

# 政策評価に関する統一研修(中央研修) 2018

「エビデンス」と「評価」は  
なぜ政策現場で疎んじられるのか？

橋本英樹

東京大学大学院医学系研究科

# いよいよEvidence-based policy making本格化

- 統計改革推進会議(2017・5月)
  - “限られた資源を有効に活用し、国民により信頼される行政を展開するためには、政策部門が、統計等を積極的に利用して、証拠に基づく政策立案(EBPM)を推進する必要がある”
- 政策、施策、事業の各段階での
  - ✓ “エビデンス”の活用
  - ✓ “因果関係”と“ロジックモデル”
  - ✓ “統計分析”と“評価”

# 期待と不安

- 昨年秋の「秋のレビュー」での試行を受けての自己点検、「公開プロセス」へ
- 模擬セッション（6-7月）
  - 事業の目的の明確さ
  - 目的・手段・成果目標指標の整合性（ロジックモデル1）
  - 施策・政策内での事業の立ち位置（ロジックモデル2）
  - 事業の有効性評価と外部要因の判別（因果検証）
  - “統計データ”とは？その使いどころは？
  - “データ”と“分析”？

そして試行本番、とりまとめ、検証（11月）

# 現場的には

- もともと「風」が吹いてきて決まった「事業」に今更後付けで数字を当てはめろと言われても
- データを示せと言うが、個々の事業設計にあわせて統計取っているものではないし
- インパクトとかアウトカムとかの指標になる良い「統計」がない
- 「統計分析」を厳密にやるには専門の技術を持った人材がない

# そもそも論として

- エビデンスとは？
- ロジックとは？
- 因果関係とは？
- 「データ」「統計」「分析」とは？
- Evidence-based policy makingは何のため？

ひもとく材料として“evidence-based XXX”の先駆例である XXX = medicineをめぐる過去の議論から

# Evidence-based medicine (EBM)

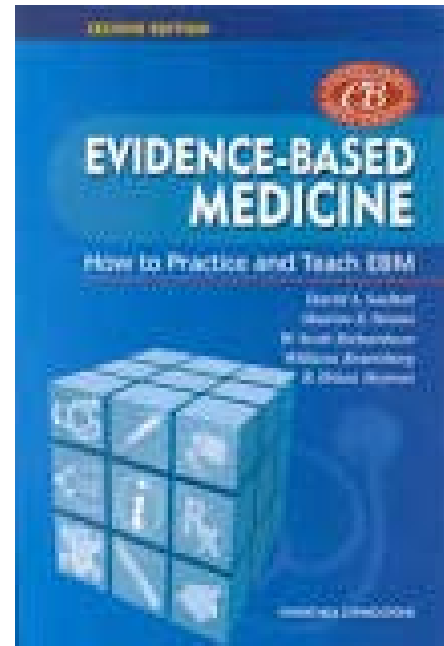
Sacket D. 1997

- “is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients”

**最良のエビデンスを慎重・明瞭・思慮深い  
意思決定のために用いる**

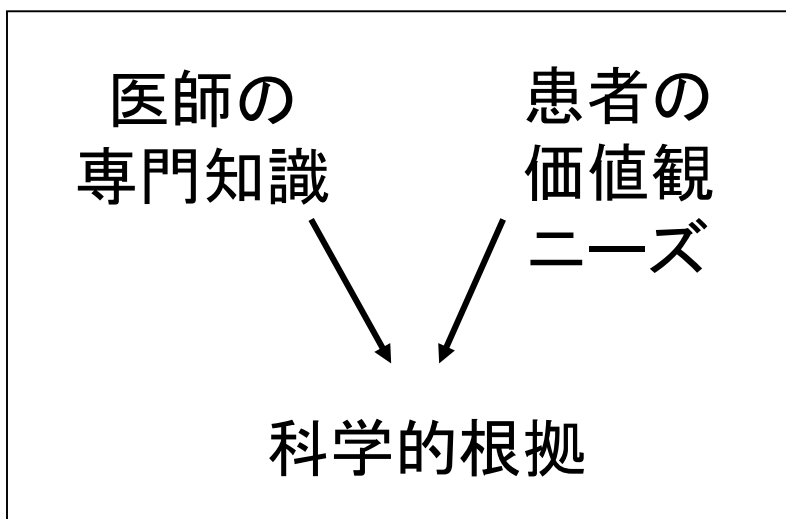
- “Good doctors use both individual clinical expertise and the best available external evidence, and neither alone is enough”

**エビデンスと経験のいずれも必要**

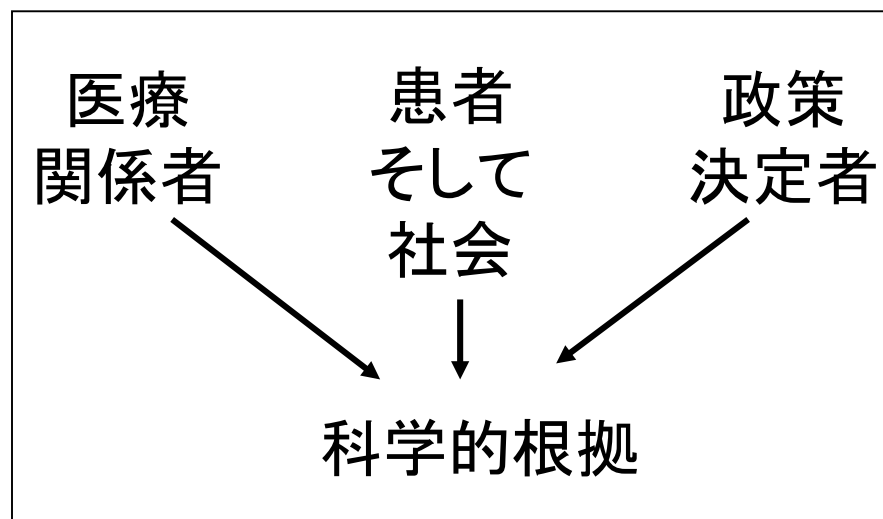


# なぜエビデンスが必要か？ = 開かれた意思決定のため

## EvidenceBased Medicine (Sacket D) 1997



## Evidence Based Healthcare (Gray M) 1997



意思決定の過程を合理的かつ開かれたものにするための「共通言語」としての「科学的手続き」

# 「統計」の種類と「データ」の性質

- 記述統計

  - 分布に関する記述統計(平均・中央値・分散)

