

連載

ヘルスサービスリサーチ(15)

「ヘルスサービスリサーチに対する計量経済学的手法の応用
—観察（非実験）データの可能性を探る—」

国立社会保障・人口問題研究所 社会保障基礎理論研究部 野口 晴子

1. はじめに —観察（非実験）データに基づく
因果推論—

計量経済学 (econometrics) とは、経済社会システムや市場、また、その中における諸個人の意思決定（行動）などを観察の対象として、経済学がアダム・スミス以来約300年間にわたって構築してきた経済理論に、「実証的な内容をもたせ、それらを立証したり、反証したりするために、経済データの分析に統計的・数学的方法を応用する」学問である¹⁾。計量経済学は、そうした意味で、分析対象は異なるものの、生物統計学、計量社会学、生体測定学、計量心理学などと目的を同じくする学問であるといえる。20世紀における計量経済学の発展は、1930年における国際計量経済学会 (Econometric Society) の創設と、当該学会によって1933年に創刊された“エコノメトリカ (Econometrica)”誌によるところが大きい¹⁾。同誌では、創刊号以来、計量経済学の目的と研究対象が、次のように記されている。「経済的な諸問題を解決するために経済理論—質的および実証的—数量的なアプローチを統合し、構造的かつ厳密な思考に基づいて研究を進めることが計量経済学の主な目的である。そして、このような方法は自然科学で支配的な方法でもある^{1),2)}。

本稿の目的に照らして、ここで着目すべきは、計量経済学が、単なる「測定」を超えて、「構造的」思考に基づいて研究を進めることを目的としている、という点である。「構造的」思考とは、とりもなおさず、因果推論 (causal inference) の手法を経済社会における諸現象に適用するということを意味しており、因果推論の手法とは因果的順序のロジックを伴う統計的推論にはかならない³⁾。さらに、もう1つの重要な点は、計量経済学が、自然科学とは異なり、無作為抽出化試験 (Randomized Control Trial: RCT) などの実験的手法を用いることが極めて困難な「経済」や「社会」を分析の対象としているということである。近年、特に米国の研究者を中心とした実験系・行動系の開発経済学において、社

会的実験がアフリカなどの発展途上地域において頻繁に実施されているが、人道的見地に照らして、その妥当性が問われることも多い⁴⁾。医療経済学の分野では、アメリカにおいて、医療サービスに対する患者の自己負担率を無作為に割り当てるという大規模な社会的実験が、1974年から1982年にかけて、RAND研究所によって行われ (Health Insurance Experiment)、本研究は、現在に至るまで医療需要分析における試金石となっている^{5,6)}。しかし、こうした経済社会全般を対象とした実験は、倫理面はもとより、費用がかかりすぎるため、とりわけ先進地域での実施が難しい。こうした背景から、計量経済学では、非実験による観察データに基づく因果推論の手法が数多く開発されてきた。

本稿の構成は以下のとおりである。次節では、RCTと観察に基づくデータを比較することによって、観察データに特有のセレクション・バイアスと、因果の方向性にかかわる内生性の問題について概観する。第3節では、その調整方法の一例として操作変数法を紹介する。最終節では、日本における医療情報のさらなる整備へ向けて若干の提言を行う。

2. 「RCT」対「非実験による観察データ」—
セレクション・バイアスと内生性—

近年、日本においても、研究者によるデータ収集に対する不断の努力もあって、レセプトや診断群分類 (Diagnosis Procedure Combination: DPC) データ、研究者独自の調査など、非実験的手法 (記録) による観察データの個票を利用した実証研究については枚挙にいとまがない。2009年における統計法改正は、こうした流れを後押しするものであり、官庁統計に対する利便性が向上することで今後さらなる実証研究の蓄積が期待される。しかし、情報が「実験」によらず「観察」データである以上、たとえば、年齢、性別、人種、家族構成、地理的環境、重症度、副疾患の有無、経済状況、保険種別など個々の患者属性に結果が大きく依存するため、治療や政策の影響を正確に推定することは、統計上困難を極め

