



RIETI Policy Discussion Paper Series 24-P-008

# 英国におけるEBPMの深化 —政策評価タスクフォースを中心として—

小林 庸平  
経済産業研究所

内山 融  
経済産業研究所

川瀬 仁志  
総務省

尾谷 祐樹  
デジタル庁



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<https://www.rieti.go.jp/jp/>

# 英国における EBPM の深化<sup>1</sup>

## －政策評価タスクフォースを中心として－

小林 庸平（三菱 UFJ リサーチ & コンサルティング/戦略国際問題研究所/経済産業研究所）

内山 融（東京大学/経済産業研究所）

川瀬 仁志（総務省）

尾谷 祐樹（デジタル庁）

### 要 旨

英国政府では、近年、政策形成や予算編成における政策評価の結果の活用や、政策の執行前からの評価計画策定の動きが進んでいる。これまで英国の EBPM では予算要求を行う際に行われる費用便益分析に基づく事前評価が中心的であったが、事前評価における因果エビデンスの参照や、事後評価の政策形成への組み込みも重視されてきているのである。本稿では、財務省や他の各省に政策評価についての助言を行う組織として 2021 年に設置された政策評価タスクフォース(Evaluation Task Force, ETF)を中心に、そうした動きを解説する。まず会計検査院(National Audit Office, NAO)の報告書に基づき英国政府における評価の実施状況等を見た上で、EBPM に関する体制の改革として、政策分析を担う専門職グループである分析ファンクションの設置と、政策評価のガイドラインであるマジェンタブックの大幅改訂を取り上げる。ETF については、同組織の幹部に対して行ったインタビューに大きく依りつつ、その導入に至る経緯、業務の概要、具体的な活動等について明らかにする。最後に、今後の課題と日本への示唆を整理する。

**キーワード:**英国、エビデンスに基づく政策形成、Evaluation Task Force

**JEL Classification:**H8、Z00

RIETI ポリシー・ディスカッション・ペーパーは、専門論文の形式でまとめられた研究成果を公開し、活発な議論を喚起することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び（独）経済産業研究所としての見解を示すものではありません。

<sup>1</sup> 本稿は、独立行政法人経済産業研究所（RIETI）におけるプロジェクト「日本におけるエビデンスに基づく政策形成の実装」の成果の一部である。本稿の作成にあたって、英国内閣府政策評価タスクフォース長であるキャサリン・ハッチンソン(Catherine Hutchinson)氏、英国会計検査院のルース・ケリー(Ruth Kelly)氏およびフィリップ・ブラドバーン(Philip Bradburn)氏、ロンドン・キングズ・カレッジのマイケル・サンダース(Michael Sanders)教授にインタビューに応じて頂き、引用についてご快諾頂いた。また大竹文雄教授、川口大司教授、富浦英一所長、正木祐輔氏、井上誠一郎氏などから、多くの有益なコメントを頂いた。ここに記して、感謝の意を表したい。本稿の内容や意見は、執筆者個人に属し、所属する組織の公式見解を示すものではない。

# 1 はじめに

日本では、2017年の統計改革推進会議最終取りまとめ以降、以下のとおりエビデンスに基づく政策形成（Evidence-Based Policy Making, EBPM）が進展してきた。

2017年5月 統計改革推進会議の最終取りまとめ

2017年8月 EBPM推進委員会発足

2018年4月～ 政策立案総括審議官の設置

2022年4月 RIETI EBPMセンター設立

2022年7月～ EBPM補佐官制度、伴走型支援ネットワーク、政策設計ラボの設置

2023年度～ 行政事業レビューにEBPMを取り込み

一方英国は、経済学を始めとした社会科学的分析を政策立案に取り込んだ先駆的な国だといえる。日本へのEBPMの導入も、英国の仕組みを参考として行われた面も大きい。もっとも、社会科学的分析やエビデンスを政策形成に取り入れてきた英国だが、ホワット・ワークス・センター(What Works Centres, WWC)はあくまでも政府外機関であり、政策立案に直接的に関与してきたわけではなかった。また英国の政策立案における事前評価のガイダンスであるグリーンブック(Green Book)もコンセプトの中心は費用便益分析であり、因果関係を示すエビデンスに力点が置かれてきたわけではなかった<sup>1</sup>。

しかしながら2020年に、内閣府(Cabinet Office)と財務省(HM Treasury)の共管として政策評価タスクフォース(Evaluation Task Force, ETF)が設置された。ETFは、施策・取組の効果に関する頑健なエビデンスを確保し、国民生活の改善につなげることを目的とした組織であり、政策評価から得られたエビデンスを政策形成に本格的に活用させるための組織であるといえる。また、英国会計検査院(National Audit Office, NAO)は、「政府の支出を評価する」"Evaluating Government Spending"というレポートを2021年に公表し、英国政府内におけるエビデンスの活用状況を分析している。すなわち、事前評価における因果エビデンスの参照や、事後評価の政策形成への組み込みが重視されてきているのである。

---

<sup>1</sup> WWCやグリーンブックについては内山・小林他(2022)を参照されたい。

そこで本稿では、政策形成や予算編成における政策評価の結果の活用や、政策の執行前からの評価計画策定の動きが進んでいる英国の動向について明らかにする。まず NAO の報告書に基づき英国政府における評価の実施状況等を見た上で、EBPM に関する体制の改革として、政策分析を担う専門職グループである分析ファンクション(Analysis Function)の設置と、政策評価のガイダンスであるマジェンタブック(Magenta Book)の大幅改訂を取り上げる。ETFについては、同組織の幹部に対して行ったインタビューに大きく依りつつ、その導入に至る経緯、業務の概要、具体的な活動等について明らかにする。最後に、今後の課題と日本への示唆を整理する。

## 2 英国政府内におけるエビデンス活用の現状

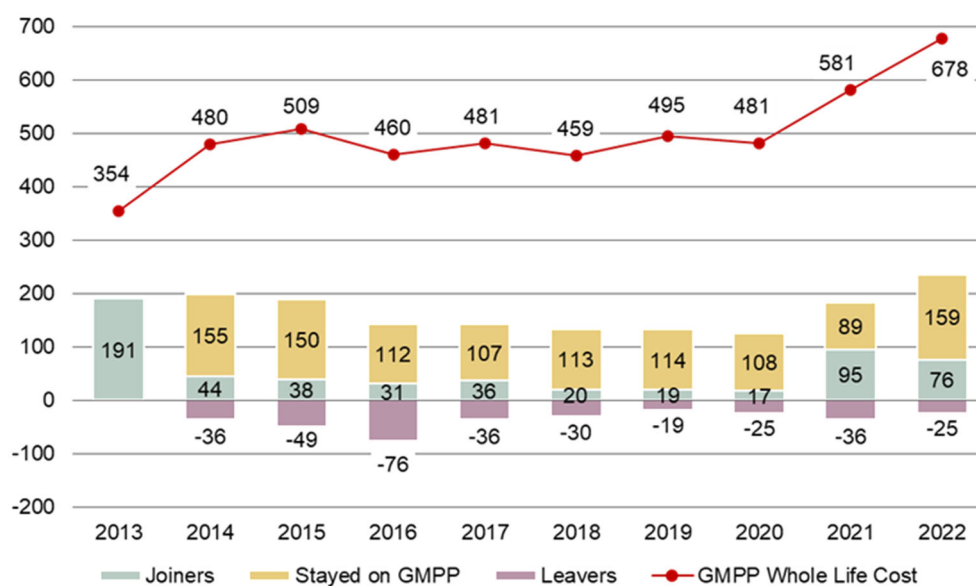
本稿では ETF 設置前後以降の英国の EBPM の深化を論じるが、その前段階として NAO の報告書（NAO 2021）をもとに、ETF が設立される頃までの英国における EBPM や評価の状況について整理する。

### 2.1 政府における評価の実施状況

2019 年 12 月に首相執行ユニット(Prime Minister’s Implementation Unit, PMIU) は、政府主要プロジェクトポートフォリオ（Government Major Projects Portfolio, GMPP）を用いて、評価の実施状況を精査した。GMPP は、英国政府が行うプロジェクト（またはプログラム）の中で特に複雑で戦略的に重要なものをマネジメントするためのデータベースであり、2013 年から作成されている。GMPP は英国政府が実施する中で規模が大きく、革新的でリスクの高いプロジェクトで構成されているため、評価が適切に実施される必要性が高いプロジェクトだといえる。2013 年から 2022 年にかけての GMPP の推移を表したものが図 2-1 である。

図 2-1 GMPP に含まれるプロジェクト数・規模の推移

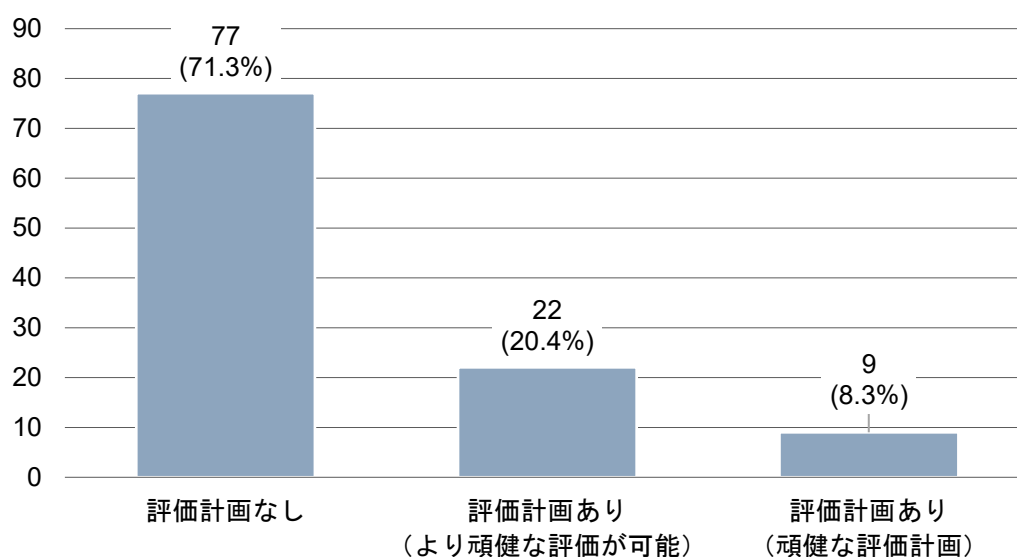
（単位：10 億ポンド（折れ線）、件（棒グラフ））



（出所）Infrastructure and Projects Authority (2022) “Annual Report on Major Projects 2021-22”をもとに作成

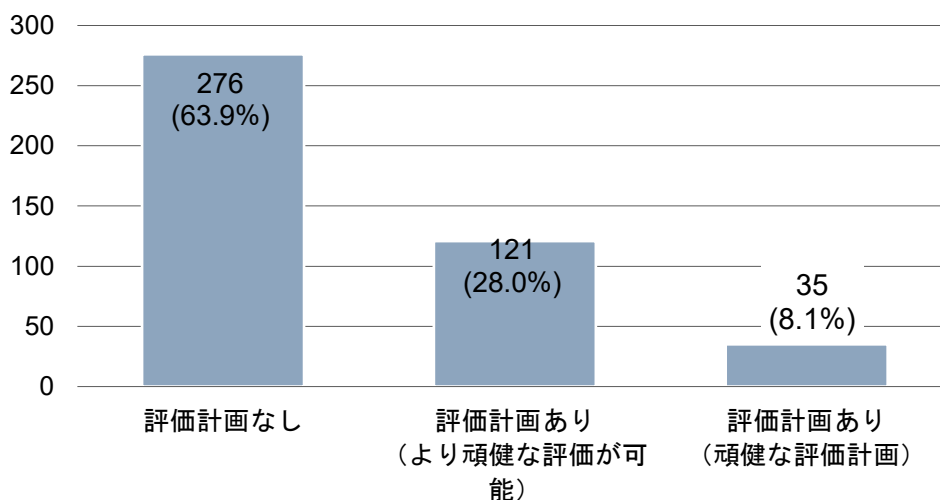
PMIU は GMPP に含まれるプロジェクトの中でも、特に複雑で戦略的に重要な 108 について、評価計画の策定状況を精査した。その結果を示したものが図 2-2 および図 2-3 である。対象となった 108 プロジェクトは革新的でリスクの高いプロジェクトであるため頑健な評価計画を持つことが期待されるが、実際に頑健な評価計画を有していたのは 9 プロジェクトのみであった。支出金額ベースでも「評価計画なし」が全体の 63.9%、「評価計画はあるが頑健な評価が可能」が 28.0%、「評価計画あり（頑健な評価計画）」は 8.1%に留まっていた。

図 2-2 評価の実施割合（件数ベース、過去内は全体占める割合）



(出所) NAO (2021)をもとに作成

図 2-3 評価の実施割合（支出額（10 億ポンド）ベース、カッコ内は全体に占める割合）



(出所) NAO (2021)をもとに作成

## 2.2 評価の実行可能性と障壁

### 2.2.1 評価の実行可能性に関するチーフアナリストの見解

「政府中枢」(centre of government)<sup>2</sup>から各省に対して評価の要請がなされているが、それがどの程度実行可能なのかについて主要 16 省庁のチーフアナリスト (Chief Analyst)<sup>3</sup>に対するアンケート結果を整理したものが図 2-4 である<sup>4</sup>。

