

エビデンスを「つかう」を学ぶ入門ワークショップ  
講義録

なぜ、エビデンスを「つかう」のか？

「つかう」とは何か？

---

対人サービス分野におけるエビデンスに基づく実践を考える

エビデンスを「つかう」を学ぶ入門ワークショップ 2018

講師 津富宏先生

主催：Early Intervention 勉強会

本書は2018年10月に開催された「エビデンスを『つかう』を学ぶ入門ワークショップ」での講義の一部を津富先生のご好意により記録させていただいたものです。

なぜ、エビデンスを「つかう」のか？「つかう」とは何か？

## なぜ、エビデンスを「つかう」のか？ 「つかう」とは何か？

静岡県立大学  
NPO法人青少年就労支援ネットワーク静岡  
一般社団法人静岡学習支援ネットワーク  
津 富 宏

では、徐々に今日の導入に入っていきます。これは私の略歴です。なぜ、そもそもここに立っているかと言うと、前は少年院の先生をしていました。畑に行ったり体育をしたり、日々はとても充実していて感動したりするんだけど、私のやっていることは、少年の人生に役に立っているのか、手応えはあるけど、意味があるのか？ないのか？など、間違っていることをやってはいないとは思いつつ、疑問にも思いながら働いていました。

### プロフィール

略歴	1983	5年かかって大学卒業 法務教官(少年院の教官)となる
	1986-88	ウィスコンシン州立大学マディソン校社会学部修士課程に留学
	2002	法務省から静岡県立大学に移る
専門		犯罪学(原因論、非行からの離脱研究) 評価研究(犯罪者処遇の効果研究) 青少年の社会参加支援
社会活動		NPO法人青少年就労支援ネットワーク静岡理事長 一般社団法人静岡学習支援ネットワーク理事長

たまたま留学をさせていただく機会があり、評価研究と言う分野を知りました。その一部に、刑事政策の効果研究と言う分野があり、1980年代当時のアメリカでは既に刑事政策の分野のランダム化比較試験(Randomized Control Trial: RCT)を紹介する授業がありました。保護観察に割付けるか、少年院に割付けるかをランダムに行った研究が1960年代にユタ州であって…といった研究を紹介する授業を受けていました。15コマ教えられ

るくらいの量の刑事政策のエビデンスを紹介する授業があったのです。日本でそんなことも発想もしたことがない私は、こういうことがやれるんだ、因果関係ってわかるんだと、とても驚いて聞いていました。帰国後、法務省の矯正研修所で30代前半働いていて、留学で学んだことを教えていたんですが、今日やる内容はそこで教えていたこととほとんど変わりません。

エビデンスを「つかう」ということに関しては、たった1つの立場があるわけではないと思います。議論はあると思いますが、こういう風に使う立場があるんだなあと、いくつかの見方・考え方を自分のものにしていくのが一番良いのではないかと思っています。

さて、今日の最初の話ですが、私は就労支援のNPOをしているんですけども、自分がやっていることに意味がないのではないかと思ってやるのは無責任だなど思ったので、自分の団体の就労支援を組み立てるときに、それなりにRCTを読み込んで支援を作りました。

少年院の教官をしている時に、どうやったら人の役に立つのかと考えたことは、おそらく一生忘れないと思います。

### 私の生涯をかけての問い

私(たち)はどのように人の役に立つのか

無意味なことをやると、相手の時間も自分の時間も、相手のエネルギーも自分のエネルギーも、無駄にしまいます。

エビデンスに基づく医療(Evidence based

Medicine : EBM)の本を読むと、おわかりになりますけれども、これまで無益どころか有害な医療介入が大量に行われて来ました。有名なのはジョージ・ワシントンが亡くなった瀉血などです。千年以上も、素晴らしい治療だと思われて来たものです。

身体を対象とする医療介入はともかく、人間を対象にしたその他の介入は物理的な反応が生まれから、受けた相手が「良い」と言うと、やったほうも「良い」と思ってしまいます。少年院のときは、子どもたちに「満足しましたか」と聞くアンケートをとったりもしていました。職員が一生懸命にやると、「はい」と言ってくれる。しかし、その介入が、それが本当にその子の将来にとっていいかがわからないのです。

犯罪学の分野では、ケンブリッジ・サマービル研究という、夏休みに非行少年をキャンプに連れて行くかどうかを RCT で割付けた研究があります。キャンプに行った子どもの方が行かなかった子どもより、早く死んだり、病気になったり、非行したりしたということが結果としてわかった研究です。キャンプに行った子どもたちの調査では、本人たちはとても満足したという結果が出ました。やった側も受けた側も、とても素敵と書いても、本人の将来に有害な影響を及ぼしている可能性があるということです。効果が出ないあるいは逆効果になることはよくあります。EBM が広まる根本となった「間違っことはしていけない」、「有害なことをしてはいけない」という理念は、犯罪学も人間を対象とするので全く同じです。

なぜ、この話を先にするかと言うと、休職預金などで介入の成果を可視化することが言われているけれども、それ以前に、そもそもその「介入」が有益であるという前提がどこから出てきて

いるのかを疑問に思うからです。私たちが人に関わる場合には、倫理性が求められます。介入を有益であると仮定して成果を可視化するよりも先に、その介入が人の人生を不幸にしていないかという問いを、まず先に持つべきであると考えています。

あらかじめお断りすべきこと

また、お断りですが、私たちが普段生きて行動しているときに1個1個エビデンスを求めているか、というとそんなことはありません。私たちのやっている大半のことは、間違っていないと思っています。

