

厚生労働科学研究費補助金（長寿科学総合研究事業）
「介護保険の総合的政策評価ベンチマークシステムの開発」
平成 22 年度分担研究報告書

介護予防事業の定量的評価のための非構造的な計量経済学的手法に関する諸考察

研究分担者 野口晴子
国立社会保障・人口問題研究所社会保障基礎理論研究部・第二室長

研究要旨

本研究の目的は、介護予防事業を定量的に評価するための非構造的な計量経済学的手法について考察を行うことにある。「非構造的」という意味は、特定の経済モデルとは直結しない形で政策評価を行う手法という意味である。

介護予防事業に限らず、何かしらの「介入」を行いその効果を測定するための実証分析のゴールド・スタンダードは、「介入」による外生的なショックを利用した「実験」的手法である。その中でも王道と言えるのが、政策の対象者がまったく関与できない外生的ショックを人為的に与えて、被験者をその外生要因の対象となるトリートメント・グループ（対象群）と対象とならないコントロール・グループ（非対象群）とに無作為に分け、グループ間の効果の差を計測する社会的実験（無作為抽出化臨床比較試験(Randomized Clinical Trial: RCT)や準無作為化比較試験 (controlled clinical trial:CCT)）である。しかしながら、こうした社会的実験は、それが社会全体に関わるマクロ政策の場合は無論のこと、ミクロ・レベルでの治療評価ですら、莫大な費用がかかる上に、高齢者や重症度の高い患者、低所得者など、リスクの高い人々が実験から除外される可能性が高く、倫理的にも問題が発生しやすいという重大な欠点がある。したがって、純粋な「実験」と比べるとセレクション・バイアスが生ずる可能性は否めないが、どうしても観察によるデータを用いた分析や仮想市場法などによる「補完」が必要となってくる。

観察によるデータを用いた場合も、「実験」に近い環境を設定することが可能である。それは、自然実験と呼ばれる方法で、たとえば、2000 年における公的介護保険の導入や、本研究が対象とする介護予防事業などの大規模な制度変更を利用して、その前後で人々の行動がどのように変容するかを観察するという手法である。自然実験に際して最も適した分析手法は、差の差の (Differences-in-Differences : DD) 推定と呼ばれる手法である。今、何らかの制度変更なり新たな政策がとられて、そうした変更に影響を受ける対象群 A と影響を受けない非対象群 B が存在するとする。DD 推定法とは、変更前後で対象群が受けた平均的な影響から非対象群に起こった平均的な変化を差し引くことで、対象群と非対象群の間にもともとある固有の違いと時間経過による経年変化からくる推定のバイアスを取り除くという方法である。

先行研究から、社会的実験によらず、自然実験による分析の手法として、参加者と非参加者を疑似的に randomize することによって、政策実施前後で、参加者と非参加者との「差」の「差」を見る手法が有効であることがわかった。今後の研究の進め方としては、本研究で得られた方法論に関する知見を実際のデータに活用し、介護予防事業の政策効果を定量的に測定することにしたい。

A. 研究目的

本研究の目的は、介護予防事業を定量的に評価するための非構造的な計量経済学的手法について考察を行うことにある。「非構造的」という意味は、特定の経済モデルとは直結しない形で政策評価を行う手法という意味である（市村,2010）。

B. 研究方法

計量経済学の文献を中心として、本研究の分析対象である介護予防事業に代表されるような、何かしらの外生的な「介入」に対する効果（広義の意味での、「政策評価」）分析に対する先行研究に対するレビューを行った。政策評価、非構造的アプローチ、識別問題、selection bias、内生性（endogeneity）、社会的実験（social experiment）、自然実験（natural experiment）、仮想市場法、操作変数法（instrumental variable）、propensity scoring matching method 等をキーワードとした systematic review を行った。

C. 研究結果及び考察

介護予防事業に限らず、何かしらの「介入」を行いその効果を測定するための実証分析のゴールド・スタンダードは、「介入」による外生的なショックを利用した「実験」的手法である。その中でも王道と言えるのが、政策の対象者がまったく関与できない外生的ショックを人為的に与えて、被験者をその外生要因の対象となるトリートメント・グループ（対象群）と対象とならないコントロール・グループ（非対象群）とに無作為に分け、グループ間の効果の差を計測する社会的実験（無作為抽出化臨床比較試験(Randomized Clinical Trial: RCT)や準無作為化比較試験（controlled clinical trial:CCT）である。

社会的実験については、後にも先にも歴史上一度だけアメリカにおいて医療需要に対する大規模な社会的実験が行われた。米国の RAND 研究所が、アメリカ国内の 6 地

域から無作為に選出された 2,000 世帯に対して 1971 年から 1982 年にかけて行った Health Insurance Experiment(HIE)である（Newhouse,1993）。近年、特に米国を開発経済学者による社会的実験のデータが蓄積されており、Socio-economic status (SES) と「健康」との複雑な因果関係（causality、もしくは、gradient）を識別し、紐解こうとする試みが盛んに行われている（Julio, Eduardo, Octavio et al ,2006; Duflo E and Rohini P, 2007; Cutler,Lleras-Muney, Vogel ,2008）。

社会に対して何かしらの「介入」を行う場合は、対象者と非対象者を無作為抽出によって選別する社会的実験を実施するのが理想的ではあるが、他方、こうした社会的実験は莫大な費用がかかる上に、それが社会全体に関わるマクロ政策の場合は無論のこと、特に高齢者を対象とした介護予防事業のような政策をミクロ・レベルで評価する場合には、健康状態があまり良くない者や低所得者など、リスクの高い人々が実験から除外される可能性が高く、倫理的にも問題が発生しやすいという重大な欠点がある。したがって、純粹な「実験」と比べるとセレクション・バイアスが生ずる可能性は否めないが、どうしても観察によるデータを用いた分析や仮想市場法などによる「補完」が必要となってくる。

C-1. 介護予防事業プログラムへの参加者の特定手順

ここでは、本研究の主眼である介護予防事業の政策効果を定量的に測定するという目的のために、まず、いかにして当該事業の対象者が選ばれているか、その手順について確認することにする。図 1 は、特定高齢者の決定から介護予防プログラムへの参加までの手順を示している。まず、特定高齢者を決定するための基本チェックリストについて、平成 20 年における高齢者への新規配布数は、14,827,663 で、うち実施者数が 8,694,702 (58.6%) であった。実施者

のうちの 25.1%(2,178,952)が介護予防事業への参加候補者として選ばれ、さらに、これら候補者のうち 31.7%(690,450)が新規に特定高齢者として決定された。平成 19 年からの継続者 361,745 を併せると、計 1,052,195 が平成 20 年における特定高齢者として決定されている。したがって、チェックリストを新規に配布した高齢者のうちの 95.3%(14,137,213)は非特定高齢者として介護予防事業の対象外となっている。平成 20 年度における最終的な介護予防事業への参加者は、128,253(特定高齢者として決定された者のうち 12.2%)、非参加者は、918,475(87.3%)であった。

