

## OLAPによるDPCデータの解析

フシミ キョヒデ  
伏見 清秀\*

**目的** 平成15年にわが国独自の診断群分類DPC(Diagnosis Procedure Combination)を用いた包括評価が導入されているが、制度設計上議論となる点が残っており今後の改善が必要とされている。そのためには、毎年7月から10月にDPC包括評価対象病院から収集される膨大な調査データの効率的で正確な解析が不可欠であるが、その手法は確立されていない。本稿では、DPC調査データの解析にOLAP(On Line Analytical Processing)を活用する方法を検討し、その実効性を検証することを目的とした。

**方法** 調査データからリレーショナルデータベースと多次元データキューブを構築し、ネットワークおよびローカルファイルを介してクライアントソフトを用いてOLAP解析を実施した。OLAPキューブは定義表の項目に沿って7から26個の集計軸を設定し、病名集計軸に関してはICD10コード、DPC傷病名分類、MDC分類レベルの3段階の粒度で集計する設計とした。副傷病は併存症と続発症に分けてその影響度を集計した。

**結果** システム構築、分析の実施、データの配布の実現可能性と有用性が確認された。対話的な分析により、手術グループの差異による在院日数への影響の違い、化学療法、放射線療法などの在院日数、診療報酬点数への影響の違いなどが示された。また、副傷病の解析では、DPC傷病分類レベルの集計によって、循環器系疾患においては呼吸不全、腎不全の影響が大きく、消化器系疾患においては肺炎の影響が大きいなど、主たる疾患によって医療資源必要度に影響を与える副傷病が異なることが示された。

**結論** DPC調査データの解析におけるOLAP法の活用の実現可能性と有用性が示された。特に現在のDPC定義表では十分に整理されていない副傷病の評価については意義が大きいと考えられる。DPCの恒常的な見直し作業にOLAP法が活用され、データに基づく医療評価の1つのツールとしての地位が確立されることが期待される。

**キーワード** 包括評価、診断群分類、医療費、医療評価、探索的分析

### I 緒 言

平成15年にわが国独自の診断群分類(DPC: Diagnosis Procedure Combination)を利用した入院医療診療報酬の包括評価が特定機能病院等に導入された<sup>1)</sup>。その後、この診療報酬支払いの仕組みは定着しつつあると考えられるが、こ

の制度は完成したものではなく、診断群分類そのものと診断群分類を用いた包括評価のあり方にはまだ改善すべき点が多くあると考えられる。DPCは当初の設計時から特定機能病院などから収集された実際のデータに基づいて構築されてきた経緯があり、また、包括評価に関しても実データに基づく分析がその基本にある。現在の

\* 東京医科歯科大学大学院医療政策学講座医療情報・システム学分野助教授

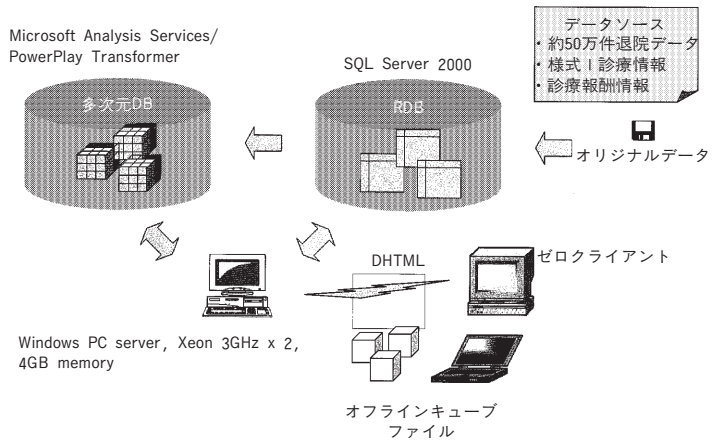
様々な改訂作業もデータ分析によるところが大きく、DPCが実際のデータに基づいて改訂が検討されている点は、診療報酬制度の透明性の確保や科学的かつ客観的な医療費支払体制の構築などの観点からは、高く評価されていいであろう。

一方、DPC包括評価のために収集されるデータは量が膨大であるばかりでなく、データの内容が数値、数字コード、文字など多様であり、また、DPC自体が非常に複雑なロジックで構成されているため、古典的で単純な統計集計のみ

