的狀態による

《図表 27》（米国アクチュアリー会 2007 年報告書における Table V.3）は、各個々人の 2003 年における所定の状態の有無により分類した場合の 2003 年の PR を示したものである。2003 年における状態で分類する場合、現状分析シナリオの方が将来予測シナリオより、一般的に、モデルのパフォーマンスが格段によいことが示されている²⁰。

（３）過去の期間のコストも組み込んだ将来予測 — 2003 年および 2004 年の医的狀態による

《図表 28》（米国アクチュアリー会 2007 年報告書における Table V.4）および《図表 29》（米国アクチュアリー会 2007 年報告書における Table V.5）は、将来予測シナリオで、過去の期間のコストも組み込んでリスク・ウェイトを調整した場合の PR を示している。《図表 28》は 2003 年における医的な状態で分類した場合の結果、《図表 29》は 2004 年における医的な状態で分類した場合の結果である。

予想に違わず、2003 年における医的な状態で分類した場合の方が、2004 年における医的な状態で分類した場合に比べて、一般的に、PR がよく、また、PR が 100% を超えることも多いことが見て取れる。これらの状態に対するコストの実績値のばらつきを考えれば、このことは予想されることであると米国アクチュアリー会 2007 年報告書では述べている。

²⁰ 米国アクチュアリー会 2007 年報告書における原文では、“The performance generally improves considerably for the current models compared to prospective results with medical conditions in 2004.”と記載されているが、現状分析シナリオは 2003 年の医的狀態に基づいて分類されているので、上記本文のとおり修正して記述した。

《図表 27》現状分析の PR（2003 年における医的状态で分類した場合）

TABLE V.3		Predictive Ratios by Medical Condition in 2003 (Recalibrated Nonlagged Concurrent without Prior Costs, 250K Truncation)					
Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	103.2%	102.5%	88.8%	91.3%	41.0%	100.6%
CDPS	Diag	104.7%	76.5%	87.1%	83.8%	80.1%	80.2%
Clinical Risk Groups*	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	92.9%	98.4%	93.0%	95.8%	83.3%	94.7%
DxCG RxGroups	Rx	85.8%	79.7%	89.4%	75.2%	67.6%	79.6%
Ingenix PRG**	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	85.8%	75.9%	90.1%	65.0%	73.2%	79.9%
Impact Pro**	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	92.5%	96.6%	93.7%	89.8%	74.8%	85.2%
ACG w/ Prior Cost***	Diag+\$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model***	Diag+\$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
MEDai***	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models do not include a concurrent option.

*** These models include prior cost as input.

(出典) 前掲注 4 の資料、p.26 の Table V.3 を引用。

《図表 28》将来予測の PR（過去の期間のコストを変数追加、2003 年における医的状态で分類した場合）

TABLE V.4		Predictive Ratios by Medical Condition in 2003 (Recalibrated Nonlagged Prospective with Prior Costs, 250K Truncation)					
Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	99.0%	91.0%	100.1%	105.6%	115.5%	99.2%
CDPS	Diag	93.2%	86.6%	99.7%	96.8%	96.4%	94.6%
Clinical Risk Groups	Diag	96.5%	110.2%	110.0%	115.8%	109.3%	101.8%
DxCG DCG	Diag	95.8%	90.2%	99.2%	96.2%	99.3%	100.3%
DxCG RxGroups	Rx	101.2%	79.6%	99.0%	97.0%	94.6%	96.8%
Ingenix PRG	Rx	97.9%	80.0%	98.4%	96.4%	93.5%	94.9%
MedicaidRx	Rx	97.9%	84.2%	98.7%	95.2%	96.3%	96.8%
Impact Pro	Med+Rx+Use	100.8%	99.9%	99.5%	98.6%	106.5%	100.0%
Ingenix ERG	Med+Rx	99.8%	92.6%	101.0%	97.8%	92.7%	97.2%
ACG w/ Prior Cost	Diag+\$Rx	100.7%	101.0%	100.5%	102.5%	119.1%	100.1%
DxCG UW Model	Diag+\$Total	99.1%	93.1%	100.7%	97.6%	107.3%	101.0%
Service Vendor	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
MEDai	All	104.4%	93.3%	102.6%	97.9%	96.1%	99.7%

(出典) 前掲注 4 の資料、p.27 の Table V.4 を引用。

《図表 29》 将来予測の PR（過去の期間のコストを変数追加、
2004 年における医的状态で分類した場合）

TABLE V.5 Predictive Ratios by Medical Condition in 2004 (Recalibrated Nonlagged Prospective with Prior Costs, 250K Truncation)							
Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	72.1%	64.2%	86.9%	62.6%	83.4%	71.7%
CDPS	Diag	68.7%	61.4%	85.7%	57.7%	66.5%	68.5%
Clinical Risk Groups	Diag	75.2%	65.4%	89.6%	59.9%	66.0%	73.1%
DxCG DCG	Diag	76.7%	57.3%	88.4%	60.9%	68.1%	72.5%
DxCG RxGroups	Rx	76.7%	57.3%	88.4%	60.9%	68.1%	72.5%
Ingenix PRG	Rx	74.6%	57.6%	88.0%	60.5%	67.8%	72.6%
MedicaidRx	Rx	74.4%	60.7%	88.1%	59.4%	69.1%	71.3%
Impact Pro	Med+Rx+Use	76.7%	71.9%	89.0%	62.6%	77.6%	71.9%
Ingenix ERG	Med+Rx	76.7%	71.9%	89.0%	62.6%	77.6%	71.9%
ACG w/ Prior Cost	Diag+\$Rx	75.3%	70.3%	88.2%	62.3%	85.6%	73.9%
DxCG UW Model	Diag+\$Total	75.3%	70.3%	88.2%	62.3%	85.6%	73.9%
Service Vendor	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
MEDai	All	79.8%	66.8%	91.1%	62.0%	70.7%	75.4%

(出典) 前掲注 4 の資料、p.27 の Table V.5 を引用。

4. コスト・カテゴリーに対する結果

以下の分析では、各個々人を 2003 年または 2004 年のいずれか（シナリオによる）におけるコストの実績値に基づいてコスト・カテゴリーに分類し、PR を算出している。コスト・カテゴリーは、実績値によるパーセンタイルの幅で設定した。また、PR は、将来予測シナリオでは、2004 年に対する予測値と 2004 年の実績値の割合で、現状分析シナリオでは、2003 年に対する計算値と 2003 年の実績値の割合で算出されている。

(1) 将来予測の場合および現状分析の場合

《図表 30》（米国アクチュアリー会 2007 年報告書における Table VI.1）は、2004 年のコストの実績値に基づいてコスト・カテゴリーを分類した将来予測シナリオにおける結果である。コスト・カテゴリーは、実績値によるパーセン

タイルの幅で設定されているが、《図表 30》の場合、99-100 というカテゴリーは、2004 年のコスト実績値のトップ 1%の人達を意味し、一方、0-20 というカテゴリーは、2004 年のコスト実績値の下から 20%までの人達を意味している。この分析により、2004 年のコストが高かった人達、中庸だった人達、低かった人達それぞれに対して、各モデルが 2004 年のコストの平均値をいかによく予測したかを示している。

《図表 30》を見ると、いずれの予測モデルにおいても、ハイ・コストの人達については予測値が過小になり、ロー・コストの人達については予測値が過大になることを示している。《図表 30》は、また、コストによる百分順位が小さいほど、PR が大きくなることも示している。このパフォーマンスの動きはモデルが異なっても極めて似通っている。Impact Pro は²¹、96 パーセント以上以上のコスト・カテゴリーで相対的に

²¹ 米国アクチュアリー会 2007 年報告書における原文では、“Clinical Risk Groups and Impact Pro performed...” と記載されているが、Clinical Risk Groups は《図表 24》において、他のモデルと一貫性のあるリスク・ウェイトの調整ができなかったことから「N/A」と表示されているので、上記本文のとおり修正して記述した。

《図表 30》 将来予測の PR (2004 年のコスト実績値で分類した場合)

TABLE VI.1	Prospective without Prior Cost (Recalibrated, Nonlagged) Predictive Ratios by Cost Percentile Groupings (Cost Groupings Defined for 2004)							
	Percentile Ranges							
Risk Adjuster Tool	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	21.8%	42.5%	67.5%	100.0%	152.2%	265.0%	570.7%	8308.1%
CDPS	18.2%	38.4%	63.6%	96.8%	154.5%	275.1%	595.3%	9335.9%
Clinical Risk Groups*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	20.5%	41.7%	67.3%	100.1%	153.3%	263.6%	558.3%	7869.0%
DxCG RxGroups	18.2%	43.8%	72.8%	105.8%	155.0%	248.8%	516.9%	7914.0%
Ingenix PRG	19.2%	44.3%	72.6%	104.2%	152.9%	247.4%	523.9%	8301.4%
MedicaidRx	15.9%	40.1%	69.9%	107.0%	163.4%	261.9%	516.9%	7374.3%
Impact Pro	26.9%	48.3%	73.3%	103.9%	152.1%	241.4%	480.9%	6605.6%
Ingenix ERG	18.0%	41.5%	71.1%	108.7%	163.6%	261.4%	509.2%	6171.7%
ACG w/ Prior Cost**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Percentile Ranges							
	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
MEDai**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input.

