



ESRI Research Note No.36

医療の質の変化を反映した実質アウトプット・価格の把握 ～方法論の整理～

杉原茂、市川恭子、今井健太郎
野口良平、岡崎康平、小池健太

January 2018



内閣府経済社会総合研究所
Economic and Social Research Institute
Cabinet Office
Tokyo, Japan

ESRI Research Note は、すべて研究者個人の責任で執筆されており、内閣府経済社会総合研究所の見解を示すものではありません（問い合わせ先：<https://form.cao.go.jp/esri/opinion-0002.html>）。

ESRI リサーチ・ノート・シリーズは、内閣府経済社会総合研究所内の議論の一端を公開するために取りまとめられた資料であり、学界、研究機関等の関係する方々から幅広くコメントを頂き、今後の研究に役立てることを意図して発表しております。

資料は、すべて研究者個人の責任で執筆されており、内閣府経済社会総合研究所の見解を示すものではありません。

The views expressed in “ESRI Research Note” are those of the authors and not those of the Economic and Social Research Institute, the Cabinet Office, or the Government of Japan.

医療の質の変化を反映した実質アウトプット・価格の把握 ～方法論の整理～

杉原茂、市川恭子、今井健太郎、野口良平、岡崎康平、小池健太

1 はじめに

物価指数の向上を巡る議論は、米国上院財政委員会の「消費者物価指数（CPI）を検討するための諮問委員会」報告（「ボスキンレポート」1996年）を契機に発展してきたと考えることができる。1990年代後半以降、米国ではCPIを巡る議論が盛んになり、「ボスキンレポート」ではCPIに関する4つのバイアス（消費者がより安価な代替品を購入しても反映されない、消費者が安売り店への購買ヘシフトする動きが反映されない、品質向上が適切に反映されない、新製品が発売されてもすぐに反映されない）が指摘された。これを契機に世界的に物価指数に対する注目が高まり、研究プロジェクト等が行われていった。特に品質や新製品といった観点を踏まえて、物価指数の動向を適切に把握するのは重要なポイントである。

物価指数に関する議論の流れは主に次のようなプロセスを経て、現在に至ると考えられる。まずは財・サービス全般の物価指数についてボスキンレポートが指摘するバイアス等への関心が高まり、研究が開始された。それに続いて特に医療については、米国 National Bureau of Economic Research (NBER) では Cutler らを中心に医療の質と価格の関係に関する研究プロジェクトが取り組まれ始めた。また、医療サービスについて、米国では市場価格が存在するため物価指数（デフレーター）自体に質を反映する観点から研究が行われたが、英国では医療は公的価格となるためデフレーターではなく実質アウトプット（産出）について直接質を調整するアプローチが取られた。英国では非市場サービス（政府サービス）にかかるアウトプットや生産性の計測手法について Atkinson(2005)を契機として、研究が行われ始めた。Atkinson(2005)や York 大学の研究プロジェクトを受けて、死亡率等の明示的な質指標を織り込んだ実質アウトプットを推計し、参考指標群として公表されている。また、欧州全体では eurostat が実質アウトプットを直接推計する手法を推奨し、欧州各国の国民経済計算（SNA）で採用されてきている。

こうした国際的な動向を踏まえ、我が国においても「統計改革の基本方針」（平成28年12月経済財政諮問会議決定）において「医療・介護、教育の質の変化を反映した価格の把握手法（中略）について研究する」こととされた。これを受けて、内閣府経済社会総合研究所(ESRI)は、2017年度から医療・介護、

教育の質の変化を反映した価格の把握手法について研究を開始した。

本稿では、各国が質の調整を行うに至った経緯・問題意識を踏まえ、基本的なアウトプット計測における質の考え方を示した上で、医療の質研究にかかる特徴を述べる。さらに、医療の質の変化を反映した実質アウトプット・価格の把握手法の概念を整理した上で、各地域・国の現状を述べ、本研究において取り組むべき今後の検討課題と使用候補となるデータの特徴を示す。

2 質の調整を行うに至った各国の経緯

医療において質の調整が特に関心を集めるのは、医療サービスに対する消費者の評価を計測することが簡単でないということによる。例えば、規制がなければ、市場における価格と需要量は消費者の評価を端的に表わすものであり、こうした市場における価格と需要量を利用してサービスの質を計測することが可能である（ヘドニック分析等）。しかし、医療においては、多くの国で、価格が公定され、また、供給量にも規制がかかっているため、市場における価格や需要量が消費者の評価を必ずしも反映していない。そうした場合には、質の計測手法や計測のためのデータ等について特別の工夫が必要となる。さらに、供給者に対する規制が少なく自由な市場において医療サービスが提供されている国でも、医療保険によって消費者が直面する市場価格と医療サービスの価値がかい離するため、消費者の評価を市場価格から単純に推計することはできず、医療サービスの結果、得られる生活の質（QOL）の向上等の便益を何らかの「命の値段」（Value of Statistical Life（VSL）あるいはQALYの価値等）で評価する等通常とは異なる直接的なアプローチが必要となる。

SNAにおける医療の質の捉え方に関しては、各国特有の問題意識に基づいて質の調整に至っている。米国、英国、EU（欧州）及び日本についてそれぞれ経緯をまず見てみよう。

米国では、1980年代以降医療デフレーター（CPI）が他の財・サービスに比べて急増し、それにともない実質アウトプットが増加しないという状況にあった。しかしながら、Cutler, et al. (1998)が1983年～1993年にかけて心疾患治療に関して質の調整を行うと、CPIは過大推計されているという研究結果を得た。つまり、質の調整をすると医療デフレーターは公表された医療CPIほどには上昇しないと指摘したのである。これは、実質アウトプットは質の調整をしない場合と比べて増大することを意味する。質の調整の動向次第で実質アウトプットの動向も大きく影響を受けるという研究結果を受けて、医療における質の調整の重要性が認識され、NBERやBureau of Economic Analysis（BEA）

といった研究機関・政府機関において研究が蓄積されていった。

英国では、国民保険サービス（NHS）として公費負担で医療が提供されている。公費負担であり医療費抑制の観点から、サービスは抑制的に提供される傾向にあり、質が低いのではないかという問題意識が生じていた。例えば、入院待ち、手術待ちが恒常化しており、医療の質が政治的議論の論点となっていた。こうした状況を踏まえて、医療費を抑制しつつも質の高いサービス（アウトプット）を提供できないかという観点から、生産性の計測が注目されるようになった。こうした中、英国政府から諮問され作成されたのが上述の Atkinson Report (Atkinson (2005)) である。本レポートでは、医療を含む政府サービスについて、サービスのアウトカム（例えば、医療サービスの治療の結果である死亡率等）を質として認識し、それをアウトプットに反映させる方法を提示している。こうした動きを踏まえて、SNA 本体では分類の細分化により実質アウトプットを計測することで質の調整を行っているが、SNA 本体とは別に York 大学を中心とした研究をベースとして国家統計局（ONS）が参考指標群として生産性指標を公表し、その中で質の明示的な調整を行っている。

欧州各国では、EU 加盟各国の統計部局のキャパシティーのばらつきを踏まえ、複雑な手法を追求していないようである。もともとインプット方式（アウトプット＝インプットと仮定）を用いていたが、英国の SNA 本体同様に細分化による実質アウトプットを直接計測する手法を採用している国が多い（フランス、イタリア等）（第1表参照）。

日本では、医療は SNA において定義上市場型サービスと分類されるが、皆保険制度のもと医療サービスには公的価格が用いられており、本来市場サービスであれば消費者（患者）の評価と一致するはずの価格が、実際には評価とかい離している可能性が高い。日本の医療の質を巡る政策課題は米国や英国等と共通する点も多々見受けられる。日本では医療費抑制の観点から入院医療における包括払い方式として DPC (Diagnosis Procedure Combination 診断群分類) が 2003 年に導入されたが、米国では先行して 1983 年に高齢者対象の入院医療支払方式に DRG (Diagnosis Related Groups 診断群分類) が採用されている。また、1900 年代初頭に始まった米国の医療の質改善運動を範として、日本でも医療の質改善の取組が活発化してきた。もう少し消費者（患者）のレベルで言えば、「3 分間診療」に代表される粗診粗療への批判は英国の状況に共通している。こうした医療に対する質改善の取組や医療における技術革新の成果が適切に反映されているのかといった評価について関心が高く、教育等の他の政府サービスの分野に比べて比較的研究が進んでいる。

