

# エビデンスと政策の間 — 評価利用の研究が示すもの —

国立国会図書館  
田辺智子

第2回 エビデンスに基づく実践と政策セミナー 2018.6.30

# 本日の発表内容

---

- エビデンス活用が進んでいる領域でも、エビデンスと実践・政策の間にはギャップが存在する。その実態や得られる示唆等を紹介する。
  - 1 医療の場合(EBM)
  - 2 社会政策の場合(プログラム評価)
- あわせて、日本のEBPM議論におけるエビデンスの理解について整理する。
  - 3 日本におけるエビデンスの理解

# 1 医療の場合 (EBM)

# 医療におけるエビデンス利用

---

- 医療領域ではEBMが普及しているから、医療機関で受ける治療はエビデンスに基づいているはず？

⇒必ずしもそうとは限らない。

# エビデンス-診療ギャップ

---

- アメリカの2003年の論文 (McGlynn et. al. 2003)  
2万人の患者を対象に、診療ガイドラインと実際の診療内容を比較。

⇒ 診療ガイドライン遵守率 : 55%

乳がん : 75.7%

高脂血症 : 48.6%

大腿骨頸部骨折 : 22.8%

# “医療の質”への関心が高まる

---

## ■ “医療の質”とは ...

「個人や集団を対象に行われる医療が、望ましい健康状態をもたらす可能性の高さ、その時々<sup>の</sup>専門知識に合致している度合い」

= エビデンスに則った医療をどのくらい行っているか、が重要な要素。

(福井 2017, p.16)

# 医療の質の評価

---

- 医療の質はQI (Quality Indicator)で評価。

ストラクチャー(構造): 施設・スタッフ等の資源

プロセス(過程): 手技・投薬等の医療内容

アウトカム(結果): 生存率、QOL等の患者の健康状態

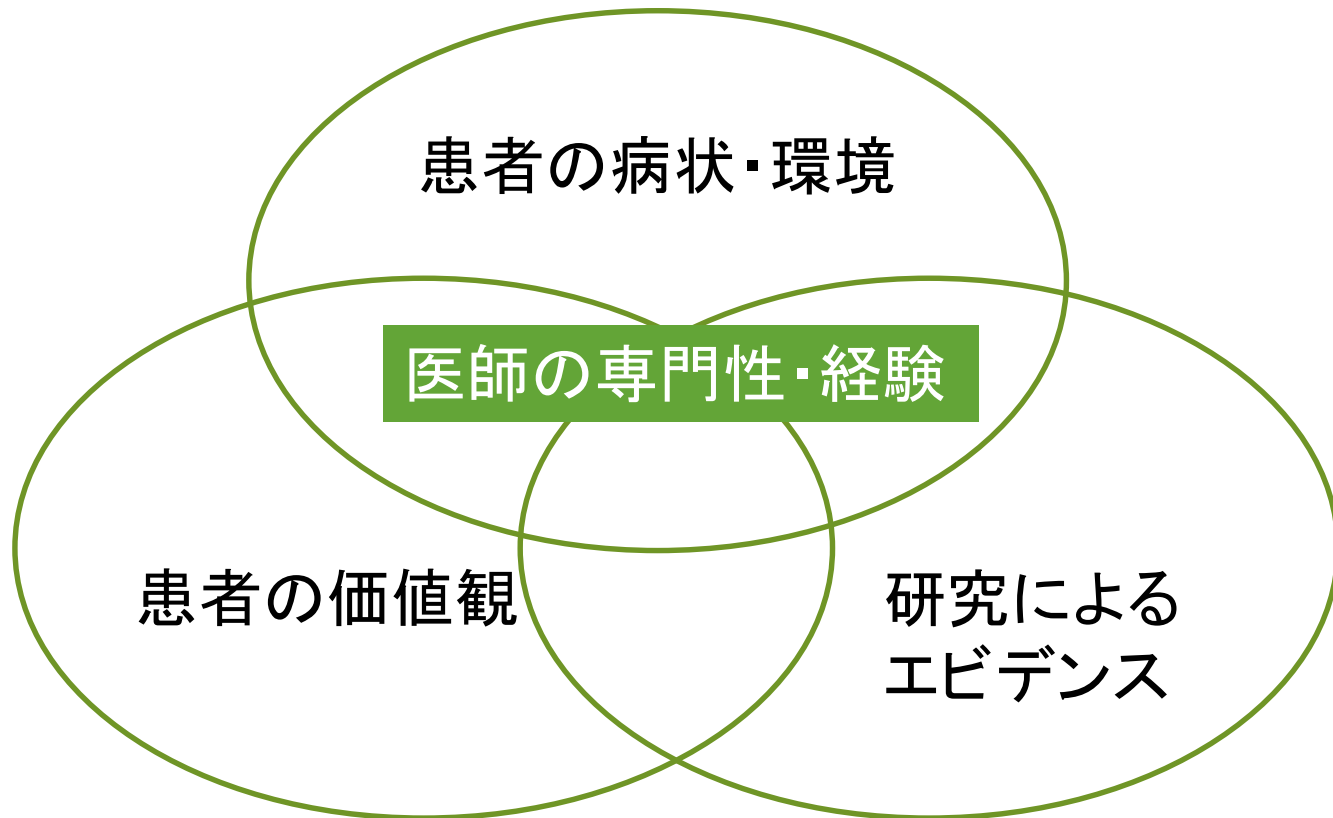
- 諸外国ではプロセス指標が基本。アウトカム指標を用いる場合は、リスク調整が必要。