厚生労働省『平成 20 年介護予防事業報告』によれば、平成 20 年度に新たに特定高齢者と決定された者の把握経路として最も比率が高かったのが、「特定健康診査等の担当部局との連携」で 62.9%、二番目が「郵送等による基本チェックリストの配布・回収」で 24.6%、その他、「医療機関からの情報提供」(3.8%)、「本人・家族からの相談」(1.3%)、「地域包括支援センターの総合相談支援業務との連携」(1.2%)、「要介護認定の担当部局との連携」0.6%、「訪問活動を実施している保健部局との連携」(0.5%)、「民生委員等からの情報提供」(0.2%)、「インターネットによる情報収集」(0.0%)、「その他」(4.9%)となっている。

C-2. サンプリング・セレクション（セレクション・バイアス）の問題

分析手法を具体的に考察する以前に、まず特定高齢者の決定経路の入口において、サンプリング・セレクションの問題があることを指摘しておく（図 1 における、外生的要因によるセレクション・バイアスの可能性①）。図 2 は、県別に見た保険者による基本チェックリスト配布状況であるが、徳島県や三重県のように 9 割を超える高齢者に配布した県もあれば、和歌山県や山形県のように 1 割前後の配布率の県もあり、居住地区によって配布状況に非常に大きなば

らつきがあることがわかる。このことから、特定高齢者、さらには、介護予防事業プログラムの参加者の選定に当たっては、まずその入口のところで、居住地区という外生的要因によるセレクション・バイアスがかかっていることがわかる。

プログラム参加者の選定に当たって、こうした外生的な違いがあると、少なくとも、横断面データで、同一時点での参加者と非参加者の違いを比較するという最も単純な手法では、介護予防事業の「純粋」な効果を評価することが非常に難しくなる。なぜならば、その違いが、果たして介護予防事業の効果なのか、それとも、地域的なさまざまな属性（財政状況、実施者の熱意、地域的な個人の社会・経済的属性な違い、疾患構造の違い等）による違いなのか、識別することが困難であるからである。

この問題は、分析者がある特定地域のデータしか対象としなかった場合はさらに深刻である。全国規模のデータが活用できる場合は、地域に特有の効果を固定効果モデルによって地域属性を調整することで、この問題はある程度緩和される。しかし、たとえば、ある特定の市区町村や県についてのみ観察してしまうと、計量経済学でいうところのいわゆる「一致性（標本数が母数に近くなるとある数値に確率的に収束する）」の条件を満たすことができず、介護予防事業の効果を過剰または過小に推定してしまう結果になる。したがって、現状では、一時点の横断面データによって介護予防事業の効果を定量的に測定することは困難を極めるが、そうした限界があることをまず認識した上で、どのような手法による分析が可能かということについて考察を行うことにする。

C-3. 分析手法

では、一時点ではなく、複数時点の時系列データが活用できると仮定した場合、政策評価の手法としては次の 3 つが考えられる。

- (1) 手法 1 : 政策実施前後で、同一参加者の「差」を見る(図 3・1)。
- (2) 手法 2 : 政策への参加者 (treatment group) と非参加者(control group)の「差」を時系列で見る (図 3・2)。
- (3) 手法 3 : 政策実施前後で、参加者と非参加者の「差」の「差」を見る(図 3・3)。

図 3・1 は手法 1 で何を見ているかについて図式化したものである。横軸は時間、縦軸は評価のターゲットとなるアウトプット（たとえば、ミクロ・レベルでの評価であれば、高齢者個々人の主観的健康観、基本チェックリストの該当項目数等、マクロ・レベルの評価であれば、保険者の財政状況等）を示している。時間軸の途中で介護予防事業が実施されたとして、手法 1 が見ているのは、事業への参加者についての平均の変化である。したがって、仮に図 3・1 のように、アウトプットが上向きに変化している場合、この変化は、介護予防事業の効果のみならず、医療技術の進歩であるとか、社会的・経済的状況の改善等の時系列による変化をすべて合わせた変化量となってしまっている可能性がある。したがって、参加者のみを対象とした分析では、「純粹」な政策効果を抽出できているとは、必ずしもいえない。参考までに、表 1 は、最も単純に平均値のみで、事業参加者の主観的健康観の変化を示したものである（厚生労働省、『平成 20 年介護予防事業報告』）。表 1 によれば、介護予防事業参加者のうち、参加後に主観的健康観が改善したものが 34,655 (29.2%)、変化がなかったものが 45,388 (38.2%)、悪化したものが 13,593 (11.5%) であるが、上記でしめしたように、これらの変化をもって、介護予防事業の効果の是非を問うことは難しい。

次に、図 3・2 は、手法 2 について図式化したものである。手法 2 では、事業実施前後におけるプログラムの参加者と非参加者の平均の違いをみるとことによって、プログラムの効果を測定しようとする手法である。

この手法では、図 1 で示したように、内生要因によるセレクション・バイアス②が結果にバイアスをかける可能性がある。これは、前節 (C-2) で論じた県別のチェックリストの配布状況の違いといった、プログラムの対象者が自分ではコントロールできない外生的なバイアスとは異なり、対象者自身の属性によるバイアスが生ずる場合である。ポイントは、プログラムへの参加が、ランダムに振り分けられているのではなく、あくまでも高齢者個人あるいは家族の同意のもとに行われているという点である。たとえば、介護予防事業へ参加する人の特徴として、教育水準高く、健康問題に対する意識が高い人、あるいは、所得や貯蓄水準の高い人、比較的健康状態が良好な人が参加する確率が高いといった場合（また、非参加者がこれとは逆の特徴を持っている場合）、参加者と非参加者の平均の差は、介護予防事業プログラムの効果と高齢者個々人の潜在的な諸属性の違いの合計値を見ることになる。したがって、手法 2 についても、こうした内生的なサンプリング・セレクションを調整することなしには、「純粹」な政策効果を抽出できているとはいはず、プログラムの評価に下方または上方の歪みが生じていることになる。

したがって、参加者と非参加者について、図 3・3 (手法 3) で示したような調整が必要となる。すなわち、参加者の平均と潜在的に同様の性質をもつ非参加者の平均とを比較するという手法を用いることによって、はじめて、介護予防事業の「純粹な」効果を測定することが可能となる。理想的には、参加者と非参加者とを無作為化実験（すなわち、社会的実験）によって選定することが望ましいが、既に論じたように、ここではそうした手法は不可能である。したがって、ここでは、既に観察されたデータから疑似的に参加者と非参加者を randomize する方法として、操作変数法(Instrumental variable : IV)法と propensity scoring matching method (PSM) 法について紹介