**あらかじめお断りすべきこと**

**エビデンスは私たちの実践の大半を支えない**

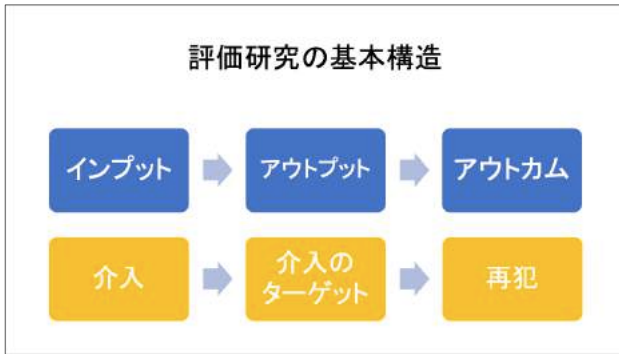
**私たちの実践の大半は所作**

**相撲の「技」について  
有効性を研究するだろうか？**

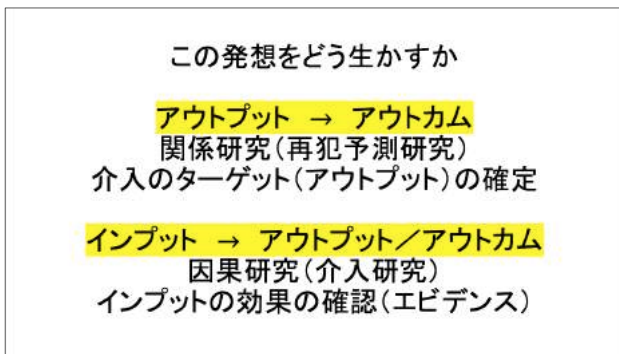
先ほど、徳永さん（参加者）が挨拶で言われていたような経験値とか暗黙知と言うものは既に存在しています。それが人を不幸にしている可能性はあるけれども、それが間違っているという前提を置く必要は、直ちには無いと思います。むしろ、うまく言語化できてないものを言語化して検証の対象にしていくということが大切です。たとえば、本人がやっている生活習慣が、他者から見ると、これはおかしいということがわかっても、本人にはわかっていなかったりします。他者がもつ経験値や暗黙知を形にして、検証の対象にしていくことは大事なことでありたいと思います。

評価研究の基本構造・どう生かすか

さて、このスライドは、インプット・アウトプット・アウトカムという評価の基本構造です。



玉突き構造で、物事を捉えるというのが前提です。



アウトプットからアウトカム (関係研究・予測研究)というのが「上」、インプットがアウトプット・アウトカムに影響を与えている (因果研究・介入研究)のが「下」です。

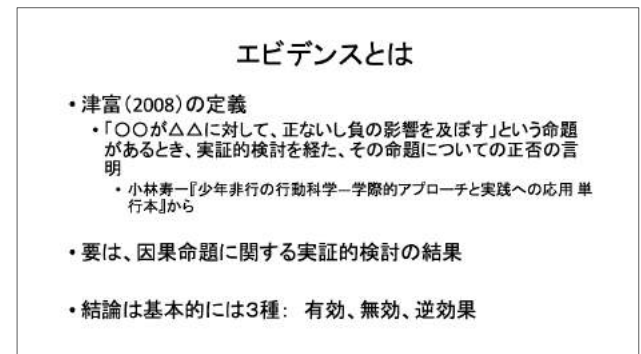
「上」はインプットがないんだけど、どういことが再犯を予測するのか？を明らかにする研究です。この研究にはインプットがなく、どうい人(=アウトプット)が病気(=アウトカム)になるのかな？といったアウトカムの予測因子を調べていく予測研究を指します。予測研究(相関研究)は、インプットという操作を加えていないので、本当の因果はわかりません。つまり、インプットが働きかけるべき候補は何か？というターゲットを確定していく研究です。

アウトカムを予測するものはいっぱいありますが、ターゲットが決まったら実際に介入してみる。そうすると、下の方(因果研究)になり、介入研究となるわけです。介入研究(因果研究)は、インプットとアウトカムが真ん中(アウトプット)を挟むという性質を持っています。インプットの

効果の確認(エビデンス)です。

先ほど、西郷さん(企画者)に、エビデンスの定義(「ある介入があるアウトカムに対して所期の効果を及ぼす」という因果命題に関する、実証的検討の結果、つまり、その因果命題が真である、あるいは、偽であるという言明(津富, 2010))に触れてもらいましたが、あくまで、今日の講義では、エビデンスとは因果関係・因果命題に関するものです。上の予測研究から分かることは、この講義ではエビデンスとは言いません。しかし、エビデンスという語は、データや関係研究の知見という意味も含めて多様に使われており、たとえば、EBPM (Evidence Based Policy Making: エビデンスに基づく政策形成)ではエビデンスという言葉を広く捉えています。しかし、今日は、エビデンスを、狭義に定義します。

エビデンスとは？



「エビデンスとは、『〇〇が△△に対して、正ないし負の影響を及ぼす』という命題があるとき、実証的検討を経た、その命題についての正否の言明」を指します。先ほど、西郷さんが使っていたのは、〇〇と△△を補足してくれたものです。要は、因果命題に関する実証的検討の結果です。結論は、基本的に3種あり、有効・無効・逆効果です。有効は良い結果、無効は英語でいうと null、逆効果は予測とは逆方向の効果であるということです。

