

原 著

医学研究において統計解析を実施する際に生じる問題点と 解析における困難の多少に影響する因子

東京女子医科大学衛生学公衆衛生学（二）教室

*東京女子医科大学総合研究所

サトウ ヤスト シミズ サトル ヤマガチ ナオヒト
佐藤 康仁・清水 悟*・山口 直人

（受理 平成21年4月8日）

Problems Concerning Statistical Analysis and Factors Influencing Their Degree and Frequency in Medical Studies

Yasuto SATO, Satoru SHIMIZU* and Naohito YAMAGUCHI

Department of Public Health, Tokyo Women's Medical University

*Medical Research Institute, Tokyo Women's Medical University

Purpose: The purpose of the present study was to reveal various problems concerning statistical analysis in medical studies, including failure to understand the analysis and uncertainty of the obtained results, and to elucidate factors contributing to the degree and frequency of these problems. **Methods:** Subjects were all researchers who participated in the Grant-in-Aid for the Center of Excellence (COE) Program for the 21st Century at Tokyo Women's Medical University as of January 2004. An anonymous questionnaire was used. At the beginning of the analysis, details of the problems concerning statistical analysis and methods corresponding to the points in question were analyzed according to the degree and frequency of the problems. Subsequently, logistic regression analysis was conducted using background information and experience related to statistical analysis as factors influencing the degree and frequency of the problems. **Results:** Of the particular problems concerning statistical analysis, the following items received high scores on the questionnaire: 'advanced analysis methods are hard to understand' (respondents who sometimes encountered this problem, 91.3%; respondents who frequently encountered this problem, 71.9%), 'choosing an appropriate analysis method is difficult' (56.5, 62.5%), and 'there is no one I can consult about statistical analysis problems' (23.9, 34.4%). From logistic regression analysis, significant differences were observed between 'duty position (OR = 4.30)' and 'reporting results of research in which a hypothesis was tested (OR = 3.52)'. **Conclusion:** The present study revealed that 'duty position' and 'reporting results of research in which a hypothesis was tested' influence the degree and frequency of the problems. All questionnaire respondents reported some degree of difficulty or frustration concerning statistical analysis. Organizational measures are steps that may serve to reduce the degree and frequency of problems concerning statistical analysis in medical studies.

Key words: medical research, medical statistics, biostatistics, research support

緒 言

わが国の医学研究者が報告した論文における統計処理には誤りが多いという問題点が指摘されている¹⁾。一方で、医学研究者の間では論文作成時の統計処理に苦慮しているという話をしばしば耳にする。

医学研究者が統計処理に苦慮することで研究の推進に支障が出るのであれば大きな問題である。

欧米における優れた医学研究では、研究チームに生物統計家が参加し、研究計画、データ管理、統計解析等の研究支援を実施している²⁾³⁾。また、研究チー

ムに生物統計家が参加することは、質の高い研究の実施を可能にするだけでなく、研究計画から結果公表までをスムーズに行うことを可能にしている。現在、欧米においては、生物統計家は医学研究に不可欠な存在になっている。

わが国においても、欧米諸国と同じように生物統計家が研究支援を実施する必要性が指摘されている⁴⁾。また、近年わが国では生物統計学に関する人材育成機関が増加しており、人材の供給が行われつつある⁵⁾。しかし、生物統計学を専門とする人材はまだ少なく、欧米諸国のように多くの医学研究を支援する段階には至っていない。

世界レベルの研究を推進するためには欧米並みの研究体制の構築が必要と考える。そこで、現在の医学研究における統計処理上の問題点を明らかにすることは、今後の研究推進に寄与すると思われる。本研究は、医学研究を進める際の統計解析において、どのような疑問点や不明点が問題になっており、どのような要因によりもたらされるのかを明らかにすることを目的とする。

対象および方法

調査の対象者は、2004年1月現在で東京女子医科大学21世紀COEプログラム（再生医学研究センター）に参加する研究者全員（114名）とした⁶⁾。調査は無記名式のアンケート調査法を用いた。調査では「背景情報」、「統計解析に関する経験」、「研究をまとめる際に統計解析を実施するか」、「統計解析においてわからない・はっきりしない等の問題を感じるか」、「統計解析における問題点の詳細」、「統計解析における疑問点への対応方法」について質問している。調査は2004年3月に実施した。

調査の回収数は87名（回収率76.3%）であり、このうち「研究をまとめる際に統計解析を実施しない」と回答した9名を除く78名を解析対象とした。アンケート調査では、「統計解析においてわからない・はっきりしない等の問題を感じるか」を質問している。これに対して「問題が多い」と答えた者は32名（41.0%）、「時々問題がある」と答えた者は46名（59.0%）であった。解析では「問題が多い」と答えた者を「困難をより多く感じている者」、「時々問題がある」と答えた者を「困難をより少なく感じている者」として2群に分け、分析した。