する。

まず、IV 法についてであるが、次のような推定式を想定する。

$$q_i = \alpha_0 + \alpha_1 d_i + \alpha_2 x_i + u_i$$

この推定式で、 q_i は i 番目の高齢者のアウトプットを、 d_i は介護予防事業プログラムへの参加状況（参加は 1、それ以外は 0）、 x_i は個人の属性、 α_0 （定数項）、 α_1 、 α_2 は推定パラメータ、 u_i は誤差項を示している。ここで問題となるのは、介護予防事業プログラムへの参加状況に関する内生性の問題である。前段で論じたように、参加状況が、 x_i では調整することのできない観察不可能な個人属性に影響を受けている場合、それらはすべて u_i として吸収されてしまう。仮に、基本統計量でみる限り、参加群の方が、アウトプットが高い傾向にあるとすると、それが純粹にプログラムへ参加したことによる効果なのか、それとも、参加群の属性によるもともとのアウトプットの高さが影響しているのかを識別する必要がある。したがって、この推定式を単に最小二乗法で推計すると、 d_i と u_i は正の相関をもつことから、 d_i の推定パラメータ α_1 に正のバイアスがかかることになる。

IV 法とは、こうした内生性問題を有する「観察」データから実験的な変動を抽出し、サンプルを疑似的にランダマイズする計量手法である（川口、2008）。操作変数 (z_i) の選定に当たっては、次の 2 つの仮定が鍵となる。1 つは、 $E(u_i|z_i) = 0$ であり、これは、操作変数 z_i は外生変数であり最終的な高齢者のアウトプットとは直接的な相関が無いということを意味している。外生変数とは、被説明変数であるアウトプットや内生変数である介護予防事業への参加状況などように、推定式の体系の中でその構造に依存して決まるのではなく、外部から所与 (given) として与えられる変数のことである。2 つめの

仮定は、 $E(d_i|z_i) \neq E(d_i)$ である。これは、 z_i が高齢者のアウトプットとは無相関であるが、介護予防事業への参加とは相関があり、 z_i によってサンプルは疑似的に、参加群と非参加群とに無作為に振り分けられるということである。この仮定は $E(x_i|z_i = 1) = E(x_i|z_i = 0)$ と、セレクション・バイアスを疑似的に調整する条件として書きかえることもできる。

仮に、この 2 つの条件を満たす IV が見つかったとすると、セレクション・バイアスの調整前後で、参加群と非参加群の属性やアウトプットについて、2 つのグループの間でほとんどその平均値に差異がなくなり、サンプルが操作変数によって疑似的にランダマイズされたことになる。適切な IV が運よく発見できれば、あとは推定式を二段階推定法なり、一般化モーメント推定法

(General Method of Moments (GMM)) なり、ヘキット推定法なりで推定すればよい。

しかし、IV 推定法は、ダートマス大学経済学部の Douglas Staiger 氏がその講義の中で“balance on the head of a pin!”と評したように、参加・非参加による差異が小さい場合、IV 推定法による推定値はしばしば頑健性に問題が生ずること、 $E(u_i|z_i) = 0$ 、 $E(d_i|z_i) \neq E(d_i)$ という 2 つの仮定をペーパーフェクトに満たす適切な操作変数を見つけることが非常に難しいこと、有限サンプルにおけるバイアスが存在することなどに留意する必要がある。

参加者と非参加者を疑似的に randomize するもう 1 つの方法は、PSM 法である。この方法は、非参加者の中から、参加者と同様の性質（潜在的能力が等しい）もの同士をマッチさせて、同様の属性をもつものの同士について平均の違いを比較するという方法である。この手法は repeated cross section でも可能だが、同一個人についてプログラムの実施前後で追跡調査を行う panel data があれば、理想的である。

参加者と同じ性質をもつ非参加者の決定

の仕方は、図4で示すように、まず、介護予防事業への参加を決定する（複数の）潜在的要因の効果を、ロジットまたはプロビットによって定量的に推計し、その推計から、各個人の参加確率を推計する。次に、各参加者を参加確率の最も近い非参加者とマッチさせて、参加者とマッチさせた非参加者の性質が平均的に等しいかどうかを統計学的に balancing test を用いて検証する。Balancing test によって、統計的に有意な「差」がない者を非参加群として、その平均値を比較する。

これは、いわゆる「自然実験」と呼ばれる方法で、たとえば、2000年における公的介護保険の導入や、本研究が対象とするような介護予防事業の実施など、大規模な制度変更やを利用して、その前後で人々の行動がどのように変容するかを観察するという手法である。自然実験に際して、平均値の差を比較する最も適した分析手法は、差の差の (Differences-in-Differences : DD) 推定と呼ばれる手法である (Ashenfelter and Card, 1985 ; Wooldridge JM(2001))。今、何らかの制度変更なり新たな政策がとられて、そうした変更に影響を受ける対象群Aと影響を受けない非対象群Bが存在するとする。これまで論じてきたように、単純に、変更後における対象群と非対象群との違いを比較したとしても、それが制度変更によるインパクトなのか、両方のグループ間にあるもともとの違いなのかという内生性の問題が生じ、それらを識別することができない。さらには、対象群のみを対象として、変更前後の違いを時系列的に観察したとしても、今度はその違いが外生的ショックによる効果なのか、それとも単に時間的経過によって自然に生じた違いなのかを判別できない。DD推定法とは、変更前後で対象群が受けた平均的な影響 ($\bar{q}_{A,2} - \bar{q}_{A,1}$) から非対象群に起った平均的な変化 ($\bar{q}_{B,2} - \bar{q}_{B,1}$) を差し引くことで、対象群（本研究の場合は参加者（群））と非対象群（非参加者（群））の間にもともとある

固有の違いと時間経過による経年的変化からくる推定のバイアスを取り除くという方法である。したがって、推計量 ($\hat{\alpha}_1$) は最も単純に、 $\hat{\alpha}_1 = (\bar{q}_{A,2} - \bar{q}_{A,1}) - (\bar{q}_{B,2} - \bar{q}_{B,1})$ のように表現することができる。

これまで、日本ではデータの制約上、ほとんどの実証研究は、DD推定法の枠組みでいうと、非対象群に対する効果は観察不可能であり、変更前後で対象群が受けた平均的な影響 ($\bar{q}_{A,2} - \bar{q}_{A,1}$) を推定するにとどまっている。

D. 結論

本研究の目的は、介護予防事業を定量的に評価するための非構造的な計量経済学的手法について考察を行うことであった。先行研究から、社会的実験によらず、自然実験による分析の手法として、参加者と非参加者を疑似的に randomize することによって、政策実施前後で、参加者と非参加者の「差」の「差」を見る手法が有効であることがわかった。

今後の研究の進め方としては、わが国における数少ない高齢者を対象とした貴重なパネルデータである日本福祉大学健康社会研究センターを中心とした『愛知老年学的評価研究 (Aichi Gerontological Evaluation Study : AGES)』を用いて、介護予防事業の定量的な効果分析を行っていきたい。

E. 健康危険情報

なし

F. 研究発表

- 1.論文発表
なし
- 2.学会発表
なし

G. 知的所有権の取得状況の出願・登録状況

- 1.特許取得
なし

2.実用新案登録

厚生労働省『平成20年介護予防事業報告』

なし

3.その他

参考文献

Ashenfelter O and Card D(1985) "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs", *The Review of Economics and Statistics*, 67:4 (1985), pp. 648-660.

