

## 科学研究費助成事業 研究成果報告書

平成 29 年 5 月 19 日現在

機関番号：12601

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26460764

研究課題名(和文) 多変量解析の不適切利用是正に向けた医学論文サーベイランスおよび研究者支援策の立案

研究課題名(英文) A surveillance of medical papers to correct inappropriate use of multivariate analysis and plan support for researchers

研究代表者

野島 正寛 (Nojima, Masanori)

東京大学・医科学研究所・准教授

研究者番号：00457699

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：「多変量解析」は、医学研究において頻用される一方で、十分な理解が進んでいるとは言いがたいのが現状である。そこで今回、学術誌に掲載された医学研究論文における実施手順の実態を明らかにすることとした。今回の検討では、不適切な実施手順として「単変量解析で有意だった変数のみを利用する」という手法に着目した。その頻度は、ロジスティック回帰の場合に8.6%、Cox回帰の場合に9.8%であった。また、筆頭著者の所属部門が疫学等の専門的な部門でない場合(臨床+その他vs. 疫学・公衆衛生部門)、不適切な手法が利用されるリスクが高くなることが示唆された(オッズ比は5.6, 95%CI: 2.8-15.2)。

研究成果の概要(英文)：While "multivariate analysis" is frequently used in medical research, it is currently difficult to say that sufficient understanding has advanced. Therefore, we decided to clarify the actual situation of implementation procedure in medical research paper published in academic journal. In this study, we focused on a method called "use only variables that were significant in univariate analysis" as an improper implementation procedure. The frequency was 8.6% for logistic regression and 9.8% for Cox regression. In addition, it was suggested that the risk of inappropriate methods being used increases when the department in which the first author belongs is not a specialized department such as epidemiology (clinical + other vs. epidemiology / public health department; the odds ratio is 5.6, 95% CI: 2.8 - 15.2).

研究分野：臨床統計学

キーワード：多変量解析 疫学 臨床疫学 医学統計 生物統計 統計専門家

## 1. 研究開始当初の背景

医学研究分野において、重回帰・ロジスティック回帰・Cox 回帰に代表される、いわゆる「多変量解析」は、観察研究における交絡のコントロールや、予後予測モデルを作成する手段として広く用いられており、統計ソフトウェアの普及により、多くの医療従事者・医学研究者にとって身近なものとなった。しかしながら、その前提となる統計学的な仮定や、作成された統計モデルの解釈など、生物統計学を専門としない研究者にとっては難解な面が多く存在することも事実である。そのため、ソフトウェアを利用した実行の容易さと理解の困難さの乖離が大きく、統計解析のブラックボックス化の弊害が最も懸念されるステップとなっている。統計を専門としない医学研究者の間で根拠のないコンセンサスが形成されている場合もしばしば見られ、正しい理解や結果の解釈を阻害する要因となっており、米国の内科系学術誌である Annals of Internal Medicine の投稿規定ではその代表例について問題を提起している。

- 1) 単変量解析において統計学的に有意であった変数のみ多変量解析に用いるのは最適な方法ではない
- 2) 交絡要因であるかどうかは、統計学的有意性ではなく回帰モデル投入時に興味ある因子へ及ぼす変化による
- 3) 検証的な場面でステップワイズ法を用いるのは避けるべきである

医学・生物統計学双方の知識を有して初めて統計モデルの適切な理解が可能となるが、研究者や査読者がこの学術的要請に応えるのは容易ではない。申請者はこれまで、臨床、基礎医学研究それぞれの経験を踏まえた上で、多くの医学研究分野の統計解析を積極的に支援してきたが、最も改善すべき問題の一つと感じたのが先に述べた「不適切な多変量解析の利用」である。

以上を背景とし、学術誌に掲載された医学研究論文における多変量解析の利用法、および結果についてなされた解釈について調査を行い、多変量解析の利用実態を明らかにする、という本研究の着想に至った。