る。さらに、医療情報の場合は、カルテなどからの情報収集の際、頻繁に起こりうる欠損値の問題も深刻である⁷⁾。したがって、さまざまな属性が、観察可能、あるいは、欠損値または脱落変数 (missing variable) や予測不可能な諸要因による影響がある場合には観察不可能な形で影響し、分析結果を偏らせるという問題が、いわゆるセレクション・バイアスである。

本節では、(旧) 米国保険財務庁 (現, Center for Medicare and Medicaid Services: CMS) と米国医学会の呼びかけにより、Collaborative Cardiovascular Project (CCP) によって収集された観察データを RCT と比較することで、セレクション・バイアスとそれに起因する内生性の実態について概観する。

1998年の *New England Journal of Medicine* 誌に、VANQWISH (Veterans Affairs Non-Q-Wave Infarction Strategies in Hospital) チームによる研究が発表された。この研究では、“stable patients with a non-Q-wave myocardial infarction” を対象とした RCT が実施され、入院後 7 日以内における心臓カテーテル検査 (catheterization: CATH) を起点とする「積極的治療」の入院後 30 日および 1 年以内死亡に与える効果が推定されている⁸⁾。本研究は、治療リスクの高い積極的治療について RCT を実施したという点で、当時、非常に貴重な研究であった。

これに対して、Geppert et al. では、CCP の観察データを用いて、方法論の面から、VANQWISH 研究との比較検証を目的とした研究を行った⁹⁾。CCP では、1994-1995年に AMI で入院した約 20 万人のメディケア受益者 (したがって、65 歳以上) 全数について、専門家のコンセンサスに基づき標準化された 100 指標以上の詳細な医療情報が、各医療施設のカルテから 1 人当たり 100 ドルという高額な予算をかけて収集された¹⁰⁾。本研究では、まず、CCP サンプルから VANQWISH と同じ臨床情報により分析対象となる患者を特定し、入院後 7 日以内の CATH の実施状況ごとにサンプルを 2 群に分け、患者属性と治療属性に関する基本統計量を検討している。表 1 はその結果を示している。

表 1 によると、VANQWISH 研究では、RCT という性格上当然ではあるが、入院後 7 日以内の CATH の実施・非実施割合はそれぞれ 462 人と 458 人とほぼ同じである。他方、観察データである CCP では、全調査対象者 180,178 人のうち 17,473 人 (10%) が VANQWISH 研究と同じ臨床的特性を持つ患者であり、うち、5,866 人 (34%) のみが当該期間内に CATH を受けている。

次に、患者属性と治療属性について、実施群と非

実施群との違いをみることにする。患者属性について、変数群 A (年齢と性別) は、観察データでも比較的収集コストが安価な情報である。なぜならば、メディケアの診療報酬データは、1984 年以来、医療政策のための行政資料として電子化されており、ルーティンワークが確立されているからである。しかし、変数群 B (リスク要因、副疾患、駆出率 (EF)、CATH 前の薬物投与) については、CCP が独自に収集した情報であり、通常の診療報酬データでは観察不可能な脱落変数となっている。

まず、年齢と性別についてみると、VANQWISH 研究では、年齢と性別ともに実施・非実施群の間で統計的に有意な差はなく、バランスしている。しかし、CCP では、年齢が若く、男性の方が積極的治療を受ける確率の高い傾向にあることがわかる。次に、CCP でのみで観察可能な入手コストの高い変数群 B をみてみると、年齢と性別同様、VANQWISH では、実施群と非実施群との間に統計的な有意差はほとんど認められないのに対して、CCP では、一部を除き、リスク要因、副疾患の有無、EF において、状態の良い患者の方が積極的治療を受ける比率が高いことを示している。最後に薬物投与をみると、VANQWISH 研究では同じくバランスしているのに対して、CCP では、カルシウム拮抗薬を除き、積極的に薬物投与をされている患者の方が 7 日以内に CATH を受ける確率が高い傾向にある。以上の結果は、観察データでは、積極的治療の実施に、明らかに、偏り (セレクション・バイアス) があることを示している。