Cutler DM, Lleras-Muney A, and Vogl T. (2008) "Socioeconomic Status and Health: Dimensions and Mechanisms". *NBER Working Papers* 14333.

Duflo E and Rohini P (2007), "DAMS," *Quarterly Journal of Economics*. May, pp. 601-646.

Julio F, Eduardo GP, Octavio GD, Miguel AL, Felicia MK(2006) "Comprehensive reform to improve health system performance in Mexico," *The Lancet*, 368(9546): pp1524 – 1534,

Newhouse JP (1993) "Free for all? Lessons from the RAND Health Insurance Experiment," Cambridge, MA: Harvard University Press.

Wooldridge JM(2001) "Econometric Analysis of Cross Section and Panel Data," The MIT Press.

市村英彦(2010)「ミクロ実証分析の進展と今後の展望」、日本経済学会(編)『日本経済学会75年史—回顧と展望—』第8章、pp.289-361.

川口大司(2008)「労働政策評価の計量経済学」『日本労働研究雑誌』569 : pp.16-28.

図1 特定高齢者の決定から介護予防プログラムへの参加までの手順

Source: [http://www.wam.go.jp/wamappl/bb05Kaig.nsf/0/61ebcd360d438363492572910007b39e/\\$FILE/20070301_2sankou.pdf](http://www.wam.go.jp/wamappl/bb05Kaig.nsf/0/61ebcd360d438363492572910007b39e/$FILE/20070301_2sankou.pdf)

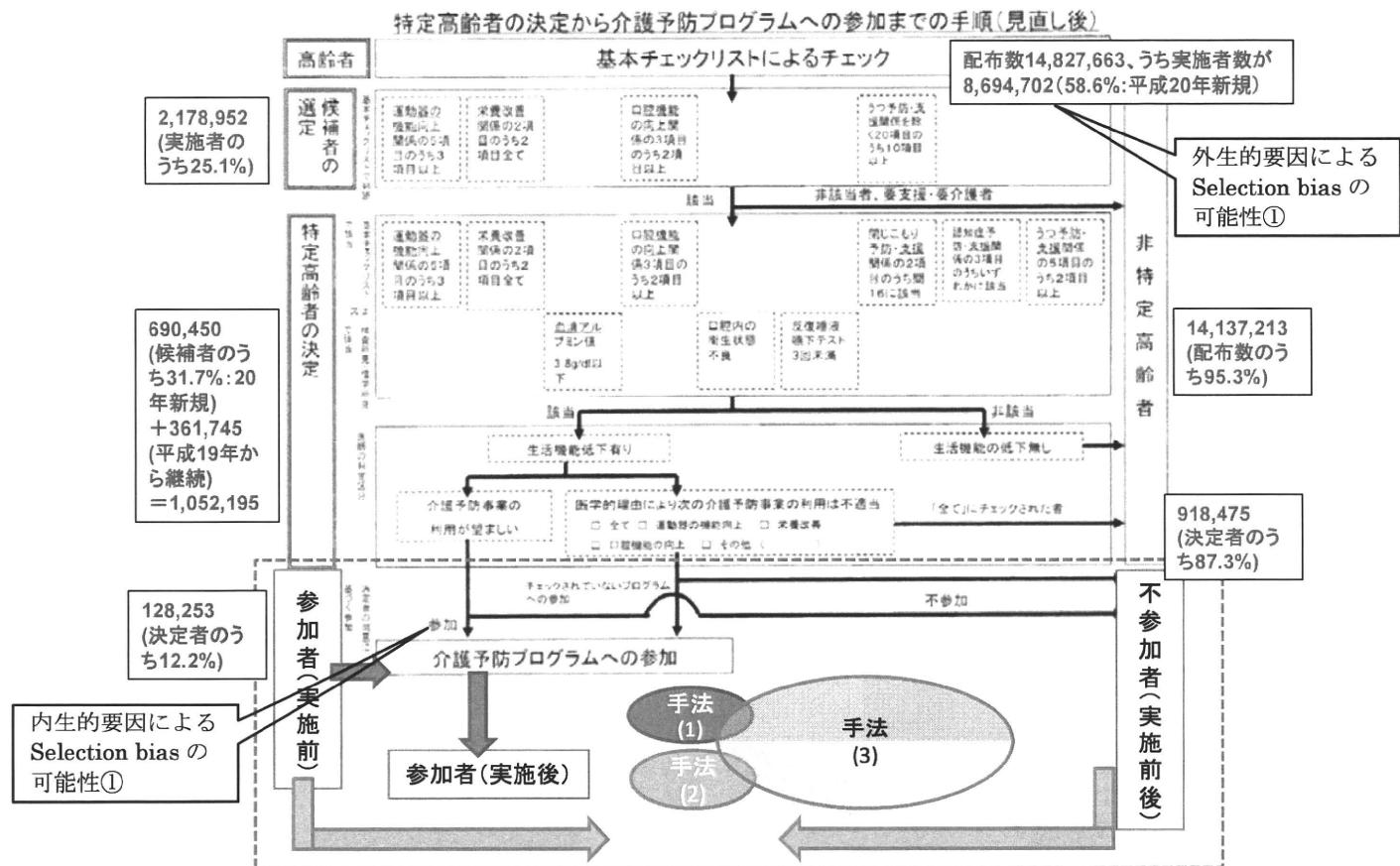
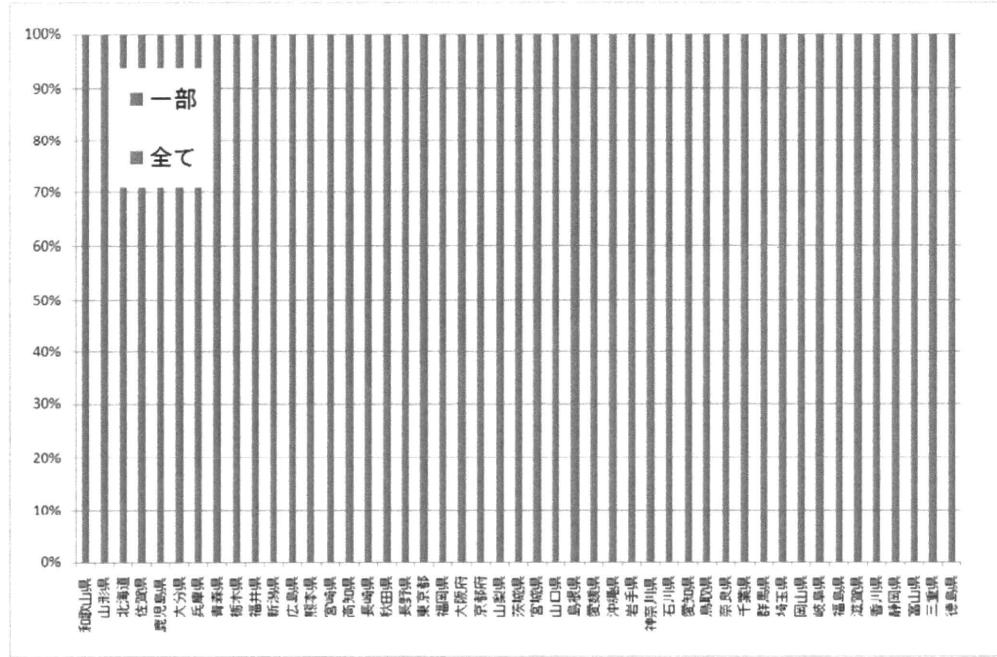