解析では、はじめに「統計解析における問題点の詳細」および「統計解析における疑問点への対応方法」について困難の多少により違いがないか分析し

表1 対象者の背景情報

		(n = 78)	
		n	%
性別	男性	61	78.2
	女性	17	21.8
年齢	20歳代	7	9.0
	30歳代	27	34.6
	40歳代	22	28.2
	50歳代以上	22	28.2
職位	教授・准教授	23	29.5
	講師・准講師	15	19.2
	助教	20	25.6
	大学院生・研究生	20	25.6
研究方法 (複数回答)	実験研究	58	74.4
	調査研究	7	9.0
	臨床研究	40	51.3
出身学部	医学部・歯学部	61	78.2
	その他の学部	16	20.5

た。続いて「背景情報」および「統計解析に関する経験」を統計解析における困難の多少に影響する因子と考え、カイ二乗検定またはFisherの正確検定を用いて有意水準5%で単変量を解析した。続いて有意差が観察された因子と統計解析における困難の多少との関連をロジスティック回帰モデルで分析した。モデリングにおいては、有意差の観察された項目および性別、年齢を説明変数とした。モデル1では説明変数を強制投入した。モデル2では説明変数をステップワイズ法で絞り込んだ。解析にはWindows版SASシステム（リリース8.02）を用いた。

結果

解析対象者の背景情報を表1に示した。性別では男性（78.2%）、年齢では30歳代（34.6%）が最も多くなっていた。職位は教授・准教授（29.5%）が最も多く、助教（25.6%）、大学院生・研究生（25.6%）が続いていた。研究方法では実験研究（74.4%）が最も多く、臨床研究（51.3%）が続いていた。出身学部は医学部・歯学部（78.2%）が最も多くなっていた。

統計解析における問題点の詳細について表2に示した。困難をより少なく感じている者においても、より多く感じている者においても、「より進んだ解析方法がわからない（困難をより少なく感じている者91.3%、困難をより多く感じている者71.9%）」、「どの解析方法を選べばよいかわからない（56.5, 62.5%）」、「わからないことを相談する相手がない（23.9, 34.4%）」の割合が高くなっていた。「より進んだ解析方法がわからない」については、困難をより少なく感じている者の方が高い割合を示したが、そ

表2 統計解析における問題点の詳細（複数回答）

	困難をより少なく 感じている者 (n = 46)		困難をより多く 感じている者 (n = 32)		p*
	n	%	n	%	
より進んだ解析方法がわからない	42	91.3	23	71.9	0.0235
どの解析方法を選べばよいかわからない	26	56.5	20	62.5	0.5975
書籍を読んでもはっきりしない	9	19.6	11	34.4	0.1406
どの書籍を読めばよいかわからない	3	6.5	9	28.1	0.0124
わからないことを相談する相手がいない	11	23.9	11	34.4	0.3125
データセットの作成方法がわからない	3	6.5	7	21.9	0.0816
統計ソフトウェアの使い方がわからない	8	17.4	10	31.3	0.1530
統計ソフトウェアの出力結果の読み方がわからない	3	6.5	8	25.0	0.0431
まとめる際の記述方法がわからない	7	15.2	9	28.1	0.1649

*カイ二乗検定, Fisherの正確検定.

表3 統計解析における疑問点への対応方法（複数回答）

	困難をより少なく 感じている者 (n = 46)		困難をより多く 感じている者 (n = 32)		p*
	n	%	n	%	
書籍を調べる	35	76.1	23	71.9	0.6752
論文を調べる	18	39.1	8	25.0	0.1929
統計ソフトウェアの解説書を調べる	24	52.2	8	25.0	0.0164
インターネットで調べる	10	21.7	5	15.6	0.5004
統計学に詳しい人（専門とする人）に相談	16	34.8	11	34.4	0.9703
統計学に詳しい人（専門としない人）に相談	20	43.5	9	28.1	0.1676
統計ソフトウェアのヘルプデスクに質問	2	4.4	2	6.3	1.0000
メーリングリストで質問	1	2.2	1	3.1	1.0000
業者に外注する	4	8.7	5	15.6	0.4749
何もしない	0	0.0	3	9.4	0.0652

*カイ二乗検定, Fisherの正確検定.

