

Cartwright, N. and Hardie, J., 2012, “Chapter I.B: The Theory that Backs up What We Say,” in *Evidence-Based Policy: A Practical Guide to Doing It Better*, Oxford, UK: Oxford University Press, 14-58.

N. カートライト & J. ハーディ, 2012, 「われわれの提案を裏づける理論」

#### レジュメ作成者による紹介

科学哲学の立場からエビデンスに基づく政策について論じた著作。ランダム化比較試験などから得られる、ある状況における政策効果についてのエビデンスのみでは、別の状況に政策を適用する場合の有効性が保証されないことなどが主張される。とりあげた章では、こうした本書全体の主張の背後にある理論が提示されている。

### 1 なぜ理論を求めるのか<sup>1</sup> (pp.14-15)

- この章では、本書の実践的な提案の背後にある理論を提示する。ここで理論を提示するのは、政策を決めることに関する本書のアイデアが一定の原理に裏づけられていることを示すため。
  - 本書のアドバイスは、エビデンス利用の理論 (theory of evidence for use) に根差している。これは、政策がある任意の人にとって有効に働くかどうかについて信頼性の高い予測をするうえで、どのような種類の知識が役に立つのかを理解するために設計された理論である。

### 2 2つの仮定 (pp.15-23)

- 仮定 1: ある確立された経験的言明 e は、それがある論証 A における前提 (premise) として提出され、かつ論証 A が仮説 h の良い論証である場合、あるいは、経験的言明 e が論証 A' の前提であり、かつ論証 A' が仮説 h の良い論証である論証 A における前提の良い論証であるなどといった場合、経験的言明 e はエビデンスである。
  - ある言明を正当化するには良い論証 (argument) が求められる。ここでの論証とは、前提に当たる諸命題と結論に当たる 1 つの命題のセットである。「私がこのワインを買う論拠はそれが安いからだ (“My argument for buying this wine is that it is cheap”）」という用例にあるような、通常の言葉でのたんなる理由とは異なる。
  - 良い論証とは、前提が信用でき (trustworthy)、ある結論に至ることを余儀なくさせるような論証のことを指す。良い論証とは、健全 (=前提が信用できる) かつ妥当 (=結論が前提から導かれる) である。

<sup>1</sup> レジュメのセクション分けは、おおむね原文の構成と対応しているが、一部省略した下位セクションがある。

- ある論証の結論を正当化するには、それぞれの前提も正当化されなければならない。前提が自明ではない場合、そのための下位の論証が必要となる（下位の論証の前提を正当化するもう一段下位の論証が必要になるかもしれない……と続く）。
  - ある言明のエビデンスとは、その言明を正当化するのに資するものであるはずだから、エビデンスは良い論証に関わるものでなければならない。このことが、ある結論（もしくは仮説）に対するエビデンスのレリヴァンスを保証する。
  - 論証の重要性を強調するのは、論証の強さがそのもっとも弱い前提と同程度であるから。9つの真である前提と1つの偽である前提からなる妥当な論証が、その結論の確からしさを90%保証することはなく、結論はまったく正当化されない。この場合、9つの前提として提出されたエビデンスは、結論が真であることとレリヴァンスがないことになる。
- 仮定 2: 「それはどこかで (somewhere) 有効に働く」というエビデンスから「それはここ (here) で有効に働くだらう」という良い論証を行うためには、「ここ」と「どこか」の因果的原則に関する事実が必要である。
    - 本書が関心を寄せる仮説は、政策の有効性の予測であり、これは因果的な言明である（「提案された政策は、もしそれが実施されれば、焦点となるアウトカムに改善をもたらすだろう」）。
    - 著者たちは、もしある原因がある結果を引き起こすとすれば、それは偶然によるものではなく、両者のあいだに信頼性の高いシステムティックな結びつきがあるからだと仮定する。このような結びつきのことを、ここでは因果的原則 (causal principle) と呼ぶ。
    - ランダム化比較試験 (randomized controlled trials: RCTs) やその他の研究は、ある政策が「どこか」で有効に働くことを示し、論証における1つの前提となる。しかし、政策が「ここ」 (=ある人が政策を実施しようとする場所) でも有効に働くであろうというこの論証には、「どこか」と「ここ」の因果的原則に関する事実が必要。さもなければ、RCTs などの研究結果が、対象となる環境において政策が実施されたときに何が生じるかの予測とレリヴァントになることは困難である。