＜プロセス指標の例＞

- ・心筋梗塞患者のうち、アスピリンを処方された割合
- ・肺炎患者のうち、来院後4時間以内に抗菌薬を投与された割合

# (参考) EBMにおける意思決定

※エビデンスだけで機械的に意思決定するわけではない。





# 医療の質の改善事例

---

## ■ 米国マサチューセッツ総合病院

- ・PDCAサイクルで医療の質を改善。
- ・院内感染予防のための医療スタッフの手指消毒の遵守率 2003年:8% ⇒ 2009年:90%

## ■ 国立病院機構

- ・系列の143病院のQIを測定・公表。
- ・急性脳梗塞で入院後4日以内にリハビリテーションを開始した割合  
2010年:20~90% ⇒ 2013年:大部分で70%以上

# 医療の質の改善策

## ① 医師を対象とした取組

---

### ■ 診療ガイドラインの配布

- ・ガイドライン遵守率はわずかに向上。

### ■ 生涯学習

- ・単純な講習会より参加型が有効。
- ・オピニオンリーダーへの講習が有効。
- ・大きな効果を得るためには資源投入が必要。

(東 2007)

# 医療の質の改善策

## ②医療機関の組織的取組

---

### ■ クリニカル・オーディット（質のモニタリングと医師へのフィードバック）

- ・コクラン・レビューで有効性が示されている。
- ・院内感染予防のための手指消毒、糖尿病患者の血糖管理等の成功事例あり。

### ■ PDCAサイクル

- ・成功事例の報告多数。救急外来待ち時間短縮、中心静脈カテーテル関連血流感染減少等。

（小松・門田・福井 2016）

# 医療の質の改善策

## ③政策的取組

---

### ■ 医療の質の公表

- ・患者の医療機関の選択にはあまり使われない。
- ・医療機関の改善活動には結びつく。

### ■ 業績に基づく支払 (P4P: Pay for Performance)

- ・医療の質向上につながったという十分なエビデンスはない。
- ・指標の数値自体は向上するが、プロセス指標と比べるアウトカム指標の改善は限定的。

(田村 2018; 津川 2016; 東 2007等)

# 医療政策のエビデンス

---

- 個人を対象とした医療（EBM）と比較し、集団を対象とした医療政策は、エビデンスの活用が遅れている。
- エビデンスが不十分なまま実施されている医療政策の例：
  - ・ 特定健診（田辺 2016a）
  - ・ がん検診（田辺2016b）

# 日本の現状

---

- 全般的な医療の質は高いが、医療機関・地域による質のばらつきが大きいと言われる。
- 医療の質評価や改善活動も医療機関によって温度差あり。
- OECDの日本への勧告（OECD 2014）
  - ・日本の医療は比較的低コストで良好な健康状態を実現しているが、今後は医療の質向上のための制度的・体系的な取組が必要。