れ以外の項目は困難をより多く感じている者の方が高い割合を示した。有意差が見られたのは「より進んだ解析方法がわからない ($p=0.0235$)」、「どの書籍を読めばよいかわからない ($p=0.0124$)」、「統計ソフトウェアの出力結果の読み方がわからない ($p=0.0431$)」であった。

統計解析における疑問点への対応方法について表3に示した。困難をより少なく感じている者は「書籍を調べる (76.1%)」、「統計ソフトウェアの解説書を調べる (52.2%)」、「統計学に詳しい人（専門としない人）に相談 (43.5%)」の割合が高くなっていた。困難をより多く感じている者は「書籍を調べる (71.9%)」、「統計学に詳しい人（専門とする人）に相談 (34.4%)」、「統計学に詳しい人（専門としない人）に相談 (28.1%)」の割合が高くなっていた。有意差が見られたのは「統計ソフトウェアの解説書を調べる ($p=0.0164$)」のみであった。

統計解析における困難の多少に影響する因子の単

変量解析の結果、困難をより多く感じている者は職位の低い者が多く ($p=0.0210$)、仮説検定を行った研究結果の報告を行ったことがない者が多かった ($p=0.0052$)。さらに、統計ソフトウェアの使用年数の短い傾向が見られたが、有意差は観察されなかった ($p=0.0812$)。統計解析における困難の多少に影響する因子の多変量解析の結果については表4に示した。モデル1では職位：大学院生・研究生 [OR(オッズ比) = 13.21, 95%CI (信頼区間) : 1.36-128.18] および仮説検定を行った研究結果の報告 (OR = 3.49, 95%CI : 1.08-11.34) に有意差が観察された。モデル2では職位：大学院生・研究生 (OR = 4.30, 95%CI : 1.36-13.57) および仮説検定を行った研究結果の報告 (OR = 3.52, 95%CI : 1.15-10.80) に有意差が観察された。

考 察

本研究は東京女子医科大学21世紀COEプログラムに参加する研究者全員を対象としている。職種

表4 統計解析における困難の多少に影響する因子（多変量解析）

		モデル1			モデル2		
		オッズ比	95%CI		オッズ比	95%CI	
性別	男性	1.00					
	女性	1.07	0.29	3.94			
年齢	50歳代以上	1.00					
	40歳代	0.51	0.11	2.44			
	30歳代	0.34	0.04	2.67			
	20歳代	0.40	0.03	6.11			
職位	教授・准教授	1.00			1.00		
	講師・准講師	2.07	0.42	10.10			
	助教	3.42	0.50	23.57			
	大学院生・研究生	13.21	1.36	128.18	4.30	1.36	13.57
仮説検定を行った研究結果の報告	あり	1.00			1.00		
	なし	3.49	1.08	11.34	3.52	1.15	10.80

別の回収率を比較すると教授・准教授 89.3%、講師・准講師 72.7%、助教 60.5%、大学院生・研究生 95.2% となっており、助教でやや低かったが、母集団を代表する標本が得られたとみなし解析した。この研究プロジェクトに参加する研究者は、臨床系および基礎系に所属する者を含んでおり、所属する部局は 26 部局にわたっていた。また、背景情報に示したように、本研究の対象者は典型的な医学部における研究者の集団であると思われる。

統計解析における問題点は、統計解析における困難の多少にかかわらず「より進んだ解析方法がわからない」、「どの解析方法を選べばよいかわからない」が多いことが明らかとなった。また、これらには「書籍を調べる」、「統計学に詳しい人に相談する」ことで対応していることが明らかとなった。一方で、統計学に詳しい人に相談している者の半数は、統計学を専門としない人に相談しているのが現状であった。また、統計解析における問題点には「わからないことを相談する相手がいない」が上位に位置づけられていた。わが国では生物統計学を専門とする人材が不足していると指摘されている⁴⁾⁵⁾。そのため個人レベルで相談相手を見つけだすことは難しいと思われ、組織レベルの対策が必要であることが示唆される。組織レベルの対策としては、欧米の大学で実施されている生物統計家による統計相談を大学医学部内に導入することが挙げられる。これが実現できれば、多くの者が指摘している「より進んだ解析方法がわからない」、「どの解析方法を選べばよいかわからない」についても改善することができる可能性がある。

統計解析において困難をより多く感じている者

は、問題点として「より進んだ解析方法がわからない」、「どの書籍を読めばよいかわからない」、「統計ソフトウェアの出力結果の読み方がわからない」を挙げている。一方で困難をより少なく感じている者は疑問点への対応方法として「統計ソフトウェアの解説書を調べる」を挙げている。さらに統計ソフトウェアの使用年数が長い傾向にあった。現在の医学研究において統計解析を行う場合、統計ソフトウェアは不可欠な存在である。ところが、現在入手できる統計ソフトウェアは非常に多くの統計手法が利用可能であり、かえって使いこなすことが難しくなっている。統計解析における困難に対処するためには統計ソフトウェアを解説書も含めて使いこなせるかどうか、ポイントになっていると考える。個人レベルでは、統計ソフトウェアを使いこなすことを目標にすることで、統計解析における困難を克服できる可能性がある。組織レベルの対策としては、大学医学部内の研究者を対象とした統計ソフトウェアを利用した統計講習会の実施を挙げることができる。