## 2. 研究の目的

このような背景の下、学術誌に掲載された医学研究論文における多変量解析の利用、および結果についてなされた解釈について調査を行い、多変量解析の利用実態を明らかにするのが本研究の目的である。さらには、不適切利用が起こりやすい状況や研究分野を特定し、効果的な支援策を立案していく。

## 3. 研究の方法

### 3.1. 対象となる学術誌の選定：

- 1) Journal Citation Reports 収載の学術誌より臨床医学・疫学を選択
- 2) インパクトファクター (IF : impact factor) に応じて階層化 (4 階層: <2, 2-<4, 4-<6, 6<)。
- 3) Article 数が 200 / 年を超えているジャーナルを選択し、IF 6 点以上の学術誌は全数抽出 (71 誌)。同数の抽出を行うよう他の各層の抽出率を設定 (71x4=284 誌)。
- 4) 選定されたジャーナルより、ロジスティック回帰 / Cox 回帰が行われている観察研究に関する論文を Pubmed にて検索 (過去 5 年以内: 2011-2015 年)。検索語は "logistic" および "hazard" とした。

本稿では中間報告として、抽出された論文のうち 2~8% をランダムサンプリングし、論文レビューを実施した時点での結果を提示する。主たる解析として、抽出率に基づき重み付けを行ったロジスティック回帰を利用した (一般化推定方程式によりロバスト分散を利用)。傾向性の検定については、カテゴリ化された変数を連続値として扱い、傾きの推定値に対し検定を実施した。

### 3.2. <主たるアウトカム>

今回の検討では、不適切な実施手順として Annals of Internal Medicine (米国) の投稿規定にも取り上げられている「単変量解析で有意だった変数のみを利用する」という手法に着目した。同手法は、交絡の影響や変数間の関連などを吟味せず機械的に実施されている場合が多く、実施者が多変量解析の意義を理解していない可能性がある。

### 3.3. 調査項目

筆頭著者の所属、発行年、IF、単変量解析の有無、有意だった変数のみを投入、P 値を利用するかどうか、イベント数、共変数の数、

etc.

#### 4. 研究成果

国別調査対象論文数を以下の図に示す（およそ 50%が米国）。



4.1. 抽出数： ロジスティック回帰 610 報（推定論文数 7424/5 年）

Cox 回帰 319 報（推定論文数 14999/5 年）

4.2. 主たるアウトカム「単変量解析で有意だった変数のみを利用」の頻度

ロジスティック回帰に対して：

8.6% (95%CI: 6.2-11.8%)

Cox 回帰に対して：

9.8% (95%CI: 5.1-17.9%)

4.3. 主たるアウトカムの要因別検討（ロジスティック・Cox を統合）

4.3.1. 推定

「単変量解析で有意だった変数のみを利用」に該当すれば「はい（不適切）」

IF	はい	イベント数	はい
6≤	9.0%	1 -	11.4%
4 to <6	4.4%	21 -	9.0%
2 to <4	14.1%	51 -	12.2%
<2	19.6%	101 -	8.9%
P trend 0.018		P trend 0.907	

筆頭著者の所属	はい
臨床/他	12.8%
疫学・公衆衛生	2.5%

P<0.001

地域	はい
Asia	9.1%
Europe	12.3%
USA	8.2%
Other	7.6%

P=0.855

インパクトファクター、筆頭著者の所属と不適切な実施手順の関連が示唆された。

筆頭著者の所属部門（臨床+その他 vs. 疫学・公衆衛生部門，N 比 約 7:3）について，IF, イベント数を調整した場合のオッズ比  
疫学・公衆衛生以外である場合  
OR 5.6, 95%CI: 2.8-15.2

#### 4.3.2 サブグループ解析

上記解析で関連が示唆される筆頭著者の所属およびインパクトファクターについて，これら 2 要因によるサブグループ解析を実施し，リスクの高い層を検討した。

筆頭著者の所属	IF	はい	95%信頼区間	
			下限	上限
臨床 / 他	6≤	12.0%	7.8%	18.0%
	4 to <6	6.9%	3.4%	13.4%
	2 to <4	17.7%	6.1%	41.6%
	<2	23.1%	10.8%	42.6%
疫学・公衆衛生	6≤	2.6%	0.8%	7.8%
	4 to <6	0.4%	0.1%	3.0%
	2 to <4	4.7%	1.3%	15.4%
	<2	9.5%	2.3%	32.5%