## 2 社会政策の場合 (プログラム評価)

# 政策評価法における評価方式

評価方式	対応する手法	性格
実績評価	業績測定 (Performance Measurement)	アウトプット、アウトカム等について、あらかじめ目標・指標を定めて達成度合いを評価。
総合評価	プログラム評価 (Program Evaluation)	施策の有効性を様々な角度から分析。施策の問題点、目的の妥当性等を分析する場合もあり。
事業評価	政策分析 (Policy Analysis)	施策の費用対効果等を分析。事前・事後いずれでも可能。



# 米国におけるプログラム評価

---

- 1960年代の米国では、「偉大な社会」「貧困との戦い」というスローガンのもと、保健・教育・住宅等の分野で多数の社会施策を導入。
- 効果検証のためにプログラム評価を実施。RCT等による厳密な評価が行われた。
- その後、プログラム評価は一つの評価方式として定着。評価を実施するコンサルタントやシンクタンクが成長し、一つの産業を形成。

# (参考) プログラム評価から業績測定へ

---

- RCT等による厳密な評価には長期間を要し、また分厚く専門的な報告書は、必ずしも政策現場で使いやすいものではなかった。



- 実用重視の評価が指向されるように。
- その流れの中で業績測定 (Performance Measurement) が誕生。(指標を用いた簡便な評価)
- 日本ではプログラム評価は定着せず、1990年代以降に導入された業績測定が主流に。

# 評価の利用・影響の理論①

- 評価研究の分野では、評価の“利用”や“影響”に焦点を当てた研究が発展。
  - 1970年～：評価利用 (evaluation use)
  - 2000年～：評価影響 (evaluation influence)
- 当初のプログラム評価が、必ずしも評価関係者が意図したように利用されなかったことが背景。
- 代表的レビュー論文：Herbert (2014)  
Johnson et al. (2009)  
Shulha and Cousins (1997)  
Cousins and Leithwood (1986)

# 評価の利用・影響の理論②

---

## ■ 評価利用の類型

「道具的利用」「概念的利用」「象徴的利用」等。

※意思決定への直接的な利用だけでなく、関係者の認識変化につながる等の間接的な利用もある。

## ■ 評価利用の決定要因

「評価の質」「施策のコンテキスト」「ステイクホルダーの評価への参加」等。

# 評価の利用・影響の理論③

評価利用の類型	性格
道具的利用 (instrumental use)	施策の改廃・見直しなどの特定の意思決定に、評価が直接的に反映される。
概念的利用 (conceptual use)	直接的な行動とは結び付けられないが、評価が関係者の認識に浸透し、意思決定の情報源の一つとなる。
象徴的利用 (symbolic use)	既存の立場の正当化や、組織の合理性を示すために評価が用いられる。

※評価の「結果」のみでなく、「プロセス」も利用対象となる。(プロセス利用)

※制度的な要求のために評価が利用される場合がある。(強制的利用)

# 評価の利用・影響の理論④

---

- 2000年代に入ると、それまでの評価利用の概念では、評価の持つ多様な影響を包括的に捉えることができないという認識のもと、“評価影響”という概念が提起される。
- 長期的・間接的な影響を含め、評価のインパクトを可能な限り広く捉えようとするもの。

# 評価の利用・影響の理論からの示唆

---

- 評価の“利用”を狭く捉えすぎると、評価の重要な機能を見落とすことになる。
- 政策過程の現実：施策は評価だけで決まるものではない。
  - ・評価は意思決定に用いられる多くの情報の一つ。
  - ・様々なステイクホルダーの存在。
  - ・既存施策の存在。

※しかし、評価にまったく影響力がないわけではない。

# GAOによるプログラム評価の利用状況に関する調査①

---

## ＜連邦政府高官へのアンケート調査＞

- 担当施策について、過去5年以内に完了したプログラム評価がある：37%
- プログラム評価の利用目的：
  - マネジメント・業績改善のための変更の実施
  - 施策の有効性・価値の評価
  - 施策についての理解の深化
  - 業績についての補足・説明
  - 有効性・その他の教訓の他者との共有
  - 業績目標の設定・改訂 等々