統計解析における困難の多少に影響する因子を検討したところ、職位と仮説検定を行った研究結果の報告が大きく影響していた。職位の高い者は長年の経験と競争により選択された者であるのに対して、職位の低い者は経験がより少なく、これから競争により選択を受ける者が多くなっていると思われる。また、職位の高い者は相談する相手が多いことが考えられる。年齢ではなく職位がモデルに取り込まれたことは、現状をうまく説明しているものと思われる。仮説検定を行った研究結果の報告については、実際に統計解析を行い、その結果を報告するという経験の重要性を示していると考えられる。わが国にお

る医学統計の教育は学部レベルでは体系的でなく、不十分であると指摘されている⁷⁸⁾。一方、大学院レベルの教育では、研究をまとめる際に統計解析を経験することが多い。この経験は、統計解析における困難の減少に大きく寄与していると思われる。

ところで本研究では、対象者のすべてが多かれ少なかれ「統計解析においてわからない・はっきりしない等の問題を感じる」と報告していた。医学研究は競争が激しく、常に国際競争にさらされている。欧米諸国では、医学研究チームに生物統計家加わることで、質の高い研究を遂行している。わが国においても研究者個人が統計処理に関する技術を向上させるとともに、研究者が生物統計家に相談ができるような研究環境の整備が必要であると思われる。

東京女子医科大学では、2008年6月より総合研究所に統計解析室を設置し、統計解析支援システムを立ち上げている⁹⁾。このシステムは総合研究所の共同利用機能であり、本大学に所属する研究者および大学院生に対して、物的および人的支援を行うものである。物的支援としては統計解析のためのコンピュータおよび統計ソフトウェアの設置、人的支援としては生物統計学を専門とする教員による研究支援を行っている。支援の内容は、共同研究、データ解析の委託、相談に分類しており、研究への関与の度合いにより生物統計家を著者とするか謝辞を受けるかを規定している。また、大学医学部内の研究者を対象に統計ソフトウェアを用いた統計解析講座を開始している。

本研究で扱った統計解析における困難は、あくまで主観的な困難度である。一方で、今まで組織的な対策が行われてこなかった大学医学部における統計解析支援について、新たな取り組みが実施されることで、研究者の困難を解消する可能性がある。今後はこのような取り組みを行う中での問題点と解析における困難の多少に影響する因子を評価する必要がある。

結 論

本研究より、統計解析における問題点は「より進

んだ解析方法がわからない」、「どの解析方法を選べばよいかわからない」が多いことが明らかとなった。また、統計解析における困難の多少には職位および仮説検定を行った研究結果の報告が影響することが明らかとなった。一方で対象者の全員が多かれ少なかれ統計解析において困難があるとしていた。これに対応するためには、組織的な対策についても考慮する必要があることが示唆された。

謝 辞

本研究を進めるにあたり、アンケート調査にご協力いただいた東京女子医科大学21世紀COEプログラム(再生医学研究センター)に参加する研究者の皆様へ感謝いたします。本研究は、東京女子医科大学21世紀COEプログラム(平成15年度)の補助を受けて実施した。

文 献

- 1) 浜田知久馬：EBMと統計学の役割 医療情報を読み解くための必須知識。医薬ジャーナル 40(1)：63-68, 2004
- 2) 佐藤俊哉：21世紀のがん臨床試験に必要な基盤 20世紀の臨床試験で統計家はなにをしてきたか 21世紀の臨床試験で生物統計家はなにをすべきか。臨研・生物統計研究会誌 22(1)：61-70, 2002
- 3) 石塚直樹：試験統計家が必要とされない、関わらない、関わってこなかった臨床試験。臨研・生物統計研究会誌 23(1)：73-78, 2003
- 4) 大橋靖雄：臨床試験の基盤整備はどう進めるべきか 生物統計学者の育成と活用。薬理と治療 28(4)：310-315, 2000
- 5) 濱崎俊光：日本における生物・医学統計家に対する専門教育と継続教育：「学」の立場から。臨医薬 23(4)：209-232, 2007
- 6) 佐藤康仁, 清水 悟, 山口直人：医学研究における統計学(1) 医学研究における統計解析の現状と問題点, および今後の方向性。東女医大誌 78(7)：300-305, 2008
- 7) 赤澤宏平, 池田 充, 本多正幸ほか：医療統計手法の開発と統計解析の実践について。医療情報学 23(3)：193-198, 2003
- 8) 田中司朗, 山口拓洋, 大橋靖雄：看護系教育課程を持つ大学における疫学・生物統計学教育の実態調査。日公衛誌 52(1)：66-75, 2005
- 9) 東京女子医科大学総合研究所：東京女子医科大学総合研究所 研究総括・推進部門。http://www.twmu.ac.jp/MRI/index.html