

評価基準：改訂版定義及び使用原則

1. 背景

1.1 改訂の根拠

1. OECD-DAC は、1991 年に「OECD-DAC 開発援助評価の基本原則」で、評価基準の 5 項目（妥当性、有効性、効率性、インパクト、持続性）を初めて示し、2002 年には、「評価及び結果重視型マネジメントにおける重要用語集」において、それらの定義を提示した。以降、評価 5 項目は、国際開発及び人道支援のプロジェクト、プログラム、政策を評価する際の重要な拠り所として機能している。評価者や評価委託者は、開発協力のみならず、公共政策の他の分野においても評価 5 項目を活用している。

2. 過去 25 年に亘って評価 5 項目の利用から得られた学びを踏まえ、持続可能な開発のための 2030 アジェンダ（持続可能な開発目標（SDGs）を含む）や気候変動に関する国際連合枠組条約締約国会議で採択されたパリ協定に沿って、国際的な評価コミュニティは評価 5 項目の見直しを開始した。2017 年 10 月 31 日のハイレベル公式声明において、OECD-DAC は、2030 アジェンダに沿って評価 5 項目の改訂について調査を行うことを決定した。評価への支援、政策決定と学習のためのエビデンスを構築する DAC 下部組織 EvalNet（開発評価ネットワーク）が、この調査を 2018 年から 2019 年にかけて実施した。

3. 今回の改訂プロセスが目指したのは、評価基準を改善するための、言わば「学びと経験の棚卸し」であった。より良い評価基準は、より良い評価につながり、より良い評価は、2030 アジェンダを進め、パリ協定や他の目標達成を目指すためのより良い政策につながる。

1.2 評価基準に関するコンサルテーション：強みと改善の余地

4. 改訂プロセス（2018 年 5 月～11 月）では、世論調査やインタビュー、国際会議での協議、文献調査などにより、広範なコンサルテーションが行われた。初案をベースに、コンサルテーションは 2019 年後半まで続いた。コンサルテーションで得られた情報は公開されている¹。

5. コンサルテーションでは、回答者たちは、評価基準が広く受け入れられ、活用されていることを踏まえて、評価基準の本体部分を大きく変えないことを望んでいることが明らかとなった。多くの回答者は、評価の専門性と評価の実践で標準化と一貫性をもたらした評価基準の重要性を指摘し、評価基準はあまり増やさず、定義の一貫性を保ちながら、引き続き簡潔さを保つことを求めている。

¹ EvalNet（2018）「OECD-DAC 評価基準：コンサルテーション回答要約（OECD-DAC Evaluation Criteria: Summary of Consultation Responses）（2018 年 11 月）」、oe.cd/criteria から入手可能。

6. 同時に、いくつかの概念について、その明確化を求める要請も寄せられた。要請の多くは、基準を実際に適用する場合の難しさを指摘している。特に問題なのは、あまりにも多くの基準の適用と問いかけをしている傾向である。「開発評価の品質基準」では、すべての基準を用いる必要はないし、その他の基準を用いてもよいと明示はしているが、実際には、評価の文脈や意図された目的について十分考慮なく機械的に適用されてしまう可能性もある。また、当初の一連の基準は、2030 アジェンダの文脈と現在の政策的優先事項に適切に対応していなかったという懸念もある。これらの基準が、プロジェクトに焦点を当て過ぎており、複雑性やトレードオフ、公平性、人権とジェンダー平等への配慮といった諸問題には十分に対処してこなかったとしている者もいる。多くの人々が、基準の利用を改善し評価の質の向上に資するために、評価基準の実施についての手引きを一層強化することを求めている。

7. さらに、今回の検討では、基準や定義の問題以外にも、評価実践と評価システムに多くの欠陥があることが明らかになった。これらの懸念に対処するために、EvalNet はよりよい評価を支援するという使命に沿って、現在グローバルな評価コミュニティと連携・協働し、今後の作業のための選択肢を追求している。

1.3. 採用された項目の重要な特徴

8. OECD・DAC ハイレベル・コミュニケの要請に応じ、上記の検討の結果得られた知見に対処するため、EvalNet は、改訂済みの一連の定義と、基準の使用に関する原則を策定した。EvalNet のメンバーとパートナーはそれぞれの草案についてコメントし、その後、定義についての、詳細な意見交換ができるウェブセミナーを開催した。また、国際的な評価専門家を招いて、これらの草案に対するコメントをもらった。