(出典) 前掲注 4 の資料、p.28 の Table VI.1 を引用。

《図表 31》 現状分析の PR (2003 年のコスト実績値で分類した場合)

TABLE VI.2	Concurrent without Prior Cost (Offered, Nonlagged) Predictive Ratios by Cost Percentile Groupings (Cost Groupings Defined for 2003)							
	Percentile Ranges							
Risk Adjuster Tool	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	57.0%	82.8%	94.8%	100.2%	107.6%	124.3%	137.9%	133.4%
CDPS	44.9%	60.9%	73.3%	86.4%	106.0%	142.9%	195.1%	283.1%
Clinical Risk Groups	62.8%	76.7%	83.8%	92.6%	105.8%	129.0%	158.9%	208.4%
DxCG DCG	75.2%	84.6%	89.0%	94.3%	102.9%	120.3%	133.4%	151.2%
DxCG RxGroups*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	43.2%	70.9%	88.1%	102.3%	116.6%	129.8%	136.3%	154.6%
Impact Pro*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	54.4%	75.2%	88.4%	101.2%	114.0%	127.6%	134.9%	131.5%
ACG w/ Prior Cost**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Percentile Ranges							
	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
MEDai**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

* These models do not include a concurrent option.

** These models include prior cost as input

(出典) 前掲注 4 の資料、p.29 の Table VI.2 を引用。

よいパフォーマンスを示している。Ingenix PRG は、96-99 パーセンタイルでは相対的にパフォーマンスがよいが、99-100 パーセンタイルではパフォーマンスがそれほどよくない。Impact Pro は、すべてのパーセンタイルのレン

ジで相対的によいパフォーマンスであったと米国アクチュアリー会 2007 年報告書ではコメントしている。

《図表 31》(米国アクチュアリー会 2007 年報告書における Table VI.2) は、2003 年のコ

ストの実績値に基づいてコスト・カテゴリーを分類した現状分析シナリオにおける結果である。

《図表 30》と比較すると、現状分析シナリオの方がコストレベルによる階層化がうまく行われることが明らかであるが、ハイ・コスト群に対しては過小予測となり、ロー・コスト群では過大予測となる点は、将来予測シナリオの場合と同様であると米国アクチュアリー会 2007 年報告書は述べている。

(2) 2003 年はロー・コストで 2004 年はハイ・コストとなる群

以下の分析は、ムーバー (“movers” : 2003 年はロー・コストだったが、2004 年はハイ・コストになった人達) について、各モデルが 2004

年のコストを如何によく予測したかを見たものである。この集団はフォローする価値のある重要な集団である。なぜなら、単に過去の期間のコストによるのではなく、予測モデルを使うことの価値の一部は、コストの変化 (すなわち、ロー・コストからハイ・コストへ、ハイ・コストからロー・コストへ) を予測する能力にあるからである。

《図表 32》(米国アクチュアリー会 2007 年報告書における Table VI.3) は、2003 年のコストの実績値が中央値未満の人達のデータを集計対象として算出した各モデルの結果を、2004 年のコストの実績値のパーセンタイルでカテゴリー分けして PR で表示してある。使用したシナリオは、将来予測でリスク・ウェイトを調整

《図表 32》 将来予測 (過去の期間のコストの変数追加なし)

TABLE VI.3		Predictive Ratios by 2004 Cost Percentile where <50th Percentile in 2003 (Prospective, Recalibrated, Nonlagged, without Prior Cost)						
		2004 Cost Percentile Range						
Risk Adjuster Tool	0-100th	70th-100th	75th-100th	80th-100th	85th-100th	90th-100th	95th-100th	
ACG	132.0%	16.9%	14.5%	12.3%	10.4%	8.3%	6.2%	
CDPS	144.8%	14.8%	12.6%	10.7%	9.0%	7.3%	5.5%	
Clinical Risk Groups*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
DxCG DCG	126.8%	17.8%	15.2%	12.9%	10.8%	8.7%	6.6%	
DxCG RxGroups	130.1%	16.2%	13.8%	11.6%	9.8%	7.9%	6.1%	
Ingenix PRG	133.5%	15.7%	13.3%	11.2%	9.4%	7.6%	5.9%	
MedicaidRx	126.5%	17.6%	15.0%	12.6%	10.5%	8.5%	6.7%	
Impact Pro	110.6%	20.1%	17.2%	14.5%	12.2%	9.9%	7.6%	
Ingenix ERG	112.1%	21.0%	18.0%	15.3%	12.8%	10.3%	7.7%	
ACG w/Prior Cost**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
DxCG UW Model**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
		2004 Cost Percentile Range						
Service Vendor	0-100th	70th-100th	75th-100th	80th-100th	85th-100th	90th-100th	95th-100th	
MEDai**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input

Note: The 0-100th percentile values were not adjusted, but all other values were normalized by 0-100th percentile values. Unadjusted predictive ratios can be calculated by multiplying shown values by 0-100th percentile values.

(注) 0-100thパーセンタイル欄の数字は、算出結果そのもので調整は加えていない。その他の欄の数字は、0-100thパーセンタイルの数字に基づいて標準化した数値を表示した。標準化する前の数値は、表示された (標準化された) 数値に 0-100thパーセンタイルの数値を乗じることによって得ることができる。

(出典) 前掲注 4 の資料、p.29 の Table VI.3 を引用。

したが、過去の期間のコストは変数として追加していないシナリオである。

《図表 32》の 0-100thパーセンタイルの欄を見ると、総じて、いずれのモデルとも、2003年のロー・コスト者についての2004年のコスト予測の平均値は過大なものとなっている（PRが100%を超えている）。これに関して、米国アクチュアリー会 2007年報告書では次のようにコメントしている。「この結果は先に掲げた諸表と一貫した（相反しない）ものである。なぜなら、予測モデルでは総じて健康な人々のコストは過大予測となるが、2003年に相対的に健康であった人々は2004年も健康であることが多いと考えられるからである。しかし、この結果をもって、モデルまたは手法の不備・欠陥であると解釈しないことが重要である。なぜなら、これらの結果は、ヘルスケアコストのもつ本来の性質や変動性、また、その定義からいって、コストの水準が大きく変わる人々のコストを予想することの困難性に起因するものであるからで

ある。」

加えて、《図表 32》は、前年はロー・コストであったが翌年は最も高額なコストとなる人を予測する階層化を各モデルがどの程度うまく行うのかを示しているとも見ることができる。Ingenix ERG は、0-100thパーセンタイルのカテゴリーを除く各カテゴリーにおいて最良のPRを示している。0-100thパーセンタイルのカテゴリーでは、Impact Pro が最良のPRを示した。

《図表 33》（米国アクチュアリー会 2007年報告書における Table VI.4）は、《図表 32》と同様の分析を、もともとは過去の期間のコストを組み込んでいないモデルについても、過去の期間のコストを独立変数として追加してリスク・ウェイトを調整したシナリオで実施したものである。DxCG UW Model は 0-100thの欄に示されているように、ロー・コストの集団全体に対してはととてもよいPRを示した。Ingenix ERG は、0-100thパーセンタイル欄を除くすべ

《図表 33》 将来予測（過去の期間のコストの変数追加あり）

TABLE VI.4	Predictive Ratios by 2004 Cost Percentile where <50th Percentile in 2003 (Prospective, Recalibrated, Nonlagged, with Prior Cost)						
	2004 Cost Percentile Range						
Risk Adjuster Tool	0-100th	70th-100th	75th-100th	80th-100th	85th-100th	90th-100th	95th-100th
ACG	127.4%	21.4%	18.3%	15.6%	13.1%	10.5%	7.9%
CDPS	126.8%	21.3%	18.2%	15.4%	13.0%	10.5%	8.0%
Clinical Risk Groups	102.1%	21.6%	18.5%	15.7%	13.1%	10.6%	7.9%
DxCG DCG	119.1%	21.6%	18.5%	15.7%	13.2%	10.6%	8.0%
DxCG RxGroups	110.8%	20.9%	17.7%	14.9%	12.4%	10.1%	7.8%
Ingenix PRG	113.9%	20.6%	17.5%	14.6%	12.2%	9.9%	7.7%
MedicaidRx	106.4%	21.6%	18.3%	15.3%	12.8%	10.3%	8.0%
Impact Pro	106.3%	21.6%	18.4%	15.5%	13.0%	10.6%	8.2%
Ingenix ERG	103.8%	22.6%	19.3%	16.3%	13.7%	11.0%	8.3%
ACG w/ Prior Cost	120.4%	20.7%	17.7%	15.1%	12.6%	10.2%	7.7%
DxCG UW Model	99.8%	21.8%	18.7%	15.8%	13.3%	10.7%	8.0%
Service Vendor	2004 Cost Percentile Range						
	0-100th	70th-100th	75th-100th	80th-100th	85th-100th	90th-100th	95th-100th
MEDai	93.5%	22.0%	18.8%	15.9%	13.4%	10.7%	8.1%

Note: The 0-100th percentile values were not adjusted, but all other values were normalized by 0-100th percentile values. Unadjusted predictive ratios can be calculated by multiplying shown values by 0-100th percentile values.

