

総 説

[東女医大誌 第 78 卷 第 7 号]
 頁 300~305 平成 20 年 7 月]

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

東京女子医科大学医学部衛生学公衆衛生学（二）教室

サトウ サトル ヤマグチ
 佐藤 康仁・清水 悟・山口 直人

（受理 平成 20 年 4 月 21 日）

Statistics in Medical Research (1)

Current Situations, Problems and Future Directions of Statistical Analysis in Medical Research

Yasuto SATO, Satoru SHIMIZU and Naohito YAMAGUCHI

Department of Hygiene and Public Health II, Tokyo Women's Medical University, School of Medicine

The problem of misusing statistical analyses in articles written by medical researchers in Japan has been brought to attention. In the field of medical research, however, we often hear of the difficulties in performing statistical work on site. If medical researchers are not completely confident about their statistical work and publish articles containing ambiguous material concerning statistical methods, this is indeed a serious problem. In this article, we have described the current status and problems of statistical analysis in medical research from questionnaires conducted at Tokyo Women's Medical University. We have also clarified the need for statistical support as a solution to be taken to solve the problem. Furthermore, based on these studies, we have presented or suggestions on the future direction of the problem: a task that our country as a whole should address and which medical departments of universities should tackle.

Key words: medical research, medical statistics, statistical consultation

はじめに

わが国の医学研究者が報告した論文における統計解析には誤用が多いという問題点が指摘されている¹⁾。一方で、医学研究の現場では、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困り、問題を残したまま論文を発表しているならば大きな問題である。

本稿では医学研究における統計解析について、東京女子医科大学で行ったアンケート調査結果によりその現状と問題点を明らかにした。また、問題点の解決策として統計支援の必要性について明らかにした。さらにこれらを踏まえた上で、今後の方向性についてわが国全体が行うべき課題と、大学医学部が行うべき課題の 2 点についてまとめた。

1. 現状と問題点

本稿では、医学研究における統計解析の現状と問題点を示すため、東京女子医科大学で実施した調査

の結果を取り上げる。この調査は、大学医学部に所属する研究者が医学研究における統計解析についてどのような問題をかかえ、どのような支援を期待しているのかについて明らかにすることを目的に行った。対象は、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）である。調査は無記名式のアンケート調査法を用いた。回収数は 87 名、回収率は 76.3% であった。回答者の背景情報は、男性 77%，年齢 30~40 歳代 62%，実験研究を行う者 75%，医学部または歯学部出身者 77% であり、典型的な医学部における研究者の集団である（表 1）。また、この研究プロジェクトに参加する研究者は、臨床系または基礎系に所属する者を含んでおり、所属する教室は複数に亘っている。

統計解析に関する現状と経験を表 2 に示した。医学研究をまとめる際の統計解析は約 90% の研究者

表1 回答者の背景情報

		n=87	
		n	%
性別	男性	67	77.0
	女性	20	23.0
年齢	20歳代	8	9.2
	30歳代	32	36.8
	40歳代	22	25.3
	50歳代	19	21.8
	60歳代	6	6.9
研究方法 (複数回答)	実験研究	66	75.9
	臨床研究	41	47.1
	調査研究	7	8.0
出身学部	医学部歯学部	67	77.0
	その他学部	19	21.8

が「ほとんど自分で行う」「しばしば自分で行う」「時々自分で行う」のいずれかと回答しており、医学研究における統計処理の必要性の高さを示している。統計解析ソフトの使用経験については、10年以下の者が74%を示しており、比較的短い傾向にあった。統計解析ソフトウェアはパソコンの普及とともに、使い勝手の良いものが普及してきている。パソコンの爆発的普及は1995年頃から始まっていることを考慮すると²⁾、パソコンの普及とともに統計解析ソフトウェアの普及も進んでいると見ることができる。また、学生時代の統計学の講義の受講は71%と高い値を示しているのに対して、学会や企業が主催する統計講習会の受講はほとんど見られなかった。一方で、仮説検定を行った研究結果の報告については67%の研究者が行っていた。これらより研究者は学生時代に統計学に関する基本的な知識を学び、その後は独学で統計解析ソフトを使いこなしている姿が読み取れる。

次に、研究をまとめる際に統計解析を行わないとする9名を除いた78名の統計解析における問題点と対応方法について表3に示した。「統計解析においてわからない、はっきりしないなどの問題を感じるか」については、全員が問題ありとしていた。医学研究における統計処理は約90%の研究者が必要としているが、全員がこのように感じているということは、医学研究を推進する上で大きな課題であると考える。実際にどのような点が問題となっているかについては、「より進んだ解析法がわからない