「評価は頑健な手法に基づくべき」、「評価は介入に対して比例的であるべき<sup>5</sup>」、「評価は質が保証されるべき」といった要請については、いずれも 87.5%のチーフアナリストが「ほとんどの場合において可能」と回答した。一方で、「評価エビデンスは政策設計に反映されるべき」、「評価は、政策設計段階から検討されるべき」、「評価結果はタイムリーに公表さ

<sup>2</sup> ここでは、内閣府、財務省および分析ファンクション幹部を示す表現として「政府中枢」(centre of government)という言葉を用いている。

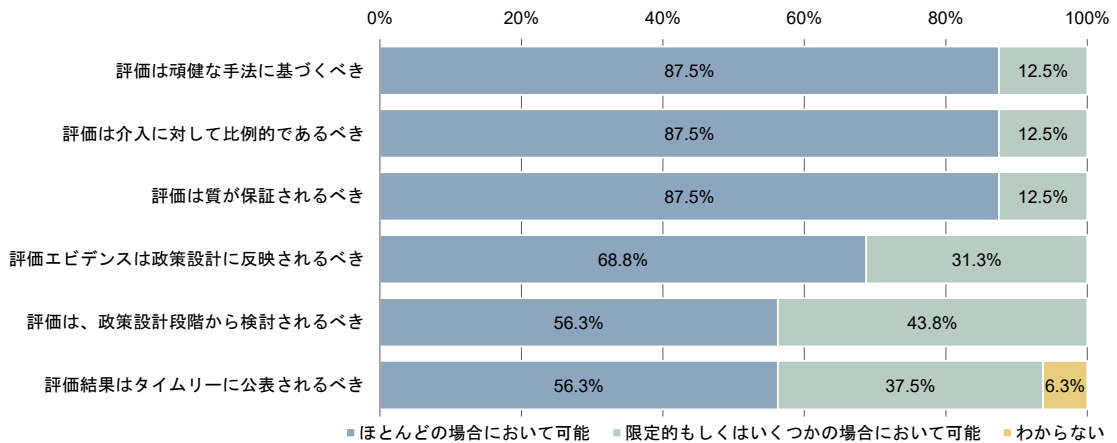
<sup>3</sup> チーフアナリストは各省の分析専門職を統括する幹部ポストである。分析専門職については後述。

<sup>4</sup> 回答率は 100%。

<sup>5</sup> 介入規模等に応じて評価リソースを比例的に配分することなどを指す。

れるべき」という要請については、3割から4割のチーフアナリストが限定的なケースでしか実行可能ではないと回答している。

図 2-4 「政府中枢」からの評価要請の  
実行可能性に関するチーフアナリストの見解 (n=16)



(出所) NAO (2021)

### 2.2.2 意思決定に対する評価エビデンス活用の障壁

図 2-5 は、主要 16 省庁のチーフアナリストおよび政策専門職の長 (Heads of Policy Profession)<sup>6</sup>を対象として、意思決定に対する評価エビデンス活用の障壁に関するアンケート調査結果を整理したものである。ここでは、各項目について「強く同意する」または「同意する」と回答した人数を合計している。また、回答率は 100%であった (n=32)。

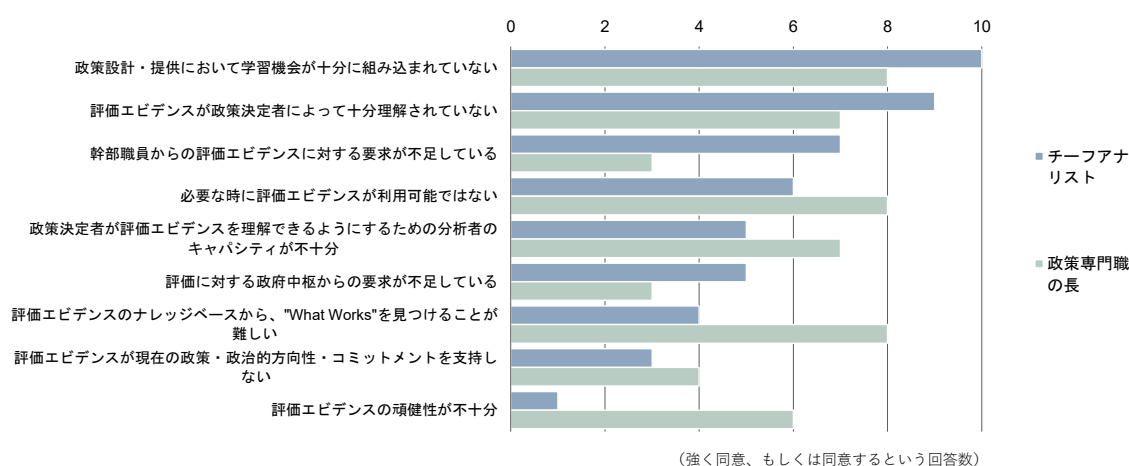
チーフアナリストおよび政策専門職の長の双方で障壁として挙げられているのは、「政策設計・提供において学習機会が十分に組み込まれていない」や「評価エビデンスが政策

<sup>6</sup> 政策専門職の長とは、各省庁の政策形成を統括する幹部職員である。英国の国家公務員はすべていずれかの専門職(profession)に属することとなっている。かつては一般的な政策形成業務を担う者はジェネラリストと呼ばれていたが、現在はそうした官僚は政策専門職に属することになっている。



決定者<sup>7</sup>によって十分理解されていない」である。チーフアナリストは障壁だと感じているものの、政策専門職の長は障壁だと感じていないのが「幹部職員からの評価エビデンスに対する要求が不足している」である。逆に、政策専門職の長は障壁だと感じているものの、チーフアナリストはそう感じていないのが「必要な時に評価エビデンスが利用可能ではない」、「評価エビデンスのナレッジベースから、"What Works"を見つけることが難しい」、「評価エビデンスの頑健性が不十分」などであり、政策専門職の長とチーフアナリストの間における認識の差が示唆されている。

図 2-5 意思決定に対する評価エビデンス活用の障壁 (n = 32)



(出所) NAO (2021)

評価の提供と利用における障害について、要因別（政治的要因、戦略的要因、技術的要因）に整理したものが表 2-1 である。戦略的要因については、プロジェクトオーナーが評価をハードルとして捉えていることや、ネガティブな分析結果が得られた場合には出版バイアスがかかることなどが障害として挙げられている。また、戦略的要因においては、評価に対する戦略的なアプローチが行われていないこと、効果的な評価を行うためのリソースが存在しないことなどが指摘されている。技術的要因では、評価のスペシャリストが不足していることなどが指摘されており、評価の提供面と需要面の両方で多くの課題が示さ

<sup>7</sup> NAO (2021) では"policy-makers"という言葉が使われている。明確な定義はされていないものの、大臣および政策専門職の官僚を指すものと考えられる。

れている。

表 2-1 評価の提供と利用における障害

リスクの分野	評価の供給に対する障害	評価への需要に対する障害
政治的要因 (Political factors)	政治的支援の不足 (Lack of political support)	<ul style="list-style-type: none"> <li>■ 評価に対する政治的関心・支援の欠如。</li> <li>・ 政治の評価実施者タイムテーブルのミスマッチ。</li> <li>・ 選挙サイクル。</li> <li>・ 上級公務員の異動。</li> <li>■ 大臣がアウトカムに興味を持っていない。</li> <li>■ プロジェクトオーナーによる評価結果の拒否。</li> </ul>
	文化的賛同 (Cultural buy-in)	<ul style="list-style-type: none"> <li>■ プロジェクトオーナーは、評価を「バリュフォーマネーを保証し、将来の最適な介入設計を行うために必要な研究開発」ではなく、「課せられたもの (imposed process)」や「ハードル」として捉えている。</li> </ul>
	透明性 (Transparency)	<ul style="list-style-type: none"> <li>■ 評価結果に関する透明性が低く、ネガティブな結果に対する出版バイアス (publication bias) がある。</li> <li>■ プログラムが失敗なることを恐れて、透明性やフィードバックに対する開放性が欠如する。</li> </ul>
戦略的要因 (Strategic factor)	政府中核からのリーダーシップとサポート (Leadership and support from centre)	<ul style="list-style-type: none"> <li>■ 不十分で効果的ではないリーダーシップやサポート。</li> <li>■ 評価結果に対して政府中核の監督がほとんどない。</li> <li>■ 政府中核と各省が評価の目的に関する共通の視点を共有していない。</li> <li>■ ビジネスケース承認後に、約束された評価が行われたかどうかについて財務省の監視が不十分。評価計画のないビジネスケースに対して意義を唱えるかどうかの基準が明確ではない。</li> <li>■ 評価を行わなかったとしても影響がほとんどない。</li> </ul>
	各省のリーダーシップと戦略 (Departmental leadership and strategies)	<ul style="list-style-type: none"> <li>■ 評価と評価に対する支出に戦略がない。</li> <li>■ 評価活動が、優先的なエビデンスギャップやリスクレベルとシステムティックな関連がない。</li> </ul>
	リソース (resources)	<ul style="list-style-type: none"> <li>■ 効果的に評価を行えるだけのリソースがない。</li> </ul>
技術的要因 (Technical factors)	政策と分析の統合 (Integrating policy and analysis)	<ul style="list-style-type: none"> <li>■ 政策や分析を効果的に行うために人材やプロセスの統合が弱い。</li> <li>■ 政策は、検証や評価のために設計されておらず、政策形成過程においてアナリストが関与するのが遅すぎる。</li> <li>■ プロジェクト承認プロセスの十分に早い段階で評価計画が組み込まれていない。</li> </ul>
	短期志向/支出見直しなどに左右される行動 (Short-termism/ behaviour driven by for example spending reviews)	<ul style="list-style-type: none"> <li>■ 短期志向。</li> <li>■ 評価計画なし支出を決定してしまう。</li> </ul>
	スキル (Skill)	<ul style="list-style-type: none"> <li>■ 評価スペシャリストの不足。</li> <li>■ スキルと能力に関する戦略的リーダーシップの欠如。</li> </ul>
	共有されたアウトカム/データ (Shared outcomes/data)	<ul style="list-style-type: none"> <li>■ データマネジメントと活用が弱い。</li> <li>■ 安全に結合された行政データを用いた、低コストで短期間のインパクト評価ができる機会が乏しい。</li> </ul>
	学習メカニズム (Learning mechanisms)	<ul style="list-style-type: none"> <li>■ 評価から得られる学びを収集・管理・アクセスできるような政府横断的なメカニズムが限られている、または存在しない。</li> </ul>

(出所) NAO (2021)

## 2.3 評価の更なる推進に向けた論点

NAO 職員へのインタビューにおいて、政策評価の更なる推進に向けた論点として、次のようなことが明らかとなった<sup>8</sup>。日本の政策評価にとっても参考となると思われるので掲げておく。

### 2.3.1 予算と評価の関係

第 1 に、予算と評価との関係である。現在英国政府では、予算配分を評価とリンクさせ

<sup>8</sup> 以下は英国会計検査院職員のルース・ケリー (Ruth Kelly) 氏とフィリップ・ブラドバーン (Philip Bradburn) 氏へのインタビューに基づく (2022 年 5 月 19 日、オンラインにて実施)。

る動きが進んでいる。まず英国の歳出見直し(Spending Review)<sup>9</sup>プロセスを概観しておく  
と、英国では1997年のブレア政権発足以降、中期的な予算編成の枠組みを導入しており、  
各省庁は数年ごとに歳出見直しに関する要求を財務省に提出することになっている。財務  
省の歳出見直しチームが各省庁からの提出を分析して、財務省の中枢に提案をすること  
になっている。その後財務省は、各省庁との合意文書(settlement)を作成する(このプロセス  
は日本における概算要求と政府原案の関係に似ている)。歳出見直しは通常、財政見直しに  
基づいて、各省の3~5年間の歳出上限を設定する。

2020年の歳出見直しに伴う合意文書では、重要な優先分野に関する評価計画が初めて  
求められた。後述する政策評価タスクフォースはそれを支援する役割を担っている。また、  
同年の歳出見直しにおいては、資本支出や資源支出を新たに行うごとに、どのようなエビ  
デンスがあるのか、エビデンスギャップがある場合にはそれをどのように埋めるかについ  
て記載するよう求められた。その上で、チーフアナリストが歳出見直しに署名しなければ  
ならなくなった。財務省の歳出見直しチームは、こうして提出されたエビデンスを考慮に  
入れて歳出の決定を行った。

とはいえ、NAOサイドから見れば、財務省の頑健な評価に対する関心はまだ低いよう  
である。財務省は大きな3つの梃子を持っている。第1に、財務省は歳出見直しの一部とし  
て、初期の財政配分を決定する。第2に、財務省は合意文書の条件を設定することが出来  
る。第3に、ビジネスケース(予算要求書)を承認するゲートウェイとなる。財務省評価  
を促進・支援するために、これら3つの梃子をもっと使えるはずだというのである。

他方で、NAOの報告書が財務省の政策評価に対する努力を促したことで、同省内の政策  
評価担当部署の活動に役立ったと思われるとの見解もNAO職員から得られた。

### 2.3.2 政策評価の法制度化

第2に、政策評価の法制度化についてである。日本では評価が法制化されている一方、  
英国では基本的には法制化はなされていない。この点についてNAO職員に問うたところ、  
以下のような答えが返ってきた。

---

<sup>9</sup> 歳出見直し(Spending Review)とは、数年にわたる予算配分の計画を記した文書であ  
る。

政策評価の仕組みを法制度化するという方向も面白いが、大事なのは各省庁がエビデンスに基づく立案・評価を行うためのインセンティブやサポートである。法制化されるかどうかよりも、それが適切に監督されるかどうかの方が重要である。包括的なシステムがどのように機能し、そのシステム内のさまざまな要素がどのような責任を負っているかが大切である。