3 アウトプット計測における質の考え方

医療の質の変化を反映した実質アウトプット・価格の把握手法について整理する前に、まず基本的なアウトプット計測における質の考え方を述べておこう。

アウトプットの計測について概念的には「数量」「質」「価値」の3要素からなると考えられ、以下の通り示される。

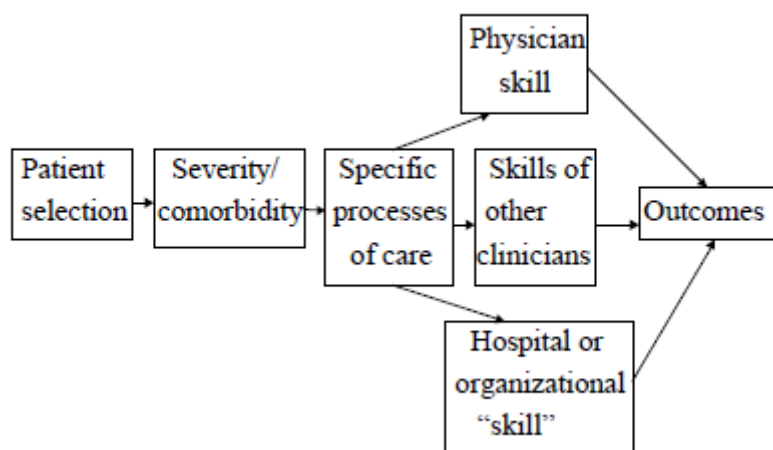
$$\text{アウトプット} = \text{数量} \times \text{質} \times \text{価値} \quad (1)$$

ここでのアウトプットは名目値である。これを医療サービスに置き換えると「数量」は疾病毎の患者数や治療件数、「質」は「医療の質」である。ここで数量を「疾病」を軸とする理由は次のとおりである。例えば、患者は体調が悪くなるとまずは診療所にかかり、それが悪化すると病院に入院し、退院してリハビリするという医療サービスの流れを思い浮かべてみよう。本来であればこうした一連の治療行為（エピソード）を一気通貫で把握して医療サービスの動きを把握したいところである。ただ、実際のデータ制約上、エピソードを包括的に把握するのはかなり難しい。その場合は次善の策として疾病といった観点（分類軸）で把握していくことが現実的な対応として考えられる。

「医療の質」指標としては、構造指標、プロセス指標とアウトカム指標がある。産業組織論では「構造」「プロセス」「アウトカム」といった3要素から質を評価するのが一般的である。医療サービスに沿って述べると、構造指標とは、例えば高度な医療体制が整備された病院等治療プロセスの基盤となる病院施設等である。プロセス指標とは、治療プロセスにおいて医療の質を決める要素の充足度合の指標である。例えば、ある疾病に効果があると判明している薬剤がある場合、その薬剤を投与する割合の高さが挙げられる。アウトカム指標とは、まさに医療による成果指標で、生活の質（QOL）や患者満足度、死亡率等がそれに当たる。

ここで、医療サービスの提供に関してプロセス、アウトカムといった概念を Halm, et al. (2000) の Figure1 に沿って以下の通り述べておこう。

第 1 図



ある治療の対象となる患者が選定され、その患者は重症度・併存疾患の状況により分類される。その次に治療プロセス（医者のスキル、他の医療関係者のスキル、病院等のスキル）を経て医療成果（アウトカム）が発生するという流れである。この場合、医療の質の変化とは、初期条件である「患者の重症度・併存疾患の状況」と「アウトカム」の差である。また、医療成果（アウトカム）は治療プロセスを経てアウトカムが生み出されるということから、アウトカムには治療プロセスの要素が包含されている点も留意が必要である。

次に「価値」であるが、医療サービスに対する評価、支払うに値すると考える価値で、例えば命の値段（余命の金銭価値）等が挙げられる。

4 医療の質研究の特徴

医療、介護、教育等の公的サービスの中でも、医療の質に関する研究は比較的研究が進んでいる分野、また、取り組みやすい分野と考えられる。その理由は主に以下の3点に整理できる。

第1に明確な質の概念が挙げられる。例えば、教育の質を定義するのは簡単なことではないが、医療の質はQOLに集約されることについて大きな異論はないであろう。その具体的な構成要素として、死亡率（生存率）や合併症、再入院率といった指標が代理変数となり得る。

第2に、比較的に進んでいる研究分野ということが指摘できる。1980年代以降、米国のNBERを中心に医療の質の計測やその価格に関する研究が進められている（乾ほか（2010）参照）。もともと医療の質改善運動が1900年代初頭から米国では盛んであるが、病院ごとに医療の質は異なるのではないかと

観点から病院の固有効果を測定する研究（プロファイリング）が行われている。Centers for Medicare and Medicaid Services (CMS) では Hospital Compare と呼ばれる各病院の医療の質の評価プログラムを実施している。その中で疾病別のリスク調整といった統計的手法も開発・利用されてきている。同様に、Pay for Performance (P4P) という医療の質に応じた医療費支払いプロジェクトではいくつかの疾病について臨床指標が設定され、それについて総合スコアを算出し病院は評価される。その際にプロセス指標（薬剤の投与）やアウトカム指標（リスク調整した死亡率）が用いられており、医療の質の評価に関する検討が蓄積されてきている。

第3は、医療データの充実である。医療の質を計測するのに有力なデータが日本では比較的充実している。例えば、DPC データやレセプトデータである。DPC データは包括医療費支払い制度に基づく入院にかかるデータであるが、分類軸が疾病・治療で分岐されてコード化され、年齢や性別といった患者の属性のほかに併存疾患や重症度、転帰（死亡情報）等の詳細情報を把握することができる貴重なデータである。レセプトデータについては、DPC データに比べると重症度情報が無い等得られる情報が限られてくるが、そのカバレッジの広さに利点がある。特に2009年以降の電子レセプト化にともないナショナルデータベース (NDB) として格納されており、国民皆保険の日本の医療制度のもとかなりのボリュームのデータが蓄積されている。

ただし、現状においては、DPC の提供は集計値に限られ個票の提供体制が整備されていない、NDB についても詳細な個票情報を入手するには厳正な審査を経る必要があり、容易に入手できる訳ではない。また、コストデータについても統計・データにより概念が不統一であったり、記入の正確性が確実ではない等の課題がある点は留意が必要である。

5 把握手法の整理と方向性

eurostat 等の国際機関や欧米各国における具体的な医療の質の取扱いや先行研究を踏まえると医療の質の調整の手法は概ね、(1) どのようなアプローチを用いるか、(2) 何を質の調整の対象とするか、という観点から整理できる。

(1) 質の調整のアプローチ

国際機関、各国においてどのようなアプローチが用いられているのかについては、概ね①細分化、②質の明示的・統計的な計測、といった2つに整理される。

①は、疾病等に着目した分類を可能な限り細分化し、分類内の医療サービスを可能な限り均質にするアプローチである。分類を可能な限り細分化することにより、同一分類内の質は同一と捉えて質の調整を行うものである。こうした細分化を基本とするという概念については、SNAの最新の国際基準である『2008年版国民勘定体系』（2008SNA）においても1つのアプローチとして位置付けられている。

以上をまとめると、細分化アプローチとは疾病や疾病・治療に着目した分類を可能な限り細分化し、分類内の医療サービスを可能な限り均質にし、質の変化を分類間の移行で捉えるアプローチである。

②は、医療の質を疾病死亡率や再入院率という明示的な指標で捉え、統計的手法を用いてアウトプットに反映させるアプローチである。

明示的な質の調整の具体的な手法としては、例えば明示的質指標である死亡率は患者の重症度や併存疾患の状況に影響を受けると考えられ、これらの変数でコントロールした統計的手法により死亡率を推計するものである。死亡率については、データから計算される死亡率（死亡者数/患者数 等）そのものを使うのか、それとも効用や金銭換算したアプローチ（Cost of Living Index（健康状態を変数に含む効用関数を仮定し、健康状態が変化した際に効用水準を一定に保つような金銭補償額を求めたうえで価格指数を作成する方法）の考え方に基づき価格指数を作成）が考えられる。

(2) 質の調整の対象

質の調整の対象には、①産出指標（実質アウトプット）、②価格指標（デフレーター）の2つの対象があり得る。①は、実質アウトプットに直接的に質の調整を施すものであり、②は、文字通りデフレーターについて質の調整を行うものである。

この考え方において、質の調整の対象を①とするか②とするかの相違は、質を「実質アウトプットに寄せて調整する」のか、「デフレーターに寄せて調整する」のかという違いである。これをもう少し具体的に考えてみよう。

上述(1)①の細分化では、明示的な質指標は用いられず、分類を可能な限り細かくして分類内の質は一定と仮定するため、細分化手法では「質」部分を表す質指標が明示的には現れず、以下のようになる。

$$\text{名目アウトプット} = \text{数量（実質アウトプット）} \times \text{価格} \quad (2)$$

細分化手法の場合は、細分化により均質化された分類をもとに数量や価格（デフレーター）を推計することになる。上記の「質を実質アウトプットに寄せて調整」とは式（2）の前半を \sum_j 数量（実質アウトプット）として細分化された疾病分類の疾病」をベースに算出し（各分類内の質が一定という点で質の調整が行われている）、それを足しあげて実質アウトプットを求めるものである。価格（デフレーター）は名目アウトプットを実質アウトプットで除して事後的に求めることになる。

