

原稿作成日： 2017年3月31日

## 研究の再現性の適正な表現と信頼性

研究発表の信頼性にはその再現性の限界を正直に伝えることが欠かせない

〈教材提供〉

AMED 支援「国際誌プロジェクト」 提供

無断転載を禁じます

### 草案

市川家國 信州大学特任教授

### 査読

西中村 隆一 (熊本大学)

黒木 登志夫 (日本学術振興会)

波多野 和男 (アステラス製薬)

太田 雅之 (日立製作所)

新田 孝作 (東京女子医科大学)

中川 敦夫 (慶應義塾大学)

Robert Geller (元東京大学)

William Currie

## 目次

### はじめに

#### 社会資源を浪費させる再現性の低さ

##### 「再現性」とは

生物学的な理由により、生命科学研究に完全な再現性は望めない

統計解析で再現性の度合いを定量的に表現する

低い再現性の原因となる研究者の問題

#### 信頼性を高めるために研究者が行うべきこと

個々の実験結果に影響を与える生物学的要素

研究の計画と実施

研究結果の発表

- A. 統計解析
- B. 画像処理
- C. 利益相反
- D. 公開・共用・提供

欧米公的機関のガイドラインと国際誌

- 1. 実験動物について
- 2. 研究プロトコール
- 3. バイアスを防止するためにとられた方策
- 4. 結果
- 5. 考察

#### 領域・地域横断的な共同研究発表

政府および研究機関等の役割

まとめ

参考文献

## はじめに

科学は多くの研究者による研究成果の積み重ねによって進歩してきました。その途中では研究手法の欠点、あるいは得られたデータの解釈の誤りから間違った結論が流布してしまうことも数多くありました。しかし、それらの誤った結論の多くは他の研究者、時に本人の手による研究によって否定され、次の進歩を生み出してきました。すなわち、科学は自浄作用を発揮しながら進歩してきたわけです。



残念なことに、自浄作用によって誤った結論が正されるまでには多くの社会資源と時間が費やされます。科学が自浄作用を持つことは過去も今日の科学も変わりはありませんが、昨今の特徴はその進歩が極めて速いことです。また、科学研究には、極めて多くの社会資源が投入されるようになりました。アメリカを例にすると、過去 50 年間で GDP に対するアメリカ国立衛生研究所 (National Institutes of Health, NIH) の予算は倍増し [1]、さらには企業による研究開発費も急増し、現在、医学領域では公的資金の約 2 倍の投資がなされています。研究機関においては、科学の進歩を促す動きは個々の研究者の業績を雇用・昇進へと積極的に結び付ける動きとして現れています。研究者は業績を積み上げるために研究費獲得に向けた激しい競争を繰り広げているのです。

科学に急速な進歩が期待され、そこに多くの社会資源が投入される一方、仮説の検証と、結論の誤りを正すためにも大きな社会資源と時間の浪費を伴うようになりました。この浪費は、科学の歴史を通して公表された多くの仮説や結論が正しくなかったと最終的に証明されたことからすると、避けることはできません。正しくなかった仮説のほとんどは、研究者の誠実な誤り (honest error) によるものです。しかしながら、誤った仮説が誠実な誤りでなく、研究不正によるものだとしたら、許されざることです。したがって、この大きな浪費は、近年研究不正の問題が明るみになる中で、一般社会においても広く認知されるようになりました。さらに最近では、意図的な研究不正以上に社会資源を浪費するものの存在が明らかになってきました。それは再現性の低い研究です。

## 学習目標

- ・ 今日、生命科学研究の再現性とその信頼性が重要視されている背景を説明できる。
- ・ 生命科学研究で生じる結果の再現性の低さについて、研究内容に応じた理由を説明できる。
- ・ 研究結果の再現性を下げる原因となる研究者の行動の例を複数挙げるができる。
- ・ 生命科学研究の再現性を上げるために、研究者ができることを説明できる。
- ・ 共同研究の発表誌をめぐる研究者間に生じる意見の相違について説明できる。

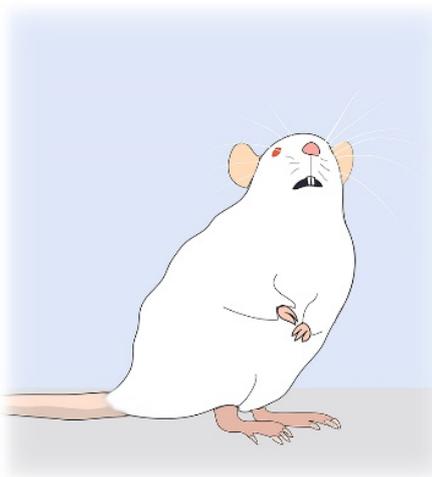
---

## 社会資源を浪費させる再現性の低さ

生命科学研究における再現性の低い研究は、社会資源を浪費させます。例えば、新薬の開発では動物実験等続いて人を対象とした治験を通じて行なわれるのが通常です。Hayらの報告によると、初期に期待された試験薬の中でも、その効果と安全性が治験で確認され、実際に市中に出回る率は平均 10%台ということです[2]。新薬開発費が 200~300 億円という事実を照らして見ると[3]、これは極めて深刻な問題であることは想像に難くありません。

動物等を用いた研究 (preclinical research) は治験を開始する上で土台となりますが、がん領域の動物実験については再現性の低さが指摘されています[4]。そうした再現性を欠く研究に投入された研究費の総額は、米国だけでも NIH 予算とほぼ同額の年間約 3 兆円と推定されています[5]。こうしたことを背景に治験を行う上で土台となる動物実験等の再現性を改善する必要性が各方面で指摘され、それに向けた具体的な動きが起きています。