# GAOによるプログラム評価の利用状況に関する調査②

---

## <評価に積極的な省庁へのインタビュー調査>

- プログラム評価の利用促進策
  - ・評価を後押しするリーダーシップ
  - ・強固なエビデンスの確立
  - ・評価過程へのステイクホルダーの関与
- 単一の評価で施策を変更することは通常ない。複数の評価が時間をかけて施策に影響する。
- RCTでなくてもよいが、評価の質が重要。

※以上、GAO 2013より（田辺 2014 で詳細を紹介）

# 3 日本におけるエビデンスの理解

# 政府の取組

---

- 2016.12 経済財政諮問会議「統計改革の基本方針」
  - ・統計改革推進会議の設置
- 2017.5 統計改革推進会議「最終取りまとめ」
  - ・政府内のEBPM推進体制の整備
  - ・政策・施策・事務事業の各段階での取組
- 2017.6 「経済財政運営と改革の基本方針 2017」
  - ・EBPM推進体制の構築
  - ・予算編成へのEBPM的視点の反映。

# EBPMの取組に係る動き

## EBPM推進体制・活動内容の検討

## 行政事業レビュー等、EBPM三本の矢の取組

【29年8月】

各府省のEBPM統括責任者等からなる  
EBPM推進委員会発足【3ページ参照】

【9月～11月】

各府省が「EBPM事例の作り込み」を実践し、  
行革事務局及び有識者がヒアリング  
⇒ヒアリング等から得られた示唆をとりまとめ、  
各府省に共有【4ページ参照】

【12月】

各府省のEBPM推進体制決定【6ページ参照】

「統計等データの提供等の判断ためのガイドライン」、  
「EBPMを推進するための人材の確保・育成等に  
関する方針」の骨子を決定【7ページ参照】

【今春以降】

EBPM推進委員会で  
「統計等データの提供等の判断ためのガイドライン」、  
「EBPMを推進するための人材の確保・育成等に  
関する方針」を決定

新たな各府省EBPM推進体制の下、EBPM推進の取組加速

EBPM三本の矢を通じ、  
政策・施策・事務事業の各  
段階で、順次EBPMを実践  
(29年11月には、秋のレ  
ビューにおける「EBPMの  
試行的検証」の実施  
【5ページ参照】)

2

※「EBPMの推進」内閣官房行政改革推進本部事務局(2018.3.6)より

# “エビデンス”はどう理解されているか①

---

- 公共政策が必要とされているという「証拠 (evidence)」の重要性はますます増している。
- そのような証拠として最も重要なものが、社会の状態を知るための統計である。

※砂原(2017)より

# “エビデンス”はどう理解されているか②

---

- 予算要求にしても法律改正にしても、組織内での優先順位の決定、予算編成過程での査定、国会審議等多くのハードルがあり、エビデンスなしに政策が立案されることは考えにくい。

※森川(2017)より

# “エビデンス”はどう理解されているか③

---

- このようなEBPMの基盤をなすのが、統計的データを始めとする各種データなどの客観的な証拠であり、政策課題の把握、政策効果の予測・測定・評価による政策の改善と統計等データの整備・改善が有機的に連動するサイクル(EBPMサイクル)を構築することが必要である。