(出典) 前掲注 4 の資料、p.30 の Table VI.4 を引用。

でのレンジで最良の PR を示した。

以上が、米国アクチュアリー会 2007 年報告書の本文に掲載された図表とそれに対する説明であるが、米国アクチュアリー会 2007 年報告書の APPENDIX では、下記の 12 のシナリオについて、《図表 19》、《図表 25》、《図表 30》の構成の図表 3 枚をそれぞれ掲載している。

APENDIX A-1. Offered, Prospective,

Nonlagged, without Prior Costs

APENDIX A-2. Offered, Prospective,

Nonlagged, with Prior Costs

APENDIX A-3. Offered, Prospective,

Lagged, without Prior Costs

APENDIX A-4. Offered, Prospective,

Lagged, with Prior Costs

APENDIX A-5. Offered, Concurrent,

Nonlagged, without Prior Costs

APENDIX A-6. Offered, Concurrent,

Lagged, without Prior Costs

APENDIX A-7. Recalibrated, Prospective,

Nonlagged, without Prior Costs

APENDIX A-8. Recalibrated, Prospective,

Nonlagged, with Prior Costs

APENDIX A-9. Recalibrated, Prospective,

Lagged, without Prior Costs

APENDIX A-10. Recalibrated, Prospective,

Lagged, with Prior Costs

APENDIX A-11. Recalibrated, Concurrent,

Nonlagged, without Prior Costs

APENDIX A-12. Recalibrated, Concurrent,

Lagged, without Prior Costs

V. 今後の課題

本稿の締めくくりとして、上記の分析を踏まえて米国アクチュアリー会 2007 年報告書で述

べられている、予測モデルの限界やそのパフォーマンスに影響を与える要素、予測モデル導入において考慮すべき諸点、および今後さらに研究すべきとされている課題について取りまとめる。

1. 予測モデルのパフォーマンスに影響を与える制限および要因

いかなる予測モデル作りのツールでもそうであるように、米国アクチュアリー会 2007 年報告書で取り上げた医療コストの各予測モデルのパフォーマンスは、データや使用上の制限といった多くのファクターの影響を受ける。これらのファクターについて、米国アクチュアリー会 2007 年報告書では、以下のように論じている。

(1) Population Specificity and Applicability

モデルは、もともとの開発時点で想定したものと異なる集団に対して使用して、完全ではないまでも満足いく程度のパフォーマンスを得るために、リスク・ウェイトの調整を行うことが可能であることが、前章の分析で示された。例えば、CDPS や MedicaidRx は、もともとはそれぞれ慢性疾患により就業障害の発生している集団またはメディケイドの集団を想定して開発されたものであるが、民間健康保険プランの加入者集団のデータセットに対応するべくリスク・ウェイトの調整を行った上で使用すれば、よいパフォーマンスが得られた。しかし、状態によるカテゴリー分けの区分や提供された情報は、分析しようとした対象に十分特異的でないかもしれない点に留意すべきであると米国アクチュアリー会 2007 年報告書では述べている。例えば、65 歳以上の集団を想定して開発された予測モデルは、妊娠関連の疾病や幼児期の疾病

に対する適切な分解は含まれていないかもしれないからである。

その予測モデルの使用目的のすべてを検討すること、およびどのような情報が集められるのかを検討することは重要である。年齢・性別の категорияや状態による категорияは、分析対象としている集団やそのツールが使用される目的に対して意味あるものとなっていることが必要である。予測モデルのベンダー、外部コンサルタントおよび自前のスタッフによるツールのカスタマイゼーションは、大きな改善をもたらし得るが、修正・改良やリスク・ウェイトの調整は慎重に行われるべきであると米国アクチュアリー会 2007 年報告書は主張している。

(2) Turnover

予測モデルを適用しようとしている集団が、失効・無効（自発的なものもあれば、非自発的なものもある）または死亡によって、予測対象の期間中の一部分しか集団に属していなかった人達を含んでいるかもしれない。同様に、新規の加入者がリスク・プールに入って来るかもしれないし、彼らに関しては、予測の基となる過去の期間について限られたデータしか存在しないか、あるいは、給付金請求データが一切存在しないということもあろう。

Milliman 社の実施した“Optimal Renewal Guideline”という研究では、最適化された予測モデルによる予測に加えて、年齢・性別だけで予測した場合のパフォーマンスを検討している。年齢・性別のみを使用して将来予測をした場合の決定係数は約 6%であった。最適化された予測モデルでの将来予測では、決定係数は約 25%であった。したがって、分析対象集団に turnover がある場合の決定係数の大まかな推計値は下記の通りと考えられることが米国アク

チュアリー会 2007 年報告書で示されている。

$$\frac{[(0.06 \times \text{turnover_rate} + 0.25 \times (1 - \text{turnover_rate})) / 0.25]}{\times \text{Pre_turnover_R_squared}}$$

例えば、15%の turnover があり（すなわち、予測対象期間における当該集団構成員の 15%の人については診断名や処方薬剤のデータがなく）、turnover を考慮しない分析における将来予測の決定係数が 27%だったとする。この場合、turnover を考慮した決定係数を上記の算式に従い推計すると約 24%となる $[(0.06 \times 0.15 + 0.25 \times 0.85) / 0.25 \times 0.27]$ 。このアプローチでは、予測期間中の一部分の期間しか集団に属していなかった人 (partial enrollment) については考慮していない。米国アクチュアリー会 2007 年報告書によれば、いくつかのベンダーは、観察期間の途中で新たに加入した人に対するリスクスコアを算出するロジックを追加しているとのことである。

上記の算式では、新規加入者に対しては年齢・性別だけで予測することが仮定されている。さらに、上記算式では、“Optimal Renewal Guidelines” 研究で示された「最適化モデル」における決定係数と「年齢・性別のみモデル」における決定係数を turnover rate で加重平均し、その加重平均値の「最適化モデル」における決定係数に対する割合に継続加入 (turnover がない) 状態での決定係数を乗じたものが予測のパワーの変化に等しいと仮定している。上記の算式では、turnover rate は、観察期間中のデータは存在しないが予測期間である保険料率適用対象期間では当該集団の構成員となる人の割合として定義されている。その意味で、これは単純化された模式的な算式であると米国アクチュアリー会 2007 年報告書では評価している。

上記算式の 0.06 と 0.25 の値は、モデリング

の条件により影響を受けるが、年齢・性別だけによる予測のパフォーマンス (0.06) は、モデリングの手法が異なっても大きくは変化しないとも考えられることから、上記の算式中の 0.25 の代わりに turnover を考慮しない場合の決定係数を用いる方が適切かもしれないという考え方も米国アクチュアリー会 2007 年報告書では紹介されている。この場合、算式は次のように簡素化される。

$$0.06 \times \text{turnover_rate} + \text{Pre_turnover_R_squared} \times (1 - \text{turnover_rate})$$

これらの算式は MAPE でも機能するし、MAPE は誤差項を二乗しないので、決定係数より外れ値の影響が小さく、より適切かもしれないと米国アクチュアリー会 2007 年報告書ではコメントしている。

(3) ラグ問題

予測モデルを実務に使用する場合、予測の対象となる期間が数か月先から始まるということはよくあることである。例えば、米国において小規模集団に対する更改契約の保険料率を決定する場合、その保険料率が適用されるべき期間の 3 から 6 か月前に作業が実施されることが一般的である。この期間のずれを予測のラグといい、前項で論じた turnover に加えて、モデルのパフォーマンスに影響を与えるものである。加入者の健康状態がその間に变化する可能性があるため、予測のラグは不確実性を生み出す。将来予測のためのいかなるモデルにおいても、個々人の健康状態は変わりうるので、将来のコストを予測するということには不確実性が入り込む。米国アクチュアリー会 2007 年報告書では、予測のラグを、データ収集期間の終わりから予測すべき対象期間の始まりの間の期間と定義

している。

多くの予測モデルは、観察期間直後から始まる一定の期間中、継続して加入していた人々の集団を対象にリスク・ウェイトが設定されている。offered リスク・ウェイトを算出したときの状態と予測モデルを適用するときの状態との間に变化があるときはいつでも、リスク・ウェイトの調整を考慮することが重要である。予測モデルのベンダーのうちいくつかの会社は、ツールのパッケージソフトの一部として、予測のラグに対するオプションを備えたモデルを提供している。モデルの使用される状況に合わせてリスク・ウェイトが調整された場合には特に、少々の予測のラグがモデルのパフォーマンスに大きな影響を与えることはない。しかし、予測のラグにより、turnover の可能性のある期間がそれだけ延びるので、予測のラグは、turnover による影響を増加させる。

データのラグは、予測のラグと関連しているが、同じものではない。予測モデルのインプットに用いられるデータは、支払い済み給付金の実績値でなければならない。医療費の給付金請求は発生後、支払が完了するまでに通常 (平均的に) 2 から 4 か月を要する。いくつかの給付金請求は、支払完了まで数年を要する可能性のあるものもある。処方薬剤の給付金請求はもっと早期に支払済みとなるが、それでも支払完了までに 1 から 2 か月を要する。したがって、多くの状況においては、潜在的に意味があり、かつタイムリーな給付金請求データは予測モデル用に利用可能ではないかもしれない。ベンダーはラグ問題を最小にするモデルを追加してきたが、データのラグはすべてのモデル、とりわけ、主として診断名データに頼っているモデルのパフォーマンスに影響を与えると、米国アクチュアリー会 2007 年報告書では指摘している。

米国アクチュアリー会 2007 年報告書では、データのラグと予測のラグのインパクトも分析している。《図表 34》(米国アクチュアリー会 2007 年報告書における Table VII.1) は、データのラグと予測のラグが合わさってモデルのパフォーマンスに与えるインパクトを示している。《図表 21》では、offered リスク・ウェイトによる「ラグあり」と「ラグなし」の比較結果を示したが、《図表 34》では、調整リスク・ウェイトによる比較結果を示している。

《図表 34》は、予測のパフォーマンスがデータのラグと予測のラグによってかなりの程度影響を受けることを示している。処方薬剤のみに基づく予測モデルの方が受ける影響は小さい。米国アクチュアリー会 2007 年報告書では、2003 年 1 月から 8 月の間に発生し支払われた給付金請求事案が、2004 年暦年の給付金コストの予測に用いられた。したがって、「ラグあり」の分析では、4 か月間のデータのラグおよび予

測のラグがモデル化されている。

データの遅れはいずれの予測モデルにしる、導入の際に問題となる米国の事情が米国アクチュアリー会 2007 年報告書では説明されている。例えば、継続して加入者であったことという条件を課すと、メディケイドの場合、現加入者の 40%から 50%の人を臨床的な状態のリスク評価の対象から除外することになる可能性があり、それゆえ、総合的な人頭払いシステムの予測のパフォーマンスを大きく損なうことになるかもしれないことが懸念される。したがって、ラグがモデルのパフォーマンスにどの程度影響を与えるかを理解することは重要であると米国アクチュアリー会 2007 年報告書では結論付けている。