### 3 因果的原則 (pp.23-29)

- ここでは、信頼性の高い政策に関する予測を支える因果的原則の種類について、近年の哲学分野の業績をもとにいくつかの仮定を置く。
  - 因果的原則は、たんに統計的な相関を記録するものではない。なぜなら、原因とは結果を生じさせるものであり、結果の産出に寄与するものであり、それのおかげで (responsible for) 結果がもたらされるから。
  - 因果的原則のなかには、あらゆる場所において有効なものもある (例: 万有引力の法則)。しかし、このような因果的原則は、医学や社会科学の領域ではまれである。

- 因果的原則は、すべて同等であるわけではない。ある原則は、他の原則よりも基礎的（fundamental）である。一般に、より基礎的な原則ほど広い範囲で有効である。
- 生物学、医学、社会科学において用いられる因果的原則は、*ceteris paribus* という条件付きの原則である。すなわち、「他のすべての条件が同じ」である場合のみ、その原則は有効である。
- 因果的原則は、決定論的もしくは確率的なものである場合もあるが、一般にはより厳格でないものである場合もある。また、因果的原則は、方程式で表される量的なものであることもあるが、量的な尺度をもたない特性と関連した質的なものであることもある。
- 政策立案において目指される種類の結果について：a) 結果に対する寄与は、さまざまなソースから、さまざまな通路を介して生まれる。b) 全体としての結果は、これら別個のソースからもたらされる寄与にシステムティックな形で依存する（最も単純なケースではそれぞれの寄与が合計される）。c) 一般に社会政策は、それ自体で目標とする結果に寄与することを保証できない。政策が結果に寄与するには、「サポートチーム」が必要である。
  - これは、哲学の用語では INUS 条件（Insufficient but Necessary part of an Unnecessary but Sufficient condition、不必要だが十分な条件の不十分だが必要な部分）と呼ばれる。
  - プログラムや政策は、ともに作用する一組の原因（チーム）の一部にすぎない。結果の産出に寄与するためには、一組のチームである原因すべてが必要である。
  - 政策を含む原因のチームは、アウトカムに寄与しうる唯一のものではない。他のチーム（政策を含むかもしれないし含まないかもしれない）もアウトカムに寄与する可能性がある。ある状況におけるアウトカムの実際の値は、その状況で作用しているすべての原因のチーム、それぞれの大きさや寄与の方向性に依存している。
- 研究の主題が何であるかに応じて、原因による寄与を組み合わせるやり方についてのルールが異なる。社会科学の場合、閾値や限界収益モデルをしばしば仮定する。計量経済学では、それぞれの要因が異なる方程式で表現されることがある。

#### 4 何が有効性の予測を真にするのか（pp.29-32）

- 有効性の予測（上述の仮定 2 の言葉でいえば「それはここで有効に働くだらう」という命題）を真にする事実が何であるかは、どのような予測を行おうとするかによる。
- 以下、人々が行いたいと考えるであろう予測の種類を列挙するが、本書の大部分で焦点を当てるのは「政策はここでポジティブに作用するだらう」という最小限の予測。
  - これは最も弱い有効性の予測であるため、この結論を導くために知るべき要因は、他のすべての予測においても必要不可欠である。