一般的に英国では政策評価に関する法的な義務は課されていないが、いくつかのガイダンスが政策評価を要請している。そのひとつが「公的資金の管理」(Managing Public Money)である。これは、各省の会計担当官(筆者注:通常は事務次官)個人に対して、自らの組織の調達、プロジェクト、プロセスが財政全体にとっての持続可能性、有効性、慎重さ、質、価値に対する信頼を提供するために、体系的に評価されるように責任を課している。またグリーンブック(Green Book)は全ての政策に求められる事前評価・事後評価のガイダンスであり<sup>10</sup>、マゼンタブック(Magenta Book)は事後評価に関する詳細なガイダンスであり、評価手法のベストプラクティスや政策サイクルにおける評価の活用方法、政策立案者や国民に対して評価エビデンスを普及する方法についてまとめている<sup>11</sup>。なお、規制インパクト評価(regulatory impact assessments)については法的な義務が課されている。規制インパクト評価では、規制的影響がビジネスにもたらされる場合、各省大臣は原則として評価結果を法律案に含めることが義務付けられている。

いずれにせよ、英国の政策評価は法律というよりも「期待(expectation)」である。NAOの報告書は、こうした期待が達成されているのか、インセンティブがあるのか等を分析したものだといえる。

---

<sup>10</sup> 詳細は、内山・小林他(2022)を参照。

<sup>11</sup> 詳細は、3.3節参照。

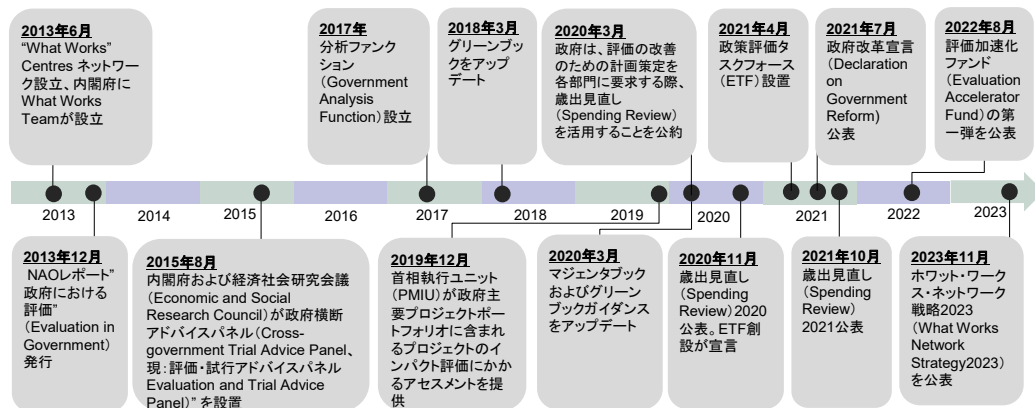
### 3 EBPM 体制の改革

#### 3.1 2013 年以降の取組と EBPM 体制

前節で扱った NAO 報告書に記されているとおり)、英国では 2010 年代終盤頃より、政策評価の結果を政策立案や予算配分とリンクさせようという動きが強まってきた。そうした動きを背景として、2010 年代末から 2020 年代初頭にかけて政策評価を含む EBPM の体制について改革が行われた。

2013 年以降の英国における EBPM の取組を整理したものが図 3-1 である。2013 年はホワット・ワークス・センター(What Works Centres, WWC)およびホワット・ワークス・ネットワーク(What Works Network, WWN)が設置され、内閣府の中にそれらを管理するホワット・ワークス・チーム(What Works Team, WWT)が設置された年である。2015 年 8 月には、内閣府および経済社会研究評議会(Economic and Social Research Council:ESRC)によって、政府横断試行アドバイスパネル(Cross-governmental Trial Advice Panel, 現 Evaluation and Trial Advice Panel) が設置された。これは政府内における政策実験の実施等を支援する専門家パネルである。2017 年には、政府エコノミストや統計職といった各分析専門職(analytical professions)を包含する制度として分析ファンクション(Analysis Function)が置かれた。2018 年 3 月にはグリーンブック(Green Book)の改訂が行われた。2020 年 3 月にはマジェンタブック(Magenta Book)およびグリーンブックが改訂された。2021 年 4 月には ETF が設置されるが、その経緯については次節で詳細を述べる。

図 3-1 2013 年以降の英国における EBPM の取組

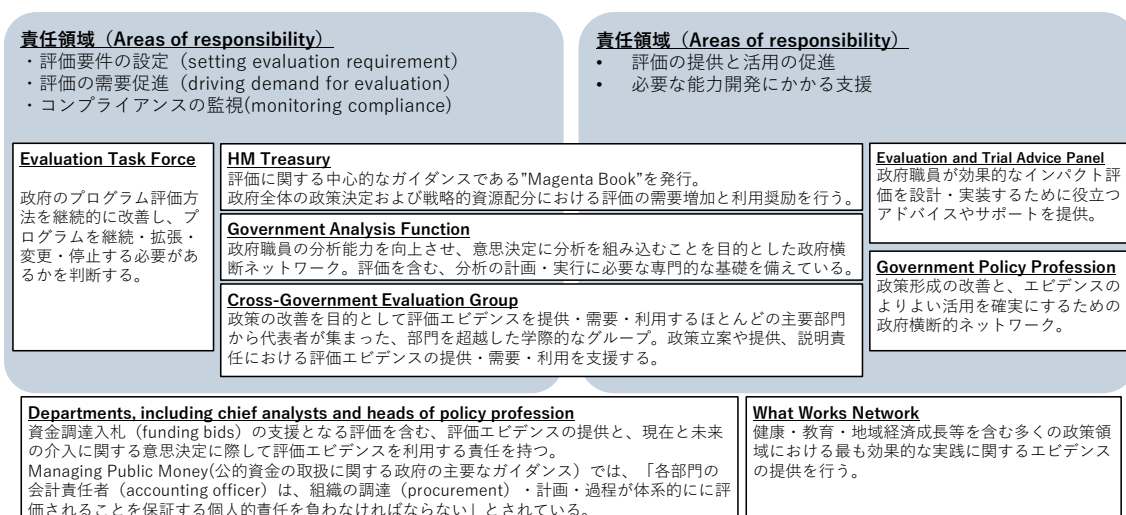


(出所) NAO (2021) などをもとに作成。

図 3-2 は評価の提供・利用に関する各省および政府横断的な責任を整理したものである。英国には、規制インパクト評価を除くと評価は法制度化されておらず、評価の提供・マネジメントに関する責任を単独で負う機関も存在しない。しかしながら各省は介入を評価し、既存の評価エビデンスから学習する責任を有しており、政府全体の評価の監視とサポートの提供に関与するいくつかの中央部門・政府横断的機能が存在する。

それらのうち、本節では分析ファンクションの設立と改訂マジェンタブックの公表を扱う。政策評価タスクフォースについては節を改めて述べる。

図 3-2 評価の提供と利用に関連する各省庁および省庁横断の責任



(出所) NAO (2021) をもとに作成。

## 3.2 分析ファンクション

### 3.2.1 分析ファンクションの概要

英国では政府エコノミスト (government economist) をはじめとする分析専門職が EBPM に大きな役割を担っている (内山・小林他 2022)。分析専門職にはエコノミスト (経済職) のほか、保険数理職 (actuary)、地理職 (geographer)、オペレーショナル・リサーチ職 (operational researcher)、社会調査職 (social researcher)、統計職 (statistician)、デジタル・データ職 (Digital and Data profession) がある。各専門職の多くはそれぞれグループ

を構成し、事務局機能も有している。主要なものを挙げると、エコノミストは「政府経済サービス」(Government Economic Service)、社会調査職は「政府社会調査」(Government Social Research)、統計職やデータサイエンティストは「政府統計グループ」(Government Statistician Group)を構成している。事務局については、政府経済サービスと社会調査は財務省、政府統計グループは国家統計局(Office for National Statistics)に置かれている。

これらの専門職グループを包括する制度として、2017年に「分析ファンクション」(Analysis Function)が設置された<sup>12</sup>。分析ファンクションの事務局は国家統計局に置かれ、ファンクションの長は国家統計官(National Statistician)が兼任している。本稿執筆時点での分析ファンクション長・国家統計官は、スコットランドのアバディーン大学で統計学教授を務めていたイアン・ダイヤモンド(Ian Diamond)氏である。

分析ファンクションは、政策分析に携わる公務員の分析能力・スキルを向上させることを目的としている。具体的には、能力・スキル向上のためのトレーニングやメンバーのキャリア開発支援などを行っている。政府においてエビデンスに基づいた政策決定がなされ、金額に見合った価値が実現されるようにすることがその使命である。

上述のとおり、分析ファンクションは7つの専門職グループを包含している。各専門職グループはそれぞれの分野での分析能力向上などに役割を果たしている一方で、分析ファンクションは、各グループを横断して、メンバーが優れた取組(グッドプラクティス)や基準を共有し、革新的な方法を開発し、インパクトのある分析を提供することを支援する役割を持つ<sup>13</sup>。また、分析ファンクションと後述するETFの関係は、制度的なものというより人的なネットワークによるものが中心のようである。例えば、ETFの業務で得られた知見が、分析ファンクションの幹部であるETF長を通じてファンクションのトレーニング

---

<sup>12</sup> 「ファンクション」とはキャメロン政権下での公務員制度改革で導入された省庁横断的なネットワーク型グループであり、分析のほか財務、人事、広報、商務、デジタルデータテクノロジーなど主要な政府機能を担うファンクションが設置されている(内山・藤田 2023)。

<sup>13</sup> 分析ファンクションウェブサイト(<https://analysisfunction.civilservice.gov.uk/>, 2024年5月14日最終アクセス)。

基準設定に反映されるといったことである<sup>14</sup>。

### 3.2.2 分析ファンクション戦略

分析ファンクションは 2022 年に今後 3 年間の戦略を示す「分析ファンクション戦略 2022-2025」を公表した。以下その内容を概観する<sup>15</sup>。

今後 3 年間を見渡しての同ファンクションの基本的なビジョンは、「政策決定を十分な情報に基づいたものにするための最良の分析を提供することを通じ、市民により良い結果を提供すること」とされている。以上のビジョンを実現するため、次のような結果を出すとの意欲が示されている。

- 政府を横断した良い分析のための、より一貫した基準
- 各省庁と各専門職を横断した活気のある分析コミュニティ
- 基準と働き方についてのより高度な調和と、それをサポートするリーダーシップ
- 規模の経済を通じた効率化の機会、変革と最先端の能力のための機会の促進
- より広いキャリア機会を享受できる、熟練した、多様で、より広く配置できる労働力
- 分析が重要な意思決定において役割を果たすこと、短期的なプライオリティや長期的な挑戦・リスク・機会に反応できるようにアジェンダを主導することを保証すること

この戦略がファンクションの構成メンバーに対して持つ意味については、すべての分析を最良のものとするためにファンクションがメンバー一人一人をサポートすることをこの戦略は保証する、とされる。その上で、メンバーが持つべき価値観として、前向き思考、

---

<sup>14</sup> ETF 長キャサリン・ハッチンソン氏へのインタビュー（注 21 参照）。

<sup>15</sup> 以下、“Analysis Function Strategy for 2022 to 2025,” 分析ファンクションウェブサイト (<https://analysisfunction.civilservice.gov.uk/about-us/analysis-function-strategy/>, 2024 年 5 月 14 日最終アクセス)。



包摂、信頼、誇りといった4つのものが挙げられている。

そして、このビジョンを実現するために、次の6つの戦略的目標が掲げられている。

- 中央からのサポート、ガイダンス、専門的助言を通じて、高品質で応答的な分析を支援する
- 革新的なツールと最先端の技術を通じて、効果的で効率的な分析を促進する
- 分析上の基準を設定しそれへの意識を高めることを通じて分析を支援する
- 多様で活発なコミュニティを形成することを通じて統合的な分析を可能にする
- 適切な時期に適切な場所で熟練した人を配置できるようにする
- 影響力があり、統合され、応答的なリーダーシップを提供する

こうした使命を達成するのに鍵となるのは、次のような者との協働(collaboration)である。

- 省庁や組織を横断した分析ファンクションの全メンバー
- 政策専門職や科学・工学専門職を含むより広い専門職グループ
- 財務ファンクションを含む他のファンクション
- 地方自治体や学界などで分析に携わる広範な外部組織

以上が「分析ファンクション戦略 2022-2025」の概要である。各専門職グループも戦略を発表しているが<sup>16</sup>、それらを包含する分析ファンクションが明確な戦略を立ててメンバーに浸透させていることは注目されよう。いわば、こうした戦略が分析ファンクションのネットワークの結節点のような役割を果たしている。各省庁を横断するとともに政府内外にわたる分析ファンクションのネットワークは、英国のEBPMにおける肝心な要素である。

### 3.3 マジェンタブック

#### 3.3.1 マジェンタブックの概要

マジェンタブック(Magenta Book: Central Government guidance on evaluation)は政策

---

<sup>16</sup> 例えば、エコノミストは「2021-2024 政府経済サービス戦略」を公表している。“GES Strategy 2021-2024” (<https://www.gov.uk/government/publications/ges-strategy-2021-2024>, 2024年5月18日最終アクセス)。

評価についての中央政府公式ガイダンスである。2011年に初版が公表されたが、前述した政策評価体制の見直しの動きを受けて2020年に大幅改訂がなされた。マジェンタブックの目的は、政策の設計と実施に評価を組み込み、政策形成の各段階を見直す手法についてのガイダンスを提供することである。またそれは、評価結果を解釈・提示する手法や、その過程において考慮されるべき事項についても解説している<sup>17</sup>。

マジェンタブックは、ジェネラリスト的な政策形成業務を担う政策専門職(policy profession)、政策の実施を担う実施専門職(delivery profession)、そして政策評価・分析を担う分析専門職のために書かれている。それらの専門職はすべて、優れたエビデンスを確保し利用することに責任を持っているのである。また、マジェンタブックには本体とは別に付録(Annex)が用意されており、そこでは分析手法の詳細が解説されている。そのほか、特定のタイプの評価ごとに補足的なガイドも用意されている<sup>18</sup>。

### 3.3.2 マジェンタブックに見る評価のあり方

以下、2020年版のマジェンタブック（本体）に基づき、英国政府における評価のあり方について述べる<sup>19</sup>。

「なぜ、どのように、いつ評価を行うか」と題する第1章では、政策評価の定義、理由、役割と時期、タイプなどが述べられている。まず政策評価は「政府の政策のデザイン、実施、結果についての体系的なアセスメント」と定義される<sup>20</sup>。評価は、次のような問いに答えるためのエビデンスを探求する。

- その介入は意図したとおりに働いているか。

---

<sup>17</sup> “The Magenta Book: HM Treasury guidance on what to consider when designing an evaluation,” 英国政府ウェブサイト GOV.UK (<https://www.gov.uk/government/publications/the-magenta-book>, 2024年5月14日最終アクセス)。

<sup>18</sup> 前掲英国政府ウェブサイト GOV.UK。

<sup>19</sup> HM Treasury, *Magenta Book: Central Government guidance on evaluation*, HM Treasury, March 2020, p.9.