では、その実態を把握するのが困難であると考えられる。特に、DPCのロジックが診断名、手術分類、処置分類、副傷病分類、重症度分類などにより多重的に構築されていることから、データ分析もこれら様々な観点から検討する必要がある。

このような多層的で大量のデータの集計解析には、OLAP(On Line Analytical Processing)が有用であるとされている<sup>2)</sup>。OLAPは直訳すると「オンライン分析処理」、つまりコンピュータと対話的、継続的に行う分析処理という意味であるが、一般的には多次元分析と訳されることが多く、近年は、企業経営分析やマーケティング解析などのビジネス・インテリジェンスの1つとして広く活用されるようになってきている。多次元分析には、文字どおり複数の次元(ディメンジョン)で分析・集計をすること、および集計の粒度に階層を持たせられることの2つの意味がある。前者は、一般的な集計表が二次元であるのに対して、集計軸を分析実行者の要求によってリアルタイムで交換することにより仮想的に多次元を実現するもので、この集計軸の交換処理を「ダイス」処理とよぶ。後者の集計粒度の可変性とは、あらかじめ集計粒度を細かいものから粗いものまで用意しておくことによ

図1 DPC解析OLAPシステム構成図



注: DB: データベース, DHTML: ダイナミックHTML, ゼロクライアント: 専用のクライアントソフトも持たない端末

り、分析実行者が概略を把握する目的の粒度の粗い集計と、詳細を検討する粒度の細かい分析を必要に応じて実行できることである。概略を見る処理を「ドリルアップ」、 「フォーカスアップ」、詳細を見る処理を「ドリルダウン」、 「フォーカスダウン」と呼ぶ。

従来は大規模なデータウェアハウスと大型で高価なコンピュータを必要とする方法であったが、近年の急速な情報技術の進歩により、研究室レベルでもOLAPを活用することができるようになってきている。そこで、本研究では、OLAPを活用してDPCの調査データを解析する方法論とその解析手法の実効性を明らかにすることを目的とする。

## II 方法

OLAP分析システムは図1のように構築した。収集されたデータはデータサーバのMicrosoft SQL Server® SP3リレーショナルデータベースに収納され、Microsoft Analysis Services®, またはCognos PowerPlay Transformerを用いて解析用のOLAPキューブが作成された。OLAP分析ソフトとして、Business Object社 Crystal Analysis® Version10, Cognos社

PowerPlay for Windows®, またはMicrosoft Excel®を使用した。多次元キューブへの接続方法として、ローカルネットワークを介してOLAPデータサーバへ接続する方法、および多次元キューブを独立したオフラインキューブファイルとして作成する方法を用いた。後者はオフライン分析、オフラインプレゼンテーション、OLAPキューブの配布として使用した。

OLAP集計軸(ディメンジョン)の設計は図2に示すように、ICD10コード, DPCコード, MDC(Major Disease Category)の階層を持つ医療資源を最も使用した病名(医療資源病名)と副傷病等の分析軸, 手術コード, 手術分類, 手術の有無の階層を持つ手術軸, および階層を持たない複数の処置等に関する軸とした。作成したディメンジョンとその要素の一覧を表1に示す。ここに示したものの以外に、医療機関のディメンジョンなどを追加することもできる。多次元データキューブは16のMDCごとに作成し、それぞれのキューブにおけるディメンジョンの総数は7から26となった。集計対象は、平均在院日数、1日または1入院当たりの全診療報酬点数、診療区別点数、包括評価部分点数(DPCごとに定められた包括評価される部分の点数で、入院料、検査料、投薬料、注射料、1000点未満の処置料などの大部分が相当する)とした。

DPCディメンジョンおよび併存症と続発症の副傷病ディメンジョンではMDCからICD10までの粒度階層を持たせて、MDC単位から傷病単位の分析ができるように設計した。DPC定義表<sup>3)</sup>に示されている手術・処置等1, 手術・処置等2は様々な組み合わせの可能性に対応するため、定義表に示された診療行為をそれぞれ独立のディメンジョンとして設定した。

OLAPを活用した副傷病の分析においては、入院時にすでにあった併存症と入院後に発症した続発症を独立して分析できることと、副傷病をICD10の病名コードのみでなくそれらをDPCの上位6けたの傷病名分類で集約したレベル(DPC6)