- 予測のタイプ 1：アウトカムの実際の値を 1 人ひとりについて予測。
  - ほとんどの場合、最も難しい予測。
  - 政策を実施した後に生じてアウトカムを産出した因果的原則、およびその原則のもとにあるすべての人々について、ある特定の個人に生じた値についての事実を知ることが必要。
- 予測のタイプ 2：母集団に含まれる人々への平均的なアウトカムを予測。
  - 政策を実施した後に生じている因果的原則、因果的原則のもとにある母集団のすべての項の平均についての事実が必要。
- 予測のタイプ 3：アウトカムの平均が、政策を適用しない場合よりも適用した場合のほうが改善したかどうかの予測。
  - 結果の産出に影響する因果的原則は、政策が実施される経過において変化する可能性がある。このような場合、アウトカムの平均的な差異は、政策が実施されていない時点の因果的原則のもとで産出された平均と、新たな原則のもとでのアウトカムの平均の差から求められる。
  - ゆえに、このタイプの予測には、過去の古い因果的原則、現在の新しい因果的原則、および古い原則のもとでのすべての項の平均と新しい原則のもとでのすべての項の平均についての事実が必要。
- 予測のタイプ 4：上記の予測のタイプ 3 において、因果的原則が変化しないと想定されるケース。
  - アウトカムの平均的な差異を規定するのは、政策が実施されなかった場合に得られるであろう平均と、政策が実施された場合に得られるであろう平均である。これらを、政策の実施前後で平均が変化するすべての項について知ることが必要。
- 予測のタイプ 5：政策が実施された場合とされなかった場合の差異ではなく、因果的原則が変わらないと想定した場合に政策そのものがもたらす平均的な差異の予測。
  - このような差異は、ある状況下における政策の効能（efficacy）と呼ばれる。
  - 効能を規定するのは、政策の前後で変化せずに保持されていると仮定した因果的原則、およびその因果的原則のもとにおける政策の補助的な要因の平均値である。

## 5 ランダム化比較試験（pp.32-40）

- 本書が焦点を当てるのはレリヴァンスである。他方、政策がどこかで有効に働くという言明の信用性（trustworthiness）に焦点を当てたランキングでは、RCTs および RCTs のメタアナリシスが、この種の言明を最もよく支持するエビデンスとみなされる。

- RCT は、J. S. Mill の差異法に基づく研究デザインである。
  - そこでは、同一の因果的原則のもとにあると仮定される個々のユニットが、想定する原因が生じるグループ（=介入群）と生じないグループ（=統制群）に無作為に割り当てられる。
  - くわえて、可能なかぎりのマスキングが行われる（例：介入を行う人物に、対象者がどちらのグループに属するかを知らせないなど）。無作為化とマスキングによって、介入の違いを除いた他の要因が、両グループ間で同じように分配されると仮定する。
- 上記の手段を用いる RCT に裏づけられた因果的言明の信用性をめぐっては議論があるが、本書の関心はレリヴァンスのほうにある。
  - それゆえ、以下では、上記のような手段が有効だと仮定し、理想的な RCT（=政策介入とその結果を除いて、介入群と統制群の間には因果的にレリヴァントな差異が存在しないような研究デザイン）に焦点化する。
- RCT は、しばしばエビデンスのランキングで最上位に位置づけられるが、さらに上位に位置づけられるのが、RCT のメタアナリシスおよび系統的レビューである。
  - 多くの RCT では小規模な集団しか対象にできない。これは、統計的な推論の妥当性を脅かすことになる。メタアナリシスとは、統計的な手法を用いて異なる RCT の対象集団を混ぜ合わせ、仮想的に巨大な母集団をつくりだす方法である。
  - 系統的レビューとは、複数の研究結果を統合することで、ある問題について利用可能な最良の研究をまとめることを指す。対象となる研究は RCT である必要はないが、「利用可能な最良の研究」であることを判断する規準として RCT を頂点とするエビデンスのランキングが用いられる。
- RCT では、介入群と統制群の間の確率的なアウトカムの差異から、介入がアウトカムを引き起こすことを演繹的に含意する。ただし、ポジティブなアウトカムから因果的な結論を演繹的に導く方法は RCT だけではない。
  - 例：因果ベイズネット、計量経済学的手法、過程追跡など。
- RCT がもつ特別な利点は、RCT の前提が、研究している状況に関与しているかもしれない要因やその状況について、ほとんど何の知識も要求しない点にある。
  - 無作為化やマスキングといった実験のデザインによって基本的な仮定を満たしていると想定するので、背後の因果的な情報について多くを知る必要がない。
  - ただし、多くのケースにおいて、良い RCT を実行することは非常に困難である。また、ときに計量経済学的方法においてもランダム化実験を組み合わせることは可能である。