(4) データに纏わる諸問題

リスクを評価するために使用するデータという観点から言えば、予測モデルの方式は、デモ

《図表 34》 将来予測シナリオ、\$250,000 で頭切り、調整リスク・ウェイト、「ラグあり」と「ラグなし」の決定係数と MAPE の変化

TABLE VII.1		R-Squared Prospective Recalibrated (Without Prior Cost, 250K Truncation)					
Risk Adjuster Tool	Inputs	R-squared			MAPE%		
		Lagged	Nonlagged	Change	Lagged	Nonlagged	Change
ACG	Diag	15.2%	19.6%	4.4%	92.8%	88.8%	-4.0%
CDPS	Diag	14.5%	17.7%	3.2%	95.1%	91.9%	-3.2%
Clinical Risk Groups*	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	16.9%	21.3%	4.4%	91.2%	87.0%	-4.2%
DxCG RxGroups	Rx	18.2%	20.5%	2.3%	87.2%	85.3%	-1.9%
Ingenix PRG	Rx	18.9%	21.2%	2.3%	87.6%	85.6%	-2.0%
MedicaidRx	Rx	15.8%	17.7%	1.9%	90.1%	88.4%	-1.7%
Impact Pro	Med+Rx+Use	21.5%	25.6%	4.1%	84.9%	81.6%	-3.3%
Ingenix ERG	Med+Rx	17.4%	20.0%	2.6%	88.4%	86.1%	-2.3%
ACG - w/ Prior Cost**	Diag+\$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+\$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	Inputs	Lagged	Nonlagged	Change	Lagged	Nonlagged	Change
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input.

(出典) 前掲注 4 の資料、p.33 の Table VII.1 を引用。

グラフィックのデータ、過去の期間の支出金額、または自己申告式の健康状態調査や検査結果数値を含む健康情報のいずれかまたはいくつか、あるいはすべてを基にするかどうかによって分類することができる。米国アクチュアリー会 2007 年報告書では、給付金請求データに基づいた健康情報を用いた方式に焦点を当てた。給付金請求データに基づくリスク調整の方式は、給付金請求または受診データに基づく診断名コードに依存するもの、診断の代用情報として処方薬剤データに依存するもの、過去の期間のコストを使用するもの、およびこれらの 3 種類のデータソースを様々な組み合わせたものにさらに区分できる。

検査結果数値や、慢性疾患や機能上の障害に関する自己申告式健康調査結果等のその他の健康情報を用いるモデルは、米国アクチュアリー会 2007 年報告書では対象としていない。この情報を用いることは、予測モデル作りの今後の注目すべき取り組み課題であると米国アクチュアリー会 2007 年報告書では述べている。また、電子カルテ (Electronic medical records, EMRs) に関する標準的フォーマットの採用が進展していることは、これら情報を用いる予測モデルの開発と実用性を加速すると期待されるとしている。

年齢、性別および保険加入状態の情報といったデモグラフィックなリスク要因のみに依存する手法は、運用が容易である。これらの方式は、ケアのプロセスの尺度ではないので、治療方法を変更したり、リスク・スコアを大きくするためにコーディングを変更したりするインセンティブは生み出さないが、残念ながら、これらの手法は、個人レベルの予測やリスクが偏った集団に対する予測では相対的に精度が悪いと米国アクチュアリー会 2007 年報告書では述べてい

る。

対照的に、個人々の過去の期間の医療費支出の総額は、将来の支出の完全ではないが満足いく程度に良い予測のための判断材料である。これらのデータは、詳細な受診データよりは扱い易い。しかし、効率的にケアを提供することに関連するインセンティブはとても弱いことが指摘されるとしている。

診断名や処方薬剤といった健康状態の尺度は、予測のためのよい判断材料であり、医的管理のための有用な情報を提供する。健康保険プランは、診断に関する情報を医療プロバイダーから取得しなければならない。いくつかの種類の健康保険プランにとってはこの種の情報の取得が難しいことがしばしばある米国の事情にも米国アクチュアリー会 2007 年報告書では簡単に触れている。その例として、健康保険プランが支払いのためのデータを要しない人頭払いの契約を医療プロバイダーと締結している場合や、出来高払いの要素がほとんどないあるいは全くないスタッフ・モデルまたはグループ・プロバイダー・モデルの健康保険プランの場合などを挙げている。外来診療における診断名のコーディングは、いくぶん信頼性に乏しいが、診断名に基づくリスク評価のソフトウェアで一般に利用可能なものは、不完全なデータにより引き起こされる問題を軽減するための予防手段が組み込まれているとのことである。

コーディングのパターンは時間の経過に伴い変化すると想定される。診断名に基づく方式では、コーディングの変化に関する主たる懸念は外来の診断名に関するものである。これらのコードは、支払や保険料率の設定には、過去広く使われてはこなかったが、今では一般的になりつつある。例えば、診断名のコードは、メディケア・アドバンテージの支払に使用されている

HCC モデルの重要な構成要素となっている。コーディング実務の変化により、主傷病として当該疾患も持った新規患者の特定や、重症度についてのコーディングに関する改善、治療に影響を与えるすべての関連する状態のコーディングの増加に繋がることがあるかもしれない。これらの変化により、予測モデルのリスク・ウェイトを算出したときの集団と比べると、リスクの高い集団と見えるようになることもあり得る。したがって、この結果、集団に対するコスト総額の見積もりを膨らませることになるかもしれない。

データに纏わるもう一つの重要な課題は、アクセスのし易さ（accessibility）である。いくつかの健康保険プランやプランの購入者は、他の者に比べ、処方薬剤のデータの方がより容易にアクセスできることもあるかもしれない。処方薬剤のデータはタイムリーであり、外来で使用される主要な薬剤についてはそのデータは比較的クリーンで不備がない。加えて、これらのデータは、医療プロバイダーから取得する必要がなく、煩わしい管理のための事務を省くことができる。健康保険プランや医療プロバイダーにとってのリスク・スコアを上げるために処方

薬剤の使用が増やされるとすれば、効率化に向けたインセンティブは貧弱かもしれない。処方薬剤に基づいたリスク評価モデルは一般的に、任意性（自由裁量性）がないと考えられている薬剤を基にしている。しかし、承認適応症外の処方が行われるとか、追加的な疾病や当該疾病でもより重症でないケース、あるいは該当するかどうかの境界線上の病態での処方に裁量がある薬剤の場合には、処方薬剤に基づく予測モデルを医療プロバイダーに対する支払いに活用しようとする場合には注意が必要である。また、短時間で新薬が導入されたり、承認適応症外使用が進んだりするので、処方薬剤に基づく予測モデルは定期的にアップデートし、リスク・ウェイトの調整を行うことが一般的に重要であるとも指摘されている。

《図表 35》（米国アクチュアリー会 2007 年報告書における Table VII.2）は、データの質や予測の正確性といったデータに関する 5 つの基準について、各リスクの尺度（データ・ソース）の特徴を質的に比較したものである。

米国アクチュアリー会 2007 年報告書で評価された諸手法は、分類のために組み込まれている医的な状態が何種類設定されているか、その

《図表 35》 リスク尺度の比較

Criteria	Risk Measures			
	Demographics	Prior Expenditures	Prescriptions	Health Diagnoses
Data Quality	High	Medium	High	Medium
Prediction Accuracy	Low	High	High	High
Administrative Burden	Low	Medium	Medium	High
Utilization Incentive	Low/None	High	High*	Low
Diagnosis Coding Incentive	Low/None	Low	Low	High

* High for prescription drugs, low for all other services.

（出典）前掲注 4 の資料、p.34 の Table VII.2 を引用。

数の点で幾分異なっている。いくつかの手法では、把握できるほとんどすべての疾病を考慮してリスク・スコアを算出している。他の手法では、リスク選択に関連しないと想定して、マイナーで急性の状態を除外している。後者のモデルでは、これらの状態は1人当たりコストの重要な要素ではなく、これらを含めることによって、逆に臨床的に必要のないコードの拡散を招くかもしれないと想定されている。しかし、プライマリーケアのプロバイダーがこれらの頻度は高いがマイナーで急性の状態をどのようにマネージしているかを評価することが目的ならば、これらの状態も組み込んだ予測モデルが相応しいだろう。

もう一つの違いは、リスク・カテゴリーに対する疾病に関する尺度の割り当て方法である。プロセスによっては、当該疾病とは余りに異質なカテゴリーを作り出すことになるかもしれない。いくつかの状態では、かかるコストが似ているからとして、関連はするが、臨床的には非常に異なった疾病を一括りにされてしまっている。加えて、重症度を示すさらに詳細なコーディングを行っても、疾病の単なる特定という目的を超えて、リスク・カテゴリーの割り付けは変更されないものもある。一方、糖尿病のような疾病は、ほとんどの予測モデルにおいて、それ自身がカテゴリーとなっており、その支払は特異的に糖尿病とコーディングされることによって、影響を受ける。

個々人にリスク・スコアを割り振るアプローチも様々である。いくつかの手法では、加算方式、すなわち、疾病カテゴリーが付加される毎に係数が加算される方式が採られ、他の手法では、乗算方式である。ほとんどすべての手法が、何らかのレベルで階層的である。支払額の決定に適用される場合には、これらのカテゴリーの

いくつかは、関連する状態を階層化して処理することが行われることがある。例えば、肺疾患は、階層構造上、最も高額なコストのカテゴリーで支払がなされる。これは、階層構造上、よりコストの低いカテゴリーは、より重篤な状態に関連した合併症を意味していると仮定していることによる。このアプローチによりダブル・カウントが回避できる。他の手法では、個々人の相互に影響しあうすべての状態により導き出された相互に排他的なリスク・カテゴリーを割り振ることとしたり、個々人にとって主要な状態に着目することとしたりして、状態間の関連性を処理している。