<sup>20</sup> Ibid., p.9.

- それは異なった集団に対しては異なった形で働いているか。それはなぜか。
- その政策は実際にどのように作用しているか。
- その政策はどこにおいて改善しうるか。
- その政策の全般的なインパクトは何か。
- それは「金額に見合った価値」をもたらしているか。
- もし同様の政策を実施するとしたら、何を異なった形で行うか。

評価を行う理由としては、学習(learning)と説明責任(accountability)が挙げられている。学習については、実施された介入から学習したことが、同じ分野または関連分野における次なる政策に適用されることが重要である。評価は、ある政策を継続するか否か、どのように改善するか、どのようにリスクを最小化するか、その政策を停止して他の分野に投資すべきかといったことについての決定を基礎づけるエビデンスを提供する。説明責任については、政府は、納税者のお金と政府の活動により得られる公的価値と結果を最大化する責任がある。そのため、政府の各省庁には、政策の結果と価値について国民に情報提供をしなくてはならないし、各省の会計担当官（筆者注：通常は事務次官）に対する説明責任と透明性を確保しなくてはならない。

評価が役割を果たすのは、ROAMEF サイクルのすべての段階においてである。ROAMEF サイクルとは、政策の根拠づけ(Rationale)→目標(Objective)→吟味(Appraisal)→モニタリング(Monitoring)→評価(Evaluation)→フィードバック(Feedback) (→根拠づけへ戻る)の各段階をたどる政策のサイクルのことである。すなわち、広義の評価は、狭義の評価（ROAMEF の E）だけにとどまるのではなく、政策実施（介入）の前、途中、後のすべての時期において利用されることが必要である。特に重要なのは、評価計画を早い段階から作っておくことである。適切なデータが効果的・効率的に収集されるようにし、評価に高い質を持たせるためには、介入をデザインする際に評価を組み込んでおくことが必要である。

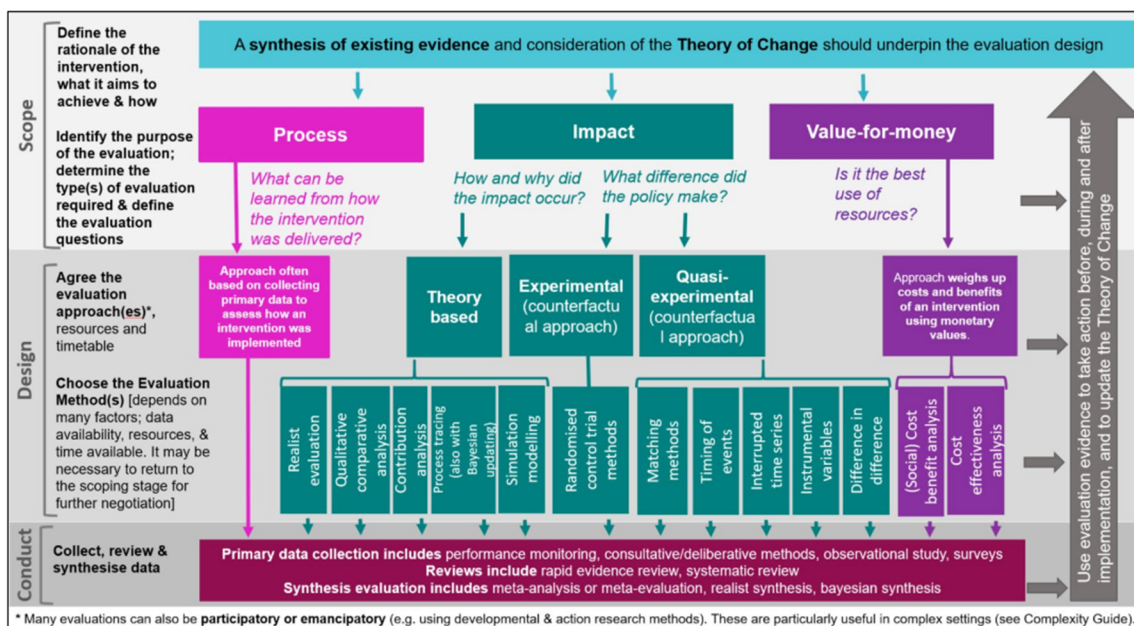
評価のタイプとしては、プロセス評価(process evaluation)、インパクト評価(impact evaluation)、「金額に見合った価値」評価(value-for-money evaluation)の3つが挙げられる。プロセス評価とは、介入の実施に関わる活動と、政策効果が国民に届くまでの経路を精査するものである。要するに、政策がどのように実施されたか、そのプロセスを評価するものである。インパクト評価は、介入によりもたらされた変化、すなわち、介入の目標

に関わる測定可能な業績に焦点を当てる。これには、その政策がどの程度の変化をもたらしたかという点（後述の実験・準実験アプローチ）と、その変化がなぜどのように起こったかという点（後述の理論に基づくアプローチ）の二つがある。「金額に見合った価値」評価は、政策の便益が費用を上回っているか、資源をもっとも効率的に利用しているかといった問題を扱う。資源がもっとも効率的に利用されたかを評価するものである。

その他本章では、評価の主体やステークホルダーなどについて述べられている。

「評価の目標設定」と題する第2章では、まず、目標設定ステージ(scoping stage)、デザインステージ(design stage)、実施ステージ(conducting stage)という評価の3ステージについて示されている（図3-3参照）。目標設定ステージとは、評価の目的、採用する評価のタイプ、評価によって答えられるべき問いを決定する段階である。デザインステージとは、もっとも適切な評価アプローチ・手法を選択する段階である。実施ステージは、様々なデータを収集し、調査・分析する段階である。デザインステージと実施ステージについては第3章以降で述べられている。

図 3-3 英国政府における政策評価の手順（目標設定、デザイン、実施）



(出所) Magenta Book, p.23

目標設定ステージにおいては、まず、「変化の理論」(Theory of Change)を明確にするこ

とが必要である。「変化の理論」とは、介入がどのように働くかについての理論である。介入が結果をもたらすまでのすべてのステップ（因果の連鎖）を明らかにすることが中心となる（筆者注：日本のロジックモデルと近似したものといえる）。採用される仮定、それを支持するエビデンス、より広い文脈的要因も含まれる。言い換えれば、これは、介入の根拠づけと、何をどのように達成するかを定義する段階である。

次に、評価の目的と評価についての問いを定義する。そして、プロセス／インパクト／「金額に見合った価値」の各評価タイプのうちからどれか（複数の場合もある）を決定する。ここで重要なのは、各評価タイプはそれぞれ弱みを持つことである。例えば、インパクト評価は、それ自体ではなぜそのインパクトが生じたかを説明することはできない。そうした場合にはプロセス評価を併せて行うことが必要となる。

次に、採用する評価タイプに基づき、適切なアプローチを選択する。インパクト評価に関するアプローチとして、実験・準実験アプローチ(experimental and quasi-experimental approach)と、理論に基づくアプローチ(theory-based approach)の2つが挙げられている。実験・準実験アプローチは、介入に影響されていない集団との統計的比較を通じて介入のインパクトを推論する。この介入に影響されていない集団は、介入を受けた集団がもしその介入がなければどうなっていたかを示すものであり、反実仮想(counterfactual)と呼ばれる。一方、理論に基づくアプローチとは、結果をもたらすと考えられる因果の連鎖が強力なエビデンスによって支持されるかを厳密に検証し、また代替的説明が棄却されることを確認することによって、介入のインパクトを評価するものである。このアプローチにおいては、上述の「変化の理論」が綿密に策定されていることが重要である。

第3章「評価の手法」は、プロセス評価、インパクト評価、「金額に見合った価値」評価において使用される評価手法を説明している。まず、プロセス評価とインパクト評価の両方で利用される手法は表 3-1 に挙げられているとおりである。

表 3-1 プロセス評価とインパクト評価の両方で利用される手法

手法	概要
インタビューとフォーカスグループ (interviews and focus groups)	関係者へのインタビューにより介入についての深い洞察が得られる。フォーカスグループは集団の見解を得るのに有益である。
事例研究(case studies)	現実世界の文脈における複雑な現象を解明できる。

サーベイ(surveys)	多数の関係者から情報を得るのに有効である。
出力・業績モニタリング (output or performance monitoring)	介入の実施や結果について低コストで知ることができる。
参与観察 (observational studies)	介入についての個人の経験を深く理解することができる。

インパクト評価に使用されるのは理論に基づくアプローチと実験・準実験アプローチである。理論に基づくアプローチに関する手法は表 3-2 のとおりである。

表 3-2 理論に基づくアプローチに関する手法

手法	概要
現実主義的評価 (realist evaluation)	介入が結果をもたらす「因果メカニズム」を解明する。
寄与分析 (contribution tracing)	「変化の理論」に基づき、介入が結果に寄与したかをステップごとに明らかにする。
過程追跡 (process tracing)	仮説的な因果メカニズムが結果を説明するか、単一の事例を体系的に調査する。
ベイズ主義的更新 (Bayesian updating)	寄与についての確立を事前に想定し、それを検証する。
寄与追跡 (contribution tracing)	結果への寄与についての主張の妥当性を検証するための参与的手法。
質的比較分析 (Qualitative Comparative Analysis)	複数の事例を体系的に比較して、結果に結びつけられる属性を明らかにする。
結果の取り入れ (outcome harvesting)	変化についてのエビデンスを収集して変化への寄与を評価する。
もっとも重要な変化 (most-significant change)	もっとも重要な変化についての経緯を収集して、複雑な介入のインパクトを評価する。

実験・準実験アプローチに関する手法は表 3-3 のとおりである。

表 3-3 実験・準実験アプローチに関する手法

手法	概要
ランダム化比較試験 (Randomised Controlled Trials (RCT))	介入があった場合の効果となかった場合の効果を、参加者をランダムに割り当てることで比較分析する。
分割時系列分析と差の差法 (interrupted time series analysis and difference-in-difference)	前者は時系列データを用いて介入に続く結果の変化を分析する。後者はコントロールグループにおけるトレンドと比較する。
回帰不連続デザイン (Regression Discontinuity Design)	対象が閾値の上か下か（あるいは境界の内か外か）の相違に基づき介入の効果を分析する。
傾向スコアマッチング (Propensity Score Matching)	介入グループにマッチする比較グループを統計的に作り出す。
合成コントロール法 (synthetic control methods)	歴史的データを使って介入を実際に受けたグループの「合成クローン」を作り出す。
操作変数法ないし自然実験 (instrumental variables or natural experiments)	前者は、結果には影響しないが介入には影響する要因（操作変数）を利用する。後者は計画することはできないが準備することはできる。
出来事のタイミング (timing of events)	介入と結果のタイミングをモデル化することにより介入の結果を推計する。

「金額に見合った価値」評価に関する手法は表 3-4 のとおりである。

表 3-4 「金額に見合った価値」評価に関する手法

手法	概要
社会的費用有効性分析 (social cost effectiveness analysis)	同じ結果指標をもたらす介入の選択肢間で費用を比較する。
社会的費用便益分析 (social cost-benefit analysis)	介入の選択肢の費用と便益を比較分析する（異なった結果指標間でも使える）。

以上の手法については、マジエンタブックの付録(Magenta Book Annex A: Analytical methods for use within an evaluation)で詳細な説明が提供されている。

第4章「データの収集、アクセス、連結」では、まず、データ収集の計画は介入の発展と並行して練られなくてはならないことが強調されている。そうすることは、データの収集とアクセスが政策のデザインや関連する法制に組み込まれることを確かにするからである。