逆に「デフレーターに寄せて調整」とは、式（2）の後半の価格（デフレーター）を分類内において一定の質となる疾病分類毎に算出した上で、足しあげて（ \sum_j 価格（単価））全体のデフレーターを算出する。実質アウトプットは名目アウトプットをデフレーターで除して算出する。

次に質の明示的・統計的手法の場合について以下の例をもとに考えてみたい。

$$\begin{aligned} \text{名目アウトプット(円)} &= \text{診療件数(件)} && \text{: 数量} \\ &\times \text{診療による QALY の改善度合(年/件)} && \text{: 質} \\ &\times \text{命の値段(円/年)} && \text{: 価格} \quad (3) \end{aligned}$$

ここで QALY (Quality Adjusted Life Year) とは質調整生存年とも呼ばれる明示的な質指標である。完全に健康な期間（年）を 1 とした場合に疾病等に罹患したり、逆に疾病を治療することで質（健康状態）がどの程度変化するかを測定する概念である。命の値段は、患者の 1 年当たりの命の価格と考えられる。ここで念のため、質の明示的・統計的手法は、細分化に対処した上で更に明示的質指標を用いて計量分析するものであることに留意が必要である。つまり式（3）も細分化された疾病ごとに成り立つのが前提である。

この場合の「質を実質アウトプットに寄せて調整」とは、まず「診療件数（件）×診療による QALY の改善度合（年/件）」として明示的質指標による質の調整を行った実質アウトプットをまず求める。デフレーターは名目アウトプットを当該実質アウトプットで除して事後的に算出される。実際の名目アウトプットはデータ制約上、

$$\text{名目アウトプット} = \text{診療件数} \times \text{保険点数} \quad (4)$$

として計測されるため、デフレーターは事後的に「保険点数÷診療による QALY の改善度合（年/件）」として求められるものである。

他方、「質をデフレーターに寄せて調整」とはまず「命の値段（円/年）」を推計する。実質アウトプット（「診療件数（件）×診療によるQALYの改善度合（年/件）」）は、名目アウトプットを「命の値段（円/年）」で除して求められると考えられる。つまり、「保険点数×診療件数÷命の値段」が実質アウトプットとなる。

「実質アウトプットに寄せて調整」した結果事後的に算出されるデフレーター（「保険点数÷診療によるQALYの改善度合」）も、「デフレーターに寄せて調整」した結果から算出される実質アウトプット（「保険点数×診療件数÷命の値段」）もともに、保険点数を含んでいる。我が国の公的価格に基づく保険点数の場合、患者（消費者）の評価（効用）を完全に反映したものとは必ずしも言い難い。従って、「寄せて調整」とは「寄せなかった変数」に生ずる歪みが含まれるものと解釈することもできよう。

上記の「2つのアプローチ」×「2つの対象」をマトリックスで整理すると第2図のように表せる。

なお、この各方式については、それぞれについて長所短所があり、各方式同士には以下のようなトレードオフの関係が見られることに留意しておく必要がある。

① 細分化と質の明示的・統計的手法

細分化については、統計的なモデルを使用しない。従って必要な仮定が相対的に少ないため、頑健であるという長所が挙げられる。他方、分類は細かくするものの、統計的なモデルにより死亡率に対するリスク要因の影響を精緻に分析する訳ではないので質の調整は不十分という短所がある。

一方、質の明示的・統計的手法については、患者属性や重症度、併存疾患等が質指標（死亡率や再入院率等）に与える影響を制御するため、精緻な質の調整が可能という利点がある。ただし、これはモデル構造に結果が大きく依存し、また疾病・治療に関するデータ上の要求も大きくなるため実際に推計するにはデータの利用・推計にかかる作業負担が重くなるという問題を伴う。

② アウトプットとデフレーター

質の調整の対象がアウトプットの場合、特に質の明示的・統計的手法を考えた場合において、余命の金銭換算額を規定する必要等がなく、実際の推計作業上の制約が少ない。他方、デフレーターの場合は我が国のSNA(JSNA)の現行方式に馴染みやすいという長所が挙げられる。現行のJSNAにおいて実質化は、産出額と中間投入額をそれぞれ実質化するというダブルデフレーション方式が

用いられており、親和性が高い。ただし、質の明示的・統計的手法を考えた場合には、余命の金銭換算額を規定する必要等があり、実際に推計するにはその点に対応しなければならない。

なお、JSNA で医療は定義上市場型サービスと位置付けられており、名目産出額をデフレーターで除して実質産出額を推計している。その意味では、質を調整したデフレーターを計測する方が現在の JSNA の推計方法と親和性が高い。

前述のとおり、市場価格や市場での需要量が消費者の評価を反映していれば、これらの情報を利用して質を調整したデフレーターを計測するアプローチが有効である。しかし、医療においては、価格が公定され、また、供給量にも規制がかかっているため、市場価格や需要量が消費者の評価を必ずしも反映していない。そうした場合には、アウトプットを直接計測する中でサービスの質を調整することが有効な可能性もある。

第 2 図

	実質アウトプットの 直接推計	デフレーター
細分化	① 細分化 × アウトプットの直接推計	② 細分化 × デフレーター
質の明示的・ 統計的	③ 明示的・統計的手法 × アウトプットの直接推計	④ 明示的・統計的手法 × デフレーター

(3) 各方式の説明

第 2 図の各方式についてより具体的に説明していこう。

① 「細分化×実質アウトプットの直接推計」方式

細分化された疾病分類や疾病・治療分類等に基づいて直接的に実質産出額（アウトプット）を求める手法である。2008SNA では output volume method（産出数量法）と呼ばれている。細分化された疾病分類等に基づき、各分類の実質産出額を直接的に求める。分類を細分化すること、つまり同質の医療サー

ビスとみなせる分類を設定することにより分類内では質は均質と考えられる。細分化された分類間での移動により質の変化（向上）を捉えていると考えられる。また、分類軸については疾病が基本である。前述の通り、本来であれば一連の治療（エピソード）を一気通貫で把握するのが理想的ではあるが、データ制約上困難であるため、個別の治療行為ではなく一連の治療を近似的に捉える分類として疾病分類が妥当と考えられるためである。疾病に加えて治療行為や合併症等でさらに細分化するという方法も考えられる。欧州各国では DRG や Healthcare Resource Group (HRG) といった疾病分類や疾病・治療分類が用いられている。

Dawson, et al. (2005) では以下の通り、アウトプット指数を定義している。アウトプット指数の考え方は、基準時点の数量（患者数、治療件数等）に対して、その時点での数量の増え方をコストウェイトで加重平均していくものである。

ここで、 x_{jt} は第 t 期における疾病 j の治療件数、 c_{jt} は第 t 期における疾病 j の治療にかかる単位コスト、 x_{kt} は第 t 期における疾病 j の治療を構成する要素 k （直接的な治療や病院の食事等）、 c_{kt} は第 t 期における疾病 j の治療を構成する要素 k の単位コストとする。疾病 j の治療コストウェイトには基準時点における NHS の総コストに占める疾病 j の治療の総コストを用いている。アウトプット指数は以下の通り示される。

$$I_{ct}^x = \sum_j \left(\frac{x_{jt}}{x_{j0}} \right) \frac{x_{j0} c_{j0}}{\sum_k x_{k0} c_{k0}} \quad (5)$$

この方式のメリットとしては、既に支払方式で利用されている DRG や DPC といった既存の分類やデータが利用可能といった点が挙げられる。他方、デメリットとしては、同分類内において時間を経ても質は一定と仮定しており learning by doing を含む技術進歩等は反映されないという点が挙げられる。また、治療方法の選択も分類軸に含まれる場合、例えば効果が薄いがコストがかかる治療（例：安定狭心症の PCI 等）が行われた場合も質の向上、アウトプットの増加とみなされる点である。また、生産性を計測する観点から見ると歪みをもたらし得るという点も挙げられる。例えば、治療法は同じでもより質の高い医療を提供すべきとする医療の質改善運動は、コストがかかるがアウトプットの上昇に寄与しないため生産性が低下するということになってしまう。

この方式は欧州各国の SNA 本体において用いられている。

② 「細分化×デフレーター」方式

細分化された疾病分類ごとの価格（Medical-Care Expenditure Index (MCE index)）を求める手法である。個々の疾病別に患者1人当たりの治療コストを把握し、これを集計することによって価格指数を求める。つまり、患者1人当たりの治療コストをデフレーターとして用いているのである。この方式では、同一分類内での治療方法が変化した場合、同一分類内では質が一定と仮定していることから、コストが変化（例えば低下）するとデフレーターも変化（低下）することになる。具体的に数式で MCE 指数を表すと以下の通り。

$$MCE_t = \sum_j \left\{ \frac{c_{j0}x_{j0}}{\sum_j c_{j0}x_{j0}} \left(c_{jt}/c_{j0} \right) \right\} \quad (6)$$