図2 OLAPの概念図

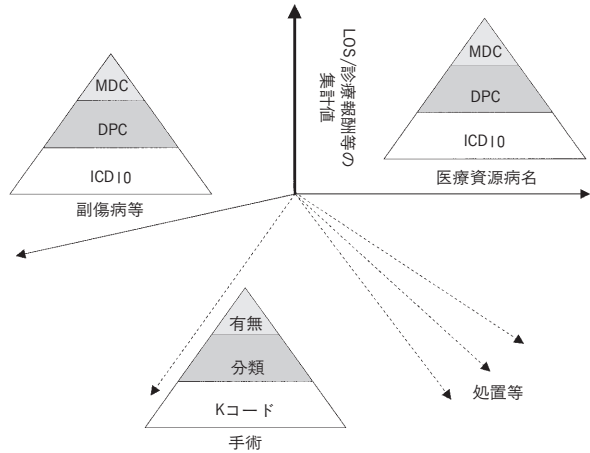
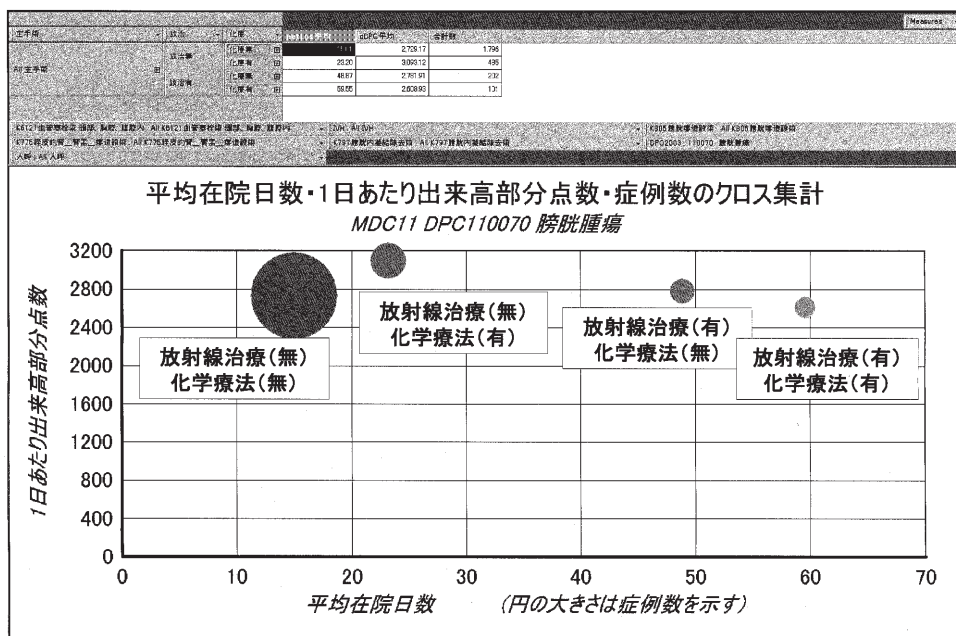


表1 DPC解析用OLAPキューブのディメンジョン構成

ディメンジョン名称	ディメンジョン数	階層数	定義
DPC	1	3	MDC/DPCコード
入院目的	1	1	検査・診断/その他診療
手術	1	2	手術グループ/手術コード
手術・処置等	1-20	1	手術・処置等名称
併存症	1	3	MDC/DPC/ICD10コード
続発症	1	3	MDC/DPC/ICD10コード
調査年	1	1	2002/2003

およびMDCで集約したレベルで分析できることの2点を満たすように設計した。DPCの調査で収集される4つの「入院時併存症名」のICD10コードを「併存症」とし、4つの「入院後発症疾患名」のICD10コードを「続発症」として2つに区分する。1症例に複数の併存症、続発症が存在することがあるため、OLAP集計・分析の目的ではそれらの中から最も主要なものを選択する必要がある。そのため次のようなデータ処理を行った。①医療資源病名DPC6(主DPC6)ごとに1入院当たり包括評価部分点数の平均値(cDPC)を求める。②「併存症」と「続発症」それぞれについて主DPC6ごと、副傷病ICD10ごとのcDPC平均値を集計し、主DPC6ごとのcDPC平均値との差分( $\Delta$ cDPC)を計算する。③主DPC6ごとに「併存症」と「続発症」について、症例数の平方根で重み付けをした $\Delta$ cDPC順に副傷病重症度順位表を作成する。④各症例の「併存症」「続発症」のうち重症度順位有意が最も上位の