米国アクチュアリー会 2007 年報告書で分析対象とした各モデルは、予測のパフォーマンスを維持しながらも、データの諸問題に対して可能な限り堅牢なデザインを採ってきている。リスクをアサインするためには、観察期間において少なくとも1回、診断または処方薬剤に該当することを条件とするのが一般的である。同じコードが何回も出てくるその回数は、関係がないとするのが一般的である。任意的な（自由裁量のある）指標や疾病によって定義された指標は除外されるかまたはコードを都合のよいように使い分けるインセンティブが最小になるようにアサインされる。これは、リスク評価に適するように、データが完全に完備であることや詳細であることを要しないことを意味している。

(5) 集団の大きさおよび職域集団の契約更改時の保険料率設定に関する州政府の規制

州政府の規制は、医的なリスクの要因に基づいて適用可能な保険料率の変更幅を制限することによって、小規模集団の保険契約を引き受ける保険者が予測モデルの予測結果に基づいて行使しうる保険料率設定上のアクションを大きく

制限していることが多い(+/-10%から無制限まで、州により異なる)。集団の大きさもモデルの予測のパフォーマンスに影響を与える。なぜなら、集団規模が大きいほど、個々人のコストのばらつきは重要でなくなり、したがって、予測の精度が増すからである。また、大規模集団の将来コストは、小規模集団に比べ、過去のコストに基づいて予測可能である傾向が強い。

予測のパフォーマンスにおける保険料率規制のインパクトを理解するための例題として、小規模集団のヘルスケアコストの予測に2つの予測モデルが用いられたとしよう。一方のモデルでは、当該集団のコストは平均より30%高いと推計され、もう一方のモデルでは、当該集団のコストは平均より35%高いと推計されたとする。事後的に給付金請求の実績値より、当該集団のコストは平均より30%高いことが判明したとする。この保険者がどの州で営業しているかにより、どちらの予測モデルが、当該集団の契約更改時の保険料率を設定するという目的に対して保険者が使用しうる有用な情報を提供するかは異なる。例えば、アイオワ州は集団の健康状態を理由として平均値から+/-25%の乖離

しか認めていない。したがって、もしこれがアイオワ州における契約更改ならば、いずれのモデルでも、当該集団に対して規制上可能な限り最も高い保険料率設定をすべきことを示唆し、その意味で完全な情報を提供している。しかし、35%以上の保険料率の乖離が認められる州においては、前者（第一のモデル）がよりよい情報を提供していることになる。

先に紹介した“Optimal Renewal Guidelines”研究では、小規模集団の保険料率設定に関する州政府の規制による制限は、予測モデルを使用する等の契約更改時における手法の実施可能な予測力に大きな影響を与えると結論付けている。加えて、手法間の有意な差は、集団規模が大きくなるほど、減少することが指摘されている。

“Optimal Renewal Guidelines”研究からの引用である《図表36》（米国アクチュアリー会2007年報告書におけるTable VII.3）は、集団の規模と保険料率設定に関する規制上の制限がMAPEに与える影響を示している。左欄のリスク・アジャスター（予測モデル）による結果というのは、過去の期間のコストを含めてリスク・

《図表 36》 集団規模と規制による料率制限が MAPE に与える影響

TABLE VII.3	Impact of Group Size and Regulatory Rating Limits (MAPE with and without Cap)				
	Group Size	Risk Adjuster +/- 25%		Manual Rate +/- 25%	
		Uncapped	With Cap	Uncapped	With Cap
1 Mbr	82.7%	16.4%	101.0%	19.7%	
1 EE	70.2%	16.7%	85.8%	21.0%	
3 EEs	50.8%	16.9%	59.9%	21.3%	
10 EEs	32.0%	16.1%	36.8%	20.0%	
25 EEs	21.3%	14.8%	24.1%	17.6%	
50 EEs	15.1%	12.6%	17.2%	14.7%	
150 EEs	9.1%	8.9%	10.3%	10.2%	

(出典) 前掲注4の資料、p.36のTable VII.3を引用。

ウェイトを調整した最適化モデルによる結果を示している。

《図表 36》によれば、マニュアル料率を適用した場合、予測モデルを用いたアプローチの場合とも、MAPE の結果は、集団の規模が大きくなるほど改善している (MAPE が小さくなっている)。また、保険料率に対する制限が導入され、厳しくなったときの方 (With Cap) が、改善している。保険料率に対する制限がある場合の MAPE の計算においては、コストの実績値は許されている保険料率の乖離幅の制限を受けており、潜在的な誤差を減少させている。過去の損害率に基づく手法は、マニュアル料率アプローチよりはパフォーマンスが良いが、その差は予測モデルを用いたアプローチに比べると小さいとされている。

(6) 健康に基づいた予測モデルの使用上の留意点

米国では、健康に基づいた予測モデルは、健康保険プランおよびその購入者によって様々に利用されている。米国アクチュアリー会 2007 年報告書では、健康に基づいた予測モデルを選択する際の留意点について、予測モデルの利用目的を次の 2 つの主要な特徴によって下記の 4 つに分類して解説している。

- 医療プロバイダーまたは健康保険プランに対する支払の決定を利用目的に含んでいるか？
- 利用目的は、標的とする (targeted) 小集団 (subpopulation) に焦点を当てたものか？それとも集団全体を対象としたものか (global)？

① 医療プロバイダーまたは健康保険プランに対する支払／集団全体の視点

この使い方には、健康保険プランによる保険

料率の設定や医療プロバイダーに対する人頭払いの金額の決定が含まれる。米国アクチュアリー会 2007 年報告書による解説は以下の通りである。この条件下では、診断名に基づくモデルのいずれかが好まれる。なぜならば、診断名はよい予測のための判断材料であり、コードを都合のよいように使い分けるインセンティブが処方薬剤に基づくモデルに比べて少ないかもしれないからである。過去の期間のコストを組み込んだモデルは使用すべきではない。医療プロバイダーレベルでのリスク選択は通常、健康保険プランにおけるリスク選択より極端であるので、医療プロバイダーに対する支払に関して、人頭払いや担当人数を目安にしたインセンティブが活用される場合には、診断名のコードを都合のよいように使い分けることや過剰治療に留意することが重要となる。疾病に対する増加した支払を示すために、処方薬剤のような実際の資源利用データ (疾病を指し示し、支払の増加を示すもの) を使用することは避けるべきである。また、使用する場合は注意を払ったアプローチを採るべきである。診断名のデータでも、コードを都合のよいように使い分けることがないわけではない。しかし、大半のというわけではないが、かなり多くの主要な状態に対して基準が設定されており、これは妥当性評価の基礎として役に立つ。人頭払いの制度下であっても健康に基づくリスク調整を使用することには、医療プロバイダーがデータを提供する強いインセンティブとなるというメリットも考えられる。

② 医療プロバイダーまたは健康保険プランに対する支払／ターゲットを絞った集団に対する視点

この使い方には、対象者の特定、高額なケースの管理または対象疾病ごとの支払といった、ディジーズ・マネジメントに対する支払の水準

を設定することも含まれる。米国アクチュアリー会 2007 年報告書による解説は以下の通りである。道理に反するインセンティブを避けるためには、選択肢は診断名に基づくモデルに限るべきである。重症度および管理対象の疾病に関連する合併症を、どのモデルが最も適切に捕捉しているかを詳細に検討する必要がある。併せて、多くの共存症をもった高額で複雑な事案を、どのモデルが最も適切に捕捉しているかも詳細に検討する必要がある。また、注目したい特定の疾病について、ある種の医的な状態の存在もしくはその重症度を示す処方薬剤に任意性（自由裁量性）がないのであれば、使用に耐えると思われる。したがって、処方薬剤に基づいたシステムまたはそうしたシステムと組み合わせたものについても考慮の対象となるかもしれない。過去の期間のコストによるモデルは使われるべきではない。ただし、病院に対する DRG (diagnosis-related groups) による契約におけるストップ・ロスの規定のように、何らかのコストの閾値を設定することは、支払に対する調整としては適切な場合もあろう。

③医療プロバイダーまたは健康保険プランに対する支払以外／集団全体の視点

この使い方には、保険料に対する事業主・従業員間の拠出レベルの設定（すなわち、小規模集団に対する保険引受判断）、医療プロバイダーの効率性に関するプロファイリング、総医療費の予想や予算策定が含まれる。米国アクチュアリー会 2007 年報告書による解説は以下の通りである。支払と関係がないならば二次的なインセンティブも弱いので、いずれのモデルもこれ

らの使い方には適用し得るだろう。したがって、データにかかる費用であったり、リスク評価情報をほかのことにも使用したりといった他の要因が手法の選択を左右することになる。小規模集団に対する保険引受判断の目的であれば、利用可能な過去の期間のコストは予測因子に含めるべきであろう。なぜなら、過去の期間のコストはモデルの予測能力を大幅に改善するからである。

保険料率設定における健康に基づくリスク調整の比較的新しい使い方は、確定拠出型の商品における従業員拠出分（従業員が負担する保険料）の調整である。CDHP²²におけるリスク調整の活用は、これらのプランがもっと広く採用されるにつれ、重要となるだろう。

④医療プロバイダーまたは健康保険プランに対する支払以外／ターゲットを絞った集団に対する視点

この使い方には、ハイ・コスト事案の特定、個人保険の引受条件判断、ディジーズ・マネジメント・プログラムのプランニングや予算設定が含まれる。標準的な選択基準に加え、どの手法が臨床的に最も意味のある個々人のカテゴリー分けを提供するかということに基づいて選択することになるだろう。