一言に再現性といっても、研究内容に応じて様々な意味合いがあります。生体内 (イン・ビボ) 実験、試験管内 (イン・ビトロ) 実験の違いは当然のことです。他にも、マウス・ラットといった動物と人間とは種が異なりますから、そもそも動物実験結果がそのまま治験に反映されなければならないというものではありません。それに加え、動物にせよ人間にせよ、生体に特有の個性がありますから、動物実験自体の再現性が低いことは避けられませんし、ある臨床研究で示された試験薬の効果が他で必ずしも再現されないということありえます。とはいえ、このような事情を踏まえ



たとしても、物理学や化学といった他の領域と比較して、生命科学領域の研究の再現性は際立って低いのです[4]。再現性の問題は医学研究、生命科学に特有のものではないものの、社会資源の浪費の問題を考慮すると、改善が求められるところです[6]。

---

## 「再現性」とは

「再現性」と言っても、生命科学的な意味合いには極めて広いものがあります。本単元では再現性を下記のように分類して定義します。

- Replicability: 同じ研究者あるいは研究室が同じ材料と方法を使った場合に同じ結果が観察される場合。
- Reproducibility: 他の研究室がほぼ同じ材料と方法を使った場合にもほぼ同様な結果が観察される場合。当然のことながら、Reproducibilityが有るためにはReplicabilityが有ることが前提です。したがって、Replicabilityを確認しておくことはReproducibilityを確保する上で重要です。
- Generalizability: 条件を変えて研究を行った場合でも、元となる研究の結果とほぼ同様な結果が得られるとして、観察された現象が一般化できる場合。Generalizabilityが有るためにはReproducibilityが有ることが前提です。したがって、Reproducibilityを確認しておくことはGeneralizabilityを確保する上で重要です。
- Translationability: 動物での実験結果が人においても同様な形で観察される場合。Translationabilityが有るためにはGeneralizabilityがあることが多くの場合必要です。そうした場合、Generalizabilityを確認しておくことはTranslationabilityを確保する上で重要です。

Reproducibilityの意味にはいろいろあり、「同じ材料」と言ってもロット番号の異なる薬品、同じ動物株でも仕入先が異なる場合があり、同じ薬品を同じ動物株に使ったつもりでも、これらの違いのために実験が再現できない、という場合もあります。Generalizabilityも同様で、その意味には幅があります。研究で得た結論が年齢によらず一般化できる場合もあれば、「若者に限って」



といった条件内での一般化が可能という場合もあります。

---

## 生物学的な理由により、生命科学研究に完全な再現性は望めない

生命科学研究では、同じ研究者が同じ材料・同じ方法を使っても以前の結果が再現できないことがあります。同じ条件と思っても、実験を行う季節、時間、温度・湿度、微妙な取扱いの違いが影響を与えることもあるでしょう。また、重要な生体内外の条件が異なっても、今日の科学では察知できない、ということもあるでしょう。生命科学研究では個体に生じる現象を定量的あるいは定性的に評価する際にしばしば抗体が用いられますが、その抗体自体も生物から作製されるものであり、質の完全な均一性を保証できるものではありません。

---

## 統計解析で再現性の度合いを定量的に表現する

近年、生命科学研究における曖昧性を回避するために、生物統計学が発達してきました。統計解析は、まさにデータの再現性の度合いを定量的に読者に伝えるための手段になります。

具体的には、研究からもたらされる結論を表現するにあたって、生物にある個体間の不均一性に由来する結論の不確かさの程度を、「P値」あるいは「信頼区間」といった数字として結論に添えるようになりました。このようにして再現性の度合いを具体的に呈示することで、生命科学研究は研究内容の信頼を得ているのです。

統計解析は用いたモデルには含まれてないシステムエラーを説明することはできません。そのため、正しく遂行した場合でも、統計解析は結果の不確かさを見積もる「ベストケース」であることを、研究者は理解すべきです。研究者には、統計を単なる「ブラックボックス」として用いるのではなく、その基盤をも理解することが望まれます。



## 低い再現性の原因となる研究者の問題

再現性を欠く研究については研究者側にも注意が必要です。例えば以下の行動は再現性を低下させると言えます。



- データの収集法および統計解析法を予め設定していない。
- 想定されるバイアスを考慮し、正しく研究を計画したり、結果のバイアスなど疫学的な視点から評価できるだけの知識がない。
- 予め設定したデータの収集および解析に関するプロトコールから逸脱して研究を行う。
- 好みの実験結果を得るまで方法を変えて実験を繰り返したり、都合の良い統計解析方法を選ぶ。
- 統計上、何か有意なものが得られるように、数多くのパラメーターを測定し、統計上有意なP値が出たものの中でも都合の良いパラメーターだけを発表する。「P ハッキング」、「データ漁り」と呼ばれる。
- 仮説検証的研究の体裁を装った研究で、統計解析も済んで、結果が明らかになってから仮説を用意するもので、「HARK (Hypothesizing after results are known)」と呼ばれる。これは「P ハッキング」に伴って行われることも多く、事前に設定した仮説が証明されたかの如く発表するもの。いわゆる「探索研究」を、「仮説検証研究」の如く発表することがこれに相当し、このように偽装して発表されたデータには他の研究者にとって再現性がない。
- ネガティブな研究データを公表しない。ネガティブな研究結果はインパクトの高い学術誌には発表しにくいため、発表を控える。また、以前からの自分の仮説を覆す研究結果の発表をはばかる。
- 特殊な条件下での実験結果をあたかも Generalizability を有するかのよう  
に報告する。例えば、雄からのみ得たデータを発表する場合、動物の性別を記載しない。
- 実験結果の再現に必要なだけの詳細な研究方法等の情報を公開しない。