図3 OLAP分析インターフェース画面



ものをキーとして、副傷病ディメンジョンを作成する。以上のプロセスによって各症例の最も重要な「併存症」と「続発症」が1つずつ決まり、副傷病の影響度を様々な視点から集計分析ができるようにOLAPキューブを作成した。

実際の分析画面を図3に示す。上部パネルの集計表と下部パネルのバブルチャートを併用して分析を進めることができる。下部のグラフは棒グラフや折れ線グラフなどにも変更することができる。リアルタイム解析では、マウスを用いた操作で集計軸を入れ替えたり、集計の粒度を変更したりする作業をほぼ瞬時に実施することができる。例えば、特定の手術での処置の影響が疑われた場合は、その手術に「ドリルダウン」し、処置の組み合わせを分析軸に移動することで、即座に情報を把握することができる。

### III 結 果

DPC110070膀胱腫瘍の解析結果を例にとり、図3に放射線治療と化学療法の有無が平均在院日数、1日あたり包括部分点数にどのような影響を及ぼしているかを、それぞれに該当する症

例数とともに示す。この例では、化学療法に比べて放射線療法の影響が大きいこと、いずれの治療も1日当たり包括部分点数にはあまり影響を与えていないことなどがわかった。これに加えて手術の有無で条件を絞る、他の処置の影響を加える、他のDPCと比較するなどの解析を対話的にリアルタイムで実施することが可能であった。

次いで、集計表を用いた解析の例として、入院中の手術、処置等が在院日数、診療報酬点数などにどのような影響を与えるかについての分析例を以下に示す。表2の膀胱腫瘍の手術に関する症例数、平均在院日数、1日当たりの包括評価部分点数の分析では、大きくみると手術グループ01と03の膀胱悪性腫瘍手術等に関する在院日数の比較的長いグループと、手術グループ06と02の経尿道的手術等の在院日数の比較的短いグループの2つがあることがわかった。

さらに、表3に示す膀胱腫瘍の処置に関する多次元集計を用いて、定義表に示されている中心静脈栄養、人工呼吸、化学療法、放射線治療のそれぞれ在院日数、診療報酬点数などへの影響と相互関係を解析した。処置が全くないグル



ープの平均在院日数約14.8日と比較すると、いずれかの処置があるグループの在院日数が30日から50日前後と大きく延びていることがわかった。一方、診療報酬点数に注目すると、多少のばらつきはあるが特定の処置が点数に大きく与える傾向は見取れなかった。これらの状況からは、現在のDPC分類で4つの処置のいずれかがある場合に「手術・処置等2あり」として評価している点で大きな問題はないと考えられた。

最後に、OLAPによる副傷病分析の実例としてMDC5循環器系疾患を例にして示す。図4はMDC5の医療資源病名と併存症の1入院当たりの医療費に与える影響を3次元棒グラフに示したものである。OLAPのドリルダウン操作で条件を絞ることにより、循環器系疾患では慢性腎不全を併存症に持つ症例が多く、またこれらの症例では併存症のない症例に比べて非常に大きい医療費がかかっていることが示された。MDC5の循環器系疾患について副傷病をDPC6単位で集計すると図5となり、呼吸不全、腎不全などの影響度が大きいことがわかった。同様の解析をMDC6消化器系疾患について行うと図6となり、肺炎、胆管炎などが上位に上がってきて、副傷