図2 県別に見た保険者による基本チェックリスト配布状況



出所:『平成20年介護予防事業報告』

図 3-1

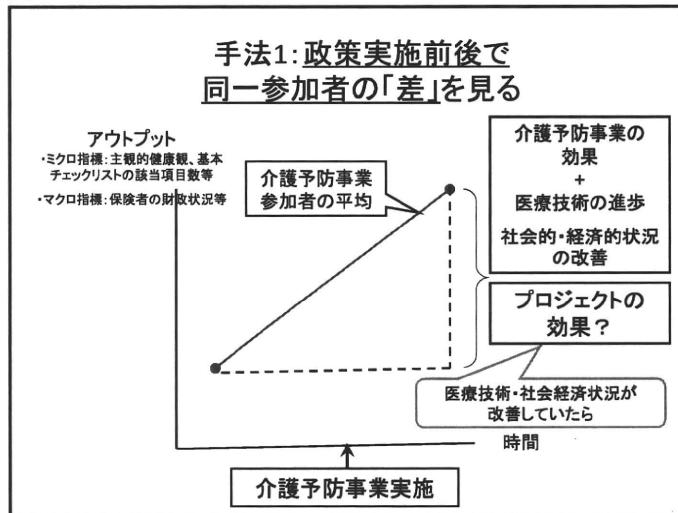


図 3-2

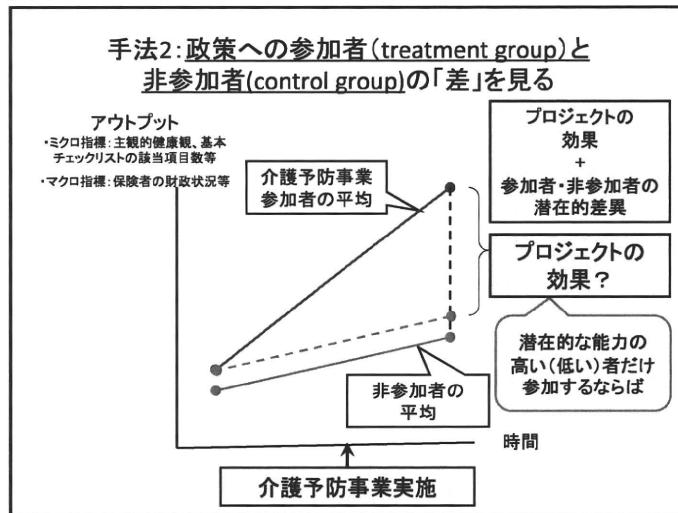


図 3-3

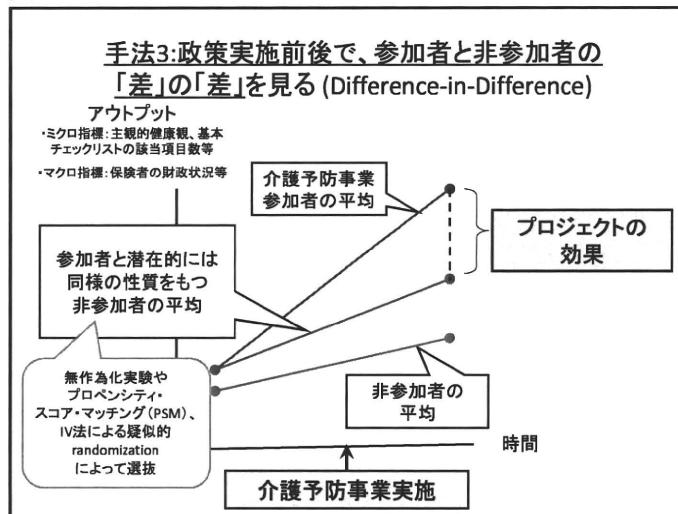


図 4

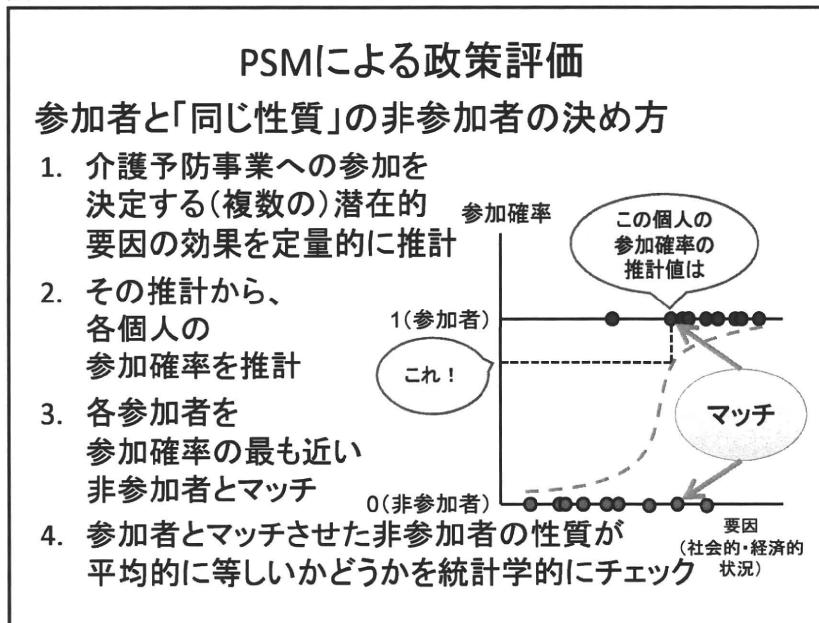


表 1

手法1(事例): 特定高齢者施策に参加した者の主観的健康感の状況

		実施後							悪化13,593 (11.5%)	
		よい	まあよい	ふつう	あまりよくな	よくない	不明	計	(単位: 人/%)	
実施前	よい	5 063 56.7%	1 431 16.0%	1 315 14.7%	330 3.7%	50 0.6%	744 8.3%	8 933 100.0%		
	まあよい	3 741 18.6%	9 078 45.3%	4 215 21.0%	1 151 5.7%	161 0.8%	1 715 8.5%	20 061 100.0%		
	ふつう	4 892 10.2%	10 823 22.6%	23 872 49.9%	3 687 7.7%	489 1.0%	4 099 8.6%	47 872 100.0%		
	あまりよくな	1 785 7.9%	3 977 17.6%	7 169 31.7%	6 749 29.9%	754 3.3%	2 151 9.5%	22 585 100.0%		
	よくない	280 8.5%	460 14.0%	697 21.3%	831 25.3%	626 19.1%	385 11.7%	3 279 100.0%		
	不明	558 3.5%	713 4.5%	1 038 6.5%	440 2.8%	87 0.5%	13 134 82.2%	15 970 100.0%		
			改善者34,655 (29.2%)		変化なし45,388 (38.2%)					