2. 予測モデルの導入において留意すべき諸点

予測モデルの導入に際して留意すべき諸点として、米国アクチュアリー会 2007 年報告書では、医療費・処方薬剤データに纏わる課題、加入者資格データに纏わる課題、リスク・スコアリングを行う時期および頻度について解説を加え

²² Consumer-Directed Health Plans : 従業員が利用する医療プロバイダーや加入する健康保険プランを自ら選択し、給付設計も自ら行うとともに、金銭面のリスクも従業員自身が相当程度負担する形態のプランのことをいう。矢倉尚典、田中健司「米国における健康保険市場と保険会社のヘルスケア事業 —2004 年を中心とする概況および職域市場・HIPAA・メディケアをめぐる最近の動き—」（損保ジャパン総研クォーターリーVo.45、2005 年）、p.41 参照。

ている。

(1) 医療費データ、処方薬剤データに纏わる課題

予測モデルの導入を計画した段階で、多少なりとも事前テストや実際のデータハンドリングをやってみるべきだと米国アクチュアリー会 2007 年報告書では薦めている。予測モデル導入予定者にとって、シミュレーションや、特に受診データといった莫大な量のデータをハンドリングすることは初めての経験かもしれない。その場合、事前の経験も積まず、いきなり導入することは、極めてチャレンジング（困難）であると助言している。

また、導入に当たって、分析に使用するデータの妥当性は確かめる必要のあることを米国アクチュアリー会 2007 年報告書は強調している。例えば、各種のデータ要素の出現頻度の分布を確認することは不完全な受診データを特定するのに役立つだろう。加入者のうちどれくらいの割合の人が給付金請求なしと予想されるかについても、標準値のようなものはないにしても、何らかの情報があるだろう。日本とは異なる米国における制度の特徴に由来するものであるが、医療コストに関する部分的な人頭前払いにより、受診データに欠損があるかもしれない。また、よくある問題として、メンタルヘルスのサービスの給付を切り出して別プログラムで運用している (carve-out) 場合に、メンタルヘルスのサービス提供者からのデータが欠損していることがあると注意を喚起している。分析対象の各人が加入している給付プランは同種のものであることが必要であるし、さもなければ、標準化のために調整も行われるべきであり、場合によっては、付加的なモデルの作成が必要になることも

あろう。集団間の差異や給付条件の差異、およびこうした差異を処理するために使用した手法は、結果を示すときには明記されるべきであると米国アクチュアリー会 2007 年報告書は述べている。

健康保険プランのタイプが異なれば、データに関する問題も様々である。出来高払いによる費用請求の実績が限られている staff model HMO²³ の場合は、受診やサービスのバンドリングに関するデータの提供に問題があるかもしれない。1 レコードあたりに記録される診断名の数を制限しているシステムを使っている健康保険プランでは、リスクの尺度が過小に表現される可能性があるかもしれない。

データの質は、健康保険プラン全体のレベルでも、個々の医療プロバイダーレベルでも問題となりうる。健康保険プランレベルでのデータに関する関心事は完全性 (completeness) に纏わるものであるが、医療プロバイダーレベルの課題は、完全性と正確性の両方である。

診断名のデータに関する健康保険プランレベルの関心事は、医療プロバイダーが記録したすべての診断名を取り込むことである。診断名のデータが欠損を起こすかもしれない事情として、次の 2 つを挙げている。

- 医療プロバイダーのうちのいくつかから提供された受診データが不備である、あるいはデータの提供が受けられない
- 医療プロバイダーから提供されたデータで、1 受診あたりに記録される診断名の数が制限されている

処方薬剤に関するデータは、健康保険プランのレベルにおいて、大半の重要な状態に関して、大抵は完全で正確である。また、データの提供

²³ HMO の組織形態の一つで、「HMO 自身が医療機関を持ち、医師を雇用している」形態の HMO のこと。前掲注 8、西村徹他、p.23 参照。

に関して、医療プロバイダーは絡まない。

医療プロバイダーレベルでの診断名のコーディングに関しては、次の3つの行為に留意する必要があると米国アクチュアリー会 2007 年報告書では述べている。これらの行為により、診断名の数と分布が変わり得るし、また、集団の基調をなす罹患率が事実としては安定的であるかもしれない場合に、その集団のリスクの計測結果を増大させることが起こり得る。

■ **Diagnostic discovery:** 診断名の数や重症度の増加が記録されること。ただし、そのいずれもが適切である。過去の報告で過少申告だったものの訂正は、診断の一貫性の欠如という問題を軽減し、集団の疾病による負荷 (illness burden) をもっと公正に表すことになる。

■ **Diagnostic creep:** 診断名が不確実であるケースにおいて診断名の数や重症度を増加させること。これは、支払請求に関して、アップ・コーディングするバイアスを示している。多くの予測モデルの基本となっている分類基準の多くは、関連した診断名を一括りにすることによって、または疾病で定義されたコードを除外することによって、この問題を極小化しよう試みている。

■ **Tentative diagnosis:** 診断手順を正当化するために診断名をそれに相応しく使うこと (rule-out)、または、治療の遅れが有害であることを理由として、あたかも罹患しているかのように、診断のための確認的検査を行わずに治療することが必要であることを知らしめるために診断名をそれに相応しく使うこと (presumptive) といった、エラーの潜在的な源である。これに対しても、多くの予測モデルの基本となっている分類基準では、暫定的に付けられた可能性が非常に大きいと

思われるコードを除外するルールを設けている。

今までのところ、医療プロバイダーレベルのコーディングのパターンにおける大きな変化は検出されていない。しかし、用心して、コーディングの実態を調べるためのモニタリングまたは監査のシステムを準備することが重要であると米国アクチュアリー会 2007 年報告書は述べている。

(2) 加入者資格データに纏わる課題

加入者資格データの鮮度も実務的には大きな問題であることが、米国アクチュアリー会 2007 年報告書で指摘されている。例えば、健康保険プランの加入者の加入資格状況の変更に関する更新データを健康保険プランの購入者である企業 (雇用主) から受け取るには2か月以上かかるかもしれない。大規模な企業 (雇用主) の場合、新規加入、加入状況 (加入プラン) の変更、脱退といった状況を事後的にデータ反映させるのにもっと長い期間を要するかもしれない。

加入資格情報が古いと、リスク・スコアリングも影響を受けることが考えられるし、大きくバイアスのかかったものとなる可能性がある。例えば、加入者の死亡が加入資格データに反映されるのに数か月かかるとすれば、死亡後のその数か月間、その死亡した加入者は健康であった (医療費が発生していない) ように見えるだろう。このことは、予測モデルを現状分析に使用する場合に最も大きな影響を与えると注意を喚起している。

(3) リスク・スコアリングを行う時期および頻度

リスク・スコアの計算・割り付けをどの程度頻繁に行うか、また、どの程度速やかに行うかは、予測モデルの利用者がコントロールすることの

できる問題である。給付金請求が支払い完了になるまでに通常かかるとされるラグ期間を考え合わせれば、この期間の幅は、6 か月から 24 か月の間となろうと米国アクチュアリー会 2007 年報告書では述べている。

データの遅れは、いずれの予測モデルにとっても導入の際の問題点である。米国における実務では、モデルで個人に対する将来予測を行う場合に、計算対象加入者に一定期間の継続加入要件（例えば、評価期間中では 6-12 か月間、給付金請求から支払までの期間中では 6-18 か月間、当該加入者のリスクに対して健康保険プランが給付金を支払った期間中では 1-12 か月間は継続して加入していることを要する等）を設定することが多い。しかし、このような継続加入期間に関する制限を行うと、メディケイド集団の場合、新規に加入した人の 40%から 50%を臨床的な状態のリスク評価から除外することになり、人頭払い制度における妥当な金額を算出するための予測のパフォーマンスを著しく低下させることが懸念される。したがって、データの遅れをほとんど含まない「実験室」的な環境下での検証結果と比較して、データの遅れがどの程度モデルのパフォーマンスを低下させるのかについて知ることは重要であると米国アクチュアリー会 2007 年報告書は指摘している。

3. 今後の研究課題

米国アクチュアリー会 2007 年報告書および米国アクチュアリー会が実施した予測モデルに関する先行研究の結果を踏まえ、今後行われるべき有益な研究課題として、下記の課題を挙げ、米国アクチュアリー会 2007 年報告書は締めくくられている。

■ turnover（すなわち、継続加入条件を満たさない人の集団）のインパクトを明示的に分析

すること。

- メディケアにおけるリスク評価に使用されている HCC を分析すること。
- 異なる集団（メディケア、メディケイド、個人契約の集団、小規模契約の集団、大規模契約の集団、HMO、PPO 等）に予測モデルを適用した場合のそれぞれの集団に対する予測尺度を分析すること。
- 予測モデルによる予測において、過去の期間のコストを追加することのインパクトを集団規模別に分析すること。予測モデルの信頼性および過去の期間のコストの要素の信頼性が集団規模によってどのように変化するかを分析すること。
- 分析対象となるデータセットが異なったり、データ収集の時期が異なったりしても、予測モデルのパフォーマンスに一貫性があるか、その堅牢さを分析すること。
- 小規模集団に対する州政府による規制のインパクトを、すべてのモデルについて明示的に分析すること。規制による一般的なインパクトはモデルが異なっても似たようなものであるという仮説を検証すること。
- 米国アクチュアリー会 2007 年報告書で分析の対象としたモデルは、予測のために各個々人の特徴を把握するために使用するデータの観察期間を基本的に 12 か月以内としているが、12 か月以上の期間のデータを使用した場合の予測の改善（または想定される改善）について分析すること。
- 検査結果数値、HRA、その他の入手可能なデータをモデルに取り込んだ場合に、予測のパフォーマンスがどの程度改善するのか、その可能性を分析すること。
- 予測モデルをディジーズ・マネジメントに使用する場合のもっと適切なモデル、および、