9. これらのインプットすべてに基づいて、改訂された諸基準は以下の特徴を有している。

- オリジナルの5つの基準の、新しく改訂した定義： 概念の明晰さという重要な強みを維持し強化するために、説明をできる限り簡潔に保ちつつ、定義を改善し、概念を説明するために注釈を用いる。
- つながり、システム思考、パートナーシップの動き、複雑性をよりよく捉えるための重要な新たな基準の追加： 整合性
- 利用を促進し混乱に対処するために、以下を追加： 基準の意図する狙いについて紹介する説明、使用原則と、それぞれの基準の利用局面とこれらを様々な評価にどのように適用するかをさらに説明するための手引き（近刊）
- 多種多様な介入と、プロジェクト以外への応用可能性の確保： 評価の対象となる、開発や人道的活動と手段が多様な範囲にわたることと、基準が開発協力

の場面以外でも用いられることに配慮し、(たとえば、従前の外部資金、プログラム/プロジェクトといった表現よりもむしろ)「介入」という用語を用いる。「ドナー」という表現も同様に削除された。

- 公平性、ジェンダー平等、「誰一人取り残さない」アジェンダなどの、現行の政策の優先事項により適切に対応： とりわけ、妥当性と有効性の定義では、公平性にかかる問題についてより詳細に分析することを奨励する。新しい評価基準は、2030 アジェンダに盛り込まれた持続可能な開発目標とパリ協定を達成することを目的とした、(国、国内地域、国家間のいずれのレベルでの) 取り組みの評価にも有用である。同時に、これらの基準は包括的なもので、広く応用可能であり、政策の優先順位や目標が変化しても引き続き適切である。
- 持続可能な開発の有する統合的な特質を反映した定義と利用手引きにより、新評価基準は今後相乗効果やトレードオフを含む相互連関的アプローチを促進する。

2. 改訂された評価基準

2.1. 評価基準の目的

10. 評価基準の目的は、評価の目的と関連している。すなわち、介入の長所、価値ないし重要性を確定できるようにすることである²。「介入」という用語は、本文書を通じて、評価の対象を意味する(ボックス1参照)。基準一つ一つは、介入をそれぞれ異なるレンズないし視角でみることができ、評価基準全体としては、介入、実施プロセス、結果のより包括的な姿を提供することになる。

11. これらの基準は規範としての役割を果たす。全体として、これらの基準は、介入の望ましい特徴を記述している。すなわち、すべての介入は文脈に適切でなければならず、その他の介入と整合し、目的を達成し、効率的に結果を出し、持続的な正のインパクトを持たなければならないという特徴である。

12. これらの基準は、評価において、以下のために用いられる。

- 一般への情報公開を含む、説明責任の確保
- 知見と教訓の創出とフィードバックを通じた、学習の支援

13. これらの基準は、モニタリングや結果重視マネジメント、また、戦略的計画や介入デザインなど、評価以外にも用いられる。

14. これらの基準は、結果(何が変化したか)だけでなく、プロセス(どのように変化が起きるのか)に着目するためにも用いられる。すべての基準は、介入前、介入中、介入後の評価に用いられる³。

² OECD DAC (1991) の『開発援助の評価原則』に述べられている価値と重要性を参照。

³ しかしながら、これらの基準の定義は主に、中間評価、最終評価、事後評価での実践を反映し

ボックス 1. 介入を評価する

我々は、介入という用語を、本文書において、評価の対象を指すものとして用いる。介入は、プロジェクト、プログラム、政策、戦略、テーマ分野、技術協力、政策助言、機関、拠出メカニズム、手段、その他の活動など、すべての種類の開発及び人道的努力を包含する。介入には、開発介入、人道的援助、平和構築、気候緩和・適応、規範的活動、政府保証のない活動などが含まれる。

これらの基準は、国際協力活動だけでなく、民間セクター、非政府アクター、国内の政策分野における中央・地方政府による介入の評価にも用いることができる。

2.2. 使用原則

15. 評価基準の使用原則を以下に記すが、詳細な活用事例や方法は近刊のガイドンスを参照されたい。また、OECD DACの「開発評価の品質標準」に評価の計画や実施に関する標準が記載されているが、質の高い有益な評価を行うには、他の標準やガイドラインを参考にしつつ、定義などをより幅広く解釈することが肝要である。

1. 評価基準は、質が高く、有益な評価のために細心の注意を払って使用されるべきである。評価基準は、それぞれの評価の性質、評価対象や関係者など評価の文脈に沿って理解されるべきである。評価のための質問（何を調べようとしているのか）やその回答の使用方法には、評価基準がどのように解釈・適用されているかを明示する必要がある。
2. 評価基準は、評価の目的に沿って活用されるべきである。評価基準を機械的に適用するのではなく、関係者のニーズや評価の性質に応じて適用すべきである。評価目的に従って、基準ごとに行われる評価分析の多寡は左右されうる。ある評価基準がどのように使用されるか（または全く使用されないか）は、データの可用性、評価資源の制約、タイミングや方法論などにも左右される⁴。