## 6 レリヴァンス (pp.40-49)

### 6-1 何が RCT の結果を有効な予測にとってレリヴァントなものにするのか

- RCT から得られる標準的な結果は「介入効果」と呼ばれるが、実験対象となった集団における介入効果は、その集団の外部についての有効な予測とは直接的なレリヴァンスがない。
  - RCT で研究した状況とその外部の任意の状況が同じ因果的原則のもとにあると仮定したケースでも、介入の補助的な要因が 2 つの状況で同じように分配されていると想定できる根拠はない。むしろ、補助的な要因の分配は、しばしばローカルな環境に強く依存する。
  - 政策がある任意の状況でポジティブに寄与することを予測したい場合、RCT のエビデンスをレリヴァントにするコツとは何か。また、RCT が「あそこで有効に働く」ことを支持する方法の 1 つでしかないとすれば、「あそこで有効に働く」ことを「ここで有効に働く」こととレリヴァントにすることはいかにして可能か。

### 6-2 「あそこで有効に働く」から「ここで有効に働く」へ

- 以下では、「ここで有効に働く」を最も弱い意味でとり、「政策 x が、状況 S の複数の個人にとって、アウトカム y の産出にポジティブに寄与する」という意味だとする。
  - これが生じるのは、政策が実施された後の S において y の産出を規定する因果的原則のなかに x が現れたとき、かつ x が y の産出に寄与するための補助的要因が、政策が実施された後の S における複数の個人に見られるときである。
  - これは、著者たちの用語で、x が S において (y に関して) ポジティブな因果的役割 (causal role) を果たすときであると言い換えられる。
- 上記を踏まえて、「あそこで有効に働く」から「ここで有効に働く」を導くことができる論証とはどのようなものかを考えると、次のようなものがそれに当たる。

<p>1: x は「あそこ」で有効に働く。 2: 「ここ」と「あそこ」は、政策が実施された後における因果的原則を共有する。 3: この因果的原則のもとで x が寄与するために必要な補助的要因が、政策が実施された後の「ここ」において見られる。 結論: x は「ここ」で有効に働く。</p>
---

- この論証は、アウトカムを産出する因果的原則が、2 つの状況において同じであることを要求する。しかし、多くの場合、時間や場所が変わるにつれて、作動する原因は頻繁かつ予測不可能な形で変化する。

- しかし、だからといって、どこかで作用している原因についての情報が他の場所で生じることとまったく関連しないわけではない。Mill がいうように、多くの原因は「安定した傾向性」、すなわち状況の違いを超えて同一の寄与をする性質をもつ。
  - ただし、すべての原因が安定した傾向をもつわけではない。安定した傾向性をもつ原因の多くも、ある一定範囲の状況では同一の寄与をするが、それ以外の状況においてはそうではない。
  - ポイントは、どのような状況がこの範囲に含まれるかを解明すること。この範囲内では、ある状況において見られた寄与が、別の状況においても見られると想定できる。しかし、とくに社会政策、経済政策、医療政策で問題となるような原因について、このような範囲を解明することは困難である。
  - $x$  が別の状況で有効に働くという言明のしばしば最も良い根拠は、 $x$  が  $y$  を産出する安定した傾向性をもつことがすでに確立されていることである。ただし、これは正当化が難しい仮定であり、正当化が弱い場合には、それがサポートするはずの有効性の予測についての正当化も弱くなることに注意が必要である。
- 政策が任意の状況で有効に働く（＝ポジティブな因果的役割を果たす）ことが正当化されたとしても、政策は一部の人にポジティブな因果的役割を果たす反面、他の人にはネガティブな因果的役割を果たしうることも肝に銘じておく必要がある。
  - どちらが支配的であるかを予測するためには、それぞれの因果的役割を果たす補助的な要因の値の分配についての情報が必要である。
- 一般に、安定した傾向性を示すのは因果のチーム全体であり、単一の原因そのものではないということも付記しておく必要がある。
  - 3 節で述べたように、個々の原因は、アウトカムに寄与するチームの一部でしかない。安定した寄与を生み出すためには、チーム全体が求められる。