これらの行為の背景としてあるのは：

- 「就職、昇進あるいは研究費を獲得しなくては」、更には「注目を浴びる研究結果を出さなくては」、というストレスに曝される中で陥る次のような自己欺瞞

(self deception) です。

- ✓ 自分の仮説に合致する所見を探し回り、矛盾した所見が目に入らない。
- ✓ 予想外の現象に専ら注目してしまう。
- ✓ どのような現象が観察されても、現象にもっともらしい理由をつける。

---

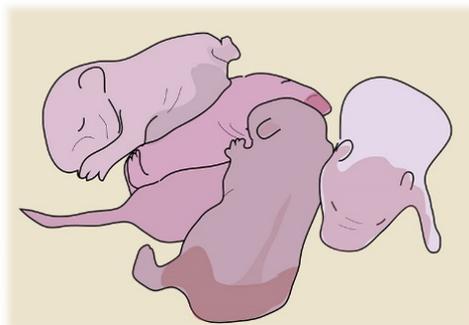
## 信頼性を高めるために研究者が行うべきこと

自らが発表した研究成果は、他の研究者によって再現されて初めて科学的価値が生じる、という原則を認識することが重要です。したがって、他の研究者が研究結果を再現できるようにするために必要な情報を適切に公開することは研究者にとって極めて重要になります。一方、再現性に難点があっても、研究の信頼性を確保することができます。それには、適切な統計解析を行い、その結果を含め、研究内容を余すことなく開示するなど、研究結果の再現性の度合いの正確な伝達に努め、また利益相反の存在など、バイアスが生じうる要因を開示しておくことです。

---

## 個々の実験結果に影響を与える生物学的要素

生命科学実験で見られる現象に影響を与える要因には多くのものがあることを認識しておく必要があります。動物実験の場合は、動物に起因する多くの要因が加わります。



### 動物

- a. 同じ動物種であっても、株（ゲノム上の違い）、性別、年齢、体重、育てられた環境（含：ラットのブリーダー）
- b. 実験時の環境：季節（例：繁殖期）、気温、湿度、明るさ、時間、食事との時間的間隔、周囲の音・におい

### 実験者

- c. 実験者、動物舎職員の動物の扱いの熟練度、扱い方

### 餌・薬品・薬剤

- d. 製造元
- e. ロット
- f. 保管方法・保管期間
- g. 投与方法

培養細胞を用いた in vitro 実験においても同様に幾多の要因が存在します。

このような各種の要素をできるかぎり均一にして実験を行えば、それだけ個々から採取されるデータも均一性が期待できるわけですが、研究目的によっては幅広い条件下で共通に見られる現象を捉えることを目的として、故意に幅広い条件下で実験を行うこともあります。いずれにせよ、再現性を高めるには、それらの条件を個々の実験について詳細に記録して発表すること、そして、ポジティブ・データ発表の前、中、あるいは後に、もし異なる条件下でネガティブな結果を得ていた場合には、それについても詳細な点まで発表する、あるいはアクセスできるようにしておくことが望まれます。

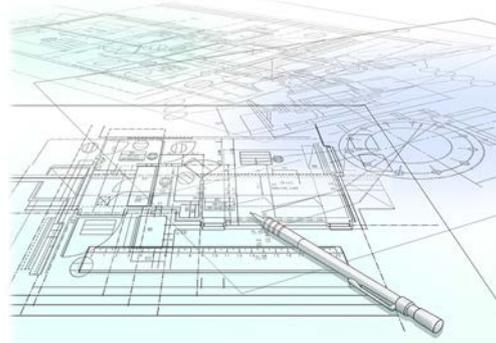
一方、実験動物の体調は飼育状況によって左右されるわけですから、動物舎職員はチェックリストなどを用いて動物の体調を常時記録して管理するとともに、研究者と動物舎の職員とは緊密な情報交換を行い、研究者からは実験目的と内容を、また動物舎職員からは日常の動物の体調を示す情報を提供することが大切です。獣医等の動物舎職員は、実験計画を立てる上で、時に有益な知識を提供してくれるものです。

動物舎においても実験室でも動物にストレスがかからぬよう、丁寧に、優しく、接することはいうまでもありません。それは価値あるデータの採取にもつながるのです。

---

## 研究の計画と実施

生命科学の多くは、母集団から一定数の個体を抽出して観察し、そこに見られた傾向を母集団にも見られるものか否かを推論することを目指しています。その推論がどの程度妥当であるかを数量的に表す学問が生物統計学です。近年、この生物統計学の飛躍的な進歩により、生命科学は、研究内容の再現性を客観的に示すことが可能になりました。人体に関する種々の研究結果が Nature、Science といった一流の自然科学誌に掲載されるようになったのは最近の事です。したがって、生命科学の研究にとって、適切な統計解析を実施できるかどうかは、死活問題です。不適切な統計解析を行うことによって、母集団の持つ性質や反応について誤った推論を行うことになれば、同じ母集団を用いて研究する他の研究者が結果を再現できないのは当然でしょう。誤った推論を最小限にとどめる上で役立つ研究者の心がけを紹介します。



- a. 自分とは異なる学説を唱える研究者を共同研究者に招く。

- b. 仮説には、対立する仮説を同時に用意し、この2つの是非をデータで検討する。
- c. 個体間・研究毎のばらつきの中で十分な検出力を得ることができるサンプルサイズをデータ採取前に決める。
- d. 最新の生物統計学に関する知識を身に着ける。
- e. 統計解析法を予め決めて公表する。不安があれば専門家に相談する。
- f. “異常値”が出た場合の処理の仕方を予め決めておく。