今ここで、入院後 30 日・1 年以内死亡を治療のアウトカムと考え、これらを被説明変数とした下記のような単純回帰の推定式を想定する。

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

この推定式で、 q_i は i 番目の患者のアウトカムを、 d_i は CATH の実施状況 (実施は 1, それ以外は 0)、 x_i は患者属性、 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。ここで最も注目すべきは、入院後 7 日以内の CATH の実施効果を示すパラメータ α_1 である。CCP で収集された詳細な医療情報 (変数群 B) は、通常のメディケア診療報酬データには存在せず、観察不可能な脱落変数であり、調整することができない。したがって、診療報酬データを利用した場合は、CATH の実施率に影響する、リスク要因、副疾患の有無や重症度などの諸属性は、 x_i では調整することができず、 u_i として吸収されてしまう。仮に、診療報酬データを使った単純回帰において、CATH 実施群の方が非実施群に比べ、有意に良好なアウトカムが得られたとしよ

表1 「RCT」(VANQWISH) 対「非実験による観察データ」(CCP) : 入院後7日以内の心臓カテーテル検査実施の有無, 及び, CATH病院への相対的距離別, 急性心筋梗塞疾患患者の患者属性と治療属性に関する基本統計量

変数の定義	「RCT」(VANQWISH)		「非実験による観察データ」(VAQWISH-Type CCP Patients)			
			セレクション・バイアスの調整前		セレクション・バイアスの調整後 ^{注)}	
	7日以内の CATH実施有 (N=462)	7日以内の CATH実施無 (N=458)	7日以内の CATH実施有 (N=5,866)	7日以内の CATH実施無 (N=11,607)	CATH病院への 相対的距離 ≤中央値 (N=8,703)	CATH病院への 相対的距離 >中央値 (N=8,770)
A. メディケア診療報酬データで観察可能 (コスト安)						
年齢 (標準偏差)	62.0(10.0)	61.0(10.0)	71.4(7.5)	77.4(9.1)	75.2(9.0)	75.5(9.1)
女性 (%)	3.0	2.2	38.2	51.3	47.4	46.4
B. メディケア診療報酬データでは観察不可能, CCPデータでのみ観察可能 (コスト高)						
<u>リスク要因 (%)</u>						
現在, 喫煙者	40.9	45.9	19.5	13.7	16.1	15.1
高血圧	56.7	51.5	64.3	66.9	67.0	65.1
インスリン糖尿病	24.9	27.3	28.3	34.5	32.0	32.8
心筋梗塞の既往	43.1	43.0	5.1	7.3	6.1	7.0
<u>副疾患 (%)</u>						
心臓病 (non-CAD 群)	11.7	9.6	11.0	29.7	23.3	23.6
神経障害	11.5	11.8	1.1	7.1	5.5	4.7
心筋梗塞以前の狭心症	42.2	46.7	55.3	49.7	53.1	50.1
<u>ベースライン駆出率 (EF) (標準偏差)</u>						
	53.0(15.0)	50.0(14.0)	48.3(12.8)	44.2(13.3)	46.2(13.2)	45.8(13.2)
<u>CATH 前の薬物投与 (%)</u>						
ベータ遮断薬	21.6	22.5	23.0	19.9	21.8	20.0
カルシウム拮抗薬	36.1	35.6	43.6	43.2	43.9	42.8
ACE 阻害薬	21.0	22.1	27.6	20.5	26.2	24.3