ここで、 MCE_t とは第 t 期の MCE 指数、 c_{jt} は第 t 期における疾病 j の治療にかかる1人当たり支出額、 x_{j0} は疾病 j の治療を受けた患者数である。すなわち、個別の疾病に関する患者1人当たりの治療コストを、それぞれの疾病の総支出額が全医療支出に占める割合でウェイト付けされて統合が行われている。

この方式のメリットとしては、質を直接計測せずにコストを把握することから計測しやすいという点が挙げられる。逆にデメリットとしては、異なる治療方法であるが同じ質のものを判定・分類することが難しい点である。ある程度大括りの分類にする必要があり、細分化の観点からは質の調整が不十分ではないかという指摘もあり得る。

この方式は米国の医療サテライト勘定（HCSPA）において用いられている手法である。

③ 「質の明示的・統計的手法×アウトプットの直接推計」方式

直接的に実質アウトプットを求める際に明示的な質指標を用いて、統計的手法により質の調整を行う手法である。①の細分化された疾病別の実質アウトプットについて、生存率や再入院率といったアウトカム指標を明示的な質指標として統計的手法を用いて質の調整をする。具体的には、個票データから求めた疾病別の死亡率や疾病別の QOL 指標を利用、さらにはリスク要因（重症度等）を考慮した疾病別の死亡率を利用することが考えられる。

Dawson, et al. (2005)に沿って計算式を定義すると以下の通りである。

ここで、 a_{jt} は第 t 期における疾病 j の治療の明示的質指標、 x_{jt} は第 t 期における疾病 j の治療件数、 c_{jt} は第 t 期における疾病 j の治療にかかる単位コストである。つまり、①のアウトプットを直接推計する式（5）に明示的質指標の変化を入れ込んだ形である。

$$I_{c0}^{xa} = \frac{\sum_j x_{jt}(a_{jt}/a_{j0})c_{j0}}{\sum_j x_{j0}c_{j0}} \quad (7)$$

この方式のメリットとしては、詳細な質の調整が可能となることである。死亡率や再入院率といった質指標を明示的に特定化した上で、患者の属性や疾病の重症度等でコントロールして推計することが可能である。つまり、死亡率等の明示的な質指標を被説明変数とし、患者属性、疾病の重症度等を説明変数として推計（リスク調整）することが出来る。他方、デメリットとしては、質指標として何を取るかということによる影響が大きいところである。正統的にはQOLであろうが、疾病別等適切なデータを入手することはかなり難しい。その代替指標として死亡率や併存疾患、再手術・再入院率等が考えられる。また、質の調整の統計的手法であるが、プロファイリング等研究の蓄積はあるものの、頑健なモデルと言い切れる訳でもない。アウトプットがモデル構造に依存するといった点や必要なデータが多くなるといった面も大きいため注意が必要である。

この方式は、英国において生産性指標を計測する際に併せて質の調整が行われており、扱いとしては参考指標群として公表されている。

④ 「質の明示的・統計的手法×デフレーター」方式

健康状態を変数に含む効用関数を用いて、健康状態の変化を補償するような金額を算出する手法である。Cutler, et al. (1998)をはじめ、米国を中心とした先行研究において用いられている。医療の質と効用の関係については、次の考え方に基づいて推計が行われている。医療の質の向上は治療にかかる支出額が同じであれば効用が増加する。効用一定の場合は、医療の質が向上すれば支出額は減少する。この効用を一定とした場合の必要な支出額の変化を価格の変化として捉えるものである。この必要な支出額は補償変分として位置付けられる。Cutler, et al. (1998)では具体的に、以下のような効用関数を仮定して補償変分(C)を求める。式の前半部分は健康状態による効用、後半部分は医療費以外の支出からもたらされる効用からなる効用関数を設定している。

$$U(t) = U(H(m(t)), Y - p(t)m(t)) \quad (8)$$

ここで、Uは効用、Hは健康状態を表す関数、mは医療サービスの数量、pは医療サービスの価格、Yは所得（一定）である。

次に第0期と第1期でmとpが変化する場合、第0期と第1期で効用を一定

と仮定すると、その差分の補償変分(C)は以下の通り求められる。

$$\begin{aligned} &U(H(m(1)), Y - p(1)m(1) - C) \\ &= U(H(m(0)), Y - p(0)m(0)) \end{aligned} \quad (9)$$

第0期から第1期へ医療サービスの数量と支払額が変化する中で、第0期と第1期の効用が一定という仮定を満たすために必要な金銭換算された補償変分がCである。Cが正であれば第0期から第1期にかけて医療の質は向上したと解釈する。Cutler, et al. (1998)では、急性心筋梗塞(AMI)に関する推計を行っており、余命の金銭換算額について先行研究の結果を参照し25,000ドル/年と仮定している。こうして得られた補償変分を用いて、Cost of Living Index ($COL_{0,1}$: デフレーター)が以下のように計算される。

$$COL_{0,1} = (Y - C)/Y \quad (10)$$

この方式のメリットは、経済学的に最も明快なアウトプット推計手法という点である。また、現行のJSNAではダブルデフレーションの枠組み(産出額と中間投入額をそれぞれ実質化)が用いられているが、この方式も名目産出額をデフレーターで除する手続きを取ることから、現行のJSNAと親和性が高い。他方、デメリットとしては、効用の金銭評価額が必要であることである。これは2つの部分に分けることができる。1つは、医療サービスの結果としてのQOL(式(9)における $H(m(0))$ ないし $H(m(1))$ に相当)を計測する必要がある。もう1つは、その効用を金銭額で評価しなければならない。前者のQOLの計測は、医学研究で活発に行われているが、すべての医療サービスについて網羅的にQOLを計測するのは相当ハードルが高い。QOLの代替的な指標として死亡率があるが、慢性期の疾病等においては、死亡率でQOLを代替して良いのかという問題もあり、出来るだけ幅広い疾病についてのQOLを計測することが望ましい。後者のQOLの金銭評価のためには「命の値段」を想定する必要がある。こうした「命の値段」の計測は、Willingness to Payの計測や賃金データを使った計測等多くの研究があるが、統計に使用する値としてコンセンサスを得ることはそれほど簡単なことではない。

6 各地域・国の現状

5において医療の質の調整のアプローチ、対象の観点から手法を整理したところで、日本、欧州、米国における医療の取扱い及び質の調整の手法を整理し

てみよう。

(1) 日本

JSNA 本体では、「医療」にかかる基本単位デフレーターに CPI を用いている。デフレーター推計に際し、医療（入院診療、入院外診療）について CPI 「診療代」等を利用している。CPI 「診療代」は、小売物価統計調査（動向編）等の価格をもとに所定のモデル式により価格指数を算出する「モデル品目」として CPI の中で扱われている。具体的には年齢区分（小児、一般、高齢者、後期高齢者）別、診療種類区分（入院、入院外、歯科）別、施設区分（病院・診療所）の別に代表的な診療行為を選定し、それぞれの区分でデフレーターを算出している。それぞれの区分（年齢・診療種類・施設。例：「病院に入院した小児の診療」「診療所の外来を受診した高齢者の診療」）ごとに質が一定と仮定し、一定の質が保たれた各区分の代表的な診療行為の構成の変化による価格の変化を捉えていると解釈出来る。実質値の算出については、『国民医療費の概況』（厚生労働省）等をもとに名目産出額を計算し、それをデフレーターで除して実質産出額を得ている。この方法では、疾病死亡率の低下等を考慮した明示的な質の変化の調整が行われている訳ではないが、24 区分（年齢 4 区分×診療種類 3 区分×施設 2 区分）内においては質が一定として調整されたデフレーターが推計されていると考えられる。

(2) 欧州

欧州では、病院サービスについて明示的な質の調整は行わずに、eurostat(2016)で示されている基本的考え方である、疾病等に基づく詳細な分類をベースに直接的にアウトプットを推計するという(3)①の「細分化×アウトプット方式」（以下、「eurostat 方式」と呼ぶ）に沿った対応をしている国が多い（第 1 表参照）。

欧州の中でも英国では、SNA 本体では HRG の疾病・治療分類ごとのコストをウェイトとして、患者数や診察件数でアウトプットを直接算出している。フランス、イタリア、オランダ等においてもアウトプットの一部は DRG や International Classification of Diseases (ICD) といった疾病分類をもとにアウトプット指数を直接的に算出している。DRG は米国で開発された疾病と合併症に基づく分類であるが、1986 年に米国の高齢者を対象とする公的医療制度の病院費用償還に導入された DRG/PPS (Prospective Payment System) において用いるため、コストを均一化することを目的とする分類である。ただし、コストが均一となることを目的とする分類においては、異なる質の医療サービスも同じ分類に属する可能性がある等、質を均一にする分類として適切かどうか