表2 DPC110070膀胱腫瘍の手術に関する多次元集計表

手術対応コード	合計数	LOS平均	dDPC平均
有	2 039	20.5	2 721
01	221	54.1	2 841
K8031 膀胱悪性腫瘍手術切除	15	32.3	2 532
K8032 膀胱悪性腫瘍手術全摘 (腸管等を利用して尿路変更を行わないもの)	33	54.3	2 812
K8033 膀胱悪性腫瘍手術全摘 (尿管S状結腸吻合を利用して尿路変更を行うもの)	3	53.7	2 781
K8034 膀胱悪性腫瘍手術全摘 (回腸又は結腸導管を利用して尿路変更を行うもの)	116	54.3	2 954
K8035 膀胱悪性腫瘍手術全摘 (代用膀胱を利用して尿路変更を行うもの)	54	59.6	2 706
02	39	17.7	2 685
K799 膀胱壁切除術	2	23.0	2 649
K800 膀胱憩室切除術	20	13.6	2 785
K802 膀胱腫瘍摘出術	9	10.2	2 749
K8171 尿道悪性腫瘍摘出術 摘出	4	48.5	2 203
K8172 尿道悪性腫瘍摘出術 内視鏡による場合	3	7.3	2 554
K8173 尿道悪性腫瘍摘出術 尿路変更を行う場合	1	63.0	2 509
03	6	43.5	2 736
K8011 膀胱単純摘除術 腸管利用の尿路変更を行うもの	1	74.0	3 941
K790 尿管皮膚瘻造設術	5	37.4	2 495
06	1 669	14.6	2 694
K8036 膀胱悪性腫瘍手術経尿道の手術	1 669	14.6	2 694
96	4	45.8	2 336
K812 回腸(結腸)導管造設術	4	45.8	2 336
97	100	43.0	2 921
無	545	22.6	3 082

注 LOS：在院日数、dDPC：1日当たり包括評価部分点数

表3 DPC110070膀胱腫瘍の処置に関する多次元集計表

IVH	人工呼吸	化学療法	放射線治療	合計数	LOS平均	LOSCV	dDPC平均	dDPCCV
無	無	無	無	1 757	14.8	0.96	2 719	0.21
		有	無	188	48.0	0.53	2 605	0.23
	有	無	無	481	22.9	0.93	3 096	0.34
		有	無	100	59.9	0.50	2 591	0.19
有	無	無	無	9	31.8	0.91	3 771	0.41
		有	無	11	59.0	0.42	5 754	1.32
	有	無	無	2	58.5	0.57	3 055	0.14
		有	無	1	29.0	0.00	4 371	0.00
有	無	無	無	6	51.7	0.38	2 672	0.25
		有	無	1	83.0	0.00	2 739	0.00
	有	無	無	2	56.0	0.00	2 322	0.09
				1	54.0	0.00	9 176	0.00

注 IVH：中心静脈栄養、LOS：在院日数、CV：変動係数、dDPC：1日当たり包括評価部分点数

病の影響は主疾患によって大きく異なることが明確に示された。

## IV 考 察

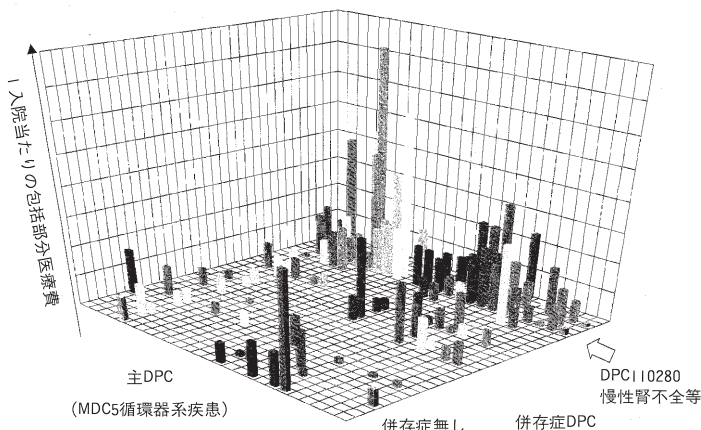
DPCと包括評価は実際の診療報酬データに基づいて決定されているが、そのデータ分析は非常に複雑であるため、DPC診断群分類の改訂作業や包括評価の点数の検討を困難にしている。本研究では、このような複雑な構造を持つ膨大なDPC関連データの分析にOLAP法を適用することの実現可能性とその有効性を示した。

OLAP分析システムの構築およびデータベースとOLAP分析キューブの生成方法には特殊な技術を要しないことを示した。

本研究では、OLAPを用いたDPCデータ解析の方法論とその結果の一部を紹介した。OLAPはここ数年の情報技術の進歩で低価格化が進むとともに、多種のソフトウェアが供給されシステムの構築が平易となって身近な存在となり、ようやく成熟期に入ってきたものと思われる。一方、OLAP自体は単なる集計表の集まりにすぎないので、得られた分析結果をどのように生かしていくのかも重要な課題である。例えば、OLAPで得られた知見をもとに仮説を立て統計的に検証するなどへの応用が考えられるであろう。多角的な解析を求められることの多い医療関連の多くの分野では、今後ますますOLAPの有用性が増してくると考えられる。