出所: “VANQWISH”については, Boden WE et al. (1998) “Outcomes in Patients with Acute Non-Q-Wave Myocardial Infarction Randomly Assigned to an Invasive as Compared with a Conservative Management Strategy”. *New England Journal of Medicine*. 338 (25): pp.1785-1792. “VAQWISH-Type CCP Patients”については, Geppert J et al (1999) “Integrating Randomized and Observational Evidence on Medical Treatment Effects”. *Stanford mimeo*, 日本公衆衛生学会, 別府市。

注) まず, 各医療施設における年間の総CATH実施回数から中央値を計算し, 中央値(3件)よりも検査数の多かった医療施設をCATH病院, それ以外をNon-CATH病院と定義した。第2に, 患者と医療施設双方の郵便番号から各患者について最も近い距離にあるCATH病院とNon-CATH病院との距離を計算し, その距離の差(近接するCATH病院との距離-近接するNon-CATH病院との距離)をCATH病院からの相対的距離として定義した。第3に, CATH病院からの相対的距離の中央値を計測し(-0.7マイル), 相対的距離が-0.7マイル以下の患者グループと-0.7マイルよりも遠い患者グループとに分けた。

う。しかし, 基本統計量でみる限り, CATHの非実施群にリスクの高い患者が多い傾向にあることから, 実施群と非実施群とのアウトカムの有意差が, 純粋に積極的治療の効果なのか, それとも, CATH実施群の患者属性の健康状態が影響しているのかを厳密に識別する必要がある。つまり, もともとの状態が悪く死亡確率が高い患者に対しては, 積極的治療を行わない確率が高いならば, 潜在的なアウトカムと積極的治療の確率との間に逆の相関が存在することになる。この場合, 最小2乗法で単純回帰すると, d_i と u_i が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。これが, いわゆる, 内生性と呼ばれる問題であり, 入院後7日以内のCATHの実施は内生変数となる。

3. 調整方法の一例 —操作変数法—

本節では, こうした内生性の問題を有する「観

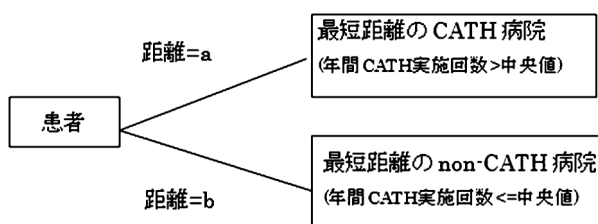
察」データから実験的な変動を抽出し, サンプルを疑似的にランダム化する計量経済学的手法として, 操作変数 (instrumental variable: IV) 法を紹介する^{11,12)}。IV (z_i) の選定に当たっては, 次の2つの仮定が鍵となる。1つは, $E(u_i | z_i) = 0$ であり, これは, IV (z_i) は外生変数であり, 最終的な患者のアウトカム (本稿では, 入院後30日および1年以内死亡) とは, 推定式の誤差項を介する相関が無いということの意味している。外生変数とは, 被説明変数である患者のアウトカム (q_i) や, 内生変数である入院後7日以内のCATHの実施などのように, 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, たとえば, 外部から一種の「ショック」のような形で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 q_i とは無相関であるが, 入院後7日以内の

CATHの実施とは相関があり、 z_i によってサンプルは「疑似的」に、積極的治療の実施と非実施へ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。

Geppert et al. によって採用されたIVの概念を示しているのが、図1である。この研究では、患者と最短距離にあるCATH病院とnon-CATH病院との相対的距離を用いている¹³⁾。CATH病院とは、年間の実施率が中央値を超えている医療施設のことであり、non-CATH病院とは実施率が中央値以下の病院をさす。実際に患者が入院した医療施設ではなく、患者から最短距離にあるCATH病院とNon-CATH病院との距離(aとb)をそれぞれ計算し、その距離の差(a-b)をCATH病院からの相対的距離として定義し、IVとした。前段で述べたIVの2つの仮定が満たされているか検証するため、相対的距離の中央値(-0.7マイル)を計測し、中央値以下の群と中央値よりも遠い群とに分け、患者属性と治療属性をみたのが、表1の右から2列までのデータである。セレクション・バイアスの調整前と比較すると、患者属性と治療属性について、これら2群でその平均値に有意差がなく、サンプルがIVによって疑似的にランダム化されていることが確認できる。