出所:『平成20年介護予防事業報告』

平成 22 年度厚生労働科学研究費補助金（長寿科学総合研究事業）
分担研究報告書

介護保険ベンチマークの健康影響評価への活用

研究分担者 藤野 善久（産業医科大学 公衆衛生学教室）

研究要旨

本項では、介護保険ベンチマークを活用した健康影響評価（HIA）の可能性について論考する。HIA は健康の社会環境モデルにもとづいて様々な政策や事業の健康影響を事前に予測し、提案された政策の適正化を試みる社会的な制度である。介護保険ベンチマークとして検討される各種の指標は、広範囲な社会的健康規定要因を含んでいること、比較的小規模地域のデータとして利用できること、また一般的に健康影響を受けやすい高齢者の特徴を捉えたものであることなど、HIA への応用という点から見た場合に豊富な情報を提供することができる。これらは、HIA を実施する際のスクリーニングやアセスメントや、健康格差対策の検討などに活用が可能である。

はじめに

Health Impact Assessment（HIA：健康影響評価）とは、影響評価（Impact Assessment）の手法の一つで、1990年代頃より諸外国において積極的に活用されてきたものであり、提案された政策・施策・事業に対して、健康影響を予測し、不利益を最小化し、便益を最大限にすることによって政策形成に貢献しようという試みである^{1,2)}。既に欧州を中心に、ダム建設や空港建設などの大型事業や地域開発、雇用政策などさまざまな分野で活用が始まっており、WHO Gothenburg paper によると HIA とは提案された政策、施策、事業によって起こりうる集団の健康影響と、集団内での影響の格差を判断するためのプロセス、方法、ツールの組み合わせであるとしている（原文 “a combination of procedures, methods, and tools by which a policy, a program or a project may be judged as to its potential”）。

effects on health of a population and the distribution of effects within the population”)³⁾。

本項は、介護保険ベンチマークが HIA にどのように応用できるかについて考察したい。

HIA の景と健康の社会環境モデル

HIA が普及してきた背景にあるのは、人々の健康を規定しているのは、医療技術の進歩や医療制度だけではなく、むしろ経済、環境、衛生、教育、住宅、雇用など「保健医療分野以外の政策」であるとの認識である。ヘルスプロモーションを提唱したオタワ憲章⁴⁾においても、健康の前提条件として、平和、住居、教育、食料、収入、安定した環境、持続可能な資源、社会的正義と公平が謳われている。個人もしくは集団の健康は、社会構造や社会経済的要因に直接・間接的に影響を受けており、そのような要因を社会的健康規定要因と呼ぶ（図 1）^{5), 6)}。その前提に立つのであれば、「保健医療分野以外の政策」の健康影響を適正化しようという試みは極めて合理的である。

そのため HIA では健康の社会環境モデルを基盤に採用している。HIA では、この社会環境モデルに基づいて広範な政策を評価の対象とする。

HIA の手順

HIA の実施方法については、各種ガイドラインがそれぞれの地域事情に沿ったモデルを提唱している^{7), 8)}。その多くは、他の影響評価（Impact Assessment）の手法と同様に、「1. スクリーニング」「2. スコピング」「3. アプレーザル」「4. レポートティング」「5. モニタリングと評価」という 5 つのステップに則り実施していくことを提唱している。

最初のスクリーニングでは、提案された政策を HIA の実施対象とすべきかどうかを選択する。提案された政策による健康影響が重大と予想され、さらに詳細なアセスメントが必要

要とされる場合には、次のスコーピングへと進む。ちなみにスクリーニングの為のチェックリストなどのツールが各種ガイドラインにおいて開発されている(21, 22)。

スコーピングのステップでは、健康影響を与える可能性があるとスクリーニングの段階で判断された政策について、HIA の全体的な実施方法について決めていく。ここでは HIA を実施するチームの設置やその権限と責任、実施スケジュール、予算などについて定める。

また、どのような健康影響に関して HIA を実施するのかなどの大体の構想を得ておく。

アプレーザルは、HIA の中で最も重要なステップである。このステップでは、政策によって起きる可能性のある健康影響を検討し、それに関するエビデンスの評価を行う。そのためには、政策の分析や、政策によって影響をうける集団や地域の特性を分析して、直接的・間接的な影響を把握する。次に、健康影響の性質、便益か不利益か、影響の大きさや範囲、時間的な経過などについて専門家意見、文献などを参考に評価を実施する。この際には、定量的な評価だけでなく質的な評価を用いることがあるのも HIA の特徴の一つである。

レポートティングのステップでは、健康影響の評価に基づいて、健康上の不利益を軽減し、また健康上の便益を促進するために、政策の修正や追加的措置などの助言を作成する。

「モニタリングと評価」のステップには、2つの意味がある。一つは、政策形成の過程において HIA が実際にどのように関与したか、またはしなかったかを評価することである。もう一つは、政策実施後に実際にどのような健康影響が生じたかについて短期、長期的に評価することである。

介護保険ベンチマークの HIA への応用

本研究が目的とする介護保険ベンチマークは、今後、自治体などにおける様々な計画作成時や事業に関する HIA の実施が求められた際に有用な情報源と成り得る⁹⁾。特に、介護保険ベンチマークは、スクリーニングおよびアプレーザルの実施において、活用が期待できる。

スクリーニングでは、自治体での計画作成時や事業の立案時などに、住民にどのような健康影響があるかを評価する。その際、多くのガイドラインではスクリーニングツールが利用される。このスクリーニングツールは、上記の「健康の社会環境モデル」に基づいて作成されており、「健康の社会環境モデル」を通して、計画や事業がどのような健康影響、もしくは健康規定要因への影響があるかを見通すことから、ヘルスレンズアプローチとも呼ばれる。

スクリーニングでは、このような「健康の社会環境モデル」に基づいて、企画されている計画や事業が、

- ・ どのような健康影響があるか？
- ・ どのような健康規定要因への影響があるか？
- ・ どのような集団へ影響するか？
- ・ 特に不利益な影響を受ける集団が存在するか？

などを評価していく。

これらの作業においては、必然的に、人口構成、所得分布、地域の社会経済状況などの地域のプロファイルは疾病や医療サービスに関する情報と合わせて最も基本となる情報である。しかしながら、これらの情報の全てが容易に取れるとは限らない。特に、日常的に利用可能な既存の行政データにおいて、「社会的健康規定要因」を含んだ指標、統計は、皆無である。さらに、実際に必要となる地区サイズ（多くは小学校区単位や、もっと小さな単位）となると、さらに限られているのが現状である。