データのソースとしては次のようなものがある。第1に、既に存在する行政データとモニタリングデータである。行政データとは各行政機関が収集したデータであり、モニタリングデータないし業績管理データは介入を通じて収集されたデータである。第2は、既に存在するサーベイデータである。国家統計局が収集したデータなどが典型である。第3は、当該評価のために特別に収集された新規データである。サーベイ、インタビューや観察などの質的手法、ウェブなどを通じて得られる。第4はソーシャルメディアデータである。介入への自然発生的な反応を調べるのに有効である。

これらのデータの取扱いは適切になされなくてはならないが、特に、EUの一般データ保護規則(General Data Protection Regulation)に配慮することが必要である<sup>21</sup>。データの連結(linking)も重要である。それは、より豊かなデータセットを作り出し、分析を高め、データ収集の重複を回避することができる。データ連結には各データを組み合わせるための識別子(identifier)が必要であるが、それは信頼される第三者機関が管理すべきである。それにより、分析者は安全な環境で特定不能な連結データにアクセスすることができる。

第5章「評価の管理」は、評価のガバナンス、資源、評価の介入とのリンク、仕様、倫理などについて述べられている。ガバナンス体制は、評価の最初の時点から有効なものが設置されていることが重要である。政策決定者やその他の関係者などのステークホルダーを巻き込んでおくことにより、状況の変化に柔軟に対応することや、評価が影響力を持つことが可能となる。

評価の資源としては財政的なものと人的なものがある。人的資源については、スペシャ

---

<sup>21</sup> 英国のEU離脱後も、一般データ保護規則は英国国内で法制化されているために有効である。



リストの評価スキルとともにジェネラリストのプロジェクト管理スキルも重要である。介入とのリンクについては、評価の計画が介入の計画と密接に関連して立てられることが必要である。例えば、評価に必要なデータをモニタリングデータに組み込んでおくなどの方法がある。評価実施の方式としては、外部委託、内部実施、それらの併用がある。どの方式をとるにせよ、評価の仕様を明確にしておくことが肝要である。倫理については、評価に関わるステークホルダーにはそれぞれに倫理的配慮が求められる。例えば、参加者には秘密の保持、資金提供者には契約上の明確性などが求められる。

第 6 章「評価によって得られた知見の利用と普及」では、まず、評価の有する価値は、それが利用され、影響力を持つことにあると述べられている。そのため、評価の利用と普及の計画をあらかじめ策定しておく必要がある。そうした計画は、評価の利用者を引き付け、インパクトを最大化することに貢献する。

関連して、評価の文化を作り出すことも重要である。評価は、独立した単体の機能ではなく、より広い組織の文化や構造の中に統合されたものでなくてはならない。評価の慣行が普段の行動様式の一部となり、すべての同僚がエビデンスについて批判的に考えるようになることが重要である。そのためには、開放性と透明性が重視されなくてはならない。

第 7 章「評価の能力」では、評価において必要な 4 つの能力(capabilities)について述べられている。第 1 は目標設定の能力である。評価の根拠を理解する、「変化の理論」を構築する、適切なアプローチを同定するといった能力である。第 2 は学習と管理の能力である。リーダーシップを発揮する、ステークホルダーと協働する、事情の変化に適応するといった能力である。第 3 は手法に関する能力である。評価の各アプローチや手法に通じていることが必要である。第 4 は利用と普及に関する能力である。データを報告・提示する、政策への示唆を検討する、評価結果を普及させるといった能力である。

改訂マジェンタブックの概要は以上であるが、ここでは以下の点に着目しておきたい。第 1 にしばしば政策評価というと定量的なインパクト分析を指すものと思われるが、英国では定性的なプロセス評価や費用便益分析も政策評価の柱となっている。RCT や準実験だけが政策評価・EBPM の手法なのではなく、定量的な手法と定性的な手法を相互補完的に使用するのが望ましい。第 2 に、政策デザインの段階で評価を組み入れるべきことが強調されている。例えば、頑健な分析の根拠となる適切なデータを取得するためには、政策執行の過程においてデータを取得できるような仕組みを政策デザインの段階で組み込ん

でおくことが有効である。第3に、マジェンタブックは分析専門職だけではなく、政策専門職や実施専門職の公務員のためにも書かれている。政策評価の基本的な考え方や論理は、企画立案や調整などを担うジェネラリスト的立場の者にも共有されるべきといえよう。

## 4 政策評価タスクフォースの取組

### 4.1 設立の経緯

2021年4月、政策評価タスクフォース(Evaluation Task Force, ETF)が設置された。ETFは内閣府と財務省の共同傘下にある組織であり、政策の有効性に関する頑健なエビデンスが政府の支出決定に当たって重視されるようにすることを主要な目的としている。具体的には、下記に詳述するとおり各省の予算要求の根拠となっているエビデンスについて財務省歳出チーム(Spending Teams、日本の主計局に相当)に助言すること、各省の評価の設計と実施に関して助言と支援を行うことなどを任務としている。またETFは、前述したマジェンタブックを財務省と共同で所管している<sup>22</sup>。

ETFの設立経緯は以下のとおりである<sup>23</sup>。ボリス・ジョンソン政権下では、政策評価に重点を置く気運が高まっていた。その中心にあったのはマイケル・ゴーヴである。ゴーヴは当時、内閣府の中でも高位の大臣であるランカスター公領相のポストに就いており、政権の重要課題を進める立場にあった。彼は2020年6月に「公務の特権」(“The privilege of public service”)と題するスピーチを行った<sup>24</sup>。その中でゴーヴは、「私たちの政策の中心に位置するのは、『何がうまく機能するのか、何が市民の繁栄に役立つのか』に焦点を当てることである。そしてこのことが意味するのは、政府のプログラムについての厳格な評価である。」と述べ、政策評価を重視する姿勢を打ち出している。

そのほかにはジョンソン首相の特別補佐官(special adviser)だったドミニク・カミング

---

<sup>22</sup> “About us,” ETF ウェブサイト (英国政府ウェブサイト GOV.UK 内)

(<https://www.gov.uk/government/organisations/evaluation-task-force/about>, 2024年5月14日最終アクセス)。

<sup>23</sup> 以下の経緯については基本的に、ETF長であるキャサリン・ハッチンソン(Catherine Hutchinson)氏へのインタビューに基づいている(2022年3月20日にオンラインにて実施)。

<sup>24</sup> “‘The privilege of public service’ given as the Ditchley Annual Lecture,” 英国政府ウェブサイト GOV.UK(<https://www.gov.uk/government/speeches/the-privilege-of-public-service-given-as-the-ditchley-annual-lecture>, 2024年5月18日最終アクセス)。

スや、財務相の特別補佐官だったマイケル・ウェブらがいる。カミングスは首相に対して適切な助言をすることを職務としていたが、彼は、ビッグデータや政府所有データを有効活用する観点からデータダッシュボードに興味を持っていた。カミングスのこのようなスタンスはエビデンスベースのアプローチを推進する役割を果たした。また、ウェブはEBPMの強力な主唱者だった。さらに、金融セクターで働いた経験を持つ大臣たちも一役買った。そうした大臣は、数字やリスクに強く、小規模な実験を大規模に広げていくといったことに慣れていて、財務省をはじめとした各省の分析専門職幹部も、政策評価を阻害しているシステムティックな問題が多々あることを指摘していた。

一方、当時は「金額に見合った価値」(value for money)の実現を求める考え方も政権内で浸透していた。例えば、2017年11月に「より良い結果を市民に届ける一公的価値を開くための現実的ステップ」と題する報告書が教育専門家のマイケル・バーバにより公表された。この報告書は、現在の英国政府が「金額に見合った価値」を十分に実現していないことを指摘し、それを克服するための手法を提案している。当時財務省首席大臣(Chief Secretary to the Treasury)であったリス・トラスは、「この報告書は、持続的に改善する公的サービスの文化を作り出す上で重要なステップである」と述べて高い関心を示している<sup>25</sup>。

さて、ETF設立の直接の契機となったのは、2020年歳出見直し(Spending Review 2020)の策定である。上記のとおり政策評価重視の気運が高まる中、2020年歳出見直し策定の時期が訪れた。そこで、同歳出見直しにETF設置の予算を盛り込もうとのアイデアが出され、多くの大臣や官僚がそれを支持した。

---

<sup>25</sup> “Sir Michael Barber report into improving value in public spending published,” 英国政府ウェブサイト GOV.UK(<https://www.gov.uk/government/news/sir-michael-barber-report-into-improving-value-in-public-spending-published>, 2024年5月18日最終アクセス)； “Delivering better outcomes for citizens: practical steps for unlocking public value,” 英国政府ウェブサイト GOV.UK(<https://www.gov.uk/government/publications/delivering-better-outcomes-for-citizens-practical-steps-for-unlocking-public-value>, 2024年5月18日最終アクセス)。

こうして、2020年11月にリシ・スナク財務相により公表された2020年歳出見直しにおいて、ETFの創設が宣言された<sup>26</sup>。すなわち、同歳出見直しでは、「結果を出すことを支援するために、政府は高品質の評価により重点を置いている。・・・この歳出見直しの期間中、何が市民に結果をもたらすかをよく知ることが将来の選択を左右するだろう。それを支えるために政策評価タスクフォースを新設する。」と記されている。

以上のとおり、2020年歳出見直しという機会、影響力を持つ者による政治的な支持、「金額に見合った価値」という論点の浸透といった複数の要因が合流したことにより、ETFの設置が実現した<sup>27</sup>。

## 4.2 政策評価タスクフォースの組織と業務の概要

ETFの組織構造および業務の概要を示したものが図4-1である。ETFには約15名のスタッフがいますが全員が内閣府の職員である。ただし組織としては内閣府と財務省の共管となっており、ETFの長(Head)の上位には内閣府のディレクターと財務省のディレクターがそれぞれ存在する。

ETFのHead<sup>28</sup>の下には担当省別に4人のアカウントマネージャーがおり、その下にはスタッフが配属されている。ETFの4人のアカウント・マネージャーが、主要な政府の

---

<sup>26</sup> “Spending Review 2020,” 英国政府ウェブサイト GOV.UK

(<https://www.gov.uk/government/publications/spending-review-2020-documents/spending-review-2020>, 2024年5月18日最終アクセス)。

<sup>27</sup> この状況は、米国の政治学者ジョン・キングドンの提唱した「複数の流れ」モデル(Multiple Streams Framework)によって説明できる。キングドンによれば、問題の流れ、政策の流れ、政治の流れという3つの流れが合流したときに「政策の窓」が開き、大きな政策変化が実現する(Kingdon 1984)。

<sup>28</sup> 現在のETF長であるハッチンソン氏のバックグラウンドは心理学である。ETFに勤務する前は、外務・英連邦・開発省(Foreign, Commonwealth and Development Office)や警察カレッジ(College of Policing, WWCの1つ)、教育省などでの勤務経験を有する。現在は、内閣府、財務省および首相官邸の社会調査専門職長も兼ねている。

省を4つに分けて担当することで、それぞれのマネージャーが政府の活動の概ね4分の1をカバーしている。アカウントマネージャーもスタッフも、社会調査(Social Research)や経済学(Economics)、統計学などのバックグラウンドを有しており、分析・評価に知見を有している。各スタッフは、評価に関する共通テストに合格した上で採用されている。またETF設立以前は、内閣府ホワット・ワークス・チーム(What Works Team, WWT)が設置され、WWCの認証や管理を行っていたが、ETF設立後は統合され、業務を継続している。

ETFのミッションは、①歳出決定に資するため各省の歳出計画の根拠となるエビデンスと評価計画について財務省歳出チーム(Spending Teams)に助言と支援を提供する、②各省における頑健なインパクト評価の設計・実施に対して助言と支援を提供する、③データ・評価計画・評価結果に対して透明であるように各省に挑戦を促す、④エビデンスに基づく政策を促進するため各省の主要なステークホルダーと関係性を構築する、の4つである。

特に重要な業務が①と②の部分である。各省は、支出計画を財務省の歳出チームに提出するが、その際、政府支出がアウトカムの改善に寄与することを示すエビデンスや、アウトカムのモニタリング・評価計画を提出する必要がある。ETFは歳出チームに対して、各省から提出されたエビデンスや反実仮定の想定が十分なものかどうかを助言するとともに、評価の質を改善するために各省に対して助言や支援を行っている。英国では、歳出決定に際して財務省と各省の間で合意条件(Settlement Conditions)が合意されることがある。これは財務省と各省の間で締結される歳出のための条件である。ETFは歳出チームと連携して合意条件のなかに評価条件を組み込むことで、評価の質の改善に取り組んでいる。ETFは、各省の評価から選定したモデルケースについて、事業概要、評価手法、評価によって得られた知見、事業のインパクトを端的に整理した上でウェブサイトに公表しているほか<sup>29</sup>、各省の評価に対して直接的に助言することもできるし、評価や実験に関する専門家ネットワークである評価・試行アドバイスパネル(Evaluation and Trial Advice Panel, ETAP)