そして、研究を実施する際は：

- g. サンプルを採取するに当たっては、無作為化 (randomize) する。
- h. 群分けして、群の間の違いを観察する研究では全てのデータの解析が終了するまで、個体の所属を明らかにしない、盲検法 (Blind test) を採用する。
- i. 予め決めた統計解析法を用い、変更しない。

すなわち：

- 1) 実験を開始する前に以下を決めておく。
  - ・ 仮説の是非を結論付ける上で重要な項目
  - ・ データの採取法と解析法
  - ・ 許容できるシグナル・ノイズ (ばらつき) の比に基づいたサンプル数
- 2) 実験に当たっては、上記の予め計画したプロトコールに厳格に従って実施する。

したがって、予定の解析法を用いた結果が自分好みのものでなかったというだけで、実験を繰り返したり、予定外の解析法を用いたりすることはNGです。そのようにする場合は、実験の報告書やレポートの中で、理由を明確に述べる必要があります。理由を明らかにしない場合は研究不正に繋がります。

---

## 研究結果の発表

学術誌の規程に準拠することは、論文が誌上発表される上での前提ですから、準拠せずに投稿すれば掲載申請を却下されます。もし掲載されてしまった場合には、規程からの逸脱の程度によっては修正ないし撤回が要求されます。それぞれの学術誌の読者は著者がその学術誌の投稿規程を準拠しているとの想定のもとに論文を理解しますから、規程に反した論文は誤解を招きます。それゆえ、そうした論文は撤回、更には不正認定に至ることがあるのです。近年、製薬企業N社の支援下で発表された降圧薬バルサルタンに関する論文が撤回に至った直接の原因は、利益相反に関する当時の国際誌の投稿規程に違反していたことにあります。当時の日本では、国や機関のガイドラインとして、そうした規程が今日のような形で周知されるには至っていませんでした。

学术论文の読者が、「準拠しているはず」、と期待するルールは、当該誌の投稿規程に書かれていますから、投稿者は投稿誌の規程に精通している必要があります。投稿規程にはオーサーシップや著作権など複数の項目について述べられていますが、研究の信頼性に直接影響を与える発表項目として次のものがあります。

- A. 統計解析
- B. 画像処理
- C. 利益相反
- D. 公開・共用・提供



#### A. 統計解析

いかに適切な統計解析の方法を選び、その結果をいかに適切に表現するかは報告する研究結果の再現性に極めて重要な役割を果たします。データが正規分布を示さない場合、サンプル数が少ない場合にどのような補正を行うか、探索研究の中で多くの項目について比較した場合、P値にどのような補正を行うか、そしてデータのばらつきをどのように表現するか、といった点について、詳細な指摘を行い、著者にそれらについての開示を求める学術誌も出てきました[7, 8]。学術誌によっては、人を対象とした研究か否か、あるいは比較する内容が質的なものか量的なものかといった研究の性格の違いによって、それぞれに共通なガイドライン[9]を提示し、これらを通じて、統計解析方法の開示など研究の透明性や再現性の確保を求めているものもあります。

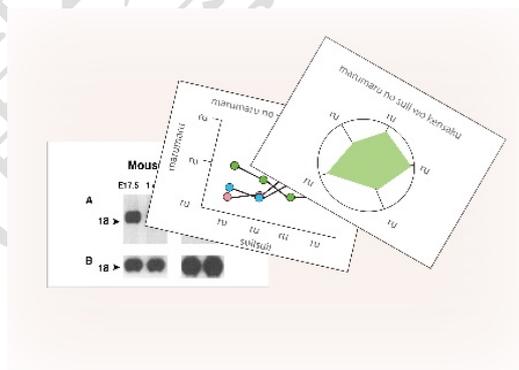


最近の *New England Journal of Medicine* に「[チョコレート消費と、認知機能とノーベル賞](#)」と題した興味ある記事が掲載されています[30]。この記事のもとになったデータは以下のものでした。それぞれの国における1000万人当たりのノーベル賞受賞者の数を縦軸に、横軸をそれぞれの国のチョコレートの消費量としてプロットしたところ、相関の程度を表すピアソンの相関係数が0.791、統計的有意差を表すP値<0.0001と、極めて高いレベルの有意差を示しました。

これをもとに、「このデータによって、チョコレートを沢山食べるほど賢くなることが証明された」と結論すれば、良識ある者であれば、それはおかしいと考えるでしょう。これはエコロジカル・ファラシー（生態学的錯誤）と称されるもので、疫学分野ではよく知られるバイアスの一つです。「エコロジカル・ファラシー」は集団レベルでは成り立っている事柄や認められた関連を、その内部のより小さな集団や個体に当てはめると誤りとなることを意味する語句です。ノーベル賞の授賞者はスイス、スウェーデン、オーストリア、デンマーク、ノルウェイなどに多く、それらの国ではチョコレートが好き人が多いのですが、集団（国レベル）でノーベル賞やチョコレートの消費量が多いからと言って、そのことを個人レベルのエビデンスとして「チョコレートを食べるとノーベル賞を取る確率が高くなる」、と言った結論に利用すると、間違った結論を導いてしまう。こうしたことは、実はよくあるのです。

## B. 画像処理

生命科学研究の発表では、口頭にせよ学術誌上にせよ、研究で得られた所見の信ぴょう性を読者に訴える手段として各種の画像が用いられてきました。そうした中で、発表者の結論をより強く読者・聴衆に訴えるために好みの画像を多くの中から選んで提示するという手法、つまり代表的な画像ではなく、意図的に選択した画像を提示するということが従来しばしば行われてきました。その行為は、「意図的に画像を選択した」として提示する限りは問題ないのですが、これを「代表的画像」のごとく提示すれば、他の研究者による研究結果の再現は期待できません。