  - 割合・程度に関する記述統計(頻度・比率)

- 分析統計

  - 関係性に関する分析統計(相関、危険比など)

  - 因果性に関する分析統計

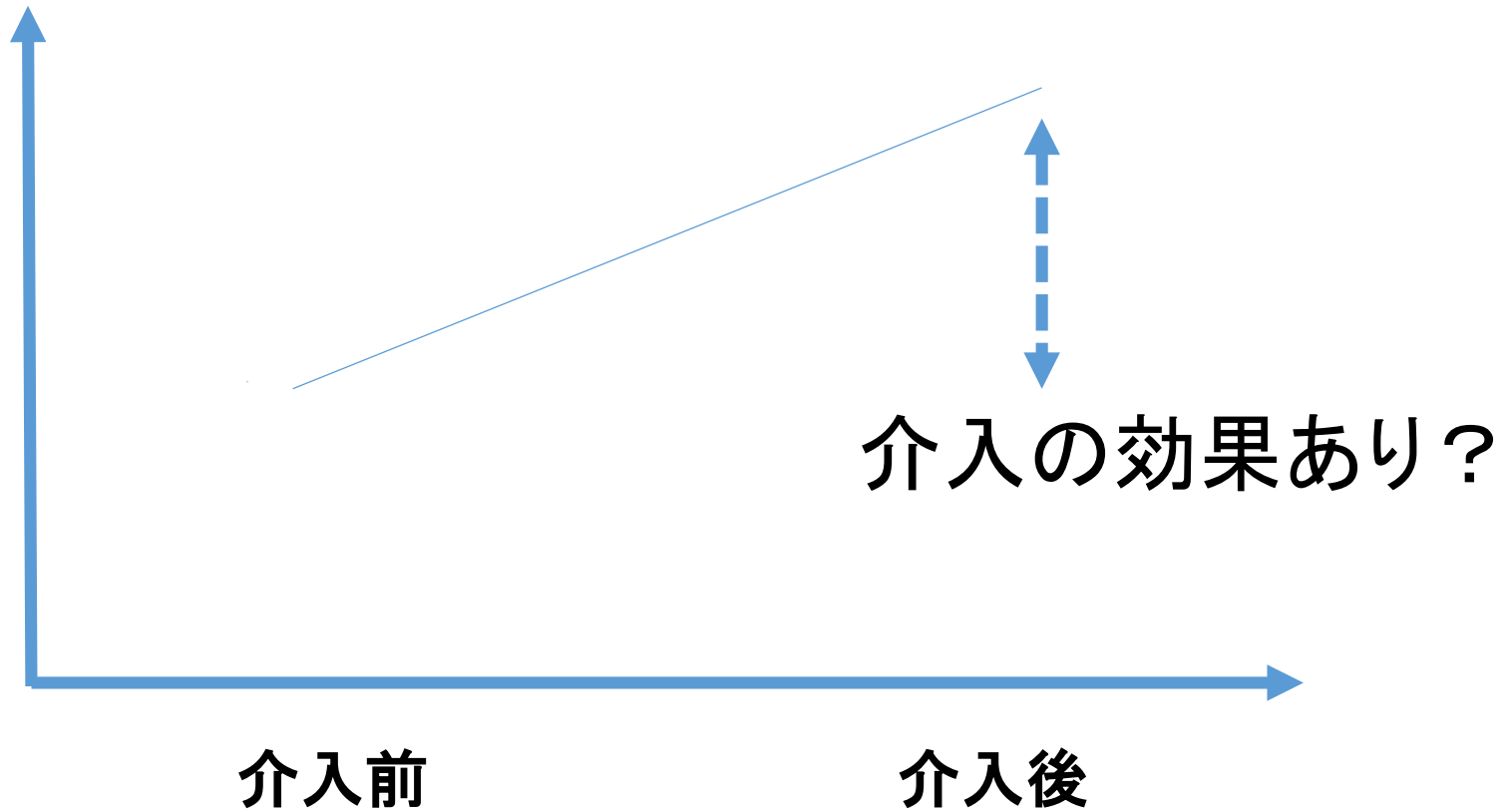
    - 横断 vs. 縦断データ

    - 観察 vs. 実験データ

因果関係を論じるときが主に問題

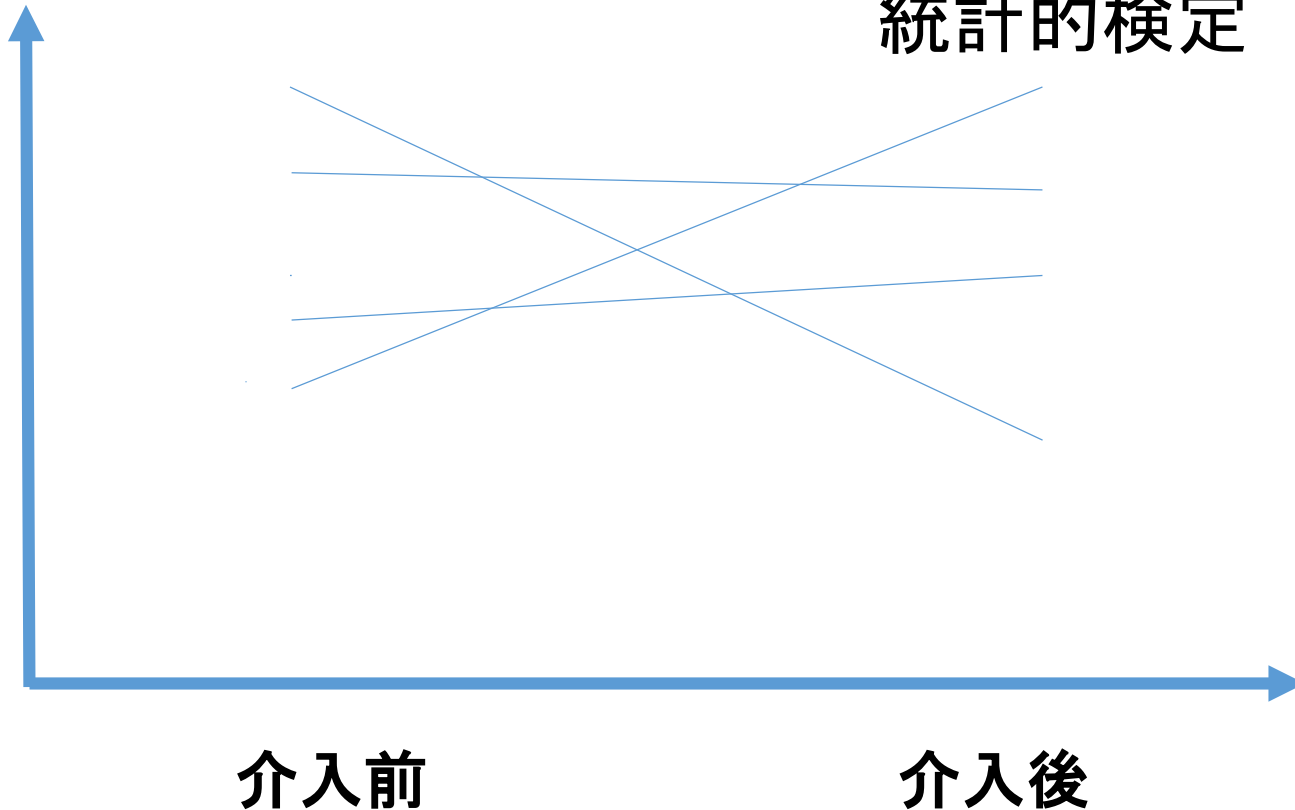


# 効果判定と因果推論



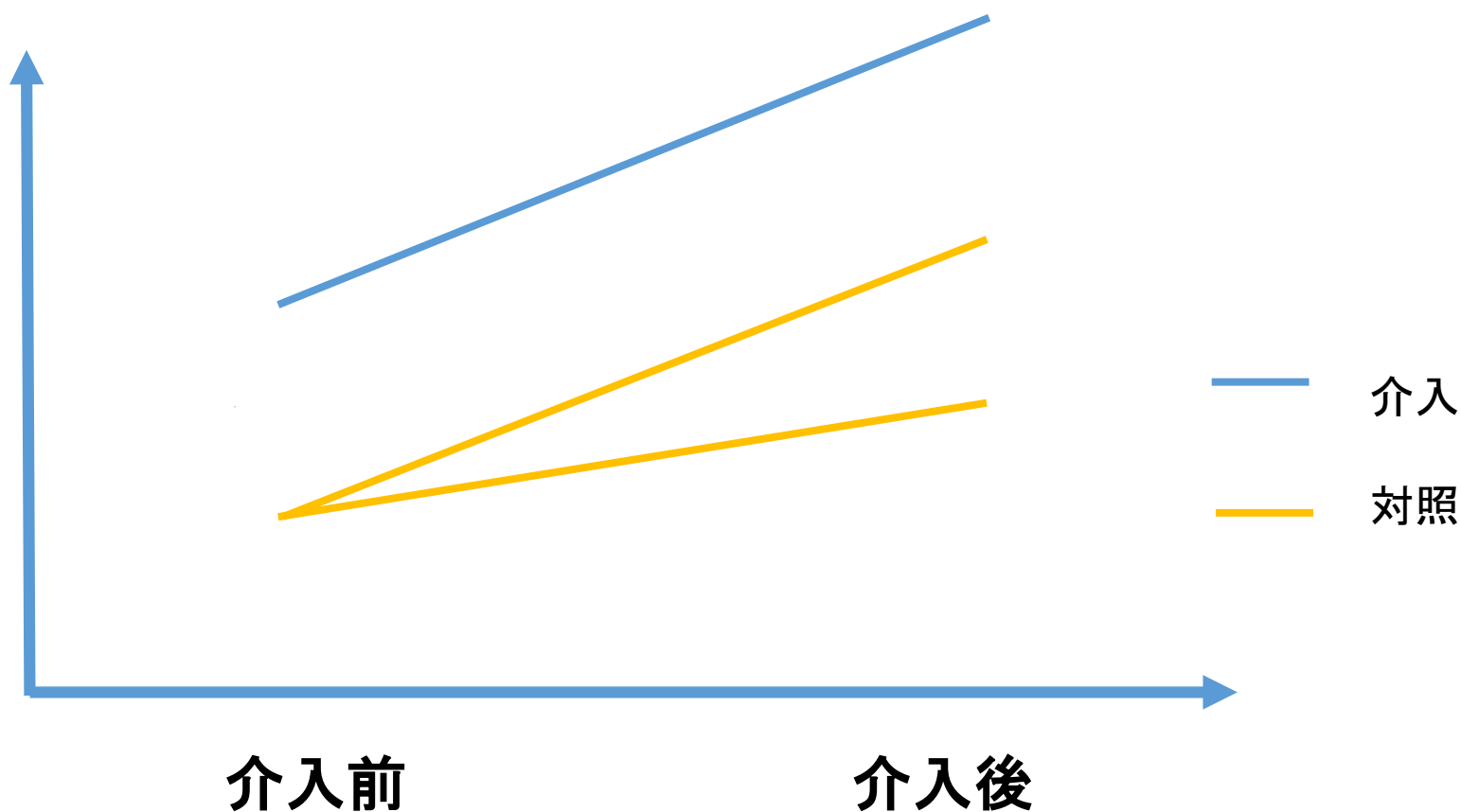
# 効果の判定 偶然性

偶然か偶然でないか？  
統計的検定



# 効果の判定

# バイアスの影響





# いまでもたまにある 無作為比較試験によるどんでん返し

- Cardiac Arrhythmia Suppression Trial (CAST)
  - 1992 心不全の突然死予防のために、当時最新の抗不整脈剤(Class III)の無作為比較試験、予想に反して治療群で死亡率が高いことがわかり中断
- Hormone Replacement Therapy
  - 2002～3 更年期症状の軽減に加え、動脈硬化予防にも役立つといわれていたが、無作為比較臨床試験の結果、むしろ心疾患死亡率が高いことがわかり中断

理論的・基礎実験的に「効果」を期待されていたものが検証の結果「無効・有害」となったものは他にも

# 無作為化比較試験の弱点

- 時間や費用がかかる
- 無作為割り付けにまつわる倫理的問題  
（社会実験がそうそう簡単にできない）
- 理想条件下での話、現実にあてはまるか？  
（社会実験できれいな結果が出にくい理由）  
アウトカムリサーチの台頭（1990年代）

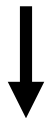
# どんなエビデンスを手に入れるのか？

理想的到達点

What we can do best?