## アメリカ留学


僕は、アメリカ留学で、ある条件を整えれば、因果関係を確定できるということに驚きました。

アメリカ留学(1986-88)

評価研究との出会い

因果関係を確定できると知った「驚き」

\* 当時はエビデンスという言葉はなかった

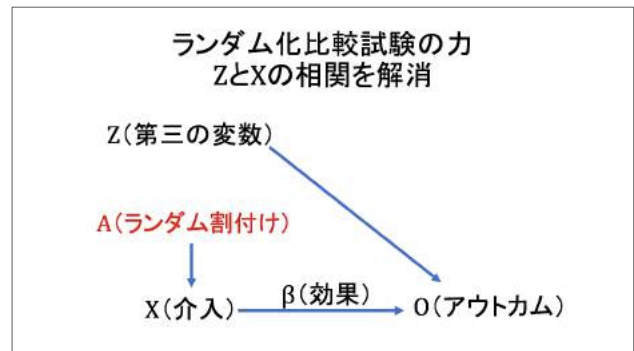
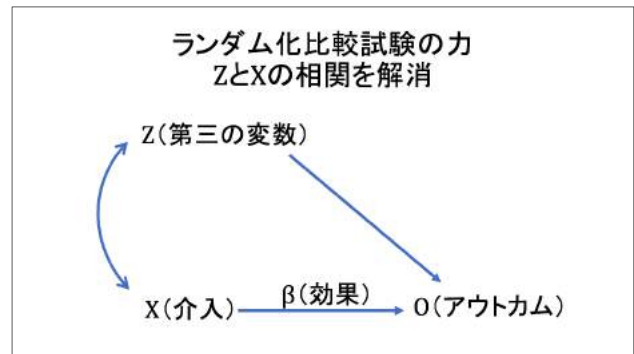


なお、当時は、アメリカでもエビデンスという言葉は使っていなかったと思います。法務省に戻ってからは、ロッシの教科書を使って矯正研修所で、評価研究を教えていました。

## ランダム化比較試験の力

さて、因果命題は確定できるという話を少ししたいと思います。RCT 以外の方法もありますが、RCT が一番クリアに因果関係を確定してくれます。さて、因果関係を確定したいと思ったときの問題は第三の変数(Z)があるということです。

例えば、ある少年院では電気工事課の訓練があるんだけど、そこに行く子は賢い子が多い。計算ができないといけない。そうすると、電気工事科に行くかどうかには、賢いかどうかという変数が影響してきます。つまり、電気工事科の子が他の子より、再犯しなかったとしても、この「賢さ」という他の変数のせいかもしれない。このように、自然状態では、因果関係は、常に Z (第三の変数) の影響を受けてしまいます。つまり、介入の効果は、自然状態では、X (介入) と相関し、O (アウトカム) に影響を与える何か=Z (第三の変数) の影響を受けているということです。



ランダム化比較試験をすると、介入を受けるかどうかは、くじ引き (ランダム割付け) によってしか決められないので、Z (第三の変数) の影響を排除できます。RCT の特徴は、Z (第三の変数) と X (介入) の相関を切ることによって、 $\beta$  (効果) の推定を正しくするということです。感動しませんか？僕は感動したんですが、因果命題を確定できるというのは、RCT の考え方が根幹にあるからです。

なぜエビデンスを使うのか？

## なぜ、エビデンスを「つかう」のか？

## 最大の目的

害を与えないため

科学にとって最も重要なのは  
対立仮説を減らしていくこと

「害を与えない」というのが最大の目的です。

あることを説明しようとするときにはいろいろな仮説がありますが、サイエンスとは何かということを考えると、それは仮説を棄却する過程をなんだと思います。間違っている介入について検証して仮説を棄却し、ダメなもの、もしかしたら人に害かもしれないという事はやらないためにサイエンスがあるのです。

メタ分析：エビデンスの集積

犯罪学における有名な介入で、スケアード・ストレートという有名な研究があります。



画像の右側にいるのは受刑者で終身刑の人などです。彼らが、やんちゃな中高生に対して、悪いことをしていると僕たちようになってしまうよ、ということ、ものすごい怖さで言う。受刑者の彼らにとっては一生懸命のクラブ活動で、自発的にやっているものです。受刑者にとってはやりがいに溢れているんです。

しかし、問題は、このスケアード・ストレートに関する RCT がこれだけあって、個々の RCT 研

究の知見を足し込んでみると、受けた子たちの方が再犯してしまうという結果が出ていることです。スケアード・ストレートは、評判のよいプログラムとして全米に広がったものです。けれども、メタアナリシス（メタ分析）という手法で、一個一個の研究の結果を足し込むと、合計値が1に触れていない。1は無効ということなのですが、1に触らないということは偶然の範囲ですらなく、明確に逆効果であるという結果です。

このような結果は出たけれども、全国に広がったプログラムを消し去るのに 20 年ぐらいかかりました。人を幸せにするためにやっているものだと思います。思い続けていたんですね。

whether は分かっても why は分からない

RCT では”whether”（効果があるか否か）は分かります。でも効果があるかどうかという”whether”がわかっても、”why”つまり、なぜなのか？ということはわかりません。実務家としては選択肢を捨てることはできるけど、次に何をやったらいいのかわからないということが起こります。