近年は画像を電子化して記録する技術が生命科学の分野に全面的に取り入れられたことで、画像に手が加えられて、読者・聴衆に誤解を誘導する行為が見られ、しばしば研究不正として表面化しています。画像の中に本来無かったものを加えたり、画像の一部のみに色を加えたり、明るさを増したり、異なる画像を張り合わせるといった手法が典型例です。

今日では、異なる画像を張り合わせを多くの学術誌が禁止しています。一方、画像の一部ではなく全体の色調や明るさに変更を加えることについては多くの学術誌が許容していますが、そうした中でも、オリジナル画像の保存が求められており、提出が求められた際にそれができない場合には掲載を却下される場合があります[10, 11, 12]。

画像とは別に、研究データをわかりやすく表現する手段として各種のグラフが用いられます。以前は、研究者が手描きで行なってきたものですが、今日では、数値データから機械的にグラフを作成するコンピュータ・ソフトができています。そのため、データの正確な表現を目的として、ソフトを用いての自動的な描画を多くの学術誌が要求するようになってきました。これは手描きによる不正を防ぐ一助と見なされています。ただし、グラフ作成ソフトにデータを手入力する際に改ざんが行われうるし、作成されたグラフ画像を編集ソフトで加工することも可能なため、ソフトによって描かれた画像が安心とは必ずしも言えません。

---

### C. 利益相反

利益相反 (COI) それ自体は非倫理的とは見なされません。利益相反の存在を適切に開示しないことで、多くの場合、研究不正へと繋がります。利益相反の典型的な例として挙げられるのが、「研究費」や「コンサルタント料」といった形で企業と経済的関係を持ちながら、当該企業製品に関する研究発表が行われるものです。利益相反の状態があること自体が問題なのではなく、発表者がそうした関係を持つことを読者・聴衆に伝え、彼らに研究内容の客観性を判断する機会を与えるか否かが問題です。その機会を与えないことで、研究は信頼性を失うのです。学術誌によっては、額にかかわらず全ての経済的関係を開示の対象としています[10]、1万ドルまでは開示の義務なし、と具体的な数値を示すものもあります[12]。一方では、「後で存在が表面化した場合、不評をかうような額については開示せよ」、といった漠然とした提示のものもあります[13]。経済関係は数年遡ったものまで記載を求めているものもあります[11]。また、研究の対象になった製品と競合関係にある企業と利害関係があれば、それを開示することを要求する学術誌もあります[10]。その一つが New England Journal of Medicine です。例えば東洋医学の寄付講座の教授は、西洋医学の典型例である抗がん剤の副作用に関する研究論文について、講座を寄付した企業名を開示する必要があります。



最近では、経済的なものばかりでなく、個人的な親しい関係といったものであっても、それが研究者の発表に影響を与えるのでは、と読者・聴衆に危惧させうるものについては、開示を求める学術誌が出てきました[14]。企業から経済的な支援を受けてない場合でも、社員から各種の情報提供や便宜を受けて親しい関係ができている場合がこれに当たります。米国では、企業からの種々の働きかけを排除するため、製薬企業の営業担当者 (い

わゆる MR (Medical Representative) ) のキャンパス内立ち入りを全面的に禁じる医学部も出てきました。

日本においても医学系研究者と企業との結びつきが強いと見られています。この傾向は臨床研究の分野で公的資金が少ないことが一因に挙げられています。また生命科学研究に必要な生物統計学の専門家が企業に偏在するなど、技術的な面で企業からの支援を必要としていることも原因です。こうした状況下で、バルサルタンの統計解析をその薬剤の販売会社に依頼したことが、スキャンダルの発端でした。その製品の販売会社の社員を論文の共著者として掲載するにあたり、その所属を客員の身分である大学とした点で国際誌ルールから逸脱していました[15]。問題なのは、大学の身分を掲載したことではなく、その製品の販売会社の社員であること、すなわち、利益相反の存在を明らかにしなかった点にあります。今日、わが国では数多くの臨床研究が行われていますが、利益相反関係にある製薬企業から必要な技術提供を受けた際には、適切に学術誌上で開示する必要があります。

---

#### D. 公開・共用・提供

再現性を確保する上で、極めて重要なのが研究の透明性です。他の研究者の手によって再現性を確認できるかどうかは、その研究が次のレベルの研究の礎として科学に貢献するかどうかの試金石です。そうした確認のため、また確認を効率よく行うため、更には不正を防止するために、研究発表にあたっては、研究の方法、用いた材料や機器、そして得られた所見の詳細を開示する必要があります。



また、詳細に関する他の研究者からの問い合わせに対しては、積極的な対応が求められています。そうした開示に対する発表者の姿勢は、論文の採択において重要な判断材料とされる一方、不正疑惑が生じた際の判断根拠となるものです。このような点で、特許権が生じる研究に関しては、研究内容の論文等での発表前に特許出願手続きを終えておく必要があります。

---

#### 欧米公的機関のガイドラインと国際誌

欧米の政府等公的機関は研究発表内容に際して、盛り込むべき詳細項目をガイドラインで具体的に提示しています[16-19]。一方、これに呼応するように、基礎医学系および臨床医学系の学術誌も、その投稿規程の中で同様な記載項目を提示しています[10-