介護保険ベンチマークは、このような状況にいくつかの大きな貢献が可能である。

- 1) 全国で統一された比較可能性のあるデータ
- 2) 行政データで利用可能なので、いわゆる業務統計として、別途にデータを収集するコストが不要

- 3) 希望とする地区サイズでの集計が可能
- 4) 比較的、広範囲の社会的健康規定要因を含む指標
- 5) 一般的に脆弱な集団（不利な健康影響を受けやすい）とされる「高齢者」のデータを提供

最後に

介護保険ベンチマークは、介護保険の事業計画や評価への利用のみだけでなく、例えば、地域開発の際に、その地区及び周辺に存在する住民、特に高齢者にどのような影響があるかを把握する際に、介護保健ベンチマークを活用することが想定される。大きな道路計画や地域の交通政策において、どのような支援が必要となるかなどのプロファイルにも活用が可能である。

介護保険ベンチマークが、将来、このようなHIAへのニーズに応えるための条件として、広範囲の社会的健康規定要因が反映されたものであること、また、介護保険行政だけではなく、多くの政策部署で日常的に活用できる形態でデータベースが運用されることが込まれる。

参考文献

- 1) Kemm J, Parry J, Palmer S eds. 健康影響評価～概念・理論・方法および実施例～監訳）藤野善久,松田晋哉. 東京：社会保険研究所, 2008.
- 2) 藤野善久, 松田晋哉. Health Impact Assessment の基本的概念および日本での今後の取り組みに関する考察. 日本公衆衛生学雑誌 2007;54:73-80.
- 3) WHO European Centre for Health Policy. Health Impact Assessment: main concepts and suggested approach Gothenburg paper. 1999.
- 4) WHO. Ottawa charter on health promotion. Geneva, 1986.
- 5) Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health. Stockholm, 1991.
- 6) Whitehead M. The concepts and principles of equity and health. Copenhagen: World Health Organization, 1990.

- 7) Commonwealth of Australia. Health Impact Assessment Guidelines. 2001.
- 8) Scott-Samuel A, Birley M, Ardern K. The Merseyside Guidelines for Health Impact Assessment. Liverpool: the International Health Impact Assessment Consortium, 2001.
- 9) 藤野善久. 【自治体中心の新たな健康政策 Health Impact Assessment の導入】健康影響評価の概要とその応用の可能性 背景,概念,定義,基盤,手法など. 公衆衛生 2009;73:483-7.

厚生労働科学研究費補助金（長寿科学総合研究事業）
分担研究報告書

介護予防対策の公正性評価のあり方に関する概念整理

研究分担者 近藤 尚己

(山梨大学大学院 医学工学総合研究部社会医学講座 講師)

研究要旨 介護予防を含む公共サービスは、あらゆる集団に対して公正であることが望ましい。本研究では、世界保健機関の報告や英国での取り組み例をレビューし、日本における介護予防対策の公正性の評価の視点のあり方について概念整理をした。整理された事項をもとに、具体的な評価指標とその運用法について現在検討を進めている。

A. 研究目的

介護予防を含む公共サービスは、さまざまな集団に対して公平に提供され、集団間でのパフォーマンスも均一であることが望ましい。健康政策評価に関して、その公正性、すなわち「公平で偏っていないこと」の重要性は近年広く認識されてきており、政策評価のあり方に関するアプローチの多くが公正の概念を取り入れた具体的な提案を行っている。ここでいう「集団の違い」を定義するものとしては、居住地域、人種、性、年齢、社会経済状況（所得、学歴、職種）など、幅広く存在する。

日本の健康政策の多くにおいては、これまでこのような集団間の違いを考慮した系統的な保健政策の評価はあまり行ってきていない。年齢（幼児・学童・成人・高齢者）、性（母子）、産業（労働衛生）といったように、保健制度自身は集団別に実施され評価されてきたが、その評価のほとんどは、対象集団全体の平均値や達成割合（たとえば塩分摂取量や喫煙率）といった、全体的指標のみで行われ、その集団間・集団内のはらつきの拡大・縮小といった観点での評価は行われてこなかった。

介護保険制度は、市区町村を保険者単位として高度に分権化することで、地域の特徴に

即したサービスの展開を期待するとともに、それぞれの自治体が互いに競い合う中で制度全体としてのパフォーマンスを向上させることも期待している。本研究班は、自治体間で比較可能な標準的なベンチマー킹・システムを構築することで、そのような介護保険制度のパフォーマンス向上に貢献しようとするものである。本報告では、そのようなベンチマーキング指標の体系において、「公平性」の観点をどのように位置づけ、それを測定する際にどのような視点に基づく評価を行うべきかについて検討するための第一歩として、先行する諸外国や国際機関の取り組みや研究をレビューし、保健医療制度における公正性の評価法やその背景理論を整理することである。

B. 研究方法

医療制度の公正性の評価を行ってきた先駆的な事例や最新の議論として、本研究班では（1）世界保健機関のWorld Health Report 2000 「Health Systems: Improving Performance」(WHO, 2000)による、国別保健システムの質の評価における公正性評価法について検討し、次に英国における政策の公正性追求のあり方をまとめた報告書（2）「Fair Society, Healthy

Lives: The Marmot Review」(Strategic Review of Health Inequalities in England post-2010, 2010)をレビューした。

上記レビューをもとに、日本の介護保険制度における公正の指標およびその運用あり方に関する考察を行った。

C. 研究結果

(1) 世界保健機関「World Health Report 2000 「Health Systems: Improving Performance」における公正性評価

本報告書は、世界保健機関加盟国の保健システムの調査時点における質について、標準化したベンチマークリング指標により横並びに評価・ランク付けすることを試みた先駆的・意欲的な試みとして知られている。同報告書が保健システムの質として評価の枠組みとしたのは、

- a. 到達度(attainment)： 健康指標が集団全体でよく(goodness)、集団間・集団内においてばらつきが少ないこと(fairness)
- b. 効率性(performance)： コストに対して結果がよいこと
- c. 責任性 (responsiveness) 医療以外の期待に応えているか：満足度)

であった。

公正性の観点で重要なのは、下線のように「集団間・集団内においてばらつきが少ないこと」を挙げている点である。家計の保護の面も強調される。つまり、必要なケアを受けるための支払いが、一気にその世帯を貧困状態に陥れるような状況(Catastrophic payment)は望まれない、ということである。

同報告書では、さらに「責任性responsiveness のばらつきも考慮すべき」としている。つまり

国民が保健医療に期待することに対する制度の対応が集団間で異なることも重要ということである。例えば、個人（患者）への人道的配慮、守秘義務の履行、患者の自主性や自己決定権への配慮、必要なケア受けるまでの待機時間、ケアに必要なアメニティ、サービス提供者の選択の自由とその幅といった観点における責任性である。