※ 統計改革推進会議「最終取りまとめ」(2017.5)より

# 日本におけるエビデンスのイメージ

## (本来の)エビデンス

- ・施策の有効性に関する信頼性の高い研究

## 評価

- ・プログラム評価
- ・業績測定
- ・費用便益分析

## 研究

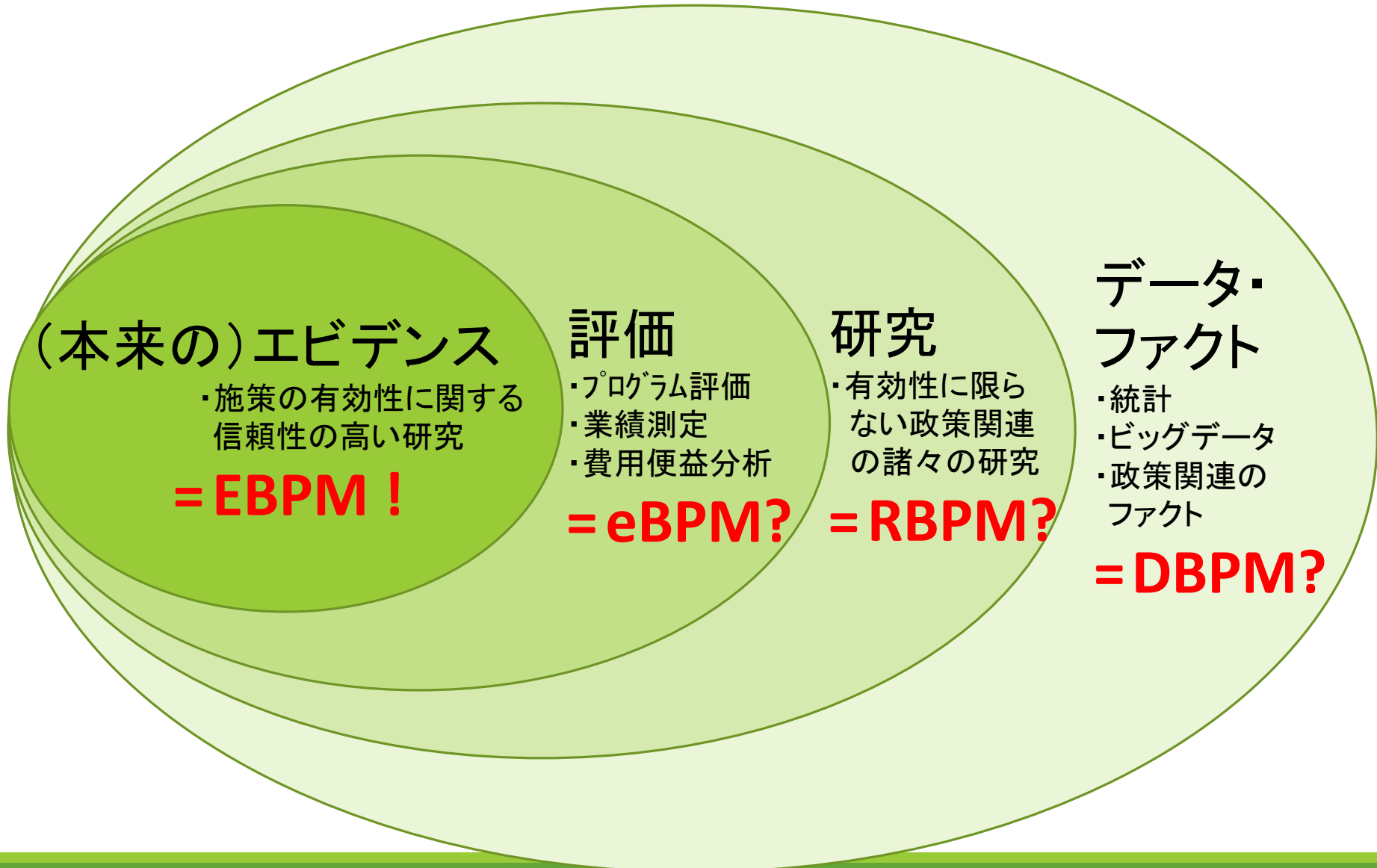
- ・有効性に限らない政策関連の諸々の研究

## データ・ファクト

- ・統計
- ・ビッグデータ
- ・政策関連のファクト



# eBPM? RBPM? DBPM?



# 国の議論におけるEBPMの整理

- 証拠に基づく政策立案(EBPM)とは、
  - (1) 政策目的を明確化させ、
  - (2) その目的のため本当に効果が上がる行政手段は何かなど、「政策の基本的な枠組み」を証拠に基づいて明確にするための取組。
- 政策の改善に繋げるため、順次、三本の矢(行政事業レビュー、政策評価、経済・財政再生計画の点検・評価)の取組を通じ、EBPMを実践。