12, 14, 20, 21]。

### 1 実験動物について

- ・ 動物種・株
- ・ 動物性別・年齢・体重
- ・ 仕入れ先
- ・ 飼育法（飼育舎環境、餌）

人の疾患モデルとして使われた場合には更に

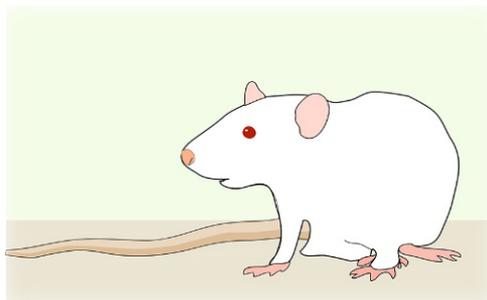
- ・ 実験条件の詳細
- ・ その動物モデルがどのような人の疾患等のモデルとして用いられたか
- ・ 動物モデルの結果を人の疾患等に当てはめて解釈する上での限界や注意点

### 2. 研究プロトコール

- ・ モデルの選択理由
- ・ コントロールの正当づけ
- ・ エンドポイントの決め方とその理由
- ・ サンプルサイズの決め方とその理由
- ・ 実験回数の定義づけ（例：n=XX とする場合、測定回数か、動物数か）
- ・ データ分析および解釈に使われた統計解析法
- ・ 使われた薬剤・処置の種類、投与・施行量、投与・施行方法、投与・施行間隔とその理由づけ

### 3. バイアスを防止するためにとられた方策

- ・ 盲検の方法、あるいはそれを行わない理由
- ・ データの採用および除外基準
- ・ ランダム化ないし階層化の方法



### 4. 結果

- ・ 動物株・種、薬物投与量、実験者など、実験条件をどの程度変えても結果は再現されるのか。関連するネガティブデータ薬物検査の場合は、投与量対反応のデータ（Dose-response results）
- ・ 薬物等の効果実験ではその薬物とそのターゲットに到達したというエビ

## デンス

### 5. 考察

- ・ データに対する異なる解釈の仕方
- ・ データの解釈を結論付ける上で、用いた実験方法の限界
- ・ 実験結果に合致する過去の文献、矛盾する文献の紹介
- ・ 実験結果が人においても実証された場合の、想定される人体への影響の大きさ

学術誌では1論文に割くことのできるスペースが限られており、本文で十分な詳細を述べることはしばしば困難です。したがって、最近では付録（Appendix）の欄を設けて詳細が記載できるようにされている雑誌が増えてきました。また、同時に責任著者（コレスポンドング・オーサー）の連絡場所を記載することが義務付けられ、文章化しえない研究上の「コツ」までを含めた研究方法の詳細を他の研究者が発表者から得られるよう配慮されています。

一方、発表された研究について他の研究者が再現性を確認する上では、研究に用いられたものと同等な材料と機器が必要です。それらが手に入りにくい場合、あるいは特殊なものである場合、費用等の負担は受取り側に求めるものの、それらを発表者側が提供することが求められます。この点に関する発表者の姿勢も論文採択、そして不正審査の際の判断材料となります。こうした材料には、抗体、細菌、ウイルス、DNA クローン、動物・植物株、細胞株などがあります。

さらに情報や材料がより多く、またより広く使われるよう、公開先を紹介し、そこへの提供を指示している学術誌もあります。

---

## 領域・地域横断的な共同研究発表

従来、「共同研究」と言うと、同じ機関内の異なる研究室間、あるいは同じ研究領域内の研究者間で行なわれるのが一般的でしたが、近年の科学研究領域の細分化、そしてそれに伴う研究技術の高度専門化とともに、研究者は共同研究する相手を広く求めるようになりました。今日では、工学系と社会学系、あるいは医学系と経済学系といった異領域の



研究者間による共同研究が機関さらには国をまたいで盛んに行われるようになりました。前者の例は社会インフラの中の交通網整備などの計画に前提として必須なものですし、後者の例は医療保険制度改革の前提として必要なものです。そうした中、学会や研究会はそれぞれに研究における行動規範を成文律あるいは不文律として設けています。一方、国ごとに法律・指針が異なります。また、同じ国内であっても、それぞれの研究機関や省庁が設けるガイドラインは同一ではありません。

そうした領域、地域、そして助成機関がもつ指針等は研究の再現性に関して要求する内容についても異なります。したがって、共同研究においては、研究者間に研究の再現性に関する姿勢の違いが存在するとの共通の認識を持ち、研究に先立って、研究計画・研究実施・研究発表に関する詳細を可能な限り確認しておくことが重要です。

今日、学術誌の種類は極めて多く、扱う研究領域も多岐にわたっています。同一領域内でも、集める注目度の大きさは千差万別で、学術誌発表を求める研究者の幅広い要望に応じています。発行元は学会等の公的なものばかりでなく、商業的なものも数多くあります。中には、論文の再現性や客観性には目をつぶり、高額な投稿費・掲載費を設定して利益追求を専らとする出版社もあります。

研究成果には、生物学的な理由にせよ、人為的な理由にせよ、再現性の度合いについては様々なレベルのものがああります。New England Journal of Medicine、Nature、Scienceなど注目度の高い学術誌は、極めて厳格な研究計画・実施・発表とともに結論に対して多くの検証を要求することで高い再現性を求めています。他方で、Journal of Medical Hypothesis and Ideasといった仮説のみを掲載の対象とする学術誌もあります。そこでは自らの研究データは要求されません。要求される再現性の点で両極端の位置にあるこれらの学術誌の間には、再現性の要求度に関してさまざまなレベルの学術誌が存在しますが、どのレベルの再現性を要求する学術誌に発表するかは、研究者の判断に任せられます。低い再現性であるにも関わらず、研究結果を粉飾してまるで高い再現性があるかのように発表した時、不正が疑われます。この場合の不正には、統計解析の手法を不適切に適用したものと、データの捏造や改ざんによるものがあります。共同研究に当たっては、どの領域の専門誌に掲載するかという点だけではなく、どのレベルの再現性を要求する学術誌への掲載を目指すかについて、研究前・中・後、特に投稿の際に研究者間で十分な合意を得ておくことが必要です。