ただし、Scared Straightのメタ分析からは  
whether は分かっても  
why は分からない

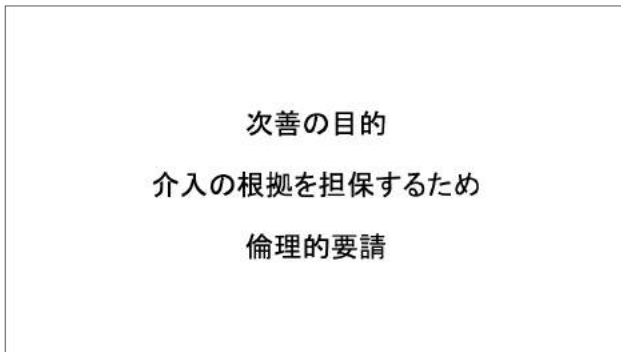
やってはいけないことは分かっても  
なぜやってはいけないかは分からない

そういう意味で、ある薬が効いたかどうかではなく、こういう理由で効かない、こういう理由で効くという理論が明確な介入を持つというのが大事だと思います。スケアード・ストレートがなぜ効かないのかを、社会心理学者が書いたものがあります。今、この本を回しますが、これはひとつ



の解釈ですが、おもしろい本です。

エビデンスを「つかう」次善の目的



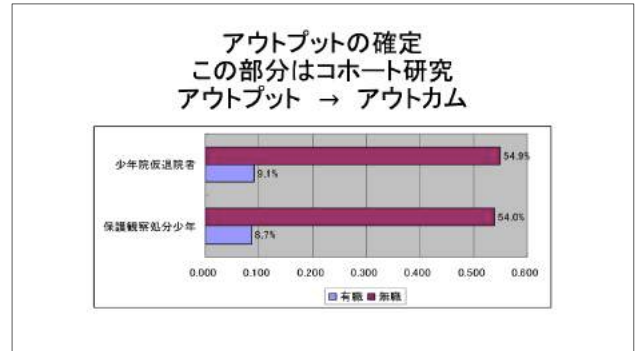
「害を与えない」に次ぐ、次善の目的は、「介入の根拠を担保するため」です。できるだけ最善のことをしたいということです。少なくとも自分たちがやっていることを間違っていないよね、ということを確認したいということです。



これは私がやっている NPO についての本で、静岡方式という就労支援について書かれているのです。私は少年院で働いていて、大学に転職したのですが、少年院で働いているときから、人の人生に1番影響を与えることは何かなあと考えて、それは働くことだと思い、就労支援を始めました。

アウトプットの確定

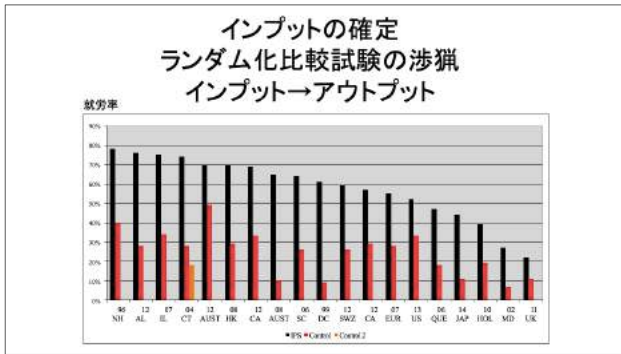
法務省はこういう統計をたくさん取っています。再犯するかしないかを、特定の属性と掛けてクロス集計したものです。



再犯に、最大の影響を与える変数は、まずは仕事、次は引受人がいるかないかだとされています。引受人がいるかないかを、他者は簡単には左右できません。しかし、仕事は、他者の介入によって何とかなると感じていました。そして、転職し、少年院の教官ではなくなったので、対象を非行少年に限らないで、就労支援を始めました。

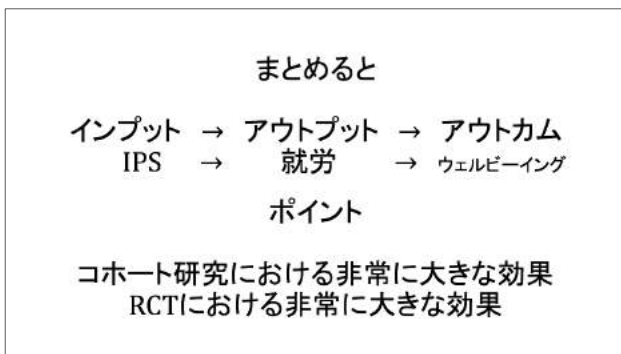
そして、就労そのものをターゲットにした介入を検討した RCT を探し出しました。効果がある RCT を、2種類見つけました。1つが精神障害を持つ人の就労支援の手法である IPS (Individual Placement and Support) でした。本人と一緒に職探しをする、本人のストレングス、つまり本人は何が好きそうか、何に興味を持つか、ということを中心に職探しをする手法です。障害者の就労支援では、障害特性に合わせて仕事を探すということが多いと思いますが、IPS はそうではありません。世界中で、IPS の RCT がやられていますが、どの RCT でも大体、就労率が倍ぐらいになっているのがわかります。





特定の介入の効果が、これだけ一貫して、ポジティブに出続けるということは滅多にありません。たとえば、認知行動療法は、総じてはプラスの結果が出ますが、ネガティブな結果を見出す研究もあります。IPSに加えて、もう一つ、RCTで効果があることを示していたのはジョブクラブでした。この2つのうちIPSを参考にすることにしました。

インプットがIPS、アウトプットが就労、そして、アウトカムはみんながハッピーになったらいということまでちょっと遠くにおいています。ひとり一人が充実した人生を送るためには、仕事が必要で、IPSは少なくとも就労に影響を与えている。そこで、IPSを採用して、静岡方式の実践を積み上げていきました。