確認する必要がある。

欧州の中でも精力的に質の調整に取り組む英国では、SNA 本体とは別に、独立した参考指標群として(3)③の「明示的・統計的手法×アウトプット方式」（以下、「英国方式」と呼ぶ）にあたる質を明示的に調整（死亡率や待ち時間等を質指標として採用）した生産性指標を論文形式で公表している。Health 部門（4 部門）のうち「病院・地域医療（hospital & community health care）」「家族医療（family health services）」2 部門の一部について明示的な質の調整を行い、残りの「調剤」「非 NHS サービス」部門については行っていない。

(3) 米国

米国の所得勘定である National Income and Product Accounts (NIPA) 本体では、実質産出額を算出する際に CPI や Producer Price Index (PPI) 等がデフレーターとして用いられている。医療については、医師・歯科医等専門家サービスや病院サービス等 Economic Census (経済センサス。基準年) や Service Annual Survey (年次サービス調査。基準年以外) や Quarterly Services Survey (四半期サービス調査) をもとに名目値を算出し、個別のサービスにそれぞれ対応するデフレーターで名目産出額を除して実質産出額を求めている。

さらに、米国では NIPA 本体とは別に HCSA を公表している。HCSA では、MCE 指数を算出して、デフレーターとして用いている。これは(3)②の「細分化×デフレーター方式」に整理される（以下、「米国方式」と呼ぶ）。考え方としては、質が一定となる（ある程度大括りの）分類を設定し、その分類の時系列の価格変化をデフレーターとして捉える考え方である。つまり、治療行為の変化は質の変化ではなく価格（デフレーター）の変化と解釈するのである。名目産出額については、NIPA 本体で公表されている名目医療支出総額をコントロールトータルとし、別途計算した疾病分類ごとの支出額のシェアを掛け合わせて、一国全体の推計値を作成している。NIPA 本体で公表されている名目医療支出総額をコントロールトータルとするのは、HCSA 推計の基礎データとなる Medical Expenditure Panel Survey (MEPS) データはサンプル数 15,000 世帯、35,000 人とサイズが小さく、全体額の推計に使用できないためコントロールトータルを用いた推計を行っている。この疾病分類ごとの名目支出額を対応する疾病分類の MCE 指数で除したものが実質産出額である。MCE 指数の算出に際しての最も細かい分類は臨床上の分類である Clinical Classification Software (CCS) による 263 分類であるが、それを 18 (15) 分類に統合して公表している。CCS とは、ICD を臨床で使用しやすいカテゴリーに分類したものであり、Dunn, et al. (2015) によると CCS のうち歯科サービスを除いて HCSA の算出に用いている。HCSA においては現行の NIPA の health 部門のうち「医師サービス」「医療関係者

サービス」「病院サービス」「調剤」部分（health部門の8割程度）をこの「米国方式」による推計値に置き換え、残りの部分については現行NIPAのGPIやPPIをデフレーターとする推計値をそのまま利用している。

その他に手法としては、(3)④で説明した「明示的・統計的手法×デフレーター方式」（以下、「Cost of Living Index 方式」と呼ぶ。）が挙げられ、先行研究はあるものの、各国・地域で実際のSNA体系や関連指標の手法として利用されている例は筆者らの知る限りでは見当たらない。ただ、前述の通り、「命の値段」の金銭評価額（VSL等）が推計上必要となり、データ制約は大きい手法である。

7 今後の検討課題

これまでの研究蓄積やデータの利用可能性からすると、第一段階として入院部分について医療サービスのアウトプット/デフレーターを計測し、第二段階に入院外部分を検討することが妥当であろう。

(1) 入院医療サービスの基本推計と感度分析

質を調整したアウトプットやデフレーターの計測の研究としては、まず、先に整理したような欧米の統計部局の取組や先行研究の次の4つの方式に基づいて医療サービスのアウトプットを計測することが適切であろう。質の調整方法としては、これが最善という方式はない現状で、日本において利用可能なデータでそれぞれの方式でどこまで正確な計測ができるかを検証することが実践的な第一歩であろう。

- ①eurostat方式（細分化された均質な疾病分類等に基づき直接的に実質産出額を求める手法で、欧州各国のSNA本体で用いられている。）
- ②米国方式（細分化された均質な疾病分類ごとの価格（1人当たりの治療コスト）を求める手法で、HCSAで用いられている。）
- ③英国方式（直接的に実質産出額を求める手法で生存率等を明示的質指標として統計的に質の調整を行う手法で、英国の参考指標群として用いられている。）
- ④Cost of Living Index方式（健康状態に注目した効用関数を用いて、医療サービスを受けて変化した健康状態から得る効用の変化を補償する補償額（Cost of Living）という概念を用いる手法で、米国を中心とした先行研究において用いられている。）

その上で、以上の各方式はそれぞれ異なる前提や異なるデータに基づいていることから、そうした前提や使用データを変更した時に、どのような違いが生じるかを感度分析的に検証する必要がある。そうした感度分析を、部分的にしか利用できないが詳細な情報を持つデータセットを使って行うことにより、前提やデータについて、より詳細な前提・データを使った場合にどのような違いが生じるのか、逆に、どこまで簡略な前提・データで統計として必要な精度が確保できるのか等を検討することが重要である。

(2) 具体的な検討課題

今後の検討課題としては、以下のような論点が挙げられる。

①、②の細分化については、a)分類軸を何にするのか、b)十分な細分化が行われているか、c)一部の試算をどのように全体に拡張するか、d)質の指標の検討である。

a)については、疾病分類とするのか、疾病・治療分類とするのか等、分類軸をどのように設定するのか検討する必要があるが、使用データ上の観点から疾病分類が有力である。その上で、DPC分類のように疾病に加えて治療方法も分類軸とした場合にどの程度の相違が生じるかを感度分析的に検証しておくことが必要であろう。ただし、DPCの分類軸に治療行為を加えたのは、臨床医療になじみやすい分類とするためとされるが、治療方法には医療供給者による選択という側面があり、計測に歪みが生じるおそれがある。治療選択のバイアスを受けないという点では、疾病分類と治療行為ではなく、ICD分類と重症度分類あるいはICD分類と併存疾患分類等を分類軸とすることも考えられる。また、DPCのようにコストが均一となることを目的とする分類においては、異なる質の医療サービスも同じ分類に属する可能性がある（例えば、不安定狭心症に対するPCIは有効であり安定狭心症に対するPCIは有効でないが、コストが大体同じなので同一分類になる）等、質を均一にする分類として適切かどうか確認する必要がある。

なお、ICD、DPC、HRG、DRG、CCSといった疾病分類の具体的な分類を、急性心筋梗塞（AMI）を例として第2表に示した。CCSでは、AMIについては更なる細分化はされていない。DRGやHRGにおいては、合併症も分類軸となっており、疾病をもとに重症度を考慮に入れた分類と言える。ICDの基本分類は、心筋梗塞が生じた部位で分類を細分化しているが、梗塞部位によってAMIの重症度が異なることに鑑みて重症度をある程度織り込んだ分類と言えよう（ただし、完全に満足できるものではないが）。DPCについては、DPC疾病分類（6桁分類）としてはAMIについては更なる細分化はされていないが、DPCとしての

分類では、治療方法も分類軸となるので、非常に細かいものとなっている。

b) 細分化をどこまで細かくする必要があるかの検証については、ICD 分類を段階的に大括りにしていった時にアウトプット/デフレーター の推計結果がどの程度異なるかを検証していくことになる。ICD10 では最も細かい基本分類で 14,000 を超える分類が定義されている。この ICD 基本分類をベースにどういった疾病分類で括りあげるのかという検討が必要である (第 3 表)。米国方式では、疾病を 263 分類とある程度大括りの分類で分類している。日本の疾病分類表や患者調査における小分類の分類数は米国 CCS よりやや多く、DPC の疾病分類 (6 桁分類) はかなり多い。社会医療診療行為別調査においては基本的に疾病分類表の中分類に従っており、かなり分類は粗い。

c) については、例えば DPC 分類を分類軸に採った場合、DPC データに含まれている病院・患者についてしかアウトプット/デフレーターを計測できない。これを入院全体の計測に拡張する必要があるが、それは容易でない問題である。正攻法としては、DPC 分類のロジックは疾病と治療行為に基づいているので、レセプトデータの疾病や治療行為の情報を利用して入院全体について DPC 分類を作成することも考えられる。しかし、この方法は、レセプトにおける治療行為のコードが通年で共通のコードではないので、その接合は大変な手間とコストを要する。

d) については、③で併せて述べる。

③の明示的な質の調整については、a) 質の指標としてどのようなものを探るか、b) リスク調整の統計的手法をどうするか (リスクモデルや推定方法等)、c) 一部の病院について利用可能な詳細なデータにより推定した医療の質を入院全体に拡張するためにはどうしたら良いか等を検討する必要がある。

a) 医療の質の指標としては、QOL が正統的な指標であるが、疾病毎の QOL 指標が網羅的に調査されている訳ではなく、データの入手可能性の観点からハードルがある。そこで、実際のデータとして QOL の代替指標として利用し得るのは、死亡率、合併症発症の有無、再手術・再入院率、ADL (食事、排泄、入浴、着替え等の日常生活動作) 等が考えられる (杉原ほか (2017a) は、AMI 患者について、死亡率以外に治療後の後遺症や再手術を QOL の計算に加えてアウトプットを計算した例である)。

b) 医療の質の計測について、問題になるのは重症度や併存疾患といった死亡率等に影響し得る要素をコントロールして推計する必要があることである。例えば、医療の質が同じでも、重篤な患者は死亡しやすく、軽症の患者は死亡する確率が低かったりする。ある年に重篤な患者が多くなった場合、素朴な推計では医療の質が低下したという結果になってしまう。

