

本日のご講演に関する コメント

広島大学・高等教育研究開発センター
EBPMの検証・検討研究チーム
令和2年11月11日

あくまで、、、、
問題を共有する立場として、
自省を込めて、
議論していきたいと思います！

(・・・批判は容易い。ただこの際だから、洗いざらいあれこれあぶり出して、潰せるものは潰して、
より健全な
e-CSTIへ展開していただければ……)

データの整備・一元化・アーカイブ化・可視化という視点から：ありがたい！おもしろい！

■高等教育研究者にとっては、こうしたデータが公開されるのであれば「助かる～」

□例：米国では…

- IPEDS(<https://nces.ed.gov/ipeds/>),
- カーネギー分類(<https://carnegieclassifications.iu.edu/>)
- 日本でもこうしたデータが欲しい…手弁当で地道にデータ整備（RIHEと東大・大学経営政策コースと連携）

■日本の公的統計はようやく公開に向けてスタートしたばかり

□人文学・社会科学データインフラ構築：https://www.jsps.go.jp/data_infrastructure/

□公的統計マイクロデータ利用：<https://www.nstac.go.jp/services/archives.html>

■公的統計整備の遅れに関する懸念：データのある国への人材の流出

□近年の中国：積極的に各種統計整備・公開 ⇐ 人材が集う（信頼性の問題はあるが…）

■可視化されたデータの分布の面白さ・ワクワク感

□スライド34など：「専門職」間での収入格差が意外にも大きい、全般的に高い「やりがい」感

■大学にまたがるデータ整備のメリット

□IR（Institutional Research）：学内データ分析で無自覚に起こしがちな交絡・セレクションバイアスの軽減

利用者・分析者“皆”が気をつけるべき点

■ 「エビデンス」「可視化」「見える化」が、本当にその先にある“はず”の目的と繋がっているのか？

□例：人材の活用状況調査（スライド33～）・・・企業において、学んだ知識と実際の業務、やりがい、年収との相関の分析

- 「興味深い」分析ではあるが・・・、この分析から**なに**（政策的含意）を導きたいのか？

利用者・分析者“皆”が気をつけるべき点

■「エビデンス」「可視化」「見える化」が、本当にその先にある“はず”の目的と繋がっているのか？

□例：「機関間移動をした研究者の方が移動していない研究者よりも、論文パフォーマンスが高い」 https://e-csti.go.jp/downloads/about/e-CSTI_20200901b.pdf

●**因果は逆の可能性もあり**：相関分析から因果を読み取ってはまずい（誰もがよくやるが）

✓反論あり（登壇者とのメールのやりとり）：「機関間移動をした人は、移動の前後のデータを採っているので、因果を考慮できている。」（…から上記解釈は正しい、と暗示）

✓**それもだめ!**：移動した人は、もともと優秀だったので、より生産性が高まる大学から声がかかった可能性を否定できない=移動の効果とは言えない ← この問題は**セレクションバイアス・交絡**

●因果推論上は…移動した人が「もし移動しなかったら」/移動しなかった人が「もし移動したら」をも想定した分析をしないと意味が無い=**反実仮想**を想定する必要がある

✓現実には不可能⇒各種の分析手法を駆使して反実仮想を擬似的に創造・分析している（因果推論を検索すればいくらでも出てきます・・・）

■**留意!**：勇み足には注意しましょう～

利用者・分析者“皆”が気をつけるべき点

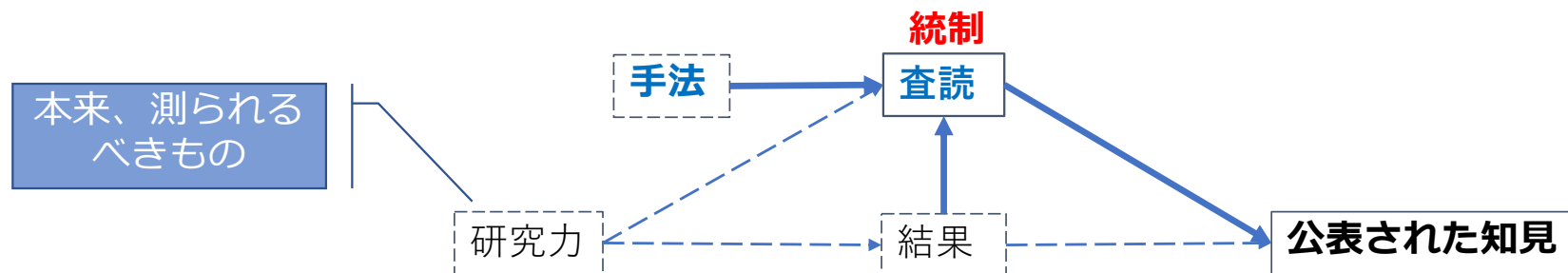
各論：

■論文“数”のデータについて

- SCI、SSCI論文至上主義を脱却し、国内誌も収集・評価してくれるようになったのは、ありがたい（泣：頑張らないと！）
- 論文数のような「数」を計量（カウント）したデータは、一筋縄ではいかない
 - そもそも正規分布しないことは、業界（科学計量学、統計学等）では既知・・・基本統計量が意味を成さない、2変量・多変量解析の応用が制限される等、限界が多々あり（村澤・立石 2017, 村澤・中尾・松宮 2019, 中尾・村澤 2018など）
 - 2次元プロット依存は問題もあり（先にも述べた相関と因果の混同）
- そもそも、「分布」のどこに注目したいかに依存するが・・・
 - 傑出した人への支援・・・はずれ値
 - 研究者集団の全体像・平均像の把握・・・
 - 研究者の底上げ・・・