推定式を最小2乗法による単純回帰とIVによる2段階推定法で推定した結果については表2に示す通りである。ここでは、脱落変数の影響の大きさをみるために、通常の診療報酬データで入手可能な変数群A(年齢、性別、患者の居住地域ダミー)のみによって調整した結果と、それらの変数に加えて、CCPが独自に収集した変数群Bによる調整を行った回帰分析の結果を併記する。まず、単純回帰の結果(1Aと1B)をみると、1Aでは、入院後7日以

図1 相対的距離の計測法



相対的距離 (IV) = a - b

相対的距離が中央値(-0.7マイル)以下の場合：比較的CATH病院に近い

相対的距離が中央値より遠い場合：比較的CATH病院から遠い

内のCATHの実施が30日以内・1年以内の死亡をそれぞれ-5.7%と-14.9%と大きく引き下げている。しかし、収集コストの高いCCP独自の情報を投入すると、いずれも効果の大きさが大幅に縮小する(-4.2%と-8.1%)ことから、単純回帰が、観測不能な脱落変数から影響を受けやすいモデルであることがわかる。

次に、IVによる調整を行った2段階推定法の結果をみると、2Aと2Bいずれも、短期間での死亡についてはCATHの実施が死亡率を引き下げているという結果であったが、いずれも統計的には有意でない。さらに、1年以内の死亡については、同じく有意性はないものの、積極的治療の効果が死亡率を引き上げている。これは、単純回帰では調整できなかった積極的治療実施の有無に対する「観察」データ特有のセレクション・バイアス、そして、それに起因するモデル内での内生性が、一定程度、解消された結果とみるべきであろう。さらに、IV法による推定結果で着目すべきもう1つの点は、とりわけ1年以内の死亡という比較的長期のアウトカムに対して、変数群Aのみで4.8、変数群A+Bで4.7と、観測不能な脱落変数の影響を推定がほとんど受けないという点である。この結果は、とりもなおさず、推定方法を工夫することには、行政コストのかかる高価な医療情報の収集を回避することができるという利点のあることを示唆している。

最後に、観察データから得られた結果を、VANQWISHを含めたいくつかの代表的なRCTの結果と比較してみる。まず、VANQWISHでは、30日・1年以内死亡ともに積極的治療が正の効果を示している。1年以内の死亡について、VANQWISHより標準誤差がかなり大きく推定されているものの、CCPデータへのIV法の応用はほぼ同程度の効果を示しており、観察データであっても、推定方法を工夫することによって、RCTと極めて類似した結果を導き出す可能性があることが確認できた。VANQWISH以外のRCTの結果をみても、おおむね、non-Q-waveの患者に対して積極的治療が有効ではない、あるいは、かえってマイナスであるという結果については共通していることから^{14,15)}、観察データに単純回帰モデルを応用することの危険性が明確にみとれる。

しかしながら、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、トリートメント(この事例では入院後7日以内のCATHの実施)による差異が小さい場合、しばしば頑健性に問題が生ずること、 $E(u_i|z_i)=0$,

表2 入院後7日以内の心臓カテーテルの30日・1年以内の死亡に与える効果

リスク調整後の入院後7日以内のCATHの効果 (標準誤差)	入院後30日以内の死亡	入院後1年以内の死亡
観察データ (VAQWISH-Type CCP Patients) による推定		
(1) 単純回帰 (最小2乗法)		
(1A) 変数群 A のみによる調整 ^{注1)}	-5.7(1.0)	-14.9(1.4)
(1B) 変数群 A+B による調整	-4.2(1.0)	-8.1(1.5)
(2) IV法 (2段階推定法) ^{注2)}		
(2A) 変数群 A のみによる調整 ^{注1)}	-1.5(4.9)	4.8(7.4)
(2B) 変数群 A+B による調整	-3.2(4.9)	4.7(6.9)
RCTによる推定事例		
VANQWISH(N=920)	3.0(1.2)	4.7(2.0)
TIMI IIIB(N=1,473)	0.0(0.8)	—(—)
DANAMI(N=1,008)	—(—)	-0.8(-1.2)