こうしたリスク調整に使える変数としては、血圧や血糖値等の検査値、高血圧や糖尿病といった疾病、Killip 分類等の重症度、Charlson Comorbidity Index のような統合指標がある。これらは、それぞれ、情報の詳細さや入手可能性が異なる。検査値は、最も精緻なリスク指標であろうが、臨床試験等ではなければなかなか入手できない。重症度は、やや大雑把なカテゴリー化がなされることになるが、特定の疾病については DPC データに含まれている。疾病は、リスク指標としてはかなり大雑把であるが、レセプトデータ等でも得ることができる。統合指標は、Charlson Comorbidity Index のように疾病から算出されるものであれば、比較的容易に計算することができる。こうしたリスクモデルは、臨床試験に伴うものや、医学研究で重症度を使ったモデル、病院のプロファイリング等で疾病をもとにしたモデル（例えば、米 CMS の Hospital Compare : 第 4 表）等の研究の蓄積がある。こうした研究を参考にリスク調整を行っていくことが考えられる。

さらに、医療の質は、単純に計測すると、毎年の変動が非常に大きくなることが知られている。これは、病院ごとの患者数が少ないことやリスク調整の不完全さ等による精度の限界によるものであるが、こうした場合には、階層ベイズモデルが有効な対応策となり得る（杉原ほか(2017a)は、階層ベイズモデルと通常の最尤法、AR モデルによる変動の調整を比較した例である）。

なお、リスクモデルの検討の際には、重症度等詳細な情報を含む DPC データではなく、重症度情報は含まれず併存疾患等基礎的情報のみ含むレセプトデータ等のみを用いたとしても、適切なリスク調整が可能かといった視点からも検討が必要である。例えば、重症度を用いた場合と併存疾患等のみを用いた場合でどの程度リスク調整の精度に違いが生じるか、あるいは、重症度には併存疾患で代替できる可能性があるもの（例えば、Killip 分類 4 とショック、CCS 分類 1~4 と安定狭心症/不安定狭心症等）もあり、そうした代替的疾患を利用してどれだけ正確なリスク調整ができるか等である。

c) については、現状ではレセプトデータをもとに入院全体を推計する方向であるが、例えば明示的な質の調整については DPC データを使用するというように組み合わせ推計する選択肢も考えられる。ただ、その場合には、DPC 病院という比較的先進的な病院のみを対象とした DPC データでは、質についての上方バイアスがかかる可能性が高い。そのため、明示的な質の調整の結果を全体に適用する際の拡張手法についても検討の必要がある。欠測値（DPC データ上の欠測、つまり DPC 病院ではないこと）についてのセレクションバイアスの調整の問題である。第 3 図のように、医療の質の高い病院は DPC へ参加しやすいが、質が低い病院は脱落しやすい（点線のように、病院数が母集団に比べて少ない）場合、医療の質の計測された平均値は点線のようになり、実線で表わさ

れる真の平均値と異なってしまふ。こうしたバイアスが発生するのは、病院が参加/脱落するかどうかに影響を与える要因が、同時に医療の質にも影響を与える場合である（これはあくまで仮説例であるが、例えば、学習効果等によりある手術の症例数が医療の質に影響するとともに、参加/脱落確率にも影響する（症例数の多い病院ほど参加しやすい）場合、症例数の少ない質の低い病院が脱落して主として症例数の多い質の高い病院が参加することから、医療の質の計測に歪みが生じる：しかし、本当に症例数と質や脱落が相関するかは、実証的に検証すべき事項である）。

脱落のメカニズムは3通りに区分され、それぞれの場合に適切な統計的手法が異なる。①Missing Completely at Random (MCAR：ランダムに脱落)の場合には、調整の必要はない。②Missing at Random (MAR：観測出来る変数のみに依存して脱落)の場合には、観測できる変数を利用して脱落によるバイアスを補正することができる。例えば、Inverse Probability Weighting という手法では、観測できる変数によって脱落する確率を推定し、その脱落確率の逆数で各変数をウェイト付けする。最近よく使われる Propensity Score Matching も同様に観測できる変数を利用した調整方法である。この場合、質 y の決定と脱落 r の決定との同時確率は $p(y, r) = p(y | r)p(r)$ という形に分解できるが、脱落確率が同程度のサンプル同士をマッチングさせることにより、脱落したサンプルの $p(y | r)$ を推定することができる。③Not Missing at Random (NMAR：脱落が観測できない変数にも依存する)の場合には、質 y と脱落の有無 r の同時確率 $p(y, r)$ を分解して、連立方程式を解くことになる。ヘックマンのセレクションモデル（同時確率を $p(r | y)p(y)$ という形に分解）やパターン混合モデル（同時確率を $p(y | r)p(r)$ という形に分解）等が代表的なモデルである（杉原ほか(2017b)は、パターン混合モデルを使って脱落の調整を行った例である）。

④の明示的な質の調整については、③の論点に加えて、「命の値段（余命の金銭換算価値）」等のデータが必要であり、データ要求にいかに対応するかといった難題に対応しなければならない。「命の値段」については、Willingness to Pay の研究（Viscusi(2014)や宮里(2010)でサーベイされている）やコスト/便益分析における想定等を参考にすることができよう。「命の値段」は、第5表にみるように、「命の値段」の推計値は大きくばらついており、統計作成上の仮定としてのコンセンサスを形成する必要がある。

以上①から④の検討に当たっては、各手法が依拠する前提やデータがどの程度不確実性をもたらすかについてシミュレーションを使って検証することが有益である。例えば、細分化の直接推計において、分類として単純に疾病を採つ

た場合に、重症度の変化がどれだけ推計結果に影響を与えるかを確認しておくことは、統計を利用する際の信頼性を評価した情報を提供することになる。このために例えば（重症度等より詳細な情報を得られるデータを使って）重症度をランダムに変化させた時にアウトプットやデフレーターがどれだけ変化するかを計算し、それを繰り返し行うことにより、統計的な不確実性の程度を把握することができる。これは、DPC データを使って計測した詳細な質の調整を入院全体の質へ拡張する際に非常に重要となる。選択バイアスの統計的な補正は、実際に実行すると、モデルが完全でない限り、大きな不確実性が存在する。この時、パラメーターに確率分布を想定するベイズ的な推定を行えば、計測されたモデルにおける係数はある分布を持って推定されるが、その分布からランダムに係数をサンプリングして、抽出された係数を使ってアウトプット/デフレーターを推計するというのを繰り返すと、やはり統計的にアウトプット/デフレーターの推計の不確実性を評価することができる。

さらには、統計実務の観点からの検討も必要になろう。継続的に医療の質を捕捉することを考えると、全て個票を用いて厳密かつ詳細に計測するというのは統計実務上負担が大きい。個票データを用いて、情報を最大限利用した精緻な分析や仮定を変えた場合の感度分析を行い、どれだけの情報が必要でどの程度まで仮定を緩めることが出来るのかといった点について検討することが必要であろう。具体的には、細分化の分類数はどの程度が妥当で、明示的な質の調整で例えばリスク調整についてはどの程度詳細に行う必要があるかも含めて考えることである。データ制約を踏まえると必ずしも DPC 情報（重症度）を用いなくてもレセプトに含まれる情報で足りるのかといった視点からの検証も必要であろう。

(3) 入院外医療サービスのアウトプット/デフレーターの計測

外来や診療所等の入院外の医療サービスの計測について簡単に述べたい。日本には特定健診という制度があるので、そこから検査値等詳細な情報と広いカバレッジを持つデータが得られる。検査値から将来の疾患の発症や合併症等が予測できれば、検査値の変化を通じて入院外の医療の質を推計することがかなりの程度できることになる（Eggleton, et al. (2009)が1つの代表的な研究である）。あるいは、カバレッジは限定されるが、国民生活基礎調査のように失明や四肢切断等の身体状態が把握できれば、それを QOL に換算することにより、医療の質を測ることも考えられる（Stewart, et al. (2007)は方法論の検討を行っている）。ただ、いずれの方法も検査値や身体状態から転帰や QOL への変換をどのように行うか更なる研究が必要である。

6 でみたように、欧米各国では様々な手法が取られており、一本化されている訳ではない。つまり、それはこれがベストであるという決め手に欠けるといふことの裏返しともいえる。医療の質研究については様々な手法を全て試した上で、感度分析等で各手法の強みと弱みを把握するというのがリーズナブルな対応であろう。