(83%)」「どの解析方法を選べばよいかわからない(59%)」「わからないことがあっても相談相手がない(28%)」の割合が高くなっていた。我が国の医学教育においては、医学統計学を体系的に習得する体制が整っていないと指摘されている³⁾。教育の体制の不備は、研究者が統計処理を行う際のわからない、はっきりしないという問題に影響していると思われる。また、最近の統計解析ソフトウェアは幅広く普及すると同時にユーザーインターフェイスの改善により比較的簡単に解析を実施することができるようになっている。さらに、多数の統計手法を選択できるようになっているため、どの方法を採用すればよいのかわからなくなっていると考えられる。また、わが国の大学医学部には生物統計学を専門とする部門をおいている施設は少ない。統計解析に問題が生じても相談相手が見つからないのが現状である。

続いて、統計解析においてわからないことがある場合にどのように対応するかについてたずねたところ、「書籍を調べる(74%)」「統計解析ソフトの解説書を調べる(41%)」「統計学に詳しい人(専門としない人)に相談(37%)」の割合が高くなっていた。統計解析ソフトの解説書を調べるについては、解説書の例を自分の研究に当てはめ理解を試みているものと想像できるが、これには問題が多く含まれる。統計解析ソフトの解説書はあくまで一つの例を示していくに過ぎない。研究者はそれぞれの分野で最先端の研究を行っている。その結果を、単純に解説書の例に当てはめて、解析や解釈を行っているとなると問題が大きい。また4割もの研究者が行っているという点も大きな問題である。また、統計学に詳しい人に聞くという方法は手っ取り早く問題を解決する方法である。しかしながら専門としない人に相談する場合、誤った知識を教わる可能性もある。医学研究における統計処理の誤用は数多く指摘されているが、その原因の一つに専門としない者が誤った知識を伝えている場合があると思われる。

2. 統計支援の必要性

以上、見てきたように医学研究における統計解析は必要性が高いにもかかわらず、現状では問題点が多く、解決の手段もなかなか見出せない状況である。そこで、アンケートでは医学研究における統計支援の必要性について質問している。

統計支援のニーズとメリットについて表4に示した。大学医学部内における統計支援を行う部門については、94%があるとよいとしていた。また、この

表2 統計解析に関する現状と経験

		n=87	
		n	%
研究をまとめる際の統計解析	ほとんど自分で行う	35	40.2
	しばしば自分で行う	20	23.0
	時々自分で行う	22	25.3
	行わない	9	10.3
	未回答	1	1.1
統計解析ソフトウェアの使用年数	経験なし	12	13.8
	1年以下	5	5.7
	2~3年	10	11.5
	4~5年	11	12.6
	6~10年	27	31.0
	11~15年	9	10.3
	16年以上	7	8.0
	未回答	6	6.9
学生時代に統計学の講義を受講した		62	71.3
学生時代に統計解析ソフトウェアを使う講義を受講した		12	13.8
学会が主催する統計講習会を受講した		2	2.3
企業が主催する統計講習会を受講した		4	4.6
仮説検定を行なった研究結果を報告した		59	67.8

統計支援に対しては「自分が利用したい(80%)」「研究チーム内で利用したい(57%)」と高い割合で利用したいとの意見が得られた。アンケートの自由記述欄には「欧米では普通に行われている統計支援のシステムについて、ぜひ実現してほしい」等の意見が寄せられた。統計支援のメリットとしては、「質の高い研究結果(87%)」を期待する者が最も多く、「労力の軽減(42%)」や「時間の短縮(41%)」という意見もあった。これらより、医学研究の激しい競争を勝ち抜くために充実した研究体制を構築したいという研究者の思いが読み取れる。

続いて統計支援を行う際には、研究者はどのようななかたちを想定しているのかについて表5に示した。研究者は、「データ解析段階(87%)」「統計手法の選択(80%)」や「統計解析結果の選択(66%)」を「対面でのやり取り(74%)」または「電子メールでのやり取り(67%)」で希望していた。生物統計に関する相談を実施する場合、データ解析段階で相談を受けても研究デザインやデータに不備があり対処できないことがある。また論文作成段階とする者が55%いたが、この時点からの相談では更に対処するのが難しくなると思われる。統計支援はなるべく早い段階、つまり研究計画段階から実施することで、研究デザインやデータ収集に関する助言をすることができ、よりよい支援を行うことができると言える。また、統計支援を行うにあたり、支援を受ける者が