### 6-3 外的妥当性

- RCT に関する通説によれば、研究が行われたときと「同じ介入」が「同じ結果」をある特定の対象にもたらすとき、その研究は外的妥当性をもつとされる。その場合のアドバイスとは、ターゲット集団が研究を行った集団と「十分に類似」していれば、研究の外的妥当性が期待できるというものである。
- 著者たちは、「あそこで有効に働く」から「ここで有効に働く」という結論を得るうえで、このアドバイスが役に立つかについて、以下のような理由で懐疑的である。

- ①アドバイスが曖昧である。
  - 「同じ介入」を行うためには、その介入の正しい記述を特定しておく必要がある。この場合の正しい記述とは、ターゲット集団と研究を行った集団において同じ因果的役割を果たすものことである。
  - 「同じ結果」の意味も曖昧。かりに「同じ介入効果」に関心があるとすると、それを得るためには、政策が 2 つの集団において同じ因果的役割を果たしうるものでなければならない。くわえて、政策の補助的な要因およびその値の分配も 2 つの集団間で同じでなければならない。これは無理な注文であり、それが外的妥当性の要件だとすれば、現実にはその見込みはないだろう。
  - 「同じ結果」は、たとえば「全体として同じ結果」と理解できるかもしれない。これは異なるタイプの予測であり、上述の 4 節で示したように、その予測を真にするには異なる種類の事実が必要。曖昧な「類似性」が必要なのではない。
- ②類似性は過度な要求である。
  - Ludwig ら（2008）<sup>2</sup>は、Moving to Opportunity（MTO）という実験の対象となった集団外部のどこで実験と同じ結果（たとえば移住した家族における女性の高校卒業率の上昇）が得られると仮定できるかを問うている。結論として Ludwig らは、同様の結果が得られるのは厳密にはごく一部の集団（1990 年代半ばに高層の公営住宅に住んでいる人々で、少なくとも移住に興味があり、移住の機会に注目すると同時に申し込みを完了することができるのに十分なほど組織化されていた人々）であるとしている。
  - これが、MTO の結果がレリヴァントである範囲の限界だとすれば、最初からその研究を行う価値はなかったかもしれない。
- ③類似性は誤った考えである。
  - 上記の MTO と同じ結果が得られる集団のリストを見ると、それが寄せ集めであることがわかる。また、そのリストは、同じ結果が得られるために何が必要なのかということを示していない。2 つの集団間で同じ介入効果を得るためには、政策が同じ因果的役割を果たし、補助的要因が同じように分配されていることが必要だが、上記のリストはそのことを保証するだろうか。
  - 列挙された要因はあくまで指標であり、指標を共有しているということが、2 つの集団が重要な事実を共有していることを保証すると反論する向きもあるかもしれない。しかし、指標のリストを示すのであれば、その指標が役に立つことを説明する必要がある。これは、研究を行った集団において実験が果たした役割やその補助的役割についての考察なしには難しいだろう。

<sup>2</sup> Ludwig, J., Liebman, J. B., Kling, J. R., Duncan, G. J., Katz, L. F., Kessler, R. C. and Sanbonmatsu, L., 2008, "What Can We Learn about Neighborhood Effects from the Moving to Opportunity Experiment?" *American Journal of Sociology*, 114(1): 144-188.



- ④類似性は無駄が多い。
  - 介入効果は、補助的要因の配列を平均化したもの。どのような配列かによって、政策がもたらす寄与は大きくもなれば小さくもなる。研究が行われた集団と同じ配列の組み合わせをターゲット集団に要求すべきではない。
  - 同じ介入効果を期待するのではなく、「政策がポジティブな寄与をするだろう」という弱い結論を得るためには、研究を行った集団と政策のターゲット集団の間の類似性を探し出す必要はない。目指している予測を正しく行うためには何が必要かを探すことが必要である。因果的役割や補助的要因といった観点から考えることで、政策が「ここで有効に働く」ことを確信するためにどのようなエビデンスが必要かを理解しはじめることができる。