8 データ

今後の研究において利用することが有力なデータは DPC データとレセプトデータであるが、その特徴について最後に述べておく。

DPC データについては、上述の通り、詳細な患者属性、重症度、併存疾患、転帰（死亡）、ADL 等豊富な調査項目を含んでいるという特徴が挙げられる（ただし、すべての該当患者について記載されているとは限らない）。しかし、カバレッジの観点からすると全ての病院が DPC を導入している訳ではない。しかも、2003 年には DPC 病院が 80 病院程度でスタートしたが、直近では参加病院は約 1500 病院と年々 DPC 対象病院が増えている状況にある。そうした状況を踏まえると DPC のカバレッジについては横断面のバイアス（各年における DPC 病院の参加割合）と時系列のバイアス（年によって DPC 対象病院のカバレッジが異なる）をいかに補正すべきかという課題がある。

他方、レセプトデータについては疾病情報が含まれるものの、重症度情報が含まれない、併存疾患についても状況が詳細に分からない（併存疾患の発症が入院前か後か等）。しかしながら、全データを入手することができれば即ち全入院・全期間をカバーすることとなりその悉皆性に特徴がある。

両者を組み合わせてアウトプットを推計していくのが選択し得る対応であろう。例えば、数量（患者数）についてはレセプト（件数）で質の調整にかかるデータは DPC でということも考えられる。また、技術的に可能であれば、レセプトデータを何らかの方法で擬似 DPC データに変換するということも対応として考えられる。

なお、DPC データの持つ詳細な情報を利用して限定した範囲の病院で精緻な推計ができたとしても、それを情報の少ない入院全体の推計に拡張しようとすると大きな不確実性が生じるので、各手法を評価する際にも、そうしたカバレッジのギャップに起因する不確実性もシミュレーション等で良く把握して評価する必要がある。

第1表 欧州諸国における医療サービスのアウトプット数量の計測方法

	色の意味:		病院		数量の直接計測		インプット価格によるデフレーション	
	救急病院	精神病院、薬物依存症治療 専門病院、その他専門病院	内科	歯科・歯科診療所	数量の直接計測	インプット価格によるデフレーション	数量の直接計測	インプット価格によるデフレーション
オーストリア	単位コスト指数によるデフレーション	単位コスト指数によるデフレーション	治療件数 【収入ウェイト】	料金単価に基づく価格指数によるデフレーション	治療件数	料金単価に基づく価格指数によるデフレーション?	治療件数	料金単価に基づく価格指数によるデフレーション
ベルギー	DRGに基づく直接数量指数 【コストウェイト】	入院日数 【収入ウェイト】	診察件数	診察件数	診察件数	診察件数	診察件数	診察件数
チエコ	収入単価に基づく価格指数によるデフレーション	収入単価に基づく価格指数によるデフレーション	治療件数	治療件数	治療件数	治療件数	治療件数	治療件数
デンマーク (2012~)	単位コスト指数によるデフレーション	単位コスト指数によるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	単位コスト指数によるデフレーション	単位コスト指数によるデフレーション
フィンランド	DRGに基づく数量指数 【コストウェイト】	治療日数	診察件数	診察件数	診察件数	診察件数	診察件数	診察件数
フランス	DRGに基づく数量指数 【コストウェイト】	DRGに基づく数量指数 【コストウェイト】	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション
ドイツ	単位コスト指数によるデフレーション	治療日数・件数 【コストウェイト】	収入単価またはCPIに基づく価格指数によるデフレーション	収入単価またはCPIに基づく価格指数によるデフレーション	収入単価またはCPIに基づく価格指数によるデフレーション	収入単価またはCPIに基づく価格指数によるデフレーション	収入単価またはCPIに基づく価格指数によるデフレーション	収入単価またはCPIに基づく価格指数によるデフレーション
ギリシャ	治療日数	治療日数	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション
ハンガリー	DRGに基づく数量指数 【単価ウェイト】	DRGに基づく数量指数 【単価ウェイト】	診察件数	診察件数	診察件数	診察件数	(医療保険の?)点数 (Number of scores)	(医療保険の?)点数 (Number of scores)
アイスランド	インプット価格によるデフレーション	NA	NA	NA	NA	NA	NA	NA
アイルランド	インプット価格によるデフレーション	NA	NA	NA	NA	NA	NA	NA
イタリア	DRGに基づく数量指数 【コストウェイト】	DRGに基づく数量指数 【コストウェイト】	処方せん数	処方せん数	処方せん数	処方せん数	CPIによるデフレーション	CPIによるデフレーション
ルクセンブルク	CPIによるデフレーション	CPIによるデフレーション	非市場産出: 診察・治療件数 【コストウェイト、質調整なし】	非市場産出: 診察・治療件数 【コストウェイト、質調整なし】	非市場産出: 診察・治療件数 【コストウェイト、質調整なし】	非市場産出: 診察・治療件数 【コストウェイト、質調整なし】	市場産出: CPIによるデフレーション	市場産出: CPIによるデフレーション
オランダ	IDCに基づく直接数量指数 【日数シェアウェイト】	治療日数・入院日数・看護時間に基づく直接数量指数	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション
ノルウェー	DRGに基づく直接数量指数 【コストウェイト】	治療日数	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション	CPIによるデフレーション
ポルトガル	DRGに基づく直接数量指数 【コストウェイト】	DRGに基づく直接数量指数	診察件数に基づく直接計測	診察件数に基づく直接計測	診察件数に基づく直接計測	診察件数に基づく直接計測	診察件数に基づく直接計測	診察件数に基づく直接計測
スウェーデン	DRGに基づく直接数量指数 【コストウェイト】	治療日数に基づく直接数量指数	診察件数に基づく直接計測 【コストウェイト】	診察件数に基づく直接計測 【コストウェイト】	診察件数に基づく直接計測 【コストウェイト】	診察件数に基づく直接計測 【コストウェイト】	診察件数に基づく直接計測 【コストウェイト】	診察件数に基づく直接計測 【コストウェイト】
スイス	インプット価格によるデフレーション	NA	NA	NA	NA	NA	NA	NA
イギリス	HRGに基づく直接数量指数 【コストウェイト】	HRGに基づく直接数量指数 【コストウェイト】	診察件数に基づく直接数量指数 【コストウェイト】	診察件数に基づく直接数量指数 【コストウェイト】	診察件数に基づく直接数量指数 【コストウェイト】	診察件数に基づく直接数量指数 【コストウェイト】	病院サービスの伸びを補助系列として延伸	病院サービスの伸びを補助系列として延伸

(出所) Schreyer, P.(2010), "Towards Measuring the Volume Output of Education and Health Services: A Handbook", OECD Statistics Working Papers, 2010/02

第3表 疾病分類の種類と分類数の例

分類の種類 (疾病分類が掲載された調査)		分類数
レセプト		約25,000
ICD	章分類	22
	基本分類	14,609
DPCの疾病分類		506
疾病分類表	大分類	85
	中分類	148
	小分類	374
患者調査	大分類	20
	中分類	38
	小分類	348
米国CCS		263
社会医療診療行為別統計	中分類	119

第4表 急性心筋梗塞に関するリスク調整モデルの例

Variable	07/2012-06/2013 OR (95% CI)
Age minus 65 (years above 65, continuous)	1.01 (1.01 - 1.01)
Male	0.90 (0.88 - 0.93)
History of Percutaneous Transluminal Coronary Angioplasty (PTCA) (ICD-9 diagnosis code V45.82; ICD-9 procedure codes 00.66, 36.06, 36.07)	0.95 (0.92 - 0.98)
History of Coronary Artery Bypass Graft (CABG) surgery (ICD-9 diagnosis code V45.81; ICD-9 procedure codes 36.10-36.16)	1.03 (0.99 - 1.07)
Congestive heart failure (CC 80)	1.19 (1.15 - 1.23)
Acute coronary syndrome (CC 81-82)	0.99 (0.96 - 1.02)
Anterior myocardial infarction (ICD-9 diagnosis codes 410.00-410.12)	1.24 (1.18 - 1.31)
Other location of myocardial infarction (ICD-9 diagnosis codes 410.20-410.62)	0.90 (0.86 - 0.94)
Angina pectoris/old myocardial infarction (CC 83)	1.07 (1.04 - 1.10)
Coronary atherosclerosis/other chronic ischemic heart disease (CC 84)	1.03 (0.99 - 1.07)
Valvular or rheumatic heart disease (CC 86)	1.13 (1.10 - 1.16)
Specified arrhythmias and other heart rhythm disorders (CC 92-93)	1.10 (1.07 - 1.13)
History of infection (CC 1, 3-6)	1.04 (1.01 - 1.07)
Metastatic cancer or acute leukemia (CC 7)	1.26 (1.16 - 1.38)
Cancer (CC 8-12)	1.01 (0.98 - 1.05)
Diabetes mellitus (DM) or DM complications (CC 15-19, 119-120)	1.19 (1.16 - 1.22)
Protein-calorie malnutrition (CC 21)	1.10 (1.05 - 1.15)
Disorders of fluid/electrolyte/acid-base (CC 22-23)	1.13 (1.09 - 1.17)
Iron deficiency or other unspecified anemias and blood disease (CC 47)	1.34 (1.30 - 1.38)
Dementia or other specified brain disorders (CC 49-50)	1.00 (0.97 - 1.04)