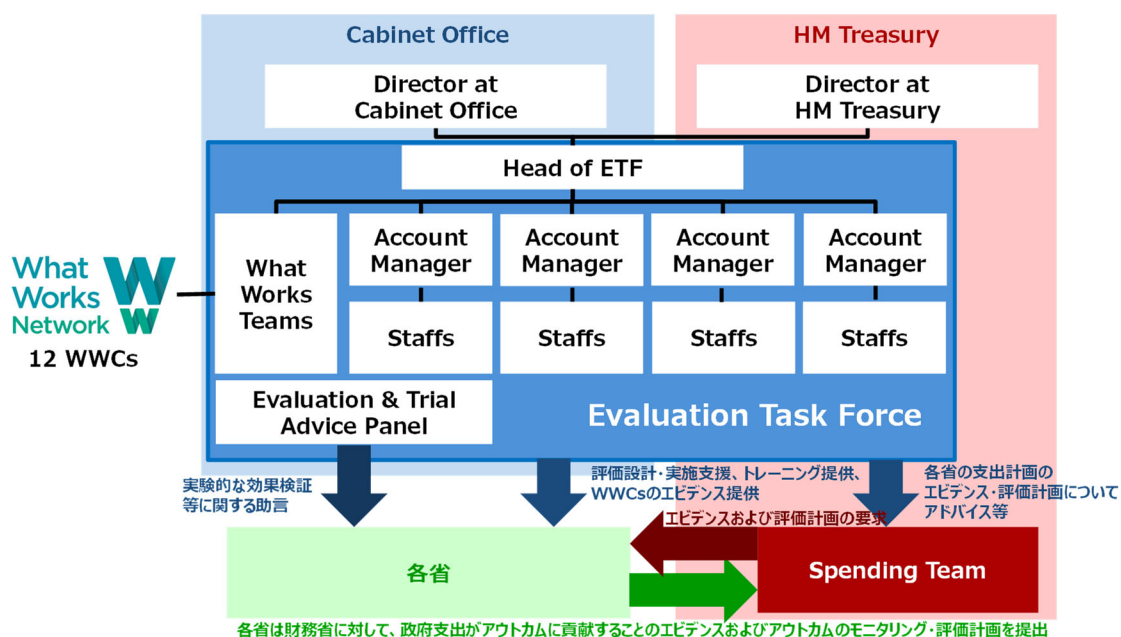
---

<sup>29</sup> “Featured Evaluations: Government,” 英国政府ウェブサイト GOV.UK (<https://www.gov.uk/government/collections/top-evaluations-government>, 2024年5月28日最終アクセス)。

を経由して間接的に支援することもできる<sup>30</sup>。

つまり、財務省の予算編成プロセスを通じて評価の改善を図るという「ムチ (Stick)」の機能と、各省の評価を支援するという「アメ (Carrot)」の機能を両方備えていると言える<sup>31</sup>。現地インタビューでも、「ETF、政治的支援が得られる内閣府と、お金をコントロールできる財務省の双方から支援が受けられるため、ETF が実際に有している権限以上に各省に与える影響は大きい<sup>32</sup>」との指摘があった。

図 4-1 Evaluation Task Force の組織構造と業務



(出所) Evaluation Task Force ウェブサイトおよびインタビューより作成

## 4.3 具体的な取組

### 4.3.1 財務省歳出チームとの連携について (2021 歳出見直しにおける取組)

歳出見直しの策定については、2.3.1 で述べた通り、まず、各省庁が数年ごとに歳出見直

<sup>30</sup> 評価・試行アドバイスパネルについては後述。

<sup>31</sup> NAO 職員ケリー氏・ブラドバーン氏に対するインタビュー。

<sup>32</sup> ETF 長ハッチンソン氏へのインタビュー。

しに関する要求を財務省に提出し、これを受けた財務省の歳出チームが各省庁からの提出を分析して、財務省の中枢に提案を行い、その後財務省は、各省庁との合意文書 (settlement) を作成する。

財務省歳出チームは、歳出見直しの策定に向けて各省の予算事業の要否を判断するにあたり、費用対効果、支出可能性等、多様な観点から検討を行っている。その中で、各省が財務省に提供するエビデンスの質や評価計画については、①事業が上手くいく見込みがあるか、②本格的に事業を実施するか、それとも試験的に事業を実施するか、③課題解決のための異なるアプローチはないか、といったことを検討する観点から重要であるが、これらの審査はすべて財務省歳出チームの担当官が担っていた。しかしながら、財務省歳出チームの担当官は、必ずしもエビデンスの質や評価計画の専門知識を持ち合わせておらず、担当官の素養によるところが大きかった。

これまでも、内閣府と財務省の共管で設置されていたインフラ・プロジェクト庁 (Infrastructure and Project Authority, IPA) が、インフラ投資の計画等について財務省に助言を行っていたが (IPA 2023) が、外部専門家からのアドバイスまでは受けていなかった。この点、ETF では、質の高さを担保するため、外部専門家からのアドバイスを受けた上で、助言を行っている。

一方で、各省庁の予算事業の要否について財務省が最終的に判断することについては ETF 設置後も変わっておらず、この点で、ETF からの助言は、財務省の裁量を狭めるものではない。

実際、十分なエビデンスがなかったとしても、財務省は予算事業の構想、設計を元に資金提供の要否を判断せざるを得ない。ETF がエビデンスが不十分だと財務省に主張したとしても、財務省が予算事業に対して肯定的な評価を下せば、資金提供が行われることとなる。

ただし、このような場合に、そのプログラムが金額に見合った価値かどうかの懸念があることから、ETF としては、当該事業の評価計画を策定し、事後的な評価を行うように財務省に提案を行った。結果として、プロジェクトを実施したい各省庁と、予算について検討を行う財務省の間で、財務省から各省庁に事後的な評価を求め、各省庁における事後的評価の実施という条件付きで予算措置が妥結した事業もあった。財務省はプログラムに資金を提供しており、各省はそれを評価する責任がある。そして最終的には、その評価のお



かげで、ETF はその便益や「金額に見合った価値」が期待に沿うものかどうかを理解することができる。ETF が作成した証拠が数年後の将来の歳出の決定に影響を与える可能性がある。

また、財務省が歳出見直しの策定に向けて有している問題意識と、ETF が有している専門知識との連携の円滑化を図るための取組として、評価について横断的な役割を担っている財務省の専門チームから、歳出チーム に対して、評価活動の重要性と ETF の助けを借りることのメリットが伝えられている。

また、2021 年の歳出見直しでは、ETF は政府全体にわたるすべての省庁の予算を調べ上げた。非常に小規模で、年間 10 億ポンド以下の予算しかない省庁もあれば、保健省のように年間 1600 億ポンド以上の予算がある省庁もあり、財務省歳出チーム においては、当然に予算規模等も踏まえた人員配置等がされており、ETF についても、それぞれの財務省歳出チームが行う査定の規模等に応じた対応が行われた<sup>33</sup>。

結果として、ETF 設立後、最初に公表された歳出見直しである 2021 年歳出見直し策定にあたっては、ETF は財務省の歳出チームと連携し、3 週間で 1000 億ポンドを超える 86 件の予算要求書をレビューした。具体的には、外部専門家からのアドバイスも受けつつ、各省庁の歳出計画の根拠となるエビデンスの質を検討し、各省庁が提案されたアイデアが機能するかどうかを検証するために反実仮想アプローチを採用しているかをチェックした<sup>34</sup>。また、各省庁の予算措置の条件として、今後 6 か月間で、評価戦略、優先度が高いエビデンスギャップを明らかにすることを約束させた。2024 年 3 月現在、ETF は 382 件の評価と 2 兆 200 億ポンドのプロジェクトについて助言している。

#### 4.3.2 ビジネス・ケースにおけるエビデンス活用・評価計画策定

英国政府では、事業（Project）や施策（Programme、複数の事業を束ねたもの）を行う場合、「ビジネス・ケース」を作成する必要がある<sup>35</sup>。それぞれの支出の要否はビジネス・ケースに基づいて判断されるため、いわば予算要求書だと言える。ビジネス・ケース

---

<sup>33</sup> ETF 長ハッチンソン氏に対するインタビュー。

<sup>34</sup> 同上

<sup>35</sup> ビジネス・ケースの詳細については、内山・小林他（2022）参照。

には、事業用と施策用の両方があるが、いずれについても戦略（介入の必要性や目標、合理性など）、経済（費用便益分析に基づく社会的価値など）、商業（サービス提供者との契約条件、リスク管理など）、財務（資本と収支、バランスシートへの影響など）、管理（プログラムの実施、モニタリングなど）という5つの観点から作成される。前述の通り、ビジネス・ケースを審査するのは財務省歳出チームの役割だが、ETF はエビデンスや評価の観点からビジネス・スペースの審査と改善を支援している。

ETF はビジネス・ケースにおいて、良いエビデンスおよび評価計画が備えるべき要件として7つを整理している（表 4-1）。これをみると、マジェンタブックの考え方がベースになっていることがわかる。

1 点目は評価対象となる介入・政策が特定されていることであり、2 点目はそれらの目的である。既存のエビデンスを参照したり、新たなエビデンス構築のための評価計画を策定するには、その対象となる介入・政策が特定され、目的（アウトカム）が定義されていなければ始まらない。

3 点目は介入・政策に関する既存のエビデンスである。ETF はビジネス・ケースにおいて、実施しようとしている介入・政策に関する既存のエビデンスを記載するように求めている。既存エビデンスの参照にあたっては、参照している既存のエビデンスの文脈と現在実施しようとしている文脈がどの程度異なるか、エビデンスレベルはどの程度か、そのエビデンスと反対の結論を得ているエビデンスはないか、エビデンスが得られて以降に政策効果に影響を与えるような外部要因はあるか、プログラムの提供方法・計画に変更はあるか、エビデンスギャップはあるかといった点を要約するように求めている。既存のエビデンスと介入・政策を検討している状況がまったく一致することはあり得ないため、既存のエビデンスがどの程度妥当性を持つのかの吟味を求めていると言える。

4 点目は変化の理論であり、それによって前提条件がどの程度成り立っているか、どの程度理論やエビデンスに基づいているか、アウトカム等を測定するにはどうすれば良いかといった点を明らかにすることができる。

5 点目の評価アプローチでは、提案されている評価の種類や、評価によって応えようとしている主たる問い、提案されている評価手法、評価実施時期などが確認のポイントとなる。加えて、評価やエビデンスは政策設計と一体不可分であることも強調されている。例えば、政策を全国で一斉に実施するのではなく、導入タイミングをずらすことで比較対照

群を作り出すことが出来れば、頑健性の高い評価が実施できる。また、政策対象者や対象地域の選定方法を工夫することで、効果検証をやりやすくすることもできる。政策実施前のベースラインデータを取得できれば、評価の質を改善できる。このように、政策デザインに工夫を凝らすことで、評価やエビデンスの質を上げることが出来る。

6点目は評価予算である。一般に、政策経費の3%程度が評価に充てられるべきだとしているが、新規性や不確実性が高い事業の場合、より多くの評価予算が必要となる。一方で、すでに頑健なエビデンスが得られている場合、通常のモニタリング以上の評価予算は不要となる。また、評価予算は政策規模には必ずしも比例しないため、大規模な政策の場合、必要となる評価予算は相対的に小さくなる。

7点目が評価結果の活用計画である。いかに精緻で用意周到な評価計画を立てたとしても、それが政策形成に活かさなければ意味がなくなってしまう。こういったエビデンスをどのようなタイミングで何のために使うかを事前に明確にしておくことによって、評価結果を有効活用できる。

表 4-1 ビジネス・ケースにおいて良いエビデンス・評価計画が備えるべき要件

項目	内容
1. 評価対象となる介入・政策	<input type="checkbox"/> 評価対象となる介入・政策の具体的な内容
2. 評価対象となる介入・政策の目的	<input type="checkbox"/> 介入・政策の目的およびアウトカム
3. その介入・政策に関する既存のエビデンス（戦略）	<input type="checkbox"/> 介入・政策に関する既存のエビデンスの有無。 <input type="checkbox"/> 既存エビデンスの適用可能性の吟味。
4. 変化の理論（戦略）	<input type="checkbox"/> 政策の前提条件やロジック等の整理。 <input type="checkbox"/> アウトカム等の測定方法の検討。
5. 評価アプローチ	<input type="checkbox"/> 評価アプローチの吟味。 <input type="checkbox"/> 政策設計を工夫することによるエビデンスの質の改善。
6. 評価予算（財務）	<input type="checkbox"/> 求められる評価計画の質と評価予算の吟味。
7. 評価結果の活用計画	<input type="checkbox"/> 評価結果を、こういったタイミングでどのように使うのかの検討。

(出所) Evaluation Task Force 提供資料

#### 4.3.3 評価加速化ファンドについて

評価加速化ファンド(Evaluation Accelerator Fund, EAF)は、2022～2025 年度において 1500 万ポンドを提供するファンドであり、以下の活動を通じ、こういった予算事業が有効かの知見を高めることを目的としている。

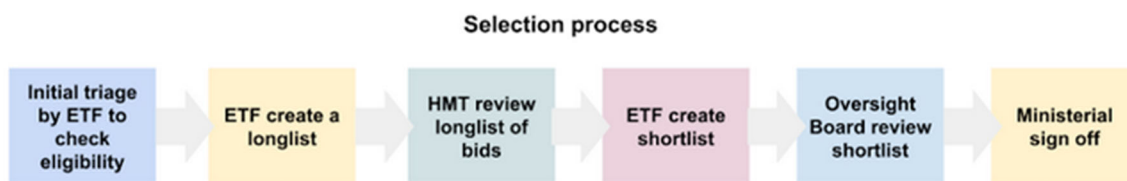
- 次の歳出見直しにおける意思決定に反映させるため、政府の優先分野における評価活動と実用的なエビデンスの作成を加速させる。
- 政府の優先分野でのエビデンスギャップに対処するため、各省庁や WWC が行っていない活動を迅速に行う。
- 新しい政策や介入、またはサービス提供への革新的なアプローチに関する頑健なエビデンスを提供する。