上記の通り，臨床系でインパクトファクターの低い論文については，不適切とされる頻度が 20%以上と高いことが示唆された(23.1%)。

#### 4.4. 他のアウトカム

4.4.1. 共変量の選択に P 値を利用（ステップワイズ法ではない。こちらは必ずしも不適切ではない）

IF	はい	イベント数	はい
6≤	20.8%	1 -	26.7%
4 to <6	13.9%	21 -	20.8%
2 to <4	28.6%	51 -	26.0%
<2	28.2%	101 -	19.7%
P trend 0.093		P trend 0.433	

筆頭著者の所属	はい
臨床/他	25.6%
疫学・公衆衛生	10.7%
P=0.001	

地域	はい
Asia	14.6%
Europe	21.9%
USA	21.2%
Other	27.8%
P=0.404	

筆頭著者の所属部門（臨床+その他 vs. 疫学・公衆衛生部門，N 比 約 7:3）について，IF, イベント数を調整した場合のオッズ比 疫学・公衆衛生以外である場合 OR 2.9, 95%CI: 1.7-5.6

#### 4.5. 考 察・まとめ

多変量解析は交絡の調整および予測モデルの作成等を目的として実施されるが，今回のアウトカム，多変量解析において「単変量解析有意だった変数のみを利用」で示されるように，いずれの目的に対しても適切でない方法が選択されている可能性がある．特に，疫学者等の専門家による研究でない場合にその頻度が高いことが示唆された．

イベント数が少なくやむを得ず共変量を減らさなくてはならない場合もあり，この手法が必ずしも不適切とは言えない場合もあるが，P 値を機械的に利用し，相関や交絡，中間変数かどうかなど，各共変量の持つ意味について検討していないとすれば重大な問題がある．逆説的ではあるが，統計の専門性を持つ（とみなせる）筆頭著者の場合に頻度が低いことを考えると，今回のアウトカムに該当する場合，多くは不適切に実行されたと考えるべきかと思われる．

インパクトファクターについては，必ずしも雑誌の価値を規定するものではないが，2 以下の比較的低い雑誌に掲載されたものについてはリスクが高いことが示唆された．インパクトのある結果をもたらす研究の場合，対象者数が多く資金も豊富で，十分な体制（各種専門家が多数関与）で計画・実施されていることが想像できる．実施体制が不十分であ

る場合には，掲載される雑誌もそれ相応のものとなり，今回のような結果となったのである．

支援体制の立案という観点では，やはり専門家へのコンサルテーションということが重要であるという結論となる．特に多変量解析に関しては，理論的背景はもちろん，利用目的や解釈についても難解な面は否めない．何が問題かを理解してもらい，専門家への相談を当たり前のものとするのが重要ではないかと考える．

また，論文抽出の段階で明らかとなったが，本研究で取り扱う臨床研究および疫学研究の圧倒的大部分は米国で実施されており，本邦の占める割合は限定的であることが確認された．これについては，すでに各所で問題視されているが，ここでも浮き彫りとなった．

現在他の調査項目についても検討を進めているほか，多角的な視点からバリデーションを実施し，エビデンスとしての信頼性を確保したのちに平成 29 年度中の論文化を目指すこととする．

#### 5．主な発表論文等

（研究代表者，研究分担者及び連携研究者には下線）

〔雑誌論文〕(計 0 件) なし – 現在作成中

〔学会発表〕(計 1 件)

野島正寛ほか．多変量解析の不適切利用是正に向けた医学論文サーベイランス．第 27 回日本疫学会学術総会, 2017/1/27, 甲府市

#### 6．研究組織

##### (1)研究代表者

野島正寛 (Masanori Nojima)  
 東京大学・医科学研究所・准教授  
 00457699