実際の分析は、OLAP本来の機能を生かしてリアルタイムに対話的に解析を行う方法と、集計結果を多次元集計表として作成した上で印刷物などの形で検討を行う方法の2つの方法が利用できる<sup>4)5)</sup>。データ分析環境と分析担当者のコンピュータリテラシーの程度に応じてこの2つの方法を使い分ける必要があるだろう。前者の対話的解析は、クライアント端末を使用してサーバーに接続して解析を行う方法と、ローカルキュー

図4 OLAPによるDPC別の併存症の影響の解析



注. 左軸に医療資源を最も消費した病名のDPC疾患分類, 右軸に併存症病名のDPC疾患分類, 上向き軸に1入院当たりの包括部分医療費を3次元棒グラフに示す。医療資源を最も消費した病名はMDC5循環器系疾患にフォーカスタウンし, 併存症はMDC11腎泌尿器系疾患にフォーカスタウンして医療費への影響を図示している。

図5 MDC5(循環器系疾患)で影響度の大きい併存症とその相対医療費

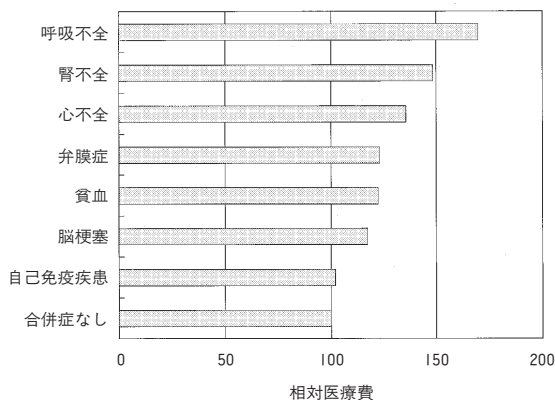
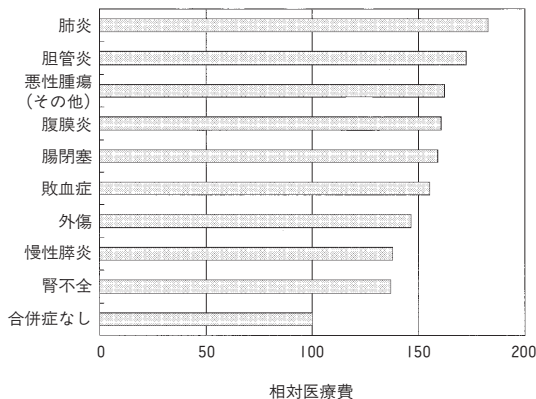


図6 MDC6(消化器系疾患)で影響度の大きい併存症とその相対医療費



ブファイルを用いて個人のノートパソコンなどでオフライン分析を行う方法の2つを状況に応じて選択しうる。オフラインでもサーバーを用いた解析とほぼ同様の解析ができるため、時間と場所に制約を受けずに効率的に分析を進めることができる。一方、集計表を印刷物の形で用意することは、せつかくのOLAP手法の長所である柔軟性を損なうものであるが、集計項目を適切に選択することで汎用性の高い集計一覧表を用意することも可能であろう。

OLAPシステムを活用した分析の進め方の一例としては、DPC上位6けたで示される、ある特定の傷病群において、手術や処置の組み合わせがどのように在院日数に影響しているかを手術分類ごとあるいは手術コードごとに解析することが可能であり、また、MDC(Major Disease Category)ごとに副傷病の影響度はどのように異なるか、DPCごとではどうかなどの検討もリアルタイムで実施することができる<sup>6)</sup>。例えば、結果に示したように、膀胱腫瘍では各処置の間に1日当たりの点数についてはあまり大きな差がないことから、在院日数と症例数を基準にDPCの再グルーピングが可能であろうと考えることができる。逆に、仮に特定の処置、例えば化学療法などが独立して在院日数、診療報酬点数などに非常に大きな影響を与えていると考えられる場合は、包括評価の設定において独立した分類を与えることが妥当になると考えられる。その他の例としては、医療資源病名が特定のICD10の場合に1日当たりの包括評価部分点数が高値であることなどが示されれば、この疾患は独立したDPCとするべきではないかなどと検討する材料となる。