理想条件下での効果

無作為化比較試験など実験研究



Efficacy (効能) としての  
エビデンス

1980年代

実状

Where we are?

通常条件下での効果

アウトカムリサーチ  
観察研究



Effectiveness (効果)  
としてのエビデンス

1990年代

大規模患者登録などを利用したビッグデータ解析 (2010~) ?

# EBMをめぐる批判と対応

- “Cook book medicine” 批判  
マニュアル診療と、エビデンスによる開かれた意思決定とは別物
- “情報過多”による実行可能性の問題  
情報の整理 (Cochrane Library) や情報技術の飛躍的発達による克服
- “数字情報”に頼り、価値・文脈(歴史ほか)・情緒などの実証データで表せない情報の軽視  
Evidence vs. Narrative based medicine 論争  
さらにはvalue-based medicineなども登場

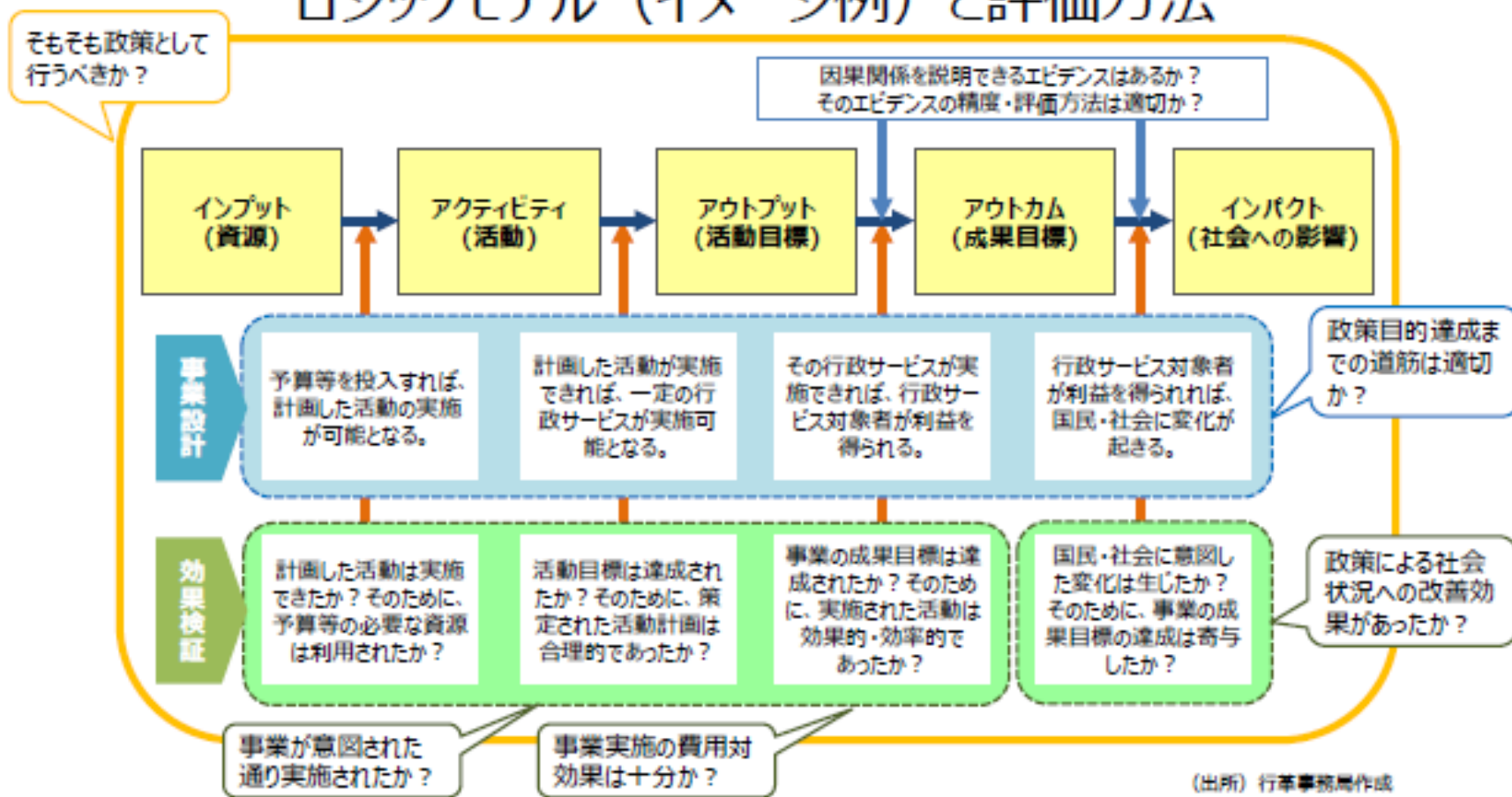


# エビデンス

- 科学的手続き＝再現可能で開かれた手続きを経ている
- 意思決定のステークホルダーが、フェアな交渉に参加できることを保証するひとつの材料
- 数字で対処方法・決定がひとつに定まる(押し切れる)、というものではない
- 科学的評価＝誤ることがあることを前提、影響を最小限にとどめ、改善のための示唆を得る
- 既存のエビデンスにただ従うのではなく、実践評価によって修正・加筆されていく情報