(2) 英国「Fair Society, Healthy Lives: The Marmot Review」

Marmot Reviewは、英国における現在の健康格差に関する実証的エビデンスを収集し、それをもとに、2010年以降の同国における行動計画を提案することを目的としている。そのキー・メッセージは以下のとおりである。

1. 健康格差の是正は公平と正義の問題である。
2. 健康には社会的勾配がある（社会的地位が低いほど不健康である）。この勾配を緩やかにするような行動が求められる。
3. 健康格差は社会格差の結果である。健康格差に関する行動はあらゆる健康の社会的決定要因に渡るべきである。
4. 最も不利な人々のみへの対策では健康格差の是正としては不十分である。健康の社会勾配を緩やかにするためには、行動は全体的（ユニバーサル）でなくてはならない。しかし、対策の強さは不利な度合いに比例させ、変えるべきである（社会経済的に不利な人にほど手厚くする）。これを比例全体主義 proportionate universalismとよぶ。
5. 健康格差の是正は社会に対して様々な形で利益を生む。健康格差に関連する不健康からくる経済的損失を減らすことがで

- きる。生産的損失、税収の減少、社会保障費の増大、そして治療コストの増大などが是正されることが、現状で明らかになっている。
6. 経済発展は国家の最高のための最重要戦略ではない。健康の公平な分配、well-being、そして持続可能性が最も重要な社会目標である。
 7. 健康格差の是正のための政策目標を達成するには、中央政府、地方自治体、NHS（英国の公的医療システム）、第三機関、民間組織、さらに地域団体による行動が求められる。
 8. あらゆる政策における健康の公平性に特化した地域における効果的なデリバリー・システムなしには、国家的政策も機能しない。
 9. 効果的な地域におけるサービス提供のためのシステムには効果的な地域参加型の決定（decision making）が必要であり、これは個人と地域コミュニティを強化することのみによりなし得る。

D. 考察

本レビューにより、保健システムの評価における公正の視点として重視されるべき事項および実際の評価のための指標のあり方、またその運用のあり方について、複数の示唆を得た。以下に記す。

1. 集団を、社会経済的な優位性という観点から明確に定義し、それらの集団の把握、あるいは個人および地域の社会経済状況の把握を行うことが必要である。居住地域・所得・職業・学歴といった面での把握が求められる。
2. 各指標について、対象者全体の平均値や割合のみで評価するのではなく、社会経済状況の異なる集団や地域それぞれごとの平均値・割合の算出とベンチマークを行った必要がある。また、集団全体の平均値についてもそのばらつきも併せて評価することが求められる。

3. サービスの質が、各集団に対してそれぞれ適切か（各集団のニーズに合致し、集団ごとの満足度が同等か）を評価する必要がある。
4. 社会経済的に弱い立場になるほどより強い制度上の保護がなされていること（proportional universalism）の考え方、介護保険制度の「公正性」を軸とした運営上の戦略として参考になる。また、ベンチマーク指標開発の際の視点としても役立つ可能性がある。
5. 介護制度の計画段階において、各集団を可能な限り代表し得る複数の住民側代表者を含めた検討組織を構成し、地域住民が参加した計画策定のプロセス（地域参加型の計画策定）が踏まれているかを確認すべきである。

E. 結論

本研究班では、以上をもとに、介護制度の評価のための指標の開発を行ってきた。具体的な項目及びその運用の仕方については、ベンチマーク・システムの開発に関する別稿に譲る。

引用文献

- STRATEGIC REVIEW OF HEALTH INEQUALITIES IN ENGLAND POST-2010 (2010)
Fair Society, Healthy Lives: The Marmot Review. IN MARMOT, M. (Ed.)
WHO (2000) *The World Health Report 2000*, World Health Organization.

厚生労働科学研究費補助金（長寿科学総合研究事業）
分担研究報告書

介護予防ベンチマークに関する先行指標の整理

研究代表者 近藤 克則（日本福祉大学社会福祉学部教授）

研究協力者 斎藤 雅茂（日本福祉大学地域ケア研究推進センター主任研究員）

研究要旨：介護予防事業ベンチマーク指標群（2010年度版）の開発に際して、基礎資料として活用した先行指標をレビューした。厚生労働省が示した介護予防事業の市町村チェックリストのほか、既存の介護給付実績情報を活用したいいくつかの有益な指標が確認された。それと同時に、介護給付実績などの行政データには、データの正確性や追加調査を要しない点などの利点がある一方で、健康の社会的決定要因において重要な変数が含まれておらず、独自調査によるデータを統合する必要があることも確認された。ベンチマーク指標の開発とともに、評価の定着にむけた実務・運用面での課題については今後さらに検討を重ねる必要がある。

A. 目的

本研究班では、介護予防に関するベンチマークシステムの開発と運用によって、介護保険制度の保険者間および保険者内の小地域間・複数時点間の相対的な比較を可能にし、当該自治体の課題と現実に即した目標設定を示すことを目指している。先進的な取り組みとして、イギリスでは、New Public Managementの一環から社会サービス業績評価フレームワーク（PSS-PAF : Personal Social Services: Performance Assessment Framework）に基づく自治体パフォーマンスに関する400指標のベンチマークが設定され、全国統一基準による自治体間の比較と自治体業績情報の公開が進められている。また、アメリカでは、ナーシングホームにおけるケアの質マネジメントシステムなどもある。

近年、国内では、人口統計だけでなく、介護保険事業状況報告（月報）がWebサイト（たとえば、Wam-net (<http://www.wam.go.jp/>)ほか）に無料公開されるなど、介護保険事業に関する保険者間の比較や経年変化を把握できるデータは比較的容易に入手できるようになってきた。しかし、介護予防事業に限定すると、地域特性を踏まえた自治体の取り組むべき課題を検討でき

るベンチマークシステムは未だ提示されていない。

とりわけ、介護予防事業に関しては、大規模な税金と介護保険料が投入されているという背景から、国を含む行政の説明責任（アカウンタビリティ）という意味で、その実績と成果を評価するベンチマークシステムが求められている。2010年11月に行われた行政刷新会議による事業仕分けでは、「科学的根拠に基づく調査・研究を行い、エビデンスを集め、費用対効果を計算し、政策評価を行ったうえで事業を継続すべきかどうか、さらに伸ばしていくかどうかについて検討するという姿が望ましい制度設計のありかた」、「事業の効果評価が見えないまま、前年度と同額の予算要求はうなづけない」とし、介護予防事業には予算の縮減という結論が出されていた。

また、保険者自らが相対的な特性を把握できるデータベースの構築は、介護予防事業の進行管理のうえでも重要な意義がある。少なくとも現時点では、介護予防を目的にした事業に関して、模範となるような効果的・効率的な事業展開や施策は明らかになっていない。このため、当該事業を試行しながら、他の事業所・自治体との比較や自らの前後比較など適切な事業評価を重ねて、