- 理論＝IPSの原則**
- ・症状が重いことを理由に就労支援の対象外としない
  - ・就労支援の専門家と医療保健の専門家で作る
  - ・職探しは、本人の興味や好みに基づく
  - ・保護的就労ではなく、一般就労をゴールとする
  - ・生活保護や障害年金などの経済的な相談に関するサービスを提供する
  - ・働きたいと本人が希望したら、迅速に就労支援サービスを提供する
  - ・職業後のサポートは継続的に行う
- [http://www8.cao.go.jp/youth/kenkyu/h19-2/html/5\\_6\\_9.html](http://www8.cao.go.jp/youth/kenkyu/h19-2/html/5_6_9.html)

IPSの実施方法は細かく定められてもいますが、大事なのは原則です。障がい者に関する制度や就労支援の制度は国によって違います。ですから、応用に当たって大事なのは、異なるコンテキストでも原則を重視することです。たとえば、IPSはアメリカでは専門職がやっていますが、静岡方式は市民がやっています。

**静岡方式**

**シンプルに言えば市民によるIPS**

エビデンスを「つかう」というとき、医療介入のように、インプット自体が確定している介入(たとえば、医薬品)を持ってきてそのまま実施するというのではなく、対人介入においては、できるだけ頑健な(robust)、つまりコンテキストによって効果が変わらない介入を探し、本質である理論部分を自分たちのコンテキストに応用すべきだというのが私の考え方です。私の考えとは異なり、エビデンスを「つかう」と言っても、プロトコルの決まった厳密なエビデンス・ベースド・プラクティスと言われるものを、現場に持ってくるという考え方もあります。私としては、このようにエビデンスを「つかう」と言っても、いろいろな立場があるので、いろいろと知った上で議論すべきだと思っています。

おそらく、現場でエビデンスは敵だ、という人がいますが、そういう人はすごく厳密なプロトコルに従ったプログラムをイメージしているのではないかと思います。例えば、学校の授業で、この教え方で教えれば効果が出ますというタイムスケジュールまで決まった教え方にするというエビデ

ンスに従い、それをやりましょうといったものです。しかし、実際は教師全員がそれをやるのはしんどいです。ヒューマンサービスの現場では、理論がしっかりしていて、コンテキストによって効果がばらつかない介入を見つけ出して、それを実践しながら、自分たちのものにしていく、というのがいいというのが私の考え方です。

エビデンスを「つかう」とは？

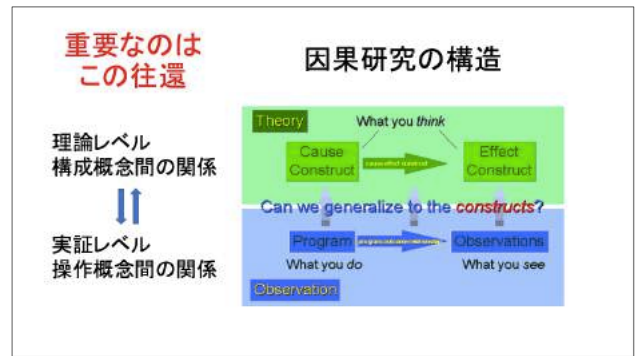
**エビデンスを「つかう」とは何か？**

エビデンスをつかうとは  
 replicability 再現可能性  
 generalizability 一般化可能性  
 を問うことである

エビデンスを「つかう」とは、再現可能性（同じことができるか）と一般化可能性（コンテキストを変えたときにうまく行くか）を問うことであると思います。

### 因果研究の構造

僕の頭の中にいつもあるのはこの図です。介入プログラムをやってみて（What you do）、目に見えるもの（What you see）を観察する。理論というのは抽象的に存在しています。こうなったらこうなるだろう、という抽象的な予測が理論です。



先ほどの例では、IPS という観念があって、これは9つの原則で書くことができますが、実際に実践してみて、実際に就労状況を観察しないと数字になってきません。実務家というのは、緑の部分（理論レベル）を頭に置いて、青の部分（実証レベル）をやる人だと思っています。青をやってみて、緑を修正していく人でもあります。緑は強固なエビデンスに支えられているとこともありますが、実際にやってみて理論を修正して行くこともあります。「理論レベル」と「実証レベル」の往還という行ったり来たり、をやるのが考える実務家。ここでは、これを反省的実践家と呼びます。緑の部分が決まっているから、それを実行しなさいということではありません。下（青の部分）をやっとうまくいっているから、上（緑の部分）は関係ないということでもありません。

自分のところの介入がうまくいっている、良いと信じるのはいいんだけど、「その介入に一般化可能性がない」のでは、意味がわかりませんよね。社会的インパクト評価で、自分のところの実践は価値がありますと事業（体）の単位での効果を主張する、つまり、青の部分の実践の世界だけで、緑の理論を抜きにして主張するというのは論理的にはよくわからないと感じています。

### 私の基本的な考え方

お茶は、かつて特定の高位の階層の人だけが飲んでいました。研究を通じてではありません

が、いいものだから、結局は、社会全体に広まっていきました。エビデンスというのは、そのときは、特別な人しか実践できないように思われても、いつかは当たり前になってくるものです。

**私の基本的な考え**

あらゆる「専門的な」介入は一般化する  
お茶の歴史のように

支援のプロトコルではなく  
原理／理論を応用する  
再現可能性／一般化可能性を抽象的に担保

反省的実践家は理論と実践を往還しながら進む  
「考える実践家」

エビデンスが、誰にでも使えるようになるように、社会の進化を進めて行くのが実務家です。

### 反省的実践家のセンスを支える道具

実証的知見、まず、エビデンスを作らなければいけない。そのための研究手法には、コホート研究や RCT、系統的レビュー（Systematic Review：SR）などがあります。