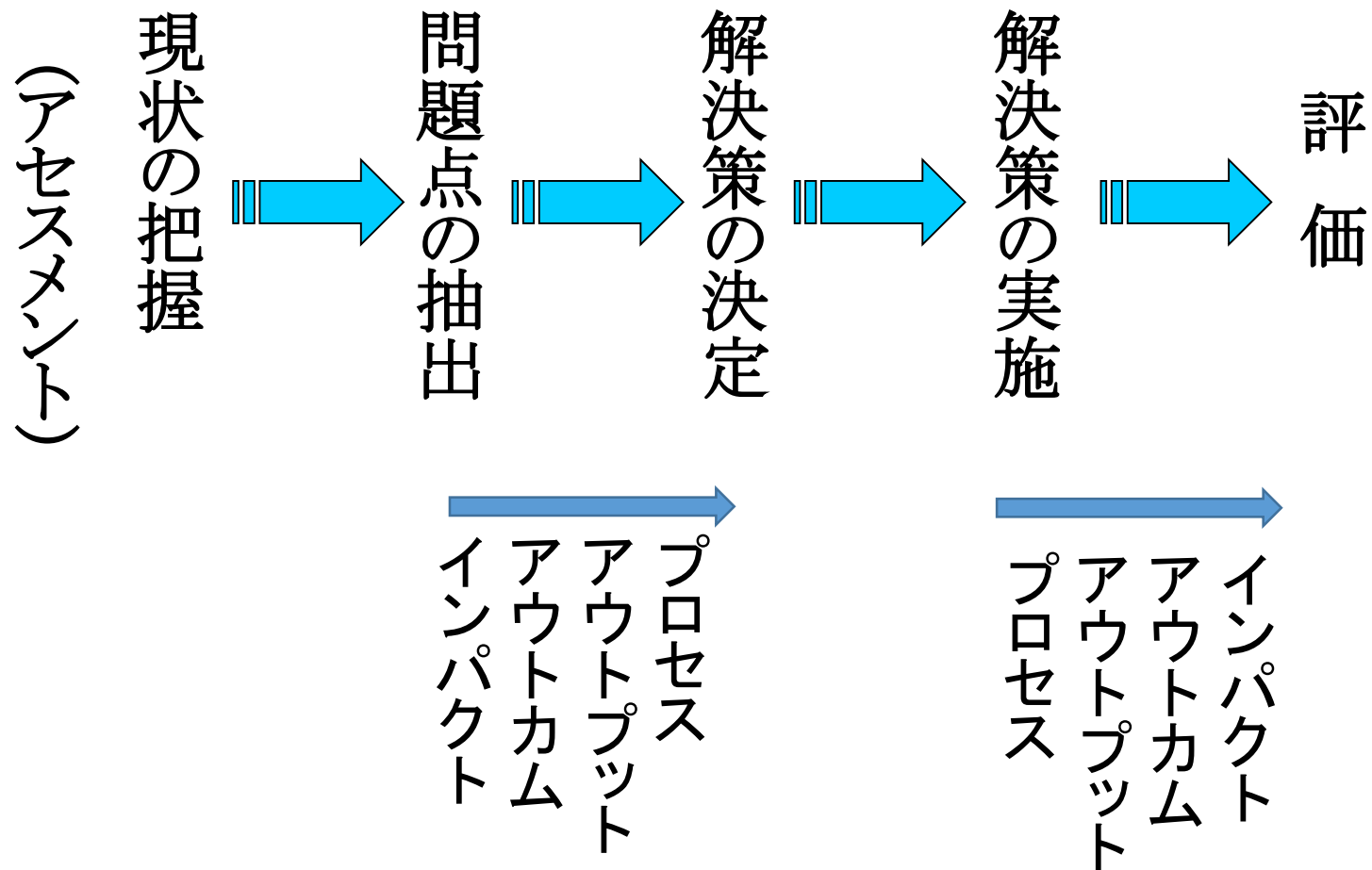
# ロジックモデルの「ロジック」とは？

## ロジックモデル（イメージ例）と評価方法



(出所) 行革事務局作成

# これとリンクしているのが



# つないで輪にしたものがPDCA もしくはProject Management Cycle

現状分析・問題特定

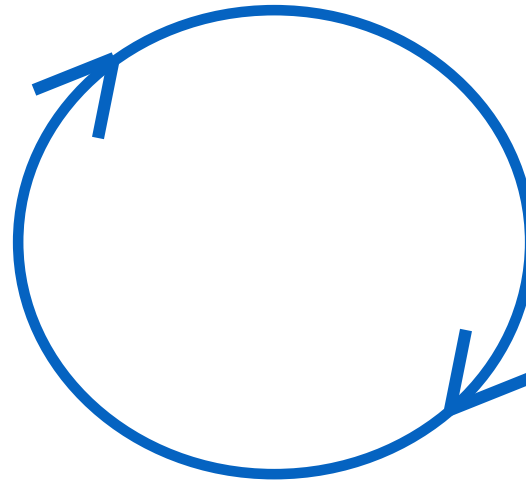
「現状」分析・「目標」設定

改善案提示

対処策作成・  
選定

結果評価・  
成果判定

実行・  
過程評価



# プログラム評価の階層

(Rossi, Lipsey, Freeman, 2004)

