

改訂にあたって

Stata を用いた分析で論文を執筆する際のポイント

2014 年に『Stata による医療系データ分析入門』を刊行してからおよそ 9 年の時が流れました。この間も Stata を使って臨床研究の論文を出し続けました。そして 2019 年、アメリカ医師会誌(JAMA)などの医学研究のトップジャーナルに論文が掲載されるようになりました。今まで共著も含め 200 編以上の英語論文を誌上発表し、他の医学専門誌に 1 万回以上引用されました。特に 2000 年以降のものは大概 Stata を使って論文化しています。

しかし、ここまでの道のりは順風満帆なものではありませんでした。今でも決して平たんな道ではありません。論文アクセプトの裏ではおそらく 1000 回くらい論文を投稿しましたがことごとくりジェクトされてきたからです。そんなとき、建築家、安藤忠雄先生が御著書「連戦連敗」の中で述べられている文章に勇気づけられました。

… これまで建築家として歩んできた 30 年余りの道程を振り返ったとき、思い通りに進んだと言える仕事がほとんどないからです。… 結局計画中断という事態も少なくはないし、コンペ（競技設計）に挑戦しても、大概が敗退に終わっている。都市への提案に至っては、実現に至るどころか、聞いてももらえないのがほとんどです。そのくり返しでここまでやってきたのですから、文字通り、連戦連敗といった状況なのです。… 厳しさゆえの緊張が、あるいは極限状態での可能性の追求が、本当の意味での創造につながると私は考えています。なぜなら、逆境は最良のチャンスだからです。… ギリギリの状態に追い

込まれば、人間は考えざるを得ない。…

英語論文を投稿すると「箸にも棒にも掛からぬ」場合、1時間くらいでリジェクトされます。これは事務レベルでリジェクトなのだと思います。そして編集会議のレベルで落選すると1週間で返信があります。一方、脈がある場合には外部に査読を依頼します。ですから、論文投稿して2週目に入ると「査読に回った」と考えてよいでしょう。

私も査読を依頼されることがありますが、通常の場合は2週間です。のりしろを含めてほしい1カ月前後に審査結果が主任研究者にeメールで返されます。しかし、編集部が最初の2~3人でアクセプトかリジェクトの判断がつけられないと2巡目の査読に回されます。このような場合、2か月、場合によってはそれ以上かかることもあります¹。

メールのタイトルには Decision Letter と書かれています。これを開くとアクセプトかリジェクトかが判るわけですから、受験の合格発表のようなものです。いつもドキドキさせられる瞬間です。パターンは4つに分かれます。

1. そのままアクセプト
2. 少し修正を加えアクセプト
3. 大幅修正をして査読者、編集委員が納得すればアクセプト
4. リジェクト：査読者のコメントを参考まで添付

依頼原稿ならいざ知らず、1あるいは2のことはめったにありません。大半は3か4です。

¹ 但し、編集部の方で適切な査読者を見つけられずに無駄に時間が過ぎている場合もあるので、要注意です。

4 の場合、'I regret to inform you that this paper cannot be considered further for publication in the Journal.'など丁寧にリジェクトであることが記載されています。「他のジャーナルに投稿しなさい」などとキツパリと書いてあることもあります。悪あがきしても無駄です。

問題は3です。例えば”interesting”という単語が多くて「2週間で再投稿せよ」といったことが書かれていればアクセプトの可能性は高まります。一方、”concern”という単語を使う査読者が1人でもいるとリジェクトになる可能性も十分あり慎重にリバイスする必要があります。

論文に対して心えぐられる査読コメントが返ってきたとき、思わず目をそむけたくになります。自分が苦勞して書いた論文を批判されて心穏やかでいられる人は少ないのではないのでしょうか？

そんな時、私は剣道の達人、千葉先生の御言葉を思い出すようにしています。大学生時代、警視庁の剣道場の朝稽古に通い、全日本剣道選手権大会で3回も優勝した千葉仁先生に稽古をつけていただいたのは今でもよい思い出になっています。千葉先生はかく語られました。

試合で勝てば相手に対して勝たせてくれて感謝、

負けても自分の弱点に気付く機会を与えてくれて感謝

何とポジティブな考えでしょう！論文で厳しい批判を受けても、「自分の弱点、論理展開の

甘さに気付く機会を与えていただきありがとうございました。しかも無料²で領域専門家に自分の論文を読んでいただき、名前は判らないけれど世界的に高名な先生のご指導を受けられて自分はなんて恵まれているのだろう。eメール上のやりとりではあるが、そんな臨床研究の神様と会話できた経験はお金で買えるものではない」と思うと不思議と気分が明るくなるものです。