**反省的実践家のセンスを支える道具**

**実証的知見**

- ・コホート研究（以下、エビデンス）
- ・RCT ランダム化比較試験
- ・SR 系統的レビュー




それらから得られるエビデンスを使っていくにあたっては、まずは、個々の論文が読める力である、**批判的吟味**が必要です。これに加えて、専門分野の**理論的知識**が必要です。たとえば、就労支援をするなら、就労支援の理論的知識を、持たなければいけません。そして、その上で、**実践**をするわけです。

**反省的実践家のセンスを支える道具**

**エビデンス・リテラシー**

- ・批判的吟味
- ・理論的知識
- ・実践



エビデンス・リテラシーというのは、批判的吟味の力だけでなく、その分野に関する理論的な知識と実際やってみてのフィーリングの3つのことだと思っています。これらを合わせて、エビデンス・リテラシーと私は考えます。今日は、このうち最初の批判的吟味をやります。

### 批判的吟味

**導入**

**研究の流れと四つの妥当性**  
キャンベル→クック→ロッシの伝統

問題の定式化 → 標本抽出 → 測定 → デザイン → 分析 → 結論

研究プロジェクトのステップ	研究のよき四つの側面(四つの妥当性)	概要
標本抽出	外的妥当性	他の時間、対象、場所へ一般化できる程度
測定	構成概念妥当性	構成概念を忠実に操作概念に移している程度
デザイン	内的妥当性	因果関係の吟味を漏れなく行っている程度
分析	統計的結論妥当性	仮説検定を正しく行っている程度

先ほど、本をお見せしましたがけれども、四つの妥当性という、研究の良さを表す観点があります。研究を読む側から言えば、研究の良さを確認するために行う批判的吟味の観点と言い換えても良いです。このテキスト「医療専門職のための研究論文の読み方」は、必ずしもこの四つの妥当性に沿

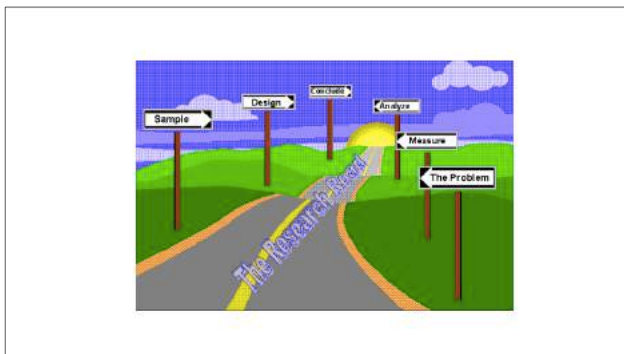
って整理をされているわけではありませんが、この四つの妥当性という観点から、研究をチェックする癖がついていれば、この四つの妥当性を下敷きにしていることがわかる本です。

外的妥当性というのは標本抽出がちゃんとできているか、構成概念妥当性というのは測定がきちんとできているか、内的妥当性というのはデザインがちゃんとできているか、統計的結論妥当性というのは分析を正しくできているかを表します。RCTはデザインがいい研究です。つまり、内的妥当性、因果命題が実証できているということです。

## 研究のステップ

僕の研究室はカメムシがたくさん入ってくるんだけど、どうやったら僕の研究室からカメムシがいなくなるか（問題設定）。じゃあ、僕の研究室を標本にしよう決めて（標本抽出）、次に、データを実際に測ります（測定）。デザインはあらかじめ考えておくものですが、ここで得られたデータがあるデザインに放り込みます（デザイン）。そして、実際に数字を回してみても（分析）、結論を出します（結論）。

例えば、就労支援についてその対象者を決めて、データを測ってみて、RCTのデザインに放り込んで、分析してみる。これは僕が大好きな、研究のプロセスを表している絵なんですけれども、こうして野越え、山越え、研究が進んで行くと、素晴らしい景色が待っています。



この標本抽出から分析までの4つのステップそれぞれに欠点（＝研究の失敗）があり得ます。その欠点のどれかが深刻な欠陥だと結論を信用してはいけないということになります。

## 研究のタイプ

研究の種類は、「記述研究」「関係研究」「因果研究」の3種類あります。

### 実証研究（とりわけ関係研究と因果研究）

理論が正しいかどうかをデータを用いて確率的に判定する研究

研究のタイプ	妥当性
記述研究: 一変数 定数 構成概念 操作概念	構成概念妥当性
関係研究: 二変数(以上) 関係仮説 仮説検定(対立仮説 帰無仮説 棄却 採択)	+統計的結論妥当性
因果研究: 二変数(以上)(独立変数 従属変数) 因果仮説 因果の方向性	+内的妥当性