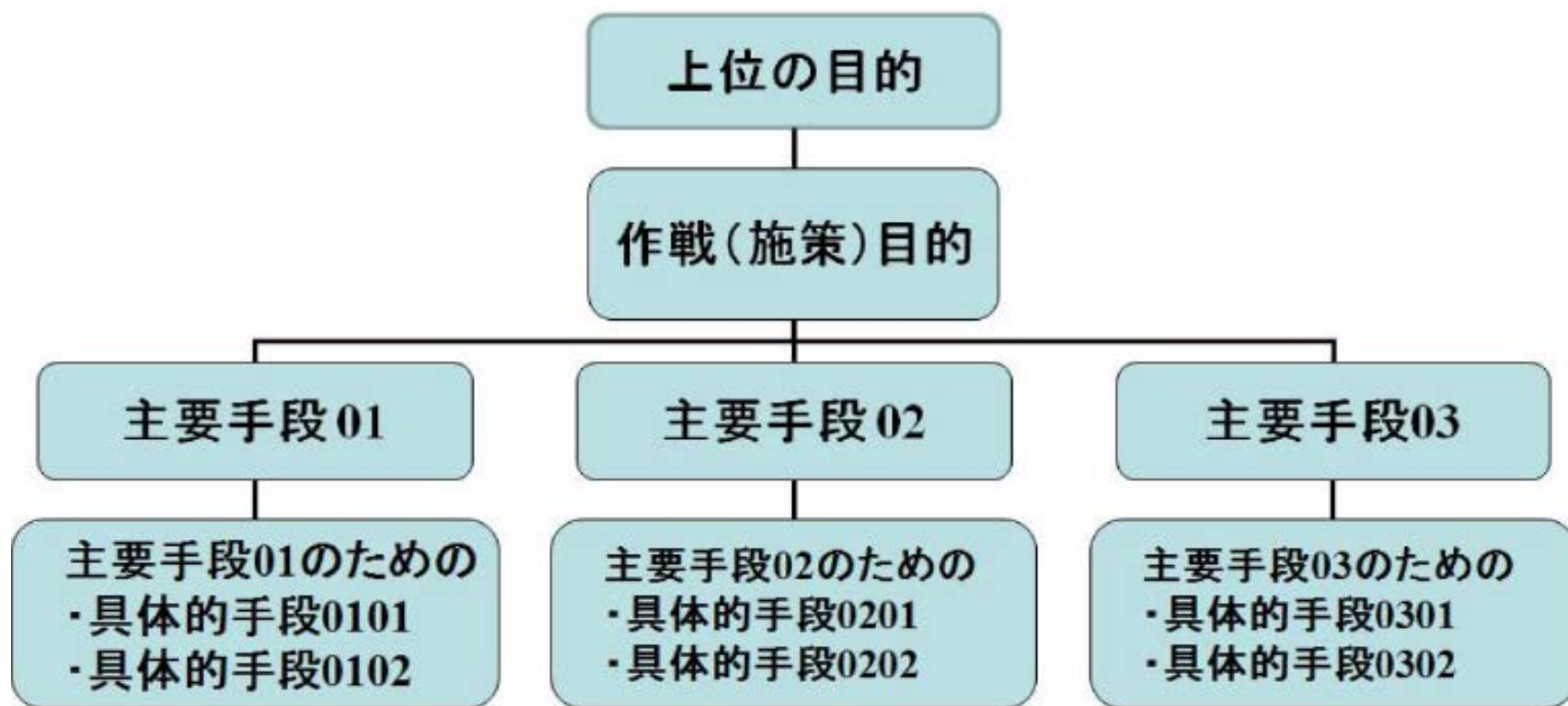
- ニーズアセスメント
- デザインと理論評価
- プロセス評価

記述統計主体

- アウトカムとインパクト評価
- 費用と効率性評価

分析統計要

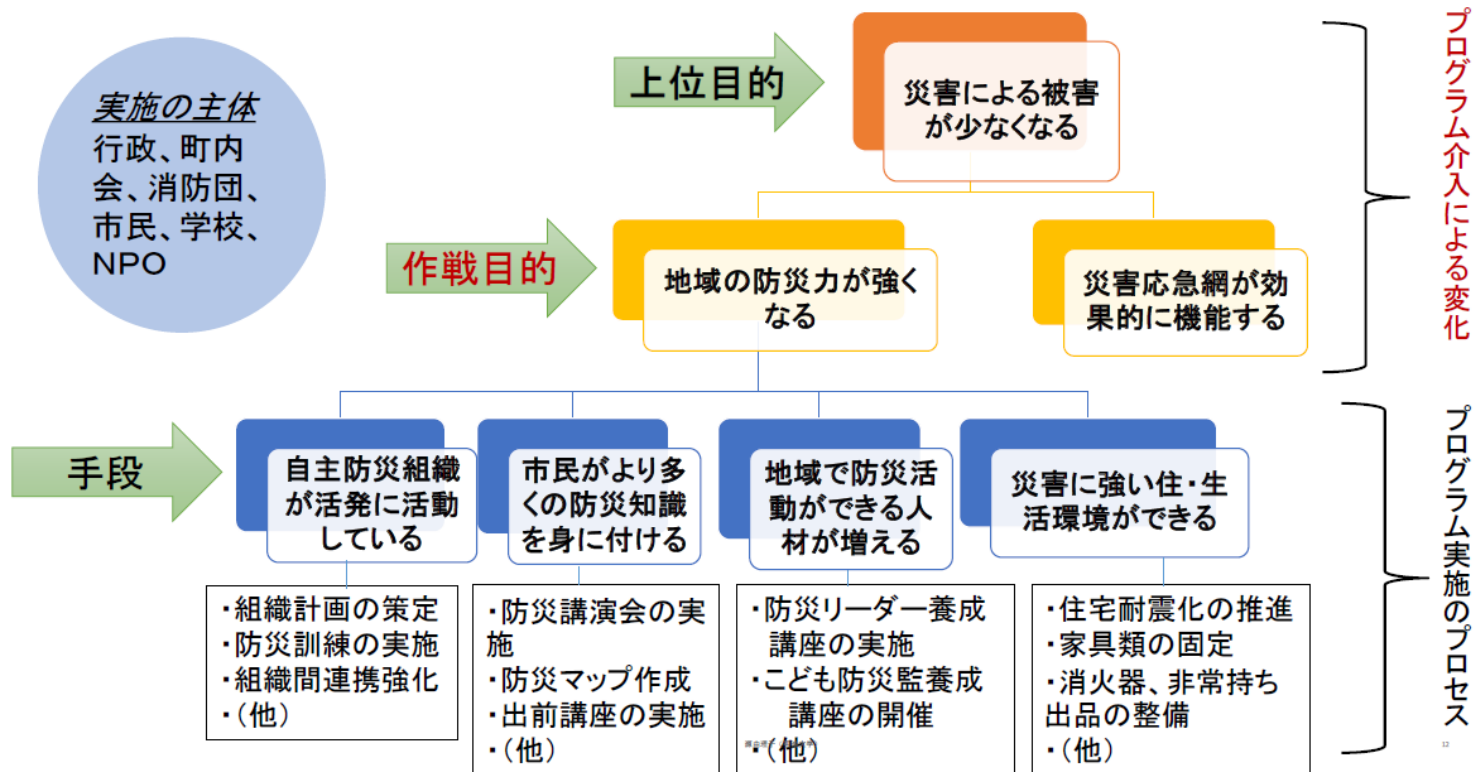
# ロジックモデルというともう一つ (ロジックツリー)