---

## 政府および研究機関等の役割

各国の公的機関は、公的資金が信頼性を欠く研究に使われることなく、再現性が高く、科学の発展に寄与する研究を支援できるよう、研究者がとるべき行動規範を提示していますが、これはごく自然なことです[16-19]。一方、欧米政府が研究者規範に近年特に注目することになった理由は、新薬開発にかかる経費とその件数の多さの中[22]、再現性を欠く研究発表がねつ造や改ざんと同様の害を及ぼしてきていることが分かってきたためです。また、その原因として人為的要因が数多く存在し、ねつ造・改ざんよりも遥かに多いと見られる[23-25]にもかかわらず、この問題は長年放置されてきました。



こうした状況に対して、各種学術誌も掲載論文の信頼性確保のために詳細な投稿規程を設けるようになりました[10-12, 14, 20, 21]。政府がこれまでに発表した研究者向けのガイドラインは、とかく過去の不正事象の再発防止に主眼を置いたものであることが多く、必ずしも今日の国際誌規程を充足しえない面がありました。研究発表こそが、再現性の低い研究が世に出て害をもたらす最後の通過点となりますから、今後、公的機関および各研究機関は公正な研究に向けたガイドライン作りにあたり、現時点の学術誌、特に注目度の高い国際誌の投稿規程を十分に視野に入れておく必要があるでしょう。

---

## まとめ

信頼性を最大限に保つ上で、研究者が行うべきことは以下です。

- バイアスの入る余地のない研究プロトコールをたてること。
- 実験は、予め立てたプロトコールに厳密に従って実施すること。
- 一方、上記に関する不完全性のため再現性に難点があれば、その度合いを適切な統計解析データで提示するなど、発表にあたって透明性を確保し、他の研究者等が研究結果を確認したり再現しようとする上で十分な情報を提供すること。

高い再現性 (Reproducibility) を得るために研究者が心がけるべきことは、偏見を持ち込む余地のない、適切な条件設定を研究計画の中で行った上で、プロトコールに厳密

(Rigor) に従って研究を実施することです。一方、これらの点で不完全さが生じうるのも生命科学研究での現実です。そうしたことによって発表するデータに難点があっても、研究内容の詳細な提供とともに適切な統計解析結果の提示を行うこと、つまり透明性 (Transparency) を持つことにより、研究成果の信頼性が保てるのです。

わが国では、欧米に比して治験の数は少なく [26]、治験の土台となる前治験研究としての動物実験等も限られているため、それらの再現性の低さに対する政府および研究機関の注目度は欧米ほどにはありませんが、今日の国際誌は研究の再現性確保のために報告事項を様々なレベルで要求していることから、わが国の研究者は、この点について努めて注目しておく必要があるでしょう [27-29]。

---

本単元は、日本医療研究開発機構：研究公正高度化モデルである「医系国際誌が規範とする研究の信頼性にかかる倫理教育プログラム」（略称：AMED国際誌プロジェクト）によって作成された教材です。作成および査読等に参加した専門家の方々の氏名は、冒頭に掲載されています。

---

## 参考文献

- [1] The Instability of Science Funding and the Failure of NIH Leadership  
<https://mikethemadbiologist.com/2013/01/28/the-instability-of-science-funding-and-the-failure-of-nih-leadership/> (最終閲覧日 2018 年 3 月 28 日)
- [2] Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. *Nature Biotechnology* 32, 40-51 (2014) doi:10.1038/nbt.2786  
<http://www.nature.com/nbt/journal/v32/n1/abs/nbt.2786.html> (最終閲覧日 2018 年 3 月 28 日)
- [3] Sertkaya A, Wong H-H, Jessup A, Beleche T. Key cost drivers of pharmaceutical clinical trials in the United States. *Clinical Trials* 13(2) 117-126, 2016.  
<https://www.ncbi.nlm.nih.gov/pubmed/26908540> (最終閲覧日 2018 年 3 月 28 日)
- [4] Begley CG, Lee M, Ellis LM. Drug development: Raise standards for preclinical cancer research. *Nature* 483:531-533 (29 March 2012) doi:10.1038/483531a  
<http://www.nature.com/nature/journal/v483/n7391/full/483531a.html> (最終閲覧日 2018 年 3 月 28 日)
- [5] Freedman LP, Cockburn IM, Simcoe TS. The economics of reproducibility in preclinical research. *PLoS/Biology* June 9, 2015  
<http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002165>  
(最終閲覧日 2018 年 3 月 28 日)
- [6] Bell M, Miller N. A replicated study on nuclear proliferation shows the critical necessity of reviewing accepted scientific results. 2013  
<http://blogs.lse.ac.uk/impactofsocialsciences/2013/11/05/reproducing-social-science-nuclear-proliferation/> (最終閲覧日 2018 年 3 月 28 日)
- [7] *New England Journal of Medicine*. Author Center. New Manuscripts Statistical Methods.  
<http://www.nejm.org/page/author-center/manuscript-submission#electronic> (最終閲覧日 2018 年 3 月 28 日)
- [8] *Nature*, Reporting requirements  
<https://www.nature.com/authors/policies/availability.html#requirements> (最終閲覧日 2018 年 3 月 28 日)
- [9] EQUATOR Reporting Guideline Decision Tree  
<http://www.equator-network.org/wp-content/uploads/2013/11/20160301-RG->