こうした使い方におけるもっと意味のある尺度（すなわち、感度²⁴）の使用について分析すること。

<参考文献>

はじめに

Ross Winkelman & Syed Mehmud, "A Comparative Analysis of Claim-Based Tools for Health Risk Assessment", April, 20, 2007, Society of Actuaries, (visited June 24, 2008) <<http://www.soa.org/research/health/hlth-risk-assessment.aspx>>

第一部第三章

田中健司、森朋也「米国における健康保険市場と保険会社のヘルスケア事業 —2005年を中心とする概況と最近の動き—」(損保ジャパン総研クォーターリーVo.47、2007年)

西村徹、江頭達政、中村岳「米国ヘルスケア市場と保険会社のヘルスケア事業 —沿革、現状および最近の動向—」(安田総研クォーターリーVol.37、2001年)

矢倉尚典「米国のディジーズ・マネジメントにおける予測モデルの活用」(損保ジャパン総研クォーターリーVo.48、2007年)

第二部第V章

矢倉尚典、田中健司「米国における健康保険市場と保険会社のヘルスケア事業 —2004年を中心とする概況および職域市場・HIPAA・メディケアをめぐる最近の動き—」(損保ジャパン総研クォーターリーVo.45、2005年)

その他

田中滋・小林篤・松田晋哉編「ヘルスサポートの方法と実践」(財団法人東京大学出版会、2007年)

Naonori Yakura, et al., "Issues and Measures for Predictive Modeling of Medical Cost", Asian Pacific Journal of Disease Management, 2, 2008, pp.23-29 (visited July 10, 2008) <http://www.jstage.jst.go.jp/article/apjdm/2/1/23/_pdf/-char/ja/>.

²⁴ 感度等の尺度については、その考え方・特徴を付録で解説したので、適宜、参考にされたい。

<付録> 感度・特異度・ROC 曲線に関する補足説明

本編第二部の最後に、今後の研究課題として掲げられた諸課題の中に、予測モデルをディジーズ・マネジメントに使用する場合における予測精度の尺度として、「感度」の使用について分析することが挙げられている。このメッセージを理解するための補足説明として、ここでは、感度およびそれに関連する諸概念について解説する。

説明のための諸前提は以下の通りである。全体の人数が T 人である集団を想定する。この集団全員に対してスクリーニングを実施し、スクリーニングの結果によりH、Lの2グループに分け、Hグループに介入を実施するものとする。スクリーニングの結果によるグループ分けとは、①各人のデータを予測モデルに入力してモデルによる各人の指標値を計算し、②その計算された指標値が、設定した選択基準（カットオフ値）に該当する人をHグループ、該当し

ない人をLグループに分類することであるとする。図示すると《図表 付1》のようになる。

個人 i に対して予測モデルにより計算される予測指標を y_i とし、その値が大きければ大きいほどハイリスクである確率が高いとする。設定したカットオフ値を y^* とすると、選択基準は次のとおりと表せる。

選択基準： $y_i \geq y^*$ なら

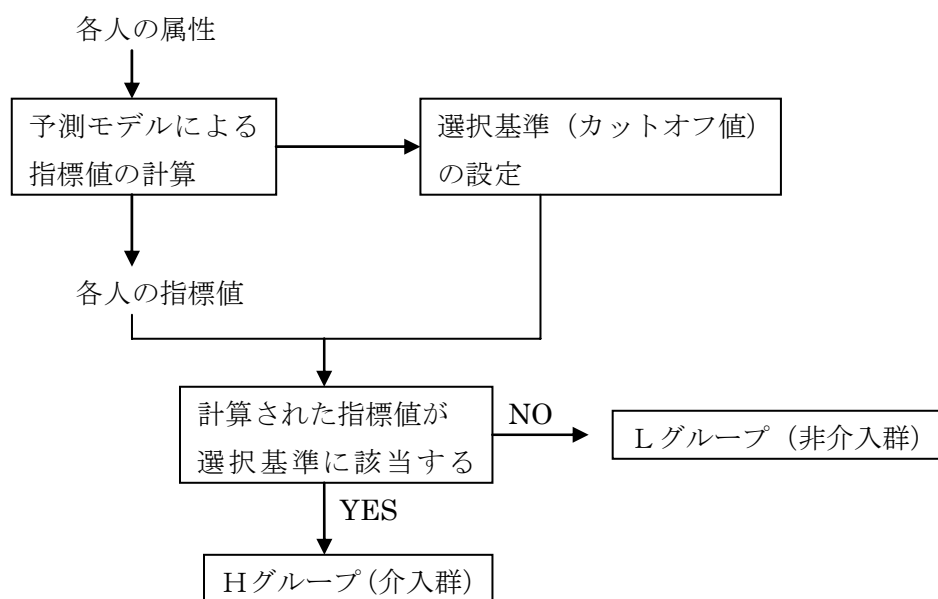
個人 i はHグループへ分類

$y_i < y^*$ なら

個人 i はLグループへ分類

全ての作用機序が解明されているわけでもなく、また将来イベントには不確実性もあるので、予測モデルによる指標値と選択基準（カットオフ値）による介入対象者の選定は、100%の確率でハイリスク者を特定することはできず、予測結果と実績（実際はどうであったのか）を把

《図表 付1》 予測モデルによる分類



(出典) 損保ジャパン総合研究所作成

握すれば、通常、《図表 付2》のようなマトリックスに人数が分布することになる。

Aはハイリスクと予測され現実にハイリスクであった（イベントの発生した）人の数、Dはローリスクと予測され現実にローリスクであった（イベントの発生しなかった）人の数を示す。予測が完全ではないので、ハイリスクと予測されたが現実にはローリスクであった人（C人）、ローリスクと予測されたが現実にはハイリスクであった人（B人）が存在する。なるべくBやCが少ない予測ほど精度の高い予測と言える。

臨床検査を例に説明すると以下のとおりである。検査結果の陽性・陰性と現実の疾病の有無で《図表 付3》のようなマトリックスを作ることができる。Cは偽陽性（疾病がないのに検査で陽性と出る人）、Bは偽陰性（疾病があるのに検査で陰性と出る人）の人数を表している。

予測精度を検討するためによく使われる指標は以下の通りである。

感度 Sensitivity : $A/(A+B)$

現実のハイリスク者のうちハイリスクと予測される者の割合

B=0 なら、Sensitivity=1

TPR (True Positive Rate) と呼ばれることもある。

特異度 Specificity : $D/(C+D)$

現実のローリスク者のうちローリスクと予測される者の割合

C=0 なら、Specificity=1

TNR (True Negative Rate) と呼ばれることもある。

偽陽性率 FPR (False Positive Rate) : $C/(C+D)$

現実のローリスク者のうち誤ってハイリスクと予測される者の割合

偽陰性率 FNR (False Negative Rate) : $B/(A+B)$

現実のハイリスク者のうち誤ってローリス

《図表 付2》 予測と実績によるマトリックス

		実績による分類	
		イベント発生	非発生
予測指標による分類	Hグループ	A	C
	Lグループ	B	D
計		A+B	C+D

(出典) 損保ジャパン総合研究所作成

《図表 付3》 臨床検査の場合のマトリックス

		疾病	
		あり	なし
検査結果	陽性	A	C
	陰性	B	D

(出典) 損保ジャパン総合研究所作成

クと予測される者の割合

陽性適中率 PPV (Positive Predictive Value) :
 $A/(A+C)$

ハイリスクと予測された人のうち現実にハイリスクである人の割合

C=0 なら、PPV=1

陰性適中率 NPV (Negative Predictive Value) :
 $D/(B+D)$

ローリスクと予測された人のうち現実にローリスクである人の割合

B=0 なら、NPV=1

同じ予測モデルを使っても、計算結果で分類する選択基準のカットオフ値を変えると、上記の感度等の数値も変わる（抽出される人数規模も変わる）ことを理解することが重要である。このことを、《図表 付4》を使って説明する。

前記のとおり、予測指標 y はその値が大きければ大きいほどハイリスクである確率が高いことを示す。カットオフ値 y^* を下げる（カットオフ値の線を右に移動させる）と選択される人の数が大きくなる（カットオフ値の線の左側にある丸の数が多くなる）。しかし、ハイリスクである確率のより低い層が含まれてくるから、陽性適中率は小さくなる。また、感度は上がるが、

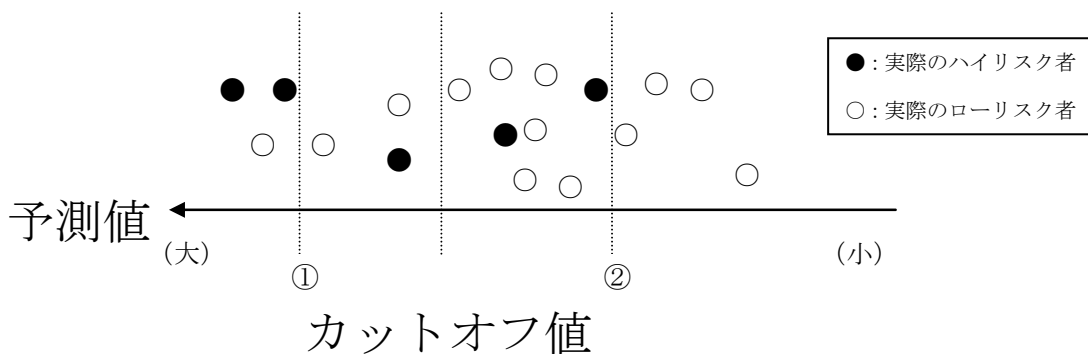
特異度は下がる。丸の数を勘定して、このことを確認する。

カットオフ値①で分類した場合の●○の数の分布は《図表 付5》の通りである。このとき、陽性適中率は 2/3、感度は 2/5、特異度は 12/13 となる。カットオフ値②で分類した場合の●○の数の分布は《図表 付6》の通りである。このとき、陽性適中率は 5/14、感度は 5/5、特異度は 4/13 となる。カットオフ値水準を①から②へ変えたことにより、陽性適中率は 2/3 から 5/14 に小さくなり、感度は 2/5 から 5/5 に上がったが、特異度は 12/13 から 4/13 へ下がっていることが分かる。