北大路信郷先生(明治大学)の平成26年度政策評価に関する統一研修スライドより

# ロジックツリー(つづき)

## 【例:安全・安心な都市:地域の防災力向上】



源由里子先生(明治大学)の平成28年度政策評価に関する統一研修スライドより

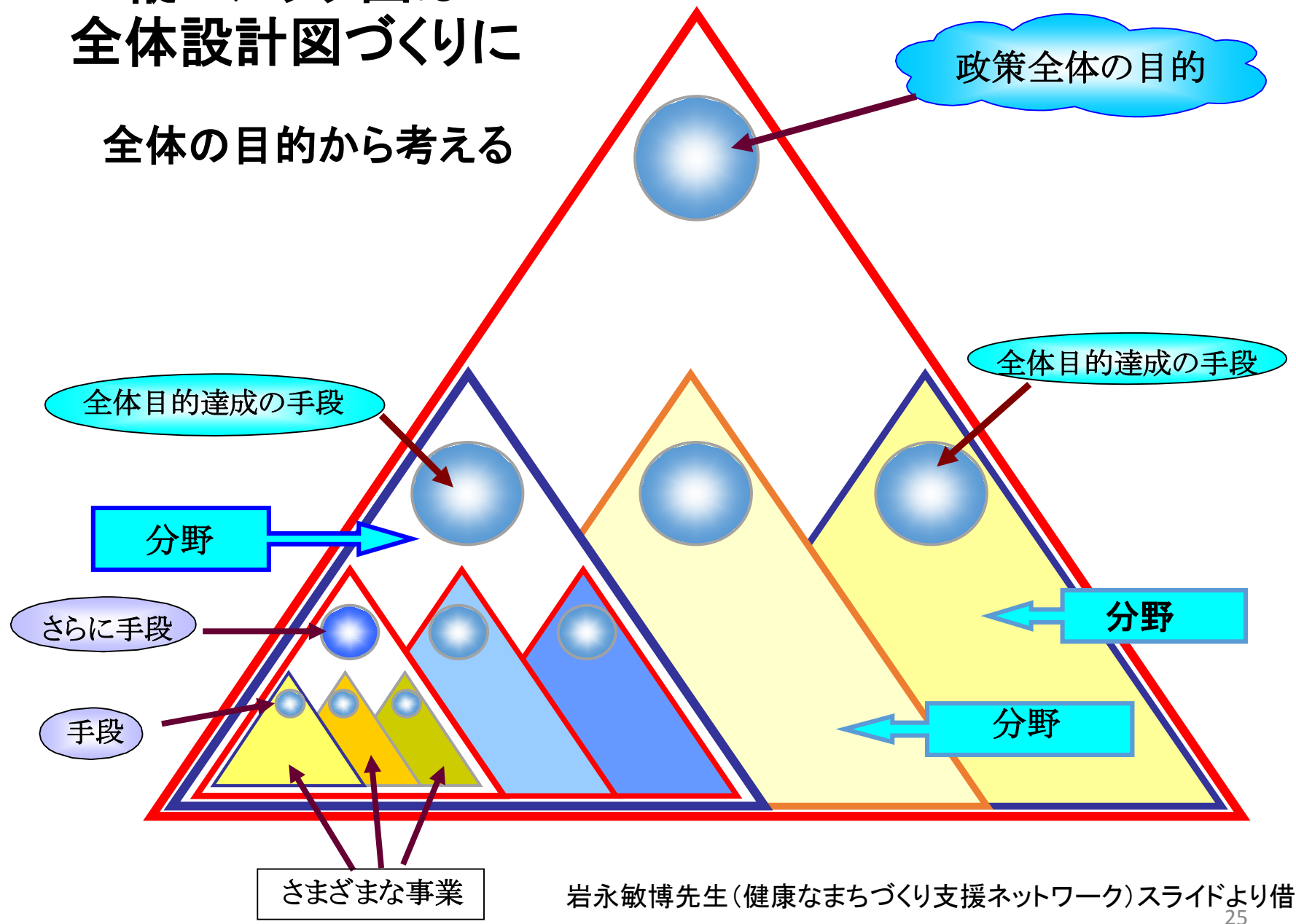
# 2つのロジックモデル図（横と縦）

- 横ロジック図は因果連鎖とその分析評価を意識  
比較的因果関係が明確なものに適用
- 縦ロジック図は目的と手段の複合的連関を意識  
上位目標を構成する下位目標群の連関  
目標に到達するさまざまな手段の連関