最後に、DPCの構築において最も検討が立ち後れている副傷病の解析について検討する。本研究で示したような解析を進めて、MDC単位あるいはDPC6単位で副傷病を分析することにより、包括評価にとって重要な副傷病がより明確になってくると考えられる。臨床病態の検討と併せてDPC定義表の改訂につなげられるであろう。この方法は、主DPCを第1DPCとし、副傷病を第2DPC(secondary DPC)として評価する

フランスなどの諸外国で用いられている方法につながるものと考えられる。副傷病を個別のICD10で評価することは細かくなりすぎて非常に難しい。しかし、副傷病をDPC6単位である程度まとめて評価すると分析を進めやすく、その結果も理解しやすくなる。現在のDPC定義表ではICD10で副傷病が定義されているため、定義表内で矛盾が生じ、もれや重複が多数認められるが、DPC6単位で整理し直すことによりこのような問題点は解消されるであろう。また、併存症と続発症を分離して評価できる点も強調したい。包括評価においてもこれらを整理して評価していく方向が望ましいであろう。現在の包括評価では、副傷病の有無で点数設定が逆転するような矛盾が発生しているが、これは定義表の副傷病の整理が十分に行われていないことに起因している可能性が高い。データ解析に基づいて影響度の大きい副傷病のみを取捨選択し、定義表を整理することでこれらの矛盾は解消されるであろう。

## V 結 語

データに基づく科学的なDPCの改訂はPDCA(Plan-Do-Check-Act)サイクルに相当する形で進められていくのが理想であろう。データの解析方法、DPCの評価方法、定義表の見直しのためのロジック、包括評価用の分類の設定などのすべてのプロセスが透明化され、そのロジックが明確に示されなくてはならないと考えられる<sup>7)</sup>。DPC包括評価に参加している医療機関は大変な労力を費やしてデータ収集に協力しているのであるから、そのデータを有効に活用するとともに、その活用のされ方やDPC改訂への反映のされ方なども明確にされなくてはならない。できれば、諸外国にならぬ、DPCデータセンターあるいはケースミックスオフィスなどが組織され、政府や利害関係者から独立した機関として公正にデータ収集、データ解析、DPCの改訂とその評価などが進められるのが望ましいが、研究班あるいは作業班でこれらの作業が実施される場合は、その過程の透明性の確保が肝要であろう。この点で、本研究で示したOLAP法は比較

的平易なデータ解析法であるとともに、詳細な集計結果をプライバシーに関する情報を完全に取り除いた上で、安全にかつ再利用しやすい形で公開できる優れた方法でもある。このような方法を活用してデータの活用を安全にかつ効率的に進めていくことが望ましいと考えられる。

なお、本研究の一部は厚生労働科学研究費補助金政策科学推進研究事業「急性期入院医療試行診断群分類を活用した調査研究」(H13-政策-034)、「診断群分類を活用した医療サービスのコスト推計に関する研究」(H16-政策-027)の成果に基づいている。

#### 文 献

- 1) 松田晋哉. DPCとこれからの医療 第1回 DPCとは何か. 社会保険旬報 2003 ; 2189 : 16-21.
- 2) Elik Thomsen. Olap Solutions : Building Multidimensional Information Systems. Hoboken : John Wiley & Sons Inc.
- 3) DPC電子点数表. 診断群分類点数表のてびき. 社会保険研究所. 2004.
- 4) 伏見清秀. 平成15年度厚生労働科学研究「レコードリネージ解析を利用した医療経済面を含めた医療関連統計調査の活用方法等に関する研究」総括研究報告書. 2004.
- 5) 伏見清秀, 松田晋哉. 平成15年度厚生労働科学研究「急性期入院医療試行診断群分類を活用した調査研究」研究報告書別冊5「特定機能病院における入院医療の包括評価OLAPデータ集」. 2004.
- 6) K. Fushimi, H. Hashimoto, Y. Imanaka, et al. Refinement of DPC classification facilitated by OLAP analysis of patient profiles and medical procedures. Proceedings of the 20th conference of PCS/E, 2004 ; 395-402.
- 7) 伏見清秀. 「DPCとこれからの医療」第9回マネジメントツールとしてのDPC~医療機関プロファイリングと疾病管理への応用の検討~. 社会保険旬報 2004 ; 2216 : 22-8.