説明内容を理解するには対面でのやり取りが一番わかりやすいと考える。また支援を行う者も研究の内容を十分理解するためには、対面でのやり取りが必要と考える。電子メールでのやり取りは、十分に対面でのやり取りを行った後ならば、機能すると考える。インターネット上にはいくつかの生物統計に関するメーリングリストがあり、統計解析に関する質問と答えが往来している。ある程度、生物統計学の知識がある者には役立つようだが、知識に不安のある者にとっては詳細におよぶやり取りは難しいようである。

最後に、統計支援を行う者、つまり生物統計学の専門家の役割については「共同研究者として名前を論文に載せる(59%)」が最も高くなっていた。しかしながら、「特に考えていない」とする者も26%いた。わが国は欧米諸国に比べ、生物統計学に関する教育および研究の伝統が短い。そのため統計支援を行う者(多くの場合、生物統計家)と支援を受ける者(多くの場合、臨床医)の関係の構築が難しいことが多い。論文の筆頭著者や知的所有権はデータを出す側(支援を受ける者)に帰属すると考えるが、論文発表の場合、統計支援を行う者を共著者に入れるのか、共著者に入れるならば何番目に入れるのか、謝辞に入る程度に止めるのか、研究費の割り当てやコンサルテーション代として金銭的報酬を与えるのかについて決める必要がある。いずれにせ

表3 統計解析における問題点と対応方法

	n=78	%
	n	%
統計解析において、わからない、はっきりしないなど問題を感じるか		
問題が多い	32	41.0
時々問題がある	46	59.0
ほとんど問題ない	0	0.0
統計解析において、どのような点が問題か（複数回答）		
より進んだ解析法がわからない	65	83.3
どの解析方法を選べばよいのかわからない	46	59.0
相談相手がない	22	28.2
本を読んでも内容がはっきりしない	20	25.6
統計解析ソフトの使い方がわからない	18	23.1
統計学的な記述方法がわからない	16	20.5
どの本を読めばよいかわからない	12	15.4
ソフトウェアの出力結果の読み方がわからない	11	14.1
データセットの作成方法がわからない	10	12.8
その他	0	0.0
わからないことがある場合、どのように対応するか（複数回答）		
書籍を調べる	58	74.4
統計解析ソフトの解説書を調べる	32	41.0
統計学に詳しい人（専門としない人）に相談	29	37.2
統計学に詳しい人（専門とする人）に相談	27	34.6
論文を調べる	26	33.3
インターネットで調べる	15	19.2
業者に外注する	9	11.5
統計解析ソフトのヘルプデスクに質問	4	5.1
何もない	3	3.8
メーリングリストで質問	2	2.6
その他	1	1.3

*研究をまとめる際に統計解析を行うとする者 (n=78) を集計

表4 統計支援のニーズとメリット

	n=87	%
	n	%
学内における統計支援を行う部門について		
あるとよい	82	94.3
どちらでもよい	4	4.6
なくてよい	0	0.0
その他	0	0.0
未回答	1	1.1
統計支援を行う部門を利用したいか（複数回答）		
自分が利用したい	70	80.5
研究チーム内で利用したい	50	57.5
利用したいとは思わない	1	1.1
その他	4	4.6
支援によりどのようなメリットがあると思うか（複数回答）		
質の高い研究結果	76	87.4
労力の軽減	37	42.5
研究時間の短縮	36	41.4
その他	2	2.3

表5 統計支援のかたち

	n	n=71	%
どの時点で支援が必要か（複数回答）			
データ解析段階	75	87.2	
研究計画段階	50	58.1	
論文作成段階	48	55.8	
その他	2	2.3	
何を支援して欲しいか（複数回答）			
統計手法の選択	69	80.2	
統計解析結果の解釈	57	66.3	
統計解析の実施	51	59.3	
実験・調査の計画・デザイン	35	40.7	
統計学的な記述のチェック	35	40.7	
臨床試験のプロトコール開発	31	36.0	
論文レフリーとのやり取りアドバイス	26	30.2	
その他	1	1.2	
どのような方法を希望するか（複数回答）			
対面でのやり取り	64	74.4	
電子メールでのやり取り	58	67.4	
紙の文書でのやり取り	7	8.1	
メーリングリストでのやり取り	3	3.5	
Webページ上のやり取り	2	2.3	
その他	1	1.2	
統計学の専門家の役割をどのように考えるか（複数回答）			
共同研究者として名前を論文に載せる	51	59.3	
研究費の一部を割り当てる	27	31.4	
特に考えていない	23	26.7	
その他	11	12.8	