EAF では、政府の優先分野として、主に以下のミッションに即した評価を支援している。これらは首相のコアミッションに関連する領域に集中している。首相のコアミッションは犯罪と正義、健康と社会的ケアのように、広い分野をカバーしており、また、対外的な説明としても、政府にとってもっとも重要な分野についてこういった予算事業が有効かについて(what works)、政府がもっともよく理解している、という状態を目指すことが重要であるということであろう。

- 英国のレベルアップ
- ネットゼロカーボン
- 教育・職・技能
- 健康
- 犯罪と司法
- その他の優先政策

EAF は、各省庁および WWC から募集を受け付けており、また、ランダム比較試験や準実験のような、頑健性の高い評価手法を用いた評価アプローチへの支援に重点を置いている。応募のあった案件は、適格性のチェックの上、ETF によってスコアリングされた。その後、評価手法の一貫性を確保した上で、採点結果について議論し、応募のあった案件の総合的な RAG(Red, Amber, Green)評価に合意するために、コンセンサス・ミーティングが開催された。その後、さらなるフィードバックを求めるため、プロジェクトのロングリストは財務省や ETF 監督委員会(Oversight Board)と共有され、具体の支援プロジェクトが決定された(図 4-2)。

図 4-2 EAF の選定プロセス



(出所) Evaluation Task Force ウェブサイト

2022 年 8 月に、プロジェクトの第一弾が発表され、選出された 16 のプログラムには、合計 1,220 万ポンドが割り当てられた。また、プロジェクトの第二陣として、2023 年 3 月に 4 つのプロジェクトが選ばれ、合計 130 万ポンドが割り当てられた。さらに、2024 年 4 月には、プロジェクトの第 3 弾として、青少年犯罪、ネット・ゼロ、高等教育機関の学生の福利厚生、公務員採用における効果的な政策介入を実現するための政府の理解を深めるための 5 つのプロジェクトに、100 万ポンドが割り当てられた。表 4-2 に掲げる EAF の交付先 25 プロジェクトのうち、9 のプロジェクトが WWC によるものである<sup>36</sup>。

表 4-2 EAF が採択したプロジェクト一覧

第一弾として採択されたプロジェクト (2022 年 8 月)

プロジェクト名	交付先	プロジェクト総額
刑務所における薬物乱用への取組	Ministry of Justice, His Majesty's Prison & Probation Service	£1,310,850
司法制度に関する行政のリンクトデータの 順実験的評価への活用	Ministry of Justice Economic and Social Research Council (ESRC)	£429,371
薬物使用に対するダイバージョン制度	The College of Policing	£2,236,892
女性と女兒に対する暴力への介入策	The College of Policing	£2,380,178
対応型の青少年オフエンディングサービス (Young Offending Service) によるロンドン の青少年の再犯防止	The College of Policing	£58,019
健康的で持続可能な食生活プログラム	Department for Environment, Food and Rural Affairs	£6,855,190

<sup>36</sup> EAF ウェブサイト (<https://www.gov.uk/government/publications/evaluation-accelerator-fund>, 2024 年 5 月 28 日最終アクセス)。

子どもの貧困削減のためのローカル・パス ファインダーの経済的・インパクト評価	The Scottish Government	£504,334
親族養育 (Kinship Care) と特別後見制度 (Special Guardianship)の有効性	What Works for Children's Social Care	£347,592
Staying Close (児童養護施設を退所する若 者への支援政策)の全国展開の影響評価	What Works for Children's Social Care	£647,534
介護離職者への現金給付	The Centre for Homelessness Impact (CHI)	£305,510
介護離職者のホームレス状態を減らすため の介入策の評価	The Centre for Homelessness Impact (CHI)	£256,850
家庭内虐待の影響を受けた子どもと家族の 支援	The Early Intervention Foundation	£853,880
Pupil Premium (恵まれない子供たちの教育 レベルを向上させるための資金援助)の 評価	Department for Education	£75,000
Education Endowment Foundation のデータ アーカイブ	Education Endowment Foundation	£256,850
コホートと実験プロジェクト	The Centre for Homelessness Impact (CHI)	£227,620
有用なパラメータ、学習、成果の開発	The Centre for Homelessness Impact (CHI)	£828,375

第二弾として採択されたプロジェクト (2023年3月)

プロジェクト名	交付先	プロジェクト総額
行政サービスの改善の評価	Cabinet Office	£938,000
電気自動車 (EV) の利用状況を評価するため のモバイルデータの探索	Department for Transport	£602,000
心理療法へのアクセス改善が労働市場に与え る影響	Office for National Statistics	£217,000
ウェルビーイングについてのエビデンスの促 進	What Works Centre for Wellbeing	£142,000

第三段として採択されたプロジェクト(2024年3月)

プロジェクト名	交付先	プロジェクト総額
“Fast Stream”のインパクト評価と公務員行政 データの活用	Cabinet Office	£199,950
ディバン作戦が若者のナイフ所持に与えた影	College of Policing	£63,000

響		
モニタリングと評価によるエビデンスの排出 量子測とネット・ゼロ評価への反映	Department for Energy Security and Net Zero	£160,000
効率的なエネルギースキームの評価	Department for Energy Security and Net Zero	£160,000
学生のウェルビーイングとメンタルヘルスを 支援するためのウェルビーイング分析に関 するエビデンスの構築	Transforming Access and Student Outcomes in Higher Education	£330,000
警察における AI と自動化	West Yorkshire Police	£20,000

実際の評価においては、時間とお金をかければかけるほど、より正確な評価ができる<sup>37</sup>という側面があるが、評価結果を次の予算に直ちに反映させようとする、時間が限られるという課題があり、政府における評価についてはこれらのバランスが重要である。

この点、EAF を活用した評価については、①原則として、政府がすでに持っているデータの利用、②プログラム全体を評価するのではなく、プログラムの一部に絞った評価、③1年のファンドで Rapid-RCT という手法を用いた評価の実施、④少数のパイロットテストを行い、規模を縮小したときの違いの検証、⑤将来の頑健な評価を行うための基礎調査として、複数年をかけて RCT を実施したり長期的なデータ整備の実施、といったそれぞれの事業に必要な評価に応じた工夫が行われている<sup>38</sup>。

#### 4.3.4 評価・試行アドバイスパネル<sup>39</sup>

評価・試行アドバイスパネル（Evaluation and Trial Advice Panel, ETAP）は、2015年に内閣府および経済社会研究評議会（Economic and Social Research Council, ESRC）に

<sup>37</sup> 例えば、子どもたちへの教育を評価しようとする場合、成果は何年も経たないと測れないなど、関心のあるアウトカムが遠い未来にしか発現しない。

<sup>38</sup> ETF 長ハッチンソン氏に対するインタビュー。

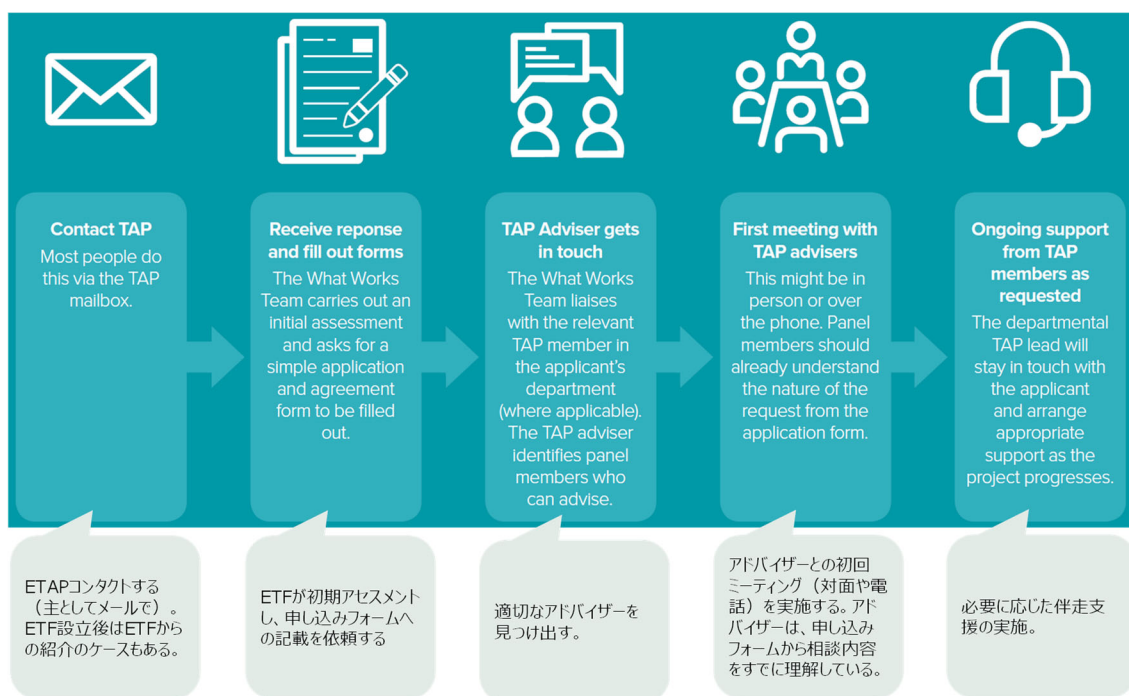
<sup>39</sup> 評価・試行アドバイスパネルウェブサイト

(<https://www.gov.uk/government/publications/cross-government-trial-advice-panel-role-and-membership/>, 2024年5月22日最終アクセス)および What Works Team

(2018) 参照。

よって設立された。当初は政府横断試行アドバイスパネル(Cross-government Trial Advice Panel, TAP) と呼ばれ内閣府の WWT が所管してきたが、同チームが ETF に統合されたことに伴って ETF 所管となり、名称も ETAP に改称された。設立当初の目的は、①行動インサイトの政策への活用、②新しいデジタル技術の発展 (A/B テストが容易に)、③WWN の発足、といった背景によって実験的政府 (Experimental Government) がトレンドとなり、それを後押しすることだった。ETAP には評価や実験に関するさまざまな専門家で 50 人以上で構成されており、公務員や WWC であれば誰でも無料で利用できる。ETAP の利用手順を示したものが図 4-3 である (ただし TAP 当時のもの)。教育、運輸、雇用、犯罪など、さまざまな領域について 170 以上のプロジェクトに助言してきた。

図 4-3 ETAP の利用手順



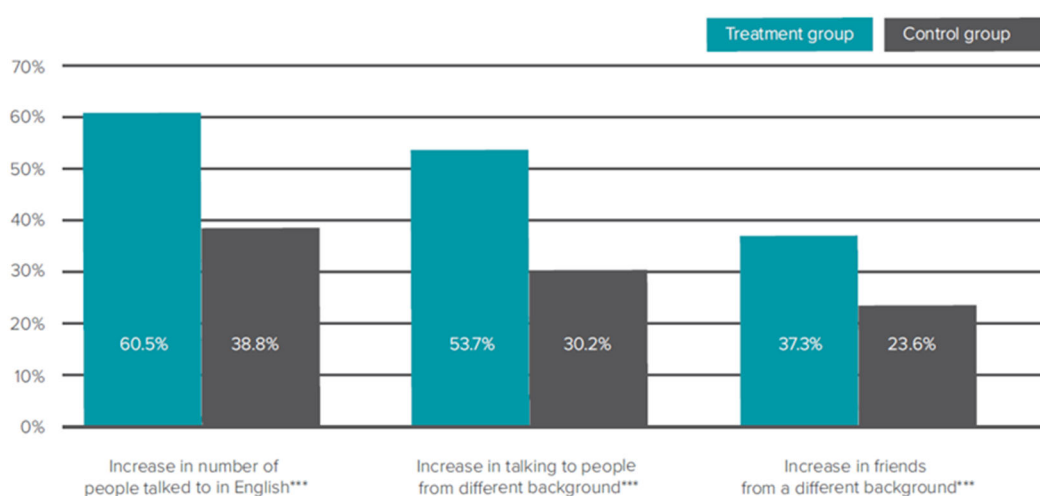
(出所) What Works Team (2018) より作成。

ETAP が支援した具体例の分析結果が図 4-4 である。住宅・コミュニティ・地方政府省 (Ministry of Housing, Communities & Local Government) は、英語学習が、社会的な統合・まとまりを強化するのかの検証を、同省初のウェイトリスト方式のランダム化比較試験として 2016 年に実施した。実験では、英語力の低い 527 人の参加者が、①即時に英語



学習支援を受けられるグループ（処置群）と、実験終了後に支援を受けられるグループ（対照群）にランダムに分けられた。処置群は、英語のライティング・スピーキングスキルが格段に上がっただけでなく、社会的統合のアウトカム（例えば異なる文化の人との友人関係の構築）も改善した。このエビデンスは省の政策立案の核になっているとともに、その後も実験的手法が活用される機運が醸成され、省のフラッグシッププロジェクトである Housing First でもパイロット実験のための予算が 2017 年に確保された。

図 4-4 英語学習の提供による社会経済効果



（出所）What Works Team（2018）より作成。

#### 4.4 ホワット・ワークス・ネットワーク戦略 2023

ETF はかつて内閣府にあった WWT を統合しているが、WWT の役割は WWC の協働や管理を担っている。WWC については家子・小林他（2016）や内山・小林他（2022）で詳述しているが、エビデンスの創出・翻訳・統合等を担う組織である。WWC は政策分野ごとに設置されており、現在 9 つのフルメンバーと 3 つの affiliates の合計 12 の組織で構成されている。WWC は多くの政策領域における政策・実践を改善してきた<sup>40</sup>。