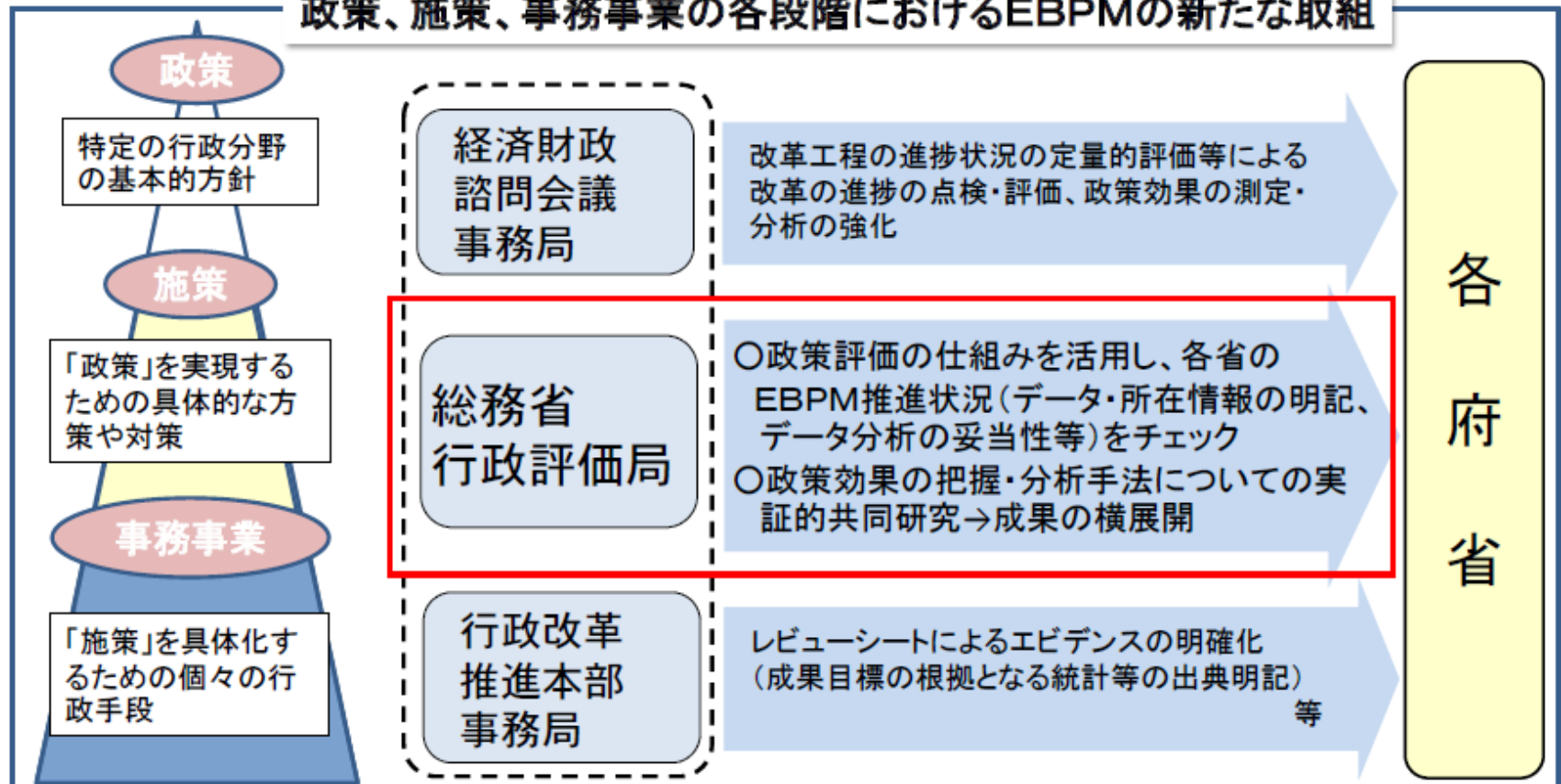
※「EBPMの推進」内閣官房行政改革推進本部事務局(2018.3.6)より

※政策評価の方法論を使うことを想定している！(eBPM?)