\* 外的妥当性はすべてにかかわる

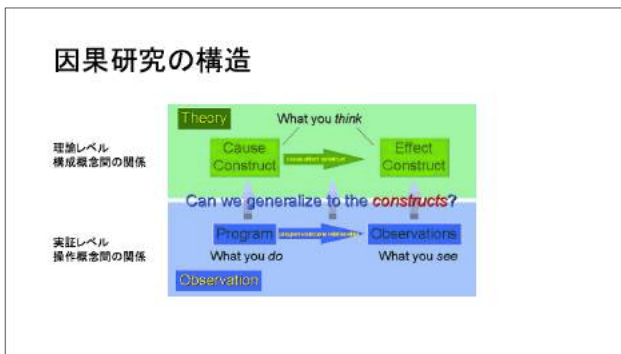
「記述研究」は、数えるもの、いわゆるデータを取る研究です。測定をしなければいけないから構成概念妥当性が問われます。モチベーションをしりたいのであれば、どうやってモチベーションを測るかということを考え、たとえば、質問紙にチェックをつけてもらうということを行います。構成概念妥当性は、その質問紙は、本当にモチベーションを測れているのかということを行います。モチベーションという抽象的な構成概念を、操作概念としての数字に落としたときに、操作概念としてのモチベーションが、適切に構成概念を反映しているという測定の良さを表します。構成概念妥当性は、理論と実証の間の縦関係を担保している非常に重要なものです。

「関係研究」というのは、AとBに関係がありますか？という研究です。たとえば、今朝ここに到着した時刻はモチベーションと関係がありますか？という問いに答える研究です。関係の有無というのは、統計的に偶然の範囲内（関係がない）か？偶然の範囲を超えている（関係がある）か？



を判断するということで、これは仮説検定の仕事です。

さらに、どっちがどっちの原因なのかを知りたい、朝早く来たからモチベーションが高いのか、モチベーションが高いから早くきたのか？という方向性を決定したいのなら「因果研究」ですから、内的妥当性が問われます。この3つの研究は、一つ一つ積み上がって行くものです。



同時に最後、外的妥当性というのは3種全ての研究にかかると書いていますけれども、これは「一般化可能か?」「特殊なサンプルではないのか?」ということです。何かを測ったら、常に外的妥当性が問われます。外的妥当性、つまり、一般化可能性は、3種の研究すべてに問われる、非常に重要な問題です。

エビデンスというのは、内的妥当性を優先するので、よく批判をされるのはこの外的妥当性が不足しているという点です。あるところでうまく行っていることが、ほかでうまくいくのかという問いです。実際のところはやってみるしかありませんが、RCTへの批判の多くは外的妥当性が低いことに関するものです。しかし、RCTでない研究も、特定のサンプルを使ってやっているわけですから、この批判には、納得できないところもあります。

### インパクト評価と、社会的インパクト評価

$$\text{Intervention (X)} \xrightarrow{\beta} \text{outcome(O)} \xrightarrow{\gamma} \text{social impact(SO)}$$

掛け算なので、 $\beta$ がゼロならば、社会的インパクトは当然ゼロ。しかし  
 ・「Outcome」1単位が達成されたら、これだけのSO(金銭換算)が得られる(はず)と主張されることが多い。  
 ・OとSOの関係は、実験的には(内的妥当性の担保を持って)証明されていない。

昨今の社会的インパクト評価を僕なりにこの文脈で定義をすると、介入(X)→アウトカム(O)という効果( $\beta$ )まではエビデンスが責任を持つのですが、アウトカムの値打ちについて、社会的にどれだけの価値を明らかにするために $\gamma$ をかけていくのが社会的インパクト評価です。社会的インパクト評価では、因果研究をきちんとやらないで、つまり、前半の内的妥当性が低いままで、 $\beta$ と $\gamma$ をかけ合わせるということが、しばしば行われます。しかし、 $\beta$ に誤差が入っていると何の意味もなくなってしまいます。つまり、内的妥当性を担保しないまま、社会的インパクト評価を推定するのは不思議だなと思っています。投資の方から入ってきた研究者の方には、因果研究を軽視する方もおられますが、評価研究をやる最大の理由は、人を不幸にしないことだと考えると、倫理性を欠いているのでは、とも感じます。

### 系統誤差と偶然誤差

#### 概念整理

	記述研究	関係研究	因果研究
外的妥当性	○	○	○
構成概念妥当性	○	○	○
内的妥当性			○
統計的結論妥当性		○	○

	系統誤差	偶然誤差
外的妥当性	○	○
構成概念妥当性	○	○
内的妥当性	○	
統計的結論妥当性		○

誤差には、2種類あります。「系統誤差」は、一貫して生じる誤差です。最近の若い方には、伝わ

らないことがある例えですが、私は、昔、体重計に乗って降りてみると針が「0」を指していないということがよくありました。自分の体重が増えたのかなと思ってんだけど増えたわけじゃない、体重計の針が最初からずれているわけです。何度乗っても偏りがある、というのが「系統誤差」です。

一方、体重計を降りてみると針は「0」になっているけれど、測るたびに微妙に違いが出るということも昔はありました。そうすると、何度か乗ってみて、子どもなりに平均値を出していました。平均値をとることで、打ち消しあってくれるのが「偶然誤差」です。「偶然誤差」はサンプルを大きくすれば、つまり、何度か測れば対処可能なので、系統誤差の方が罪は重いです。

## テキスト2章 論文の構造

さて、論文の構造は、基本的にタイトル・要旨・はじめに・方法・結果・考察から成っています。

### 2章 論文の構造

- タイトル／要旨
- はじめに
- 方法
- 結果
- 考察

西郷さんが訳してくれた論文は、方法と評価尺度が同じ字の大きさになっていますが、英語を見ていただくと、METHOD (方法) という項目の中に Measure や Procedure といった下位項目があるのが分かると思います。論文の各部分と、どの妥当性をどこでチェックするかはだいたい対応しています。