2013 年に WWN が設置され 10 年経過した 2023 年に、「ホワット・ワークス・ネットワーク戦略 2023」（What Works Network Strategy2023）を公表している。5 つの戦略とその

<sup>40</sup> What Works Network（2018）、Sanders and Breckon（2023）など。

具体的内容を整理したものが表 4-3 である。戦略 1 では WWC が政府の重要政策を今まで以上にカバーできるようにすることが盛り込まれており、戦略 2 では政府側からのエビデンスの需要を高めることや政府の政策形成におけるエビデンスの活用や WWC との連携強化が謳われている。また戦略 4 では、期待水準に到達していない WWC の廃止も視野に入っている。これらの部分から感じ取れるのが、WWC を今まで以上に政策形成に近づけたという意図である。WWC の特徴は、政府と連携しながらも一定の距離を確保して、独立性と信頼性を高めていることだった。また、WWC の主なユーザーは、中央政府の政策立案者というよりも、現場の実務家を中心だった。しかしながらホワット・ワークス・ネットワーク戦略 2023 を見ると、WWC の基本骨格は維持しつつも、政策形成への貢献度をより高めたいという狙いが感じられる。

子ども社会ケア WWC(What Works Center for Children's Social Care)の創設者で長く責任者も努めてきたロンドン・キングズ・カレッジ(Kings College London)のマイケル・サンダース(Michael Sanders)教授も、「ここ 10 年で WWC は増えているが、政府の優先課題も時勢によって変わる中、すべての WWC を存続すべきか議論の余地がある」、「新しい戦略は政府の考え方、つまりメインのオーディエンスは中央政府であるべきということが反映されているように思う」と指摘している。同時にサンダース教授は「この戦略はすばらしいとは思いますが、WWC は独立性が高く、内閣府以外の省庁から資金提供されているため、個々の WWC の活動を大きく変えるものにはならないのではないか」と指摘している<sup>41</sup>。

---

<sup>41</sup> 本段落の引用箇所はサンダース教授へのインタビューによる。

表 4-3 ホワット・ワークス・ネットワーク戦略 2023

大項目	内容
<b>戦略1：政府の重要政策をWWC (What Works Centres) がカバーできるようにする。</b>	<ul style="list-style-type: none"> <li>❑ 社会保障など他分野へのエビデンス提供を測る。</li> <li>❑ 地方政府との連携を強化する。</li> <li>❑ 新しいWWCの設置・運営は時間もお金もかかるため、政策ニーズにアジャイルに対応できるようにWhat Works Programme membership categoryを創設する。これは特定の政策プログラムに即応するためのもの。</li> </ul>
<b>戦略2：公共サービスにおける意思決定にWhat Works evidenceをもっと活用する。</b>	<ul style="list-style-type: none"> <li>❑ 良いエビデンスがなければ、将来の意思決定に活用できるエビデンスを生み出す。</li> <li>❑ 政府とWWCの関係を強化すると共に、政策立案者や実務家向けの教育プログラムを実施し、政府内でのエビデンスに対する需要を高める。</li> <li>❑ 政策の優先順位付けへのWWCの貢献を奨励する。</li> </ul>
<b>戦略3：WWC間および国際パートナーとの協働を増やす。</b>	<ul style="list-style-type: none"> <li>❑ WWC間の知識共有を促進し、アプローチの統一化と解決策の模索を図る。</li> <li>❑ 海外の関連機関と協力し、エビデンス活用の促進と国際的なベストプラクティスの導入を図る。</li> </ul>
<b>戦略4：WWCの質を向上させる。</b>	<ul style="list-style-type: none"> <li>❑ WWNには、full、affiliate、associateの3つのメンバーシップがあるが、わかりにくいのでassociateは廃止してfullに移行させる。</li> <li>❑ WWCのパフォーマンスを向上させるため定期的なセルフアセスメントを導入し、期待水準に到達していない場合はWWNから除外することも検討する。</li> </ul>
<b>戦略5：長期的資金とWWCの独立性の重要性を求める。</b>	<ul style="list-style-type: none"> <li>❑ 長期的な資金計画を策定する。</li> <li>❑ 資金調達先を多様化する。</li> <li>❑ 透明性を強化する。</li> </ul>

(出所) Evaluation Task Force (2023) "What Works Network Strategy"

## 5 まとめと日本への示唆

### 5.1 まとめ

以上本稿では、政策評価の結果を政策形成に活用しようとする動きを中心として、近年における英国の EBPM の深化について述べてきた。第 1 に、NAO の報告書に基づいて、英国政府における評価の実施状況等について見た。EBPM の先進国とされている英国においても評価への障壁や課題が多く存在しており、政府はそれらへの対応に注力していることが確認できた。第 2 に、EBPM 体制の改革として、分析ファンクションの設置とマネジメントブックの大幅改訂について見た。様々な分析専門職が横割りのネットワークを形成していることや、政策評価のガイダンスが幅広く共有されていることを確認した。第 3 に、ETF の設立経緯や具体的な取り組みについて見た。ETF の活動はまだ序盤であり、今後の成果が期待されよう。

### 5.2 日本への示唆

最後に、英国の経験から得られた日本の EBPM への示唆について論じたい。理解の便のため、エビデンスの供給面、すなわちエビデンスを構築し提供する側面と、需要面、すなわちエビデンスを利用して政策形成に活用する側面に分けて述べる。

#### 5.2.1 エビデンスの供給面

エビデンスの供給面で重要なのは、良質のエビデンスを構築し、それを需要側に適切な形で提供することである。その点で、第 1 に、ETF には政策評価を専門とする精鋭部隊が配置されており、各省におけるエビデンス供給を支援している。また、WWC や EAF、ETAP など、エビデンス供給力強化のための支援メニューの充実も図っている。日本でもこれを見習い、内閣府（内閣官房）、総務省、財務省の共管で政策評価のエキスパートを集めた司令塔を作ることが考えられよう。

第 2 に英国政府には多数の分析専門職がおり、それらが政府横断的なネットワークー分析ファンクションと各分析専門職グループを形成していることが注目される。前述のとおり、英国ではエコノミストや社会調査職、統計職などが政府内に多数配置されており、

政策分析を担当している<sup>42</sup>。各省庁を横断するとともに政府内外にわたる分析ファクションのネットワークは、分析専門職員の能力向上やキャリア開発に大きな役割を果たしている。日本においても、分析を専門とする職員を登用するとともに、それらがネットワークを形成することで能力向上やキャリア開発を支援する体制を構築することが検討されるべきと思われる。

第3に、マジェンタブックというガイダンスの存在が重要である。こうしたガイダンスが政府内で浸透していることが、政策評価の意義やそれについての知識・スキルの普及に役立っている。日本でも、EBPM・政策評価についての詳細かつ明瞭なガイダンスを政府内に普及させることが望まれよう。

マジェンタブックについては以下の点も注目される。まず、英国では定量的評価やRCTだけでなく、定性的なプロセス評価や費用便益分析も政策評価の柱となっている。日本でも、定量的な手法と定性的な手法を相互補完的に使用するのが望ましいだろう。またマジェンタブックでは、政策デザインの段階で評価を組み入れるべきことが強調されている。日本でも、頑健な分析の根拠となる適切な比較対照群設定とデータ取得のために、政策実施前からの効果検証のデザインや政策執行の過程におけるデータ取得を政策デザインの段階で組み込んでおくことが有効であろう。

### 5.2.2 エビデンスの需要面

エビデンスの需要面については、政策決定者に対して、政策評価を実施してその結果を政策形成に活かすインセンティブを与えることが重要である。第1に、エビデンスの需要面においてもETFの役割が注目される。ETFは新設されて間もないものの、内閣府と財務

---

<sup>42</sup> 例えば、政府エコノミストの総数は2021年時点で3,103名、社会調査職は2,400名超、統計職は9,000名超、分析ファクション全体で約17,000名である（出典はそれぞれ、<https://www.gov.uk/government/publications/ges-strategy-2021-2024>, <https://www.gov.uk/government/organisations/civil-service-government-social-research-profession/about>, <https://analysisfunction.civilservice.gov.uk/government-statistical-service-and-statistician-group/the-gss-government-professions-and-analysis-function/>, いずれも2024年5月26日最終アクセス）。

省をバックとしているため、各省に大きな影響力を持っている。特に英国では、財務省が歳出決定プロセスにおいて大きな影響力を有しているが、財務省を通じて各省に対して既存のエビデンスの利用や適切な評価計画策定を求めることは、エビデンスの需要を高める上で大きな効果を持っている。

第2に、EBPM・政策評価の意義や方法について、政策形成を担うジェネラリスト的な職員にも一定の知識を持ってもらうことが重要である。既述のとおり、マジエンタブックは分析専門職だけではなく、政策専門職や実施専門職の公務員のためにも書かれている。日本でも、職員研修等においてEBPM・政策評価の基本的な考え方や論理を学ぶことを必須とすること等を通じ、知識の共有が図られるべきといえよう。

第3に、NAOの報告書でも強調されているように、評価と予算とのリンクを強化することが政策評価を各省庁に徹底する上で大きな推進力となろう。例えば、予算措置にあたり適切な評価計画の策定・実施を条件づけること、頑健な政策評価を行い、有効性が証明された施策に予算を優先的に配分すること、EAFのように政策評価を行うための資金を潤沢に配分すること等が検討されてもよいのではないか。

### 5.2.3 供給面と需要面の両方に関わる点

エビデンスの供給面と需要面の双方に関わる点としては、第1に、リーダーシップが肝要である。例えばETFの設立には与党重鎮議員のマイケル・ゴーフ大臣の役割が決定的であった。適切な政治的サポートがなければ、どんなに立派な制度を導入しても魂が入らないだろう。そこで日本においても、政官のリーダーが政策評価の体制構築や評価の実施においてリーダーシップを発揮することが望まれる。何よりも、リーダー自らがエビデンスを重視する態度を持つことが重要である。

第2に、英国においては、2013年に公表されたNAOの報告書「政府における評価」(Evaluation in Government)がEBPM・政策評価に関する関心を高めてその体制改革を促す役割を果たした<sup>43</sup>。NAOは議会に附属する機関であり、その意味で官僚制の外部にある。

---

<sup>43</sup> NAOのブラドバーン氏とケリー氏によれば、2013年に発行された報告書の影響は大きなものがあり、ビジネス・エネルギー・産業戦略省（現ビジネス貿易省等）をはじめと

すなわち、いわば「外圧」が一定の効用を発揮したのである。その他、英国では議会の公会計委員会(Public Accounts Committee)や各常任委員会(Select Committees)において、政策決定が的確なエビデンスに基づいたものであったかどうか精査されることが多い。日本でも、会計検査院や議会がEBPM・政策評価をより重視してもよいのではないだろうか。

以上本稿では、近年の英国におけるEBPM・政策評価の深化について明らかにした。本稿が日本のEBPM・政策評価の発展にいささかでも貢献すれば幸いである。

---

しいくつかの省庁は、この報告書に促され評価計画を作成することになった。もっとも大きな影響力は議会の公会計委員会を通じたものである。各省庁は同委員会からの指摘に応答する必要があるからである。

## 参考文献

- Cabinet Office (2015) *Government –guidance What Works Network*
- HM Treasury (2018) *The Green Book Central Government Guidance on Appraisal and Evaluation*
- HM Treasury (20220) *The Magenta Book: Central Government guidance on evaluation*
- IPA (2023) *Infrastructure and Projects Authority Annual Report 2022-23*
- Kingdon, John W. (1984) *Agendas, Alternatives, and Public Policies*, Boston: Little, Brown.
- National Audit Office (2013) *Evaluation in Government*
- National Audit Office (2021) *Evaluating Government Spending*
- What Works Network (2014) *What works? Evidence for decision makers*
- What Works Network (2018) *The What Works Network: Five Years On*
- What Works Team (2018) *The Rise of Experimental Government: Cross-Government Trial Advice Panel Update Report*
- 家子直幸・小林庸平・松岡夏子・西尾真治 (2016) 「エビデンスで変わる政策形成 ～イギリスにおける「エビデンスに基づく政策」の動向、ランダム化比較試験による実証、及び日本への示唆～」『MURC 政策研究レポート』
- 内山融・小林庸平・田口壮輔・小池孝英 (2022) 「英国における EBPM と日本への示唆」大竹文雄・内山融・小林庸平編著『EBPM エビデンスに基づく政策形成の導入と実践』日本経済新聞出版
- 内山融・藤田由紀子 (2023) 「政治主導と官僚制の行方 英、官僚の中立性を守る工夫」2023年10月3日付日本経済新聞朝刊「経済教室」

## 略語一覧

- EAF Evaluation Accelerator Fund 評価加速化ファンド
- EBPM evidence-based policy-making エビデンスに基づいた政策形成
- ETAP Evaluation and Trial Advice Panel 評価・試行アドバイスパネル



ETF Evaluation Task Force 政策評価タスクフォース  
IPA Infrastructure and Project Authority インフラ・プロジェクト庁  
NAO National Audit Office 英国会計検査院  
WWC What Works Centres ホワット・ワークス・センター  
WWN What Works Network ホワット・ワークス・ネットワーク  
WWT What Works Team ホワット・ワークス・チーム