*統計支援を利用したいとする者 (n=71) を集計

よ、統計支援を行う際には、支援を受ける者は、支援を行う者がどのような位置づけで研究にかかわるのかを事前に明らかにする必要があると考える。この点をあいまいにしていると、支援を行う者と受けた者との思惑の違いから、研究の遂行に支障が生じる可能性がある。

3. 今後の方向性

東京女子医科大学におけるアンケート調査により、医学研究の現場では統計支援に関するニーズが非常に高いことが明らかとなった。このニーズを満たすためには、組織的な対策が必要と考える。組織的な対策として、以下、わが国全体が行うべき課題と、大学医学部が行うべき課題の2つについてまとめた。

わが国全体が行うべき課題としては、生物統計家の育成が考えられる。アメリカでは、この20年間に公衆衛生大学院内に生物統計学の講座を多数設立し、人材の供給を行ってきている。これにより現在

のアメリカの研究チームには統計の専門家が加わることが多くなっている⁴⁾。また、論文投稿の際には条件を満たした生物統計家が統計に関するレビューを行うことが必須になっている場合があり、論文の査読者側にも統計の専門家が参加している。アメリカでは生物統計学を専門とする人材の豊富さを背景に医学研究における統計解析に関して質の維持を達成している。

わが国においては、生物統計学を専門とする人材が極めて少なく、研究を支援する段階には至っていないのが現状である^{5)~7)}。わが国でも世界トップレベルの研究を推進するために、欧米並みの研究体制の構築を行う必要がある⁸⁾。近年、わが国では生物統計学の専門家を大学の修士課程または博士課程で育成することが始まっている。北里大学、東京理科大学、東京大学、京都大学、久留米大学などがその例である。しかしながら十分な数の人的資源を確保するまでには、かなりの時間がかかるものと考える。

次に大学医学部で行うべき課題として、生物統計学に詳しい研究者が各分野の研究者と協力する体制を作ることである。これが実現できるならば、より効率的にそして質の高い研究を行うことができるところになる。

わが国では、いくつかの大学にて統計支援の実践を行っている例がある。京都大学では、薬剤疫学分野にて、医学部における研究の統計的側面を支援している⁹⁾。対象は、大学院生を含む医学部内の研究者であり、電話またはメールでの申し込み後に初回相談を行い、その後に依頼書を提出することで支援を開始している。相談の内容は、①統計解析や試験デザイン、症例数設計に関するコンサルテーション、②プログラムに関するアドバイスや統計解析の実施、③共同研究となっている。支援を行う側では「応用研究の場」と位置づけており、支援を受ける側も行う側も共に有益になるように運営を行っている。帝京大学では、医学部研究用コンピュータ室に相談室を配置している¹⁰⁾。ここでは、コンピュータネットワークやプログラミングに関する相談とともに、統計解析ソフトや生物統計学に関する相談を受け付けている。

一方で、このような体制を構築するには、どのように人件費を確保するのかという大きな問題がある。また、スキルのある人的資源を確保することが難しい中で、どの程度のスキルの人的資源を必要とするのかについても検討する必要がある。

おわりに

本稿では医学研究における統計学について、東京女子医科大学で行ったアンケート調査結果により現状と問題点を明らかにした。また、今後の方針についてわが国全体が行うべき課題と、大学医学部が行うべき課題の2つについてまとめた。医学研究の

現場は常に国際競争に曝されている。この厳しい競争に勝ち残るためには、欧米諸国並みの研究体制の構築が急務であると考える。

謝　　辞

本稿で取り上げたアンケート調査は、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) 総務省情報通信政策局：通信利用動向調査報告書世帯編。総務省情報通信統計データベース <http://www.johotsusintokei.soumu.go.jp/>
- 3) 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
- 4) 丹後俊郎：研究デザイン—無作為割り付けの重要性。「統計学のセンス—デザインする視点・データを見る目」, pp39-40, 朝倉書店, 東京 (1998)
- 5) 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・生物統計研会誌 **22** (1) : 61-70, 2002
- 6) 渡辺 亨：第三者的支援基盤での研究者主導型臨床試験の可能性。癌臨研・生物統計研会誌 **21** (1) : 79-87, 2001
- 7) 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
- 8) 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
- 9) 京都大学大学院医学研究科 社会健康医学系専攻 健康解析学講座 薬剤疫学分野：臨床試験デザイン/生物統計相談の実施について。 <http://square.u-min.ac.jp/kyotoupe/>
- 10) 帝京大学医学部研究用コンピュータ室：研究用コンピュータ室 2003年度運営状況等報告。 <http://www.med.teikyo-u.ac.jp/~kenkon/>