## 統計改革推進会議

【目的】 政府全体におけるEBPM(証拠に基づく政策立案)の定着、国民のニーズへの対応等の統計行政部門を超えた見地から推進するため、関係閣僚等で構成する統計改革推進会議を設け、改革の大きな方向性を取りまとめるとともに、改革の進捗状況をチェックする。

### 政策、施策、事務事業の各段階におけるEBPMの新たな取組

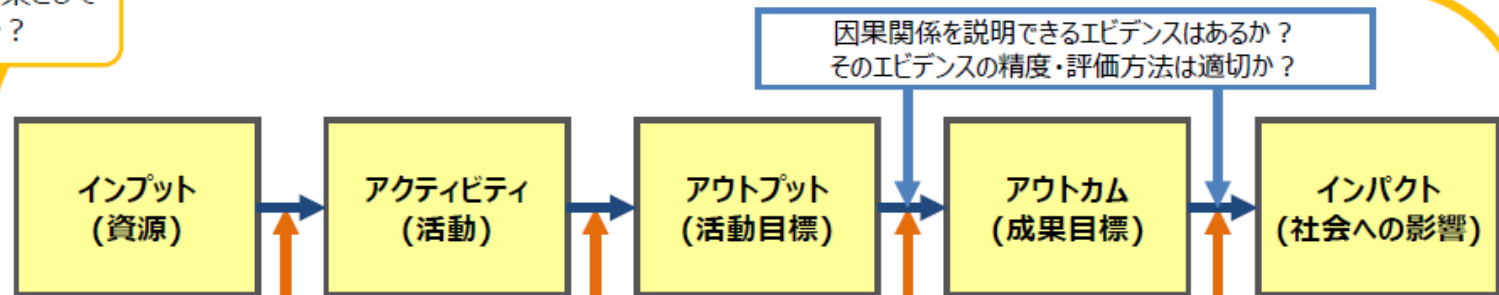


※ 総務省ウェブサイトより

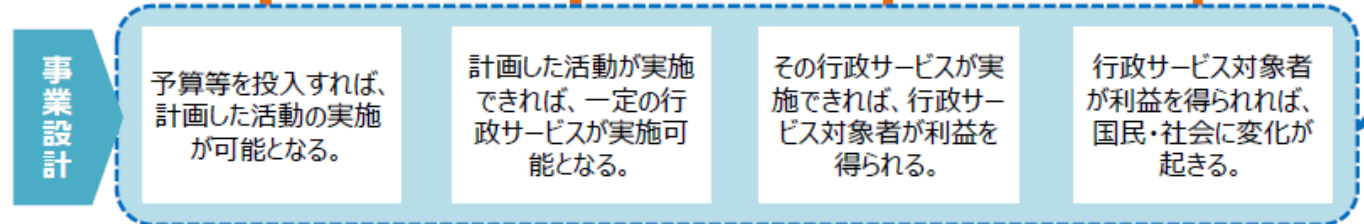
# ロジックモデル（イメージ例）と評価方法

未定稿

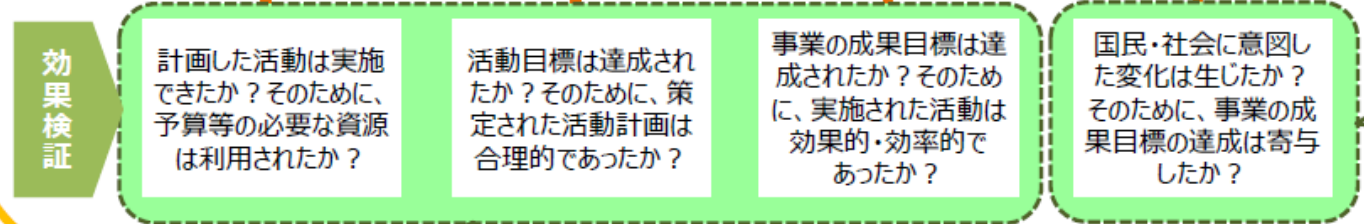
そもそも政策として  
行うべきか？



因果関係を説明できるエビデンスはあるか？  
そのエビデンスの精度・評価方法は適切か？



政策目的達成までの道筋は適切か？



政策による社会状況への改善効果があったか？

事業が意図された通り実施されたか？

事業実施の費用対効果は十分か？

(出所) 行革事務局作成

※「EBPMの試行的検証」行政改革推進本部資料, 2017.11.15より

## 参考文献① (\*を付したものはネットで入手可能)

- 田辺智子(2018a)「医療における「質に基づく支払(P4P)」の動向 — 諸外国の事例を中心に — (仮題) 『レファレンス』(執筆中).
- 田辺智子(2018b)「評価の利用と影響に関する実証分析 — 公共図書館を事例として —」日本評価学会春季第15回全国大会発表, 2018.5.26.
- 益田直子(2018)「「評価の利用・影響」に関する理論研究の概説」日本評価学会春季第15回全国大会発表, 2018.5.26.
- 田村誠(2018)「バリュー・ベース・ヘルスケア(価値に基づく医療)の意味するところ」『社会保険旬報』No.2700, 2017.1.21, pp.34-40.
- 福井次矢監修(2017)『Quality Indicator〈医療の質〉を測り改善する — 聖路加国際病院の先端的試み 2017 —』インターメディカ.
- 砂原庸介(2017)「公共政策と統計 — 証拠に基づく政策をめぐって —」(御厨貴『公共政策』放送大学教育振興会, pp.187-201.)
- 森川正之(2017)「「エビデンスに基づく政策形成」に関するエビデンス」RIETI Policy Discussion Paper Series (17-P-008).\*

## 参考文献② (\*を付したものはネットで入手可能)

- 小松康宏・門田美和子・福井次矢(2016)「医療の質改善の概念と手法—PDCA, six sigma など—」『日本内科学会雑誌』105(12), pp.2353-2357.\*
- 津川友介(2016)「医療における業績に基づく支払い方式(P4P)のエビデンスは極めて弱い」(医療政策学×医療経済学ブログ)2016.2.15.\*
- 田辺智子(2016a)「エビデンスに基づく医療政策の必要性—医療の質と費用対効果—」『調査と情報 -ISSUE BRIEF-』No.907.\*