- [Decision-Tree-used-for-EQUATOR-wizard-vn-1.pdf](#) (最終閲覧日 2018 年 3 月 28 日)
- [10] New England Journal of Medicine. Author Center. New Manuscripts.  
<http://www.nejm.org/page/author-center/manuscript-submission> (最終閲覧日 2018 年 3 月 28 日)
- [11] JAMA. Instructions for authors.  
<http://jamanetwork.com/journals/jama/pages/instructions-for-authors> (最終閲覧日 2018 年 3 月 28 日)
- [12] Cell. Information for authors.  
<http://www.cell.com/cell/authors> (最終閲覧日 2018 年 3 月 28 日)
- [13] EMBO Conflicts of interest  
<http://emboj.embopress.org/authorguide#conflictsofinterest> (最終閲覧日 2018 年 3 月 28 日)
- [14] Lancet. Information for authors.  
<http://www.thelancet.com/lancet/information-for-authors> (最終閲覧日 2018 年 3 月 28 日)
- [15] Japan investigation alleges misconduct in large scale clinical studies. New Blog. Nature.  
<http://blogs.nature.com/news/2013/08/japan-investigations-allege-misconduct-in-large-scale-clinical-studies.html> (最終閲覧日 2018 年 3 月 28 日)
- [16] Improving the Quality of NINDS-Supported Preclinical and Clinical Research through Rigorous Study Design and Transparent Reporting. National Institute of Neurological Disorders and Stroke. Notice Number:NOT-NS-11-023  
<http://grants.nih.gov/grants/guide/notice-files/NOT-NS-11-023.html> (最終閲覧日 2018 年 3 月 28 日)
- [17] Enhancing the Reliability of NIMH-Supported Research through Rigorous Study Design and Reporting. National Institute of Mental Health. Notice Number: NOT-MH-004.  
<http://grants.nih.gov/grants/guide/notice-files/NOT-MH-14-004.html> (最終閲覧日 2018 年 3 月 28 日)
- [18] NIDA Notice: Improving Reporting of Research Methods and Results in Translational Addiction Research Involving Animals is a NIDA Commitment. National Institute on Drug Abuse, Notice Number: NOT-DA-14-007  
<http://grants.nih.gov/grants/guide/notice-files/NOT-DA-14-007.html> (最終閲覧日 2018 年 3 月 28 日)
- [19] ARRIVE Guidelines. National Centre for the Replacement & Reduction of

Animal Research.

<https://www.nc3rs.org.uk/arrive-guidelines> (最終閲覧日 2018 年 3 月 28 日)

[20] Nature. Authors & Referees. Policies.

<http://www.nature.com/authors/policies/index.html> (最終閲覧日 2018 年 3 月 28 日)

[21] Science. Editorial Policies.

<http://www.sciencemag.org/authors/science-editorial-policies> (最終閲覧日 2018 年 3 月 28 日)

[22] Trends, Charts, and Maps. ClinicalTrials.gov

<https://clinicaltrials.gov/ct2/resources/trends> (最終閲覧日 2018 年 3 月 28 日)

[23] 文部科学省の予算配分又は措置により行われる研究活動において特定不正行為が認定された事案（一覧）. 文部科学省

[http://www.mext.go.jp/a\\_menu/jinzai/fusei/1360839.htm](http://www.mext.go.jp/a_menu/jinzai/fusei/1360839.htm) (最終閲覧日 2018 年 3 月 28 日)

[24] Search Case Closeout Memoranda. Office of Inspector General. National Science Foundation

<http://www.nsf.gov/oig/case-closeout/> (最終閲覧日 2018 年 3 月 28 日)

[25] Case Summaries. The Office of Research Integrity.

[http://ori.hhs.gov/case\\_summary](http://ori.hhs.gov/case_summary) (最終閲覧日 2018 年 3 月 28 日)

[26] 治験計画届出数の推移. 独立行政法人 医薬品医療機器総合機構

<https://www.pmda.go.jp/review-services/trials/0014.html> (最終閲覧日 2018 年 3 月 28 日)

[27] Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein Rbi, Bradley EW, Crystal GR, Darnell RB, Ferrante RJ, Fillit H, Finkelstein R, Fisher M, Gendelman HE, Golub RM, Goudreau JL, Gross RA, Gubits AK, Hesterlee SE, Howells DW, Huguenard J, Kelner K, Koroshetx W, Krainc D, Lazic S, Levine MS, et al.

A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 490:187-191. (11 October 2012)

doi:10.1038/nature11556

<http://www.nature.com/nature/journal/v490/n7419/full/nature11556.html> (最終閲覧日 2018 年 3 月 28 日)

[28] van der Worp, H. B. & Macleod, M. R. Preclinical studies of human disease: time to take methodological quality seriously. *J. Mol. Cell. Cardiol.* 51, 449-450 (2011)

<http://www.ncbi.nlm.nih.gov/pubmed/21549125> (最終閲覧日 2018 年 3 月 28 日)

- [29] Hackam, D. G. & Redelmeier, D. A. Translation of research evidence from animals to humans. A study reporting that a large fraction of high-impact publications in highly reputable journals lack important information related to experimental design. *J. Am. Med. Assoc.* 296, 1727-1732 (2006)

<http://www.ncbi.nlm.nih.gov/pubmed/17032985> (最終閲覧日 2018 年 3 月 28 日)

- [30] Messerli, F.H. Chocolate Consumption, Cognitive Function, and Nobel Laureates. *N. Engl. J. Med.* 367; 16

<http://www.nejm.org/doi/full/10.1056/nejmon1211064> (最終閲覧日 2018 年 3 月 28 日)

東京大学医学部附属病院