Hemiplegia, paraplegia, paralysis, functional disability (CC 67-69, 100-102, 177-178)	1.08 (1.02 - 1.13)
Stroke (CC 95-96)	1.00 (0.96 - 1.06)
Cerebrovascular disease (CC 97-99, 103)	1.07 (1.03 - 1.10)
Vascular or circulatory disease (CC 104-106)	1.10 (1.07 - 1.13)
Chronic Obstructive Pulmonary Disease (COPD) (CC108)	1.29 (1.25 - 1.33)
Asthma (CC 110)	1.00 (0.95 - 1.05)
Pneumonia (CC 111-113)	1.15 (1.11 - 1.19)
Dialysis status (CC 130)	1.40 (1.31 - 1.49)
Renal failure (CC 131)	1.14 (1.11 - 1.18)
Other urinary tract disorders (CC 136)	1.07 (1.04 - 1.10)
Decubitus ulcer or chronic skin ulcer (CC 148-149)	1.10 (1.05 - 1.15)

(出典) CMS(2016) Table4.2.3 より抜粋

第 5 表 我が国の先行研究における VSL の推計値

	VSL (億円、2002価格)	研究の特徴
ヘドニック賃金法		
Kniesner and Leeth(1991)	5.85	集計データを用いて分析
岡 (1999)	2.99	集計データを用いて分析
久米 (2010)	0.16	個票データを用い、個人のリスクに対する態度を考慮
ヘドニック価格法		
古川・磯崎 (2004)	7.79-9.74	自動車の価格と死亡事故率から分析
仮想市場法 (Contingent Valuation Method: CVM)		
山本・岡 (1994)	23.1-36.2	飲料水リスク削減の便益を分析
今長 (2001)	4.52	交通事故の人的損失費用について分析
松岡他 (2002)	4.05-5.57	大気汚染による健康リスクの削減効果を分析
Tsuge et al. (2005)	3.5	事故、癌、心臓病などの死亡リスク軽減のWTPからVSLを推計
Itaoka et al. (2007)	1.01-3.36	中・高年にアンケート調査し死亡リスク削減の効果を分析
内閣府 (2007)	2.09-5.11	交通事故の人的損失費用について分析

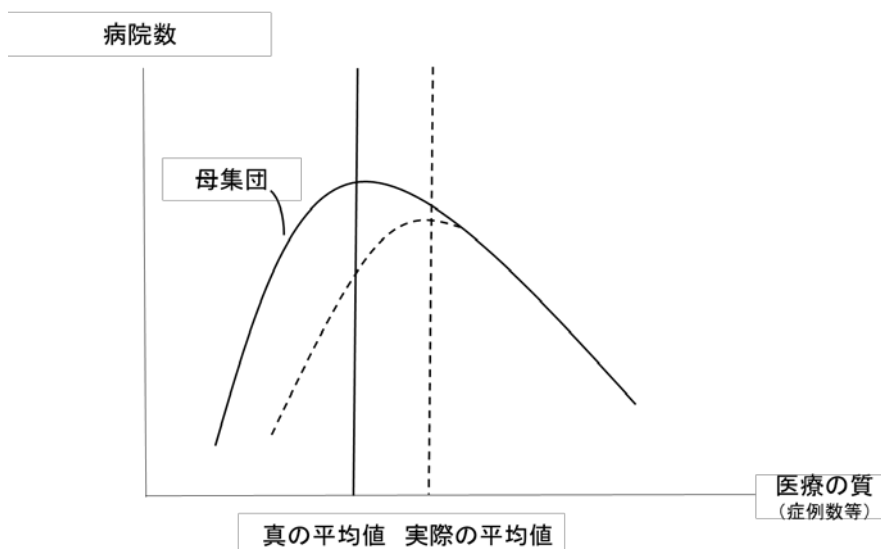
注 1) Kniesner and Leeth(1991)で死亡リスクに対する賃金プレミアムが正となるケースをもとに VSL を算出すると 5.85 億円 (2002 年価格) となった。

注 2) 岡 (1999) では死亡リスクに対する賃金プレミアムは正となっているが統計的に有意に 0 と異なる結果となっている。

注 3) 松岡他 (2002) では広島市とクアラルンプール市とでアンケート調査を行なっている。

(出典) 宮里 (2010) 表 9 より抜粋

第 3 図



【参考文献】

- 深尾京司・亀田泰佑・中村光太・難波了一・佐藤正弘, 2017, 「サービス業におけるデフレーターと実質付加価値の計測」『経済分析』第194号、pp.11-pp.44
- 乾友彦・杉原茂・川渕孝一・空閑信憲・池本賢悟・石川知宏, 2010, 「非市場型サービス産業のアウトプット計測に関する研究のサーベイー医療、教育、金融ー」ESRI Research Note, No. 12
- 宮里尚三, 2010, 「労働市場のデータを用いたValue of a Statistical Lifeの推計」日本経済研究会, No. 63
- 杉原茂・川渕孝一・池本靖子・今村育未, 2017a, 「Measuring Health Care Output」 ESRI Research Note No. 33
- 杉原茂・池本靖子・今村育未・吉田裕人・谷口優・村山陽・西真理子・藤原佳典・北村明彦・新開省二, 2017b, 「State Dependence of Long-Term Care and Preventive Effects of Care Expenditures」 ESRI Working Paper No. 42
- Atkinson, T., 2005, “Atkinson review: final report. Measurement of government output and productivity for the national accounts”
- Bureau of Economic Analysis, 2016, “Updated summary of NIPA methodologies”, *Survey of Current Business*, Vol.96, No.11
- Centers for Medicare & Medicaid Services, 2016, “Condition-Specific Measures Updates and Specifications Report Hospital-Level 30-Day Risk-Standardized Readmission Measures”
- Cutler, D., McClellan, M., Newhouse, J. and Remler, D., 1998, “Are medical prices declining? Evidence from heart attack treatments,” *The Quarterly Journal of Economics*, Vol. 113, No.4, pp.991-pp.1024
- Cutler, D, McClellan, M., Newhouse, JP., 1999. “The Costs and Benefits of Intensive Treatment for Cardiovascular Disease.” *Measuring the Prices of Medical Treatments*, edited by Jack Triplett, 34-71. Washington D.C.: The Brookings Institution, 34-71.
- Cutler D., McClellan M., Newhouse JP., 2000, “Prices and productivity in managed care insurance” *How does managed care do it?* edited by Rand J Econ. 526-48
- Dawson, D., Gravelle, H., Mahony, M., Street, A., et al., 2005, “Developing New Approaches to Measuring NHS Outputs and Activity”, *Centre For Health Economics*, CHE Research Paper No . 6

- Dunn, A., Rittmueller, L. and Whitmire, B., 2015, "Introducing the new BEA Health Care Satellite Account," *Survey of Current Business*, Vol.95, No.1
- eurostat, 2016, "Handbook on prices and volume measures in national accounts"
- Eggleton, K., Shah, N., Newhouse, JP., et al. ,2009, "The Net Value of Health Care for Patients With Type 2 Diabetes, 1997 to 2005" , *Annals of Internal Medicine* , Vol.151 , No.6 , pp.386-392
- Halm, E., Lee, C. and Chassin, M., 2000, "How is volume related to quality in health care? A systematic review of the research literature? In "Interpreting the volume-outcome relationship in the context of health care quality: Work shop summary" by Maria Hewitt for the Committee on quality of health care in America and the National Cancer Policy Board. Institute of Medicine, National Academy Press.
- Office for National Statistics, United Kingdom, 2012, "Quality adjustment of public service health output: current method"
- Office for National Statistics, United Kingdom, 2013, "Sources and methods public service productivity estimates: healthcare"
- Office for National Statistics, United Kingdom, 2015, "Public Service Productivity Estimates: Total Public Services"
- Schreyer, P., 2010, "Towards measuring the volume output of education and health services: a handbook"
- Stewart, S., Woodward, R., Rosen, A., Cutler, D., 2007, "A Proposed Method for Monitoring U.S. Population Health: Linking Symptoms, Impairments, and Health Ratings", NBER Working Paper 11358
- Viscusi, W., 2014, "The Value of Individual and Societal Risks to Life and Health", in MARK J. MACHINA , W. KIP VISCUSI , (ed) , *Handbook of the Economics of Risk and Uncertainty*, Vol 1 , Elsevier , pp.386-446