## テキスト3章 研究の方法を見分ける

### 3章 研究方法を見分ける

- サーベイ: 現状がどうなっているのか
- コホート研究: 患者に何が起きるのか
- 臨床試験: ある介入が他の介入より優れているか
- ケース・コントロール研究: ある個体群が他の個体群と異なるのはなぜか

3章は研究の種類についてです。サーベイ、コホート研究などがありますが、簡単にいうと因果研究をちゃんとできるのは臨床試験（ランダム化比較試験）だけです。ケース・コントロール研究はある前提を満たすと、因果関係を明らかにすることは不可能ではありません。しかし、いずれにせよ、臨床試験以外は関係研究です。

今日はこのテキストの4章・6章・9章をやりますが、4章・6章は一般的にすべての種類の研究に関わる内容で、9章は臨床研究のみに関わる内容です。4章・6章・9章をやると、臨床試験の批判的吟味ができるようになります。

では、テキストを読みながら進めていきます。

(講義記録はここまで)

## 參考資料



## 批判的吟味 臨床試験のチェックリスト

### ○本質的な問い

- ・治療をランダムに割り当てているか
- ・研究に含めた、すべての患者について説明しているか
- ・アウトカムは、盲検法によって評価しているか

### ○詳細な問い

#### デザイン

- ・目的を明確に述べているか
- ・サンプルサイズは正当か
- ・測定に妥当性と信頼性がありそうか
- ・被験者の選択方法が治療の効果サイズに影響を与えているか
- ・治療および治療の運営についての記述はあいまいでないか
- ・統計手法を記述しているか
- ・盲検法を使わなかったことで、バイアスを発生させていないか
- ・アウトカムは臨床上の意義があるか

#### 実行

- ・どのようにランダム化しているか
- ・研究中、予期せぬできごとが発生しなかったか

#### 分析

- ・ベースライン時点で、治療群同士は比較可能か
- ・計画した治療からの逸脱を報告しているか
- ・治療の意図にしたがって結果を分析しているか
- ・統計的優位性を判定しているか
- ・基本的なデータを適切に記述しているか
- ・数字はつじつまが合うか
- ・副作用を報告しているか

#### 解釈

- ・主たる知見は何を意味するか
- ・有意でない知見をどのように解釈しているか
- ・重要な効果を見過ごしていないか
- ・結果は先行研究と比べてどうか
- ・自分の実務にとって、研究結果はどのような意義があるか

研究のステップ

研究のステップ	各ステップの良さ*	概要	記述研究 x,y,z	関係研究 x⇔y	因果研究 x→y
1.標本抽出	外的妥当性 (母集団を推定できるか?)	他の時間、対象、場所へ一般化できる程度	○	○	○
2.測定	構成概念妥当性 (測るべきものを測っているか?)	構成概念(抽象)を忠実に操作概念(具体)に移している程度	○	○	○
3.デザイン	内的妥当性 (インパクトを正しく求められているか?)	因果関係の吟味を偏りなく行なっている程度			○
4.分析	統計的結論妥当性 (O と X の関係の有無に関する結論が正しいか?)	仮説検定を正しく行っている程度		○	○

\*「良さ」とは、誤差の小ささを指す。誤差が少ないか見ていくために4ステップある。

誤差の種類

誤差の種類	説明	対処	具体例
系統誤差 ( <i>systematic error</i> )	本来の値からの一定方向へのずれ	確率抽出(等しい確率での抽出)によって対処  (例)無作為抽出	体重計を下りても針が0 kgに戻らないようなずれ
偶然誤差 ( <i>random error</i> )	相当回数計測すると、打ち消し合う誤差	標本サイズを大きくすることによって対処	正しい体重計でも測る度に差が出る

## 結びにかえて

Evidence based Medicine という言葉が使われ始めたのは、1990 年代初めのカナダだったとされています。それから 20 年以上が経ち、日本では政策形成において「エビデンス」を活用する動きが盛り上がりつつあります。

しかし、「エビデンス」の捉え方、「エビデンス・ベースド」への理解、評価の使い方など様々な受け取り方があり、今後現場に混乱を招く可能性もあるように感じています。実践をより良くするためにエビデンスをいかに活用すべきかという議論は、特に対人サービス分野の実務家においてはまだ不十分です。

本書の元となったワークショップは、2017 年の日本子ども虐待防止学会における津富先生のご講演をきっかけに、効果のある支援プログラム自体を広めるのではなく、理論レベルと実践レベルの往還ができる反省的実践家（reflective practitioner）の育成に貢献したいと考えて企画しました。本書は、その一部だけでも広く多くの方々に知っていただきたいと考え、まとめたものです。

実践を良くしていくためにどうしていきべきかという問いに対して、唯一の答えがあるわけではありません。しかし、目の前のクライアントがより良く生きるために、実務家がこれまで積み上げられてきた知見から学び、最善の取り組みを考え続けることは不可欠であるように思います。それを支えるための道具としてエビデンスがあり、それを応用できる実務家と研究者によって、さらに新たな知見が積み上げられ、実践が少しずつよくなっていくという循環が続いていくことを願っています。

最後に、貴重な学びの機会を提供して下さった津富先生、協力団体である三菱 UFJ リサーチ&コンサルティング、エビデンスに基づく教育研究会、エビデンスに基づく政策と実践センターの皆さま、森先生、石井先生、徳永さん、家子さん、そして参加・協力いただいた皆さまに心より御礼申し上げます。

Early Intervention 勉強会 西郷 民紗

2018年12月制作

発行：Early Intervention 勉強会  
early.intervention.japan@gmail.com

ご使用・転載ご希望の方はご一報お願いします。