2.3. 定義と変更点の説明

16. 次のセクションで、各評価基準を定義する。定義をするにあたり、概念を明確にするために注釈を加えている。ボックス内に、「評価および結果重視マネジメントにおける基本用語集(OECD, 2002)」に記載された定義の変更点を説明している。用語集の第2版を準備中であり、本書で使用した介入、結果、アウトプット、アウトカム、目標などを含む多くの用語を定義するので参考にしていきたい。

ている。

⁴ 評価を行う前に評価可能性を検討すると、どのような情報が得られて、どのようなエビデンスが集められて、どのような結果が得られるのかに対して現実的な予測をすることができる。

妥当性：介入は正しいことをしているか？

介入※の目的やデザインが、受益者・世界・国・パートナー／機関のニーズ、政策、優先事項に対応し、状況の変化に応じて対応し続ける度合い

注釈：「対応」とは、介入の目的やデザインが対象地域の経済、環境、公平性、社会、政治経済（ポリティカルエコノミー）、および能力状況に対応していることである。「パートナー／機関」とは、国や地域・地方政府、市民社会、民間団体、および国際機関のことで、資金提供や介入の実施や管理を行う組織を含む。妥当性の評価は、異なる優先順位やニーズ間の差異やトレードオフの考察を含む。妥当性の評価においては、介入が妥当であり続けられる（又はあり続けた）度合いについて、あらゆる変化を分析する必要がある。

※【訳注】「介入（Intervention）」は、これまでの旧定義では「開発援助」と訳されてきた。本文書では、「開発援助・協力、平和構築・人道支援、気候変動に関する取組、民間の協力等」を指す。

ボックス 2. 妥当性の定義に関する変更点

まず、介入の妥当性に質的な要素を入れるために「デザイン」を加えた。例えば、「その介入は、関連する優先事項やニーズ、目的、セオリー・オブ・チェンジ、セオリー・オブ・アクション、運用方法、リスク分析、文脈解析などに対して、適正にデザインされていたかどうか？」などである。これは、妥当性の分析を深めるため、従来通り目標の達成度は有効性で評価し、どのように介入が実施されたかは効率性の下で評価する。

ここでは、より深い状況分析を推進するために、介入が配慮すべき要素を説明する。

何が重要で緊急性が高いのかを政府やドナーだけが決定するのではなく、影響を受ける人々の決定も重要であることを明示するために「受益者の優先事項」を加えた。受益者を重視したのは、現在の「誰一人取り残さない」という新たな国際政策を反映したからである。「ニーズ」という言葉を使っているのは、時々法的条件と間違えられる「必要条件」という言葉に比べて解りやすいためである。

新評価基準に関するコンサルテーションの中で、「受益者」という言葉に懸念を示す意見も出された。なぜなら、「受益者」とは、受動的に利益を受ける役割、あるいは便益を実際には受けていないのに便益を受けているかのような印象を与えるからである。しかしながら、この用語を引き続き使ったのは、対象となる人々を明確にし、誰にとってその介入が妥当であるべきかを評価できるようにするためである。また、混乱を避けるために、脚注に説明を加え、代替案も提示している。

2030 開発アジェンダやパリ協定に代表される「国際的な優先事項」という文言も残した。しかしながら、優先事項は変わる可能性があり、妥当性の評価基準はさまざま

まな評価対象に用いられるため、定義の中に固有名詞を明記していない。

他方、以前は注記であった時間軸を明記することにした。なぜなら、妥当性は介入の計画時だけではなく、その後も評価できるからである。複雑で常に変化している状況に適切に対応することは、有効性を高めるために重要である。これまでは、妥当性は、政策の政策上の優先事項に対する一度限りの評価として扱われることが多かった。

開発協力分野以外での評価基準の活用の実態をよりよく反映させるため、「ドナー」の単語を削除し、注釈に「パートナー／機関」とは資金提供、実施、管理を行う協力者／組織（国内、国際機関、民間など）を含むとの説明を追加している。

参加とオーナーシップに関する概念は妥当性に影響する要因であるため、評価基準の利用局面の説明とは異なり、妥当性の定義には含めていない（効果、持続性についても同様）。これらの概念はガイダンスにおいて議論する予定である。

用語集における妥当性の定義：「開発援助の目標が、受益者の要望、対象国のニーズ、地球規模の優先課題及び援助関係者とドナーの政策と整合しているかどうかの程度を示す。注：結果として、妥当性に関わる問題としては、状況が変化した後でも、援助の目標あるいは計画そのものが依然として適切であるかどうかという点を問われることが多い。」