利用者・分析者“皆”が気をつけるべき点

- 査読論文を扱う時の問題点・・・「**出版バイアス**」 = 査読がもたらすセレクションバイアス：真の研究力や質の高い研究を測定し切れていない危険性あり。DAGによる検討はRohrer (2018) を参照。
- 出版バイアスは、分析結果だけではなく、**研究手法によるセレクションバイアス**ももたらしているとの報告も (Brodeur, Cook, and Heyes 2020)
- ……実は暗黙の裡に、査読というフィルターを通すと、そもそも「優れた手法」「優れた研究知見」「優れた研究力」との関係に**疑似相関・内生性**問題が発生（下図を参照）→優れた発表成果は本当は何を意味しているのだろうか。
- 研究手法に影響する研究資金もまた同じ問題に遭遇…（以上、以下の図を参照）
- EBPMの文脈において、可視化という前提（? 称賛される趨勢?）がもたらすセレクションバイアスはないか？



利用者・分析者“皆”が気をつけるべき点

■ Evidence-Basedを謳いながら、Evidenceが喪失・・・？

- BI等による「可視化」が、むしろ「エビデンス」の信頼性を喪失させる
 - (スライド36頁) : グラフ化することにより、誰でも直感・主観による解釈が可能のような錯覚を与える ⇨ 科学的根拠が乏しい・認知バイアスを招く
 - Policy-Based Evidence Making (先に政策ありき、政策に都合のいいエビデンスの後付け構築) とも？

■ 当該文脈・専門領域に関する専門知・経験知に基づいた解釈の必要性

- Evidence-Based (エビデンスに基づいた) から・・・
- Evidence-Informed (エビデンスに通暁する知に基づいた) (加納・林 2019)
- Research-Based (研究知に基づいた) (佐藤 2018)

・・・への転換が必要？

最後にお問い合わせと期待・・・

- データは、“ある”にこしたことはない、です。データの整備と公開、本当にありがたいことです。
- データ大好きな研究者に、データを提供して、じゃんじゃん分析させれば良いのでは？
 - 「大学等研究機関の関係者」には公開いただけるんですよね？
- データを作ることが目的化しないように・・・
 - 過去の大学評価・IRの文脈では多発：「**エビデンス必要！**」⇨「**データベース必要！**」⇨**情報処理の専門家**主導で膨大なヒトモノカネつぎ込んで「データベースできた！」（どや顔）⇨・・・**分析に使えないんですけど！！**（分析の専門家）
・・・**仏作って魂入れず**・・・現況のEBPMは「因果推論」が必須であるだけに困る！
- 可視化・直感・経験値・専門知のほどよい融合による、多角的分析と解釈、そして議論の継続をしていければ、幸いです！！！！

ご講演者・オーディエンスの皆さんと有意義な議論をしましょう～。

■参考文献

- Brodeur, Abel, Nikolai Cook, and Anthony Heyes, 2020, "Methods Matter: p-Hacking and Publication Bias in Causal Analysis in Economics", *American Economic Review*, 110 (11): 3634-60.
- 加納寛之・林岳彦, 2019, 「環境分野へのEBPMの導入に向けての概念整理- "EBPM"概念の耐えられない狭さ -」第144回関西公共政策研究会（2019年5月11日、於：京都大学）発表資料.
- 村澤昌崇・立石慎治, 2017, 「計量分析の新展開:過去10年間の経験を振り返って」『高等教育研究』第20集,135-156頁.
- 村澤昌崇・中尾走・松宮慎治, 2019, 「大学の生産性とガバナンス」『名古屋高等教育研究』19,153-169頁.
- 中尾走・村澤昌崇, 2018, 「大学教員の生産性：再考—新たな計量分析の試み：欠損値補完・サンプリングバイアスの補正・"ゼロ"の意味の解釈」『ディスカッションペーパーシリーズ』No.9., 1-21頁.
- Rohrer, Julia M., 2018, "Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data", *Advances in Methods and Practices in Psychological Science*, Vol1(1):27-42.
- 佐藤学, 2018, 「教育研究におけるデータ基盤の必要性」日本学術振興会；人文学・社会科学データインフラストラクチャー構築プログラムワークショップ（2018年7月1日、於：一橋大学一橋講堂）発表資料.