出所：“VAQWISH-Type CCP Patients”については、Geppert J et al (1999) “Integrating Randomized and Observational Evidence on Medical Treatment Effects”. *Stanford mimeo*, 日本公衆衛生学会, 別府市。“VANQWISH”については、Boden WE et al. (1998) “Outcomes in Patients with Acute Non-Q-Wave Myocardial Infarction Randomly Assigned to an Invasive as Compared with a Conservative Management Strategy”. *New England Journal of Medicine*. 338(25): pp.1785-1792.

“TIMI IIIB”については、TIMI IIIB Investigators. (1994) “Unstable angina/myocardial infarction: effects of tissue plasminogen activator and an comparison of early invasive and conservative strategies in unstable angina and non-Q-wave myocardial infarction: results of the TIMI IIIB trial”. *Circulation*. 89: 1545-1556. “DANAMI”については、Madsen JK, Grande P, Saunamaki K, et al. (1997) “Danish multicenter randomized study of invasive versus conservative treatment in patients with inducible ischemia after thrombolysis in acute myocardial infarction (DANAMI)”. *Circulation*.

注1) 表1の基本統計量には示さなかったが、通常の診療報酬データでは、患者の住居地域も特定することができる。したがって、推定式には、年齢と性別に加えて、地域ダミー変数も説明変数として投入した。

注2) 2段階推定法では、まず、第1段階で、入院後7日以内のCATH（積極的治療）の実施確率を被説明変数、(2A)では変数群AとIV（CATH病院への相対的距離）、またそれに加えて(2B)では変数群Bを説明変数として、積極的治療の実施確率を推定する。第2段階では、患者のアウトカム（30日・1年以内の死亡）を被説明変数、第1段階から得られたリスク調整後の積極的治療の実施確率、加えて(2A)では変数群A、加えて(2B)では変数群Bを説明変数として投入する。

$E(d_i|z_i) \neq E(d_i)$ という2つの仮定を完全に満たすような適切なIV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。

4. 結語 —より精緻なヘルスサービスリサーチ (HSR) のために—

計量経済学では、その分析対象が経済社会全体であるために、非実験的なデータに頼らざるをえない状況の中で、因果推論のためのさまざまな手法が開発されてきた。RCTは、医科学におけるゴールドスタンダードであることには違いないが、VANQWISHなどのようなRCTを実施するためには、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。したがって、純粋な「実験」と比べると統計上の諸問題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による補完が必要

となってくる。これは、ヘルスサービスにおいてアウトカムを比較する場合にもそのままあてはまる問題であろう。

紙幅の関係上、詳細は避けるが、本稿で紹介したIV法の他にも、計量経済学では、差の差の推定 (Differences-in-Differences) に代表されるように、個人を複数時点で追跡するパネルデータを利用して、たとえば、介護保険法の導入や診療報酬の改定など大規模な制度変更があった場合、それを「自然実験」としてとらえ、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段がとられることも多い^{11,16)}。計量経済学が構築してきたこうした手法を医療情報へ応用するためには、患者自身も含め、医療情報に携わるすべて組織や関係者との協力体制を築きながら、医療情報の全国規模での整備とデータベースの作成に向けて、さらなる努力を重ねる必要がある¹⁷⁾。