整合性：介入が良く適合しているか？

国、セクター、組織に対する当該介入と他の介入との適合性

注釈：他の介入（特に政策）による当該介入への支援あるいは阻害の度合い、あるいはその逆の関係。内部的整合性と外部的整合性がある。このうち**内部的整合性**では、同じ組織／政府が行う当該介入とその他の介入の相乗効果と相互連関、及び組織／政府が従う国際規範や基準と当該介入との一貫性を扱う。**外部的整合性**では、同じ文脈のもとでの他者による介入との整合性を考慮する。それは他介入との補完、調和・調整、及び努力の重複を回避した当該介入の付加価値の度合いを含む。

ボックス 3. 整合性追加の説明とその定義

この基準はこれまでの基準では包含されていなかった観点を明らかにする。整合性という基準の欠如は活動の重複をもたらし、全体の進捗を妨げる。整合性を加えることは、これらの重要な問題の分析の質を高めることに役立つ。整合性を含めることにより、排他的にその介入や組織を中心に置く視点とは対照的に、評価者が特定のシステム（組織、セクター、課題、国）内での当該介入の役割を理解しようとする動機づけにもなる。

今日の状況では、より多くの注意を整合性に向ける必要があり、その際には政策領

域と政府をまたぐ協調に注目が高まっている分野、とりわけ紛争下や人道上の対応における相乗効果（あるいはトレードオフ）に一層焦点を当てたり、気候非常事態にも対処したりする必要がある。これに加えて、持続可能な開発のための（国際及び国内）資金が多様化していることにも留意が必要である。

「国際的規範・基準」を参照することにより、汚職行為防止法、人権条約といった国際法や国際的合意に対する政府自らの公約と介入との一貫性についても分析を促すことになる。また内部的整合性により、国家がすでに実施を約束している合意についても検討できる。以前はこのタイプの整合性についての分析は、必ずしも十分ではなかった。国際規範と基準については、グローバルな優先事項への対応という視点に立って妥当性の基準でも評価できるが、それはあくまでも補完的なものである。

有効性：介入は目的を達成しているか？

介入の目的と結果の達成又は達成見込みの度合い（諸集団の異なる帰結を含む）

注釈：有効性の分析は、短期的目的や直接的結果の相対的な重要性の考慮を含む。

ボックス 4. 有効性に関する定義変更の説明

有効性では結果の連鎖／因果関係の道筋に沿って、目標達成に向けた進捗を分析する。有効性は、介入のより上位の効果と広範な変化に対する貢献を見るインパクトとは対照的に、介入がより直接的に寄与した結果に関係する。

異なるグループへの結果の配分に関して重要な質問を投げかける契機となるよう、また意図した目的以外の影響を把握できるよう「結果」と「差異のある結果」を追加した。これは誰一人取り残さないための政策的優先事項に沿ったものである。これにより、公平性の達成が介入の目的ではない場合でも、評価者が公平性の問題や疎外されているグループに及ぼす結果について検証するよう促すことになる。

この定義では、最も関心の高い目的及び／あるいは結果に評価の焦点を当てられるよう、柔軟性を認める表現を用いている。

全体的な有効性についての結論を導く際には、達成された事柄／達成されなかった事柄／期待される目的と結果に関する重要性を比較することが必要となるため、注釈に「相対的な重要性」への言及を残した。一方、妥当性では、主要な利害関係者に対する介入の優先度合いを検証し、介入効果の全体的な重要性についてはインパクトで扱う。

用語集における有効性の定義：「開発援助の目標が実際に達成された、あるいはこれから達成されると見込まれる度合いのことであり、目標の相対的な重要度も勘案しながら判断する。注：

「有効性」は、ある活動の利点または価値を総合的に測る尺度（もしくは判断）としても用いられる。すなわち、当該援助が持続的な方法で制度の発展にプラスのインパクトをもたらすと共に、主要な目標を効率的に達成した度合いを示す。」

効率性：資源がいかにうまく使われたか？

経済的かつタイムリーな方法で結果を生む又は生むような介入実施の度合い

注釈：「経済的」とは、同じ条件下で実現可能な代替手段との比較で、インプット（資金、専門技術、天然資源、時間など）からアウトプット、アウトカム、インパクトへの変換で費用対効果が最大限となる方法を指す。「タイムリー」は、計画期間内又は状況変化へ合理的に調整した期間内での実施を指す。これは、実施運営の効率性の評価を含むこともある（介入がいかに上手く管理されたか）。