- 田辺智子(2016b)「エビデンスに基づくがん検診はなぜ実現しないのか—アイデア理論を用いた—考察—」『日本評価研究』17(1), pp.3-18.\*
- 田辺智子(2014)「業績測定を補完するプログラム評価の役割—米国GP RAMAの事例をもとに—」『日本評価研究』14(2), pp.1-16.\*
- OECD(2014)『OECD医療の質レビュー:日本 スタンダードの引き上げ—評価と提言—』.\*
- 東尚弘(2007)「エビデンス-診療ギャップを埋める研究」『医学のあゆみ』221(6), 2007.5.12, pp.5452-546.

## 参考文献③

- Herbert, J.M. (2014) “Researching evaluation influence: A review of the literature,” *Evaluation Review*. Vol.38, pp.388-419.
- GAO. (2013) *Program Evaluations: Strategies for facilitate agencies’ Use of Evaluation in Program Management and Policy Making*, GAO-13-570.\*
- Johnson, Kelli et. al. (2009) “Research on evaluation use: A review of the empirical literature from 1986 to 2005,” *American Journal of Evaluation*, Vol.30, pp.377-410.
- Mark, M.M. and Henry, G.T. (2004) “The Mechanisms and outcomes of evaluation influence,” *Evaluation*. Vol.10. No.1, pp.35-57.
- Shulha, L. M. and Cousins, J. B. (1997) “Evaluation use: Theory, research, and practice since 1986,” *Evaluation Practice*, Vol.18, pp.195-209.
- Cousins, J. B. and Leithwood, K. A. (1986) “Current empirical research on evaluation utilization,” *Review of Educational Research*, Vol.56, pp.331-364.
- McGlynn E. A., et al. (2003) “The quality of health care delivered to adults in the United States,” *New England Journal of Medicine*, (348), pp.2635-2645.