第1に、医学研究者や臨床医、そして、統計学者

など専門家による学際的な研究グループを組織し、医療評価に対するコンセンサスに基づく医療情報の標準化を促すため、社会全体の理解と協力が得られるよう働きかけること、また、標準化された情報を常時見直すことのできるような物理的システムと人的協力体制をつくりあげること、そして、そうした取り組みを、国家レベルで総合的かつ持続的に行うことである。第2に、ある特定の医科学研究あるいは政策課題に対する問題設定を行うに当たって、スタディ・デザインを企画する段階から、実際の分析結果のまとめを行う最終的な作業にいたるまですべての段階におけるクオリティー・チェックと、政策の結果と政策過程自体を常時見直す手続きを怠らないことである。第3に、患者個人を含めデータの提供者と被提供者の双方が共通の利益に向け、長期的な信頼関係を築くことができるような環境を整備するため、情報の運用と管理システムにおける透明性のあるルールを構築し、そのための人材育成を促進することである¹⁷⁾。

本稿の執筆にあたり、田宮菜奈子氏、ならびに、西晃弘氏より多数の有益なご助言を頂いた。記して御礼申し上げる。本稿における残る全ての誤りは著者に帰するものである。

文 献

- 1) 和合 肇. 計量経済学. 経済科学 2002 ; 49別冊 : 77-85.
- 2) The Econometric Society: An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics. Aims and Scope. <http://www.econometricsociety.org/aims.asp> (2011年10月1日アクセス可能)
- 3) Coleman RD. What is Econometrics? ハーバード大学経営大学院「econometrics」シラバス. 2006.
- 4) Duffo E, Glennerster R, Kremer M. Using randomization in development economics research: a toolkit. BREAD Working Paper No. 136. 2006.
- 5) Newhouse JP. Free for All? Lessons from the RAND Health Insurance Experiment. Cambridge, MA: Harvard University Press, 1993.
- 6) 井伊雅子, 大日康史. 医療サービス需要の経済分析. 東京: 日本経済新聞社, 2001.
- 7) Meng XL. The EM algorithm and medical studies: a historical link. *Statistical Methods in Medical Research* 1997; 6(1): 3-23.
- 8) Boden WE, O'Rourke RA, Crawford MH, et al. Outcomes in patients with acute non-Q-wave myocardial infarction randomly assigned to an invasive as compared with a conservative management strategy. *New England Journal of Medicine* 1998; 338(25): 1785-1792.
- 9) Geppert J, McClellan M, Noguchi H. Integrating Randomized and Observational Evidence on Medical Treatment Effects. Stanford mimeo. 第58回日本公衆衛生学会総会 (大分県別府市). 1999 (口頭発表).
- 10) Ellerbeck EF, Jencks SF, Radford MJ, et al. Quality of care for Medicare patients with acute myocardial infarction. *The Journal of the American Medical Association* 1995; 273(19): 1509-1514.
- 11) Wooldridge JM. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass.: The MIT Press, 2001.
- 12) Angrist JD, Pischke JS. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, N. J.: Princeton University Press, 2008.
- 13) McClellan MB, Newhouse JP. The marginal cost-effectiveness of medical technology. *Journal of Econometrics* 1997; 77(1): 34-64.
- 14) TIMI IIIB Investigators. Effects of tissue plasminogen activator and a comparison of early invasive and conservative strategies in unstable angina and non-Q-wave myocardial infarction: results of the TIMI IIIB Trial. *Circulation* 1994; 89(4): 1545-1556.
- 15) Madsen JK, Grande P, Saunamaki K, et al. Danish multicenter randomized study of invasive versus conservative treatment in patients with inducible ischemia after thrombolysis in acute myocardial infarction (DANAMI). *Circulation* 1997; 96(3): 748-755.
- 16) Ashenfelter O, Card D. Using the longitudinal structure of earnings to estimate the effect of training programs. *The Review of Economics and Statistics* 1985; 67(4): 648-660.
- 17) 野口晴子. 保健医療行政が EBM に対して果たすべき役割. *EBM ジャーナル* 2002; 3(4): 507-513.