ボックス 5. 効率性に関する定義変更の説明

評価の好事例に倣い、効率性は結果の連鎖（アウトプット、アウトカム、インパクト）全体に関連したインプットを見ることでも検証できるとした。結果の連鎖全体の分析、特にインパクトに対するインプットの効率性の検証は手法としての難易度が高い。ベンチマークとなるデータはいつも存在しているとは限らない。にもかかわらず、ここでインパクトへの言及を残しているのは、この分野への関心が高いためである。こうして、評価基準のユーザーが、効率性の分析において結果の連鎖部分のうち当該評価との関係性が最も高い部分に焦点を当てて評価する柔軟性が認められている。

「同じ条件下で実現可能な代替手段の比較」の追加により、有効な費用対効果分析は実際に実施した介入の価値と介入がなかったと仮定した状況との比較を必要とするという考えを補強するものである。すなわち、最も安価な介入は、適切な便益をもたらさない場合には、必ずしも最良の選択肢ではないということである。

「タイムリーさ」を追加したのは、この重要な観点が、往々欠落することがあるからである。

「実施効率性」について明示的に言及したのは、実施過程と管理上の課題が評価関係者の最大の関心事であるからである。この視点は、例えば調整メカニズムが整合性を高めた一方で取引コストも増えていないかといった事例を検証する際に役立つ。

用語集における効率性の定義：「資源及び（または）インプット（投入）（資金、専門技術（知識）、時間など）がいかに経済的に結果を生み出したかを示す尺度。」

インパクト：介入はどのような違いをもたらすか？

介入により生じた又は生じると予期される、重要な正又は負の、意図された又は意図されない、高次の効果の度合い

注釈：インパクトは、介入の最終的な意義と潜在的に変容しうる効果を扱うものである。インパクトは有効性の基準で既に捕捉されているものよりも、長期的または広範囲で、介入の社会的、環境的、経済的な効果の特定を求めるもの。短期的な結果を超えて、この基準は、間接的、二次的、潜在的な介入の結果の把握を求める。それは、システムや規範の全体的かつ永続的な変化、及び人々の幸福、人権、ジェンダー平等、環境への潜在的な影響の調査によるものである。

ボックス 6. インパクトの定義の変更の説明

この注釈は、効果の重要性、範囲、および変革的性質を捉えるために「高いレベル」を導入したことを説明している。これは、2030 アジェンダとパリ協定が革新的変化を要求している文脈において政策的な妥当性がある。

この定義は、意味のある、または重要な変化を表す「インパクト」という言葉の一般的な使用法と一致している。

「インパクト」という用語は、一部のユーザーが採用する有効な帰属の概念と混同される場合があることを認識している（「インパクト評価」で一般的に使用されているように）。しかし、有効な帰属はすべての基準に織り込み済みであると考えられる。

直接的および一次的効果は有効性に含まれるため、「直接的・間接的」および「一次的・二次的」を削除する。この注釈はさらに、有効性とインパクトの違いを明確にするのに役立つ。

用語集におけるインパクトの定義：「正および負、一次および二次、開発介入で生じた長期的効果、及び直接的または間接的、意図的または意図的でない効果」

持続性：便益は持続するか？

介入の純便益が継続する又は継続する可能性の度合い

注釈：長期的に純便益を維持するために必要なシステムの財政的、経済的、社会的、環境的、制度的能力の調査を含む。レジリエンス、リスク、潜在的なトレードオフの分析を含む。評価タイミングに応じ、純便益の実際のフローを分析したり、中長期にわたる純便益の継続見込みの計算を含め得る。

ボックス 7. 持続性の定義に対する変更の説明

定義は簡潔であり、注釈は、持続性がさまざまな側面（財務、経済、社会、環境）を持っていることを明らかにしている。

従前の定義はドナー中心であり、外部資金（主要な開発援助）のみに焦点を当てていたため、今日の開発の評価の文脈を十分に反映していなかった。加えて、資金支出や活動を実施中においても持続性の評価を行うことは有用だと思われる。外部資金への言及を削除することは、さまざまな状況での持続性の評価の助けとなる。

また、潜在的なトレードオフの分析、および便益の継続の基礎となる能力／システムの回復力の分析も促す。これとは対照的に、以前の定義には「継続的な長期的利益」と「リスクに対するレジリエンス」といった表現が重複していた。

介入に関連する継続的なコストを考慮して、介入の全体的な価値の継続に焦点を当てるために、「純」便益という用語を残す。

用語集における持続性の定義：「主要な開発援助が完了した後の開発介入からの便益の継続。長期的な便益の継続の可能性。純便益のリスクに対する回復力は時間とともに変化する。」