# 縦ロジック図は 全体設計図づくりに

全体の目的から考える



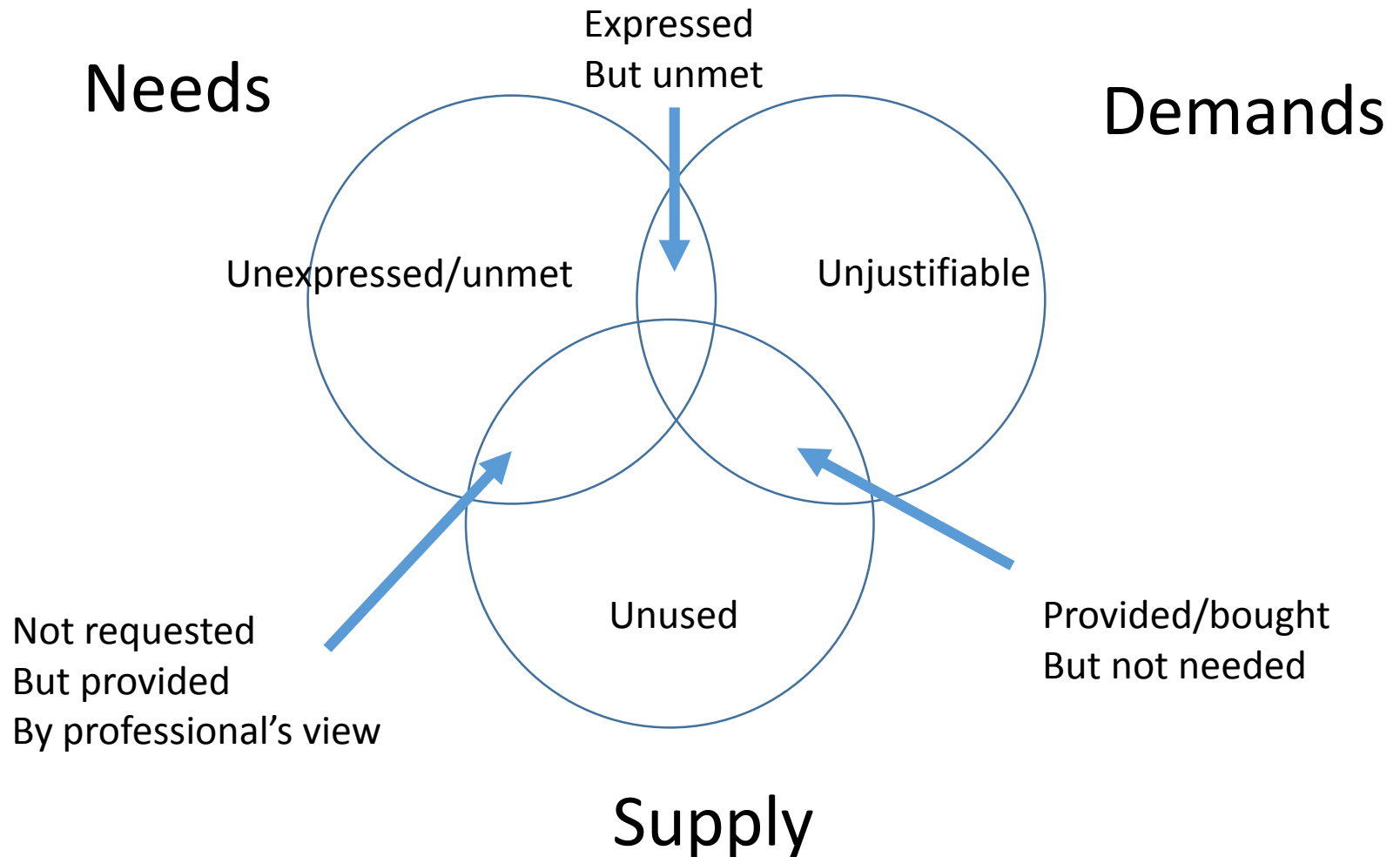
岩永敏博先生(健康なまちづくり支援ネットワーク)スライドより借用

# 現状分析と目標設定の2つのアプローチとも呼応

- 現状の「問題」はなにか？この「問題」をどう解決したいか？  
問題抽出・分析・解決アプローチ（経済分析型）  
「結果」としての問題  
「要因」としての問題  
両者の因果関係の同定が必要
- 現状はどこか、理想はどこか？理想に到達するにはなにを克服すべきか？  
目標指向・戦略的アプローチ（ビジネスマネジメント型）  
達成したい価値の同定  
動かせるリソース・機会の同定  
両者をつなぐ戦略・戦術のマッピング（順序・タイミング）

# Needs, Demands, and Supply

(Naidoo&Willis, pp258)



## 因果推論 (reasoning) と論理 (logic)

推論 = 真実であるかどうかを議論する

そのお作法として異なる論理 (logic) がある

- 演繹 (deductive)
  - 一般論 (理論) から個別の結論を正当化
- 帰納 (inductive)
  - 個別・特殊な事象から一般論を正当化

# 分野によって違う

- 経済学系の論文（主に演繹ロジック）

導入→理論→定式化→データ→

実証結果→考察・結論

検証結果が外れたら理論が外れたと結論

- 医学系の論文（主に帰納ロジック）

導入→データ・分析手法→分析結果

→考察（理論化）→結論

検証結果が有意なときは議論しやすいが外れたときは議論しにくい

# 帰納的因果推論

- なにをもって因果関係ありとするか？

Hill's criteria

関連の強固性	-----	関連性指標の計算
関連の時間性	-----	縦断データ
関連の一貫性	-----	他の分析結果との照合
関連の特異性	-----	単一原因単一効果？
生物学的説得性	-----	政策だと経済学的？
量反応関係	-----	ものによりけり？
類似性		
既存知見との整合性		
実験的証拠		

# 「たばこは肺がんの原因」論争を事例に

- 疫学的な常識として疑問の余地ない「因果関係」
- 「関連の特異性」はあてはまらない
- 「たばこ」のなにが肺がんになるのか分子レベルのメカニズムは？
  - 遺伝子、他の環境要因との相乗効果
  - ベンツピレン？それともダイオキシン？
  - 細胞癌化の分子メカニズムのどこと関係している？
- 「原因である」というstatementに「絶対的な基準」があるわけではない

## では政策介入(原因)の評価は？

- 他の要因(同時に動く社会経済状況、自然傾向ほか)の影響をどう除くか？
- 比較群がないときに計量化できるか？(regression discontinuityにいつも頼れるわけではない)
- やはりなんらかの前提をどうしても伴う「条件付き評価」
- 計量以外の情報(質的評価)も重要な評価



# 要約

- ロジックモデルも、アプローチも、因果ロジックもひとつではない
- 「現状」分析、政策目標設定、プロセス評価、アウトカム評価それぞれでもちいるデータの性質、求める「統計値」、分析手法もさまざま
- つまりEBPMのやり方に「これがhow to EBPMです」というものは多分存在しない
- でもEBPMが求められる理由はなにか？
  - 政策に「絶対的正解」や「絶対的失敗」はない
  - 政策は多数多様のステークホルダーを抱えている事業

だから、うっとおしいけど必要とされている