カットオフ値の設定によって予測精度を表す感度・特異度の値が変わること、また、感度を上げようとする、特異度が下がってしまうというトレード・オフの関係にあることを説明したが、カットオフ値をどのように設定すればよいのかについて、以下説明する。

一般的によく使われる方法として、カットオフ値をいろいろな値に変えて感度、特異度を算出し、縦軸を感度、横軸を（1-特異度）とした平面上に、カットオフ値を変化させたときの感度と（1-特異度）の組み合わせをプロットした ROC 曲線（Receiver Operating Characteristic Curve）と呼ばれる図を描いて、最も効

《図表 付4》 カットオフ値の水準による感度等の変化



(出典) 損保ジャパン総合研究所作成

《図表 付5》カットオフ値①で分類した場合の●○の数の分布

		実績による分類		計
		●	○	
予測指標に よる分類	ハイリスク	2	1	3
	ローリスク	3	12	15
計		5	13	18

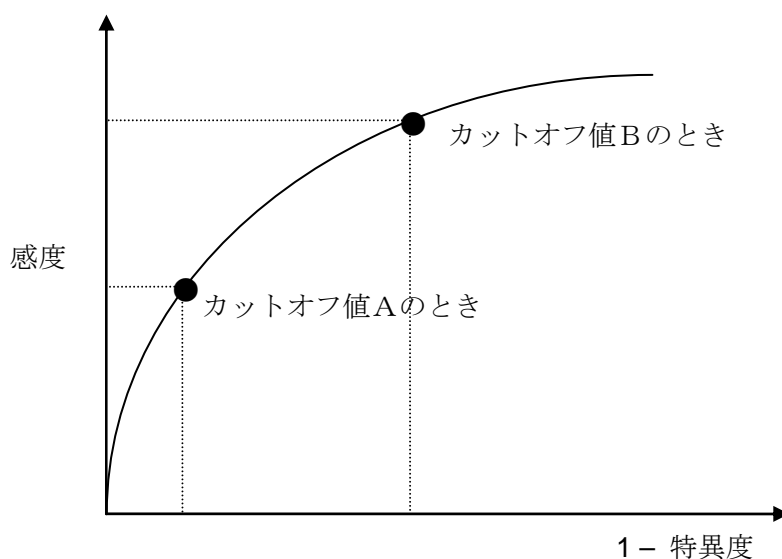
(出典) 損保ジャパン総合研究所作成

《図表 付6》カットオフ値②で分類した場合の●○の数の分布

		実績による分類		計
		●	○	
予測指標に よる分類	ハイリスク	5	9	14
	ローリスク	0	4	4
計		5	13	18

(出典) 損保ジャパン総合研究所作成

《図表 付7》ROC 曲線



(出典) 損保ジャパン総合研究所作成

率的でバランスの取れたカットオフ値水準（最も左上の点）を求める方法がある。ROC 曲線を図示すると、《図表 付7》のとおりである。通常、感度が大きくなると、特異度は小さくなる〔(1 - 特異度) は大きくなる〕ことから、ROC 曲線は右肩上がりの曲線となる。

最も効率的でバランスの取れたカットオフ値

水準は ROC 曲線上の最も左上の点である。言い換えると、これは左上からの距離が一番短い点ということである。縦軸である感度の定義域は0から1までであるから、座標 (0, 1) からの距離が最短となる点を ROC 曲線上で探せばよい。ROC 曲線上の点と座標 (0, 1) の間の距離は $[(1 - \text{感度})^2 + (1 - \text{特異度})^2]^{1/2}$ で

求められる。(1 - 感度)は偽陰性率、(1 - 特異度)は偽陽性率であるから、偽陰性率の二乗と偽陽性率の二乗の和が最小となる点を求めることになる。

数値例で確認しよう。《図表 付8》は松田晋哉他「日本型疾病管理モデルの実践」に掲載されているROC曲線である。定期健康診断データを用いて、性・年齢、血圧、脂質、糖尿病の有無、喫煙状況のリスクファクターからなるフラミンガムスタディ冠疾患発生予測モデルにより10年間疾患発生リスク(CHDリスク)を計算し、いろいろなカットオフ値でハイリスク群・ローリスク群に分類し、実際の冠疾患発生を基に感度・特異度を評価してROC曲線を描

いている。

数値例の作成を容易にするため、図示された3点の数値は《図表 付9》の通りとする。

イベントの発生割合を10%、全体人数を1,000人として数値例を示すと《図表 付10》の通りである。

カットオフ値が14.5%の場合が、座標(0,1)との距離が一番小さいことが確認できる。

カットオフ値が8.0%の場合は、感度が0.8と高いため、イベント発生者の80%を抽出できるが、一方、偽陽性率(1 - 特異度)も0.6と高いため、この基準では多くの偽陽性者も抽出されることとなり、抽出人数規模は620人と大きくなる。その結果、100人中80人は抽出できるものの、偽陽性者を540人抽出することとな

《図表 付8》ROC曲線の実例

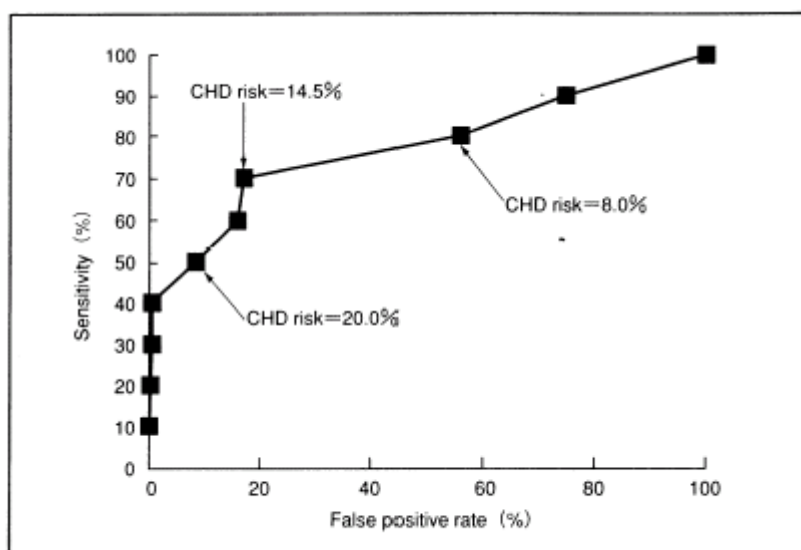


図2-1-11 The ROC curve of the CHD risks (CHDリンクにおけるROC曲線)

(出典) 松田晋哉他「日本型疾病管理モデルの実践」(じほう、2004年)

《図表 付9》カットオフ値水準別感度等の数値例

カットオフ値	感度	1 - 特異度
8.0%	0.8	0.6
14.5%	0.7	0.2
20.0%	0.5	0.1

(出典) 損保ジャパン総合研究所作成

《図表 付 10》カットオフ値水準別人数分布の数値例

		実績による分類			
		イベント発生	非発生	計	
予測指標による 分類	H	0.8 (注1) 80	0.6 (注2) 540	620	PPV: 0.129 (0,1)と(0.6, 0.8) の距離: 0.632456
	L	0.2 20	0.4 360	380	
	計	0.1 (注3) 100	0.9 900	1,000	
		実績による分類			
		イベント発生	非発生	計	
予測指標による 分類	H	0.7 (注1) 70	0.2 (注2) 180	250	PPV: 0.280 (0,1)と(0.2, 0.7) の距離: 0.360555
	L	0.3 30	0.8 720	750	
	計	0.1 (注3) 100	0.9 900	1,000	
		実績による分類			
		イベント発生	非発生	計	
予測指標による 分類	H	0.5 (注1) 50	0.1 (注2) 90	140	PPV: 0.357 (0,1)と(0.1, 0.5) の距離: 0.509902
	L	0.5 50	0.9 810	860	
	計	0.1 (注3) 100	0.9 900	1,000	

(注1) 感度

(注2) 1 - 特異度

(注3) イベント発生率

(出典) 損保ジャパン総合研究所作成

り、陽性適中率は 0.129 に留まる。

カットオフ値が 20.0% の場合は、偽陽性率 (1 - 特異度) が 0.1 と低いため、陽性適中率は 0.357 と大きな値を示す。しかし、感度が 0.5 と低いため、イベント発生者の 50% が抽出漏れとなる。

カットオフ値が 14.5% の場合が、座標(0,1)との距離が一番小さく、偽陰性者・偽陽性者の出

方が最も効率的であると考えるのである。ただし、陽性適中率は 0.280 に留まる。

なお、感度、特異度の値が同じでも、イベント発生割合が変われば陽性適中率の値そのものは変わるものである点にも留意する必要がある(《図表 付 11》参照)。つまり、感度、特異度の値が同じでも、イベント発生割合が大きいと、それだけ偽陽性者の混入が小さくなるから陽性

《図表 付 11》 イベント発生率 0.2 の場合の数値例

		カットオフ値 : 8.0%		
		実績による分類		
		イベント発生	非発生	計
予測指標による 分類	H	0.8 160	0.6 480	640
	L	0.2 40	0.4 320	360
	計	0.2 200	0.8 800	1,000

PPV: 0.250
 《図表 付 10》の最上段の例の PPV より大きい。

(出典) 損保ジャパン総合研究所作成

適中率は大きくなるのである。