査読者のコメントは坐禅修行中に背中を警策で叩かれるようなものです。自分では気づかない姿勢の悪さを指摘してもらい、「ハッ」と我に返るのです。厳しい批判であれば「喝」かもしれません。時間を置いて冷静な気持ちになったら再度コメントを一つ一つ丁寧に読み返し回答していきます。最近はこの問答を楽しむことができるようになりました。厳しいコメントに対しても論理的に矛盾なきよう回答を返します。査読者が降参すれば' We have completed our review of your manuscript and are pleased to accept it for publication in XXX journal.'といった返事が編集部より寄せられます。これは著者側の勝利です。

このように世界の研究者と議論を重ねることで、自分の医学研究者としての成長を実感できます。また、データが集まり Stata で解析して統計学的有意差を検知したとき思わずガッツポーズをしまいます。世界中の誰も知らない事実を自分だけが知った瞬間、これこそが研究の醍醐味だと感じています。

そして、研究の楽しさを知ると、日々の臨床が楽しくなります。何故なら患者さんやご家族の話に耳を傾け、丁寧に診察することで次の研究のヒントや仮説が見えてくるからです。教科書にはわかっていることしか書いてありませんし、他者により誌上発表されたエビデ

² 私が知る限り査読の報酬はありません。

ンスを検証しても nothing new です。

そこでこの改訂版では私の実体験も交え「Stata を用いた分析で論文を執筆する際のポイント」を加えました。また、研究の裏話もコラムとしてまとめてあります。論文だけ読むとドライなものですが、その裏には沢山の物語があり、実は研究の過程と発表、そして社会実装までの間に喜びや苦悩、いろんなドラマやナラティブな世界があるものです。

初版では Stata でできる医療系データ分析を広くカバーしようと心がけました。しかし、その後、東京図書より「医師が知りたい医学統計」と「臨床研究論文の読み方」を出版することができました。そこで、初版の流れはある程度踏襲しながらも、上記2つの著書に譲れる理論部分は省き、Stata で使うデータの多くを最近私が手掛けたものに入れ替えました。そして医学研究で最も使われる生存解析により多くの紙面を割きました。

第1章 Stata しよう！

基本的には同じです。Stata ver. 13 を ver. 17 にしました。変数名に日本語が使えるようになった点が大きな変化かもしれません。個人的に dofile は使わないのでここは割愛しました。

第2章 「研究デザインのエレメント」「ケース・コントロール研究」「コホート研究」☞「疫学のエレメント」にまとめました。そして2011年にドイツで発生した病原性大腸菌のエピデミック事例を入れました。標本分布やなどの統計学の基本原理は拙著「医師が知りたい医学統計」を参照いただければと思います。

第3章 「統計解析のエレメント」「仮説設定」「Stata 解析落穂ひろい」☞ 「統計のエ

レメント」

確率分布の部分や中心極限定理、ロジスティック回帰などを「医師が知りたい医学統計」と「臨床研究論文の読み方」に委ね、Mann-Whitney test やカイ二乗検定など医学研究論文を書く上で頻繁に用いる検定の部分を中心にもってきました。ここを理解できれば臨床研究論文の患者背景因子の比較の表を作ることができる構成となっています。また Stata で作った図をパワーポイントに張り付けて学会あるいは投稿用に綺麗に仕上げるための裏技を説明してあります（Windows だけで使えるようです。Mac ユーザーの方には申し訳ありません）。

第4章 「線形回帰」 + 「繰り返し測定」 ⇄ 「線形回帰」

最近、JAMA Network Open に掲載された拙著論文に差し替えました。編集委員や査読者とのやり取りが読者の皆さんの参考になると考えたからです。新型コロナの超過死亡の話題です。また、繰り返し測定も新型コロナの PCR Ct 値の比較という話題に差し替えました。

第5章 「ケース・コントロール研究」「マッチング」 ⇄ 「ケース・コントロール研究」

初版と同じ「乳児突然死症候群」を使ってあります。

第6章 「生存分析」 ⇄ 「コホート研究」

結果が術後30日以内の死亡率といった短期間でできるものはリスク比やオッズ比で解析します。初版では生存解析を扱う章はたった1章分しかありませんでしたが、今回は9章と10章に基礎と応用に分けて詳述しましたが、本章では神経芽細胞腫レジストリーのコホート研究を紹介します。

第7章 「ポアソン回帰」 初版と同じ内容です。

薬剤疫学の分野です。

第8章 「二重盲検ランダム化プラセボ比較試験」 ⇄ 「ランダム化臨床試験」

JAMA Pediatrics に掲載された ABC 試験に刷新しました。

第9章 「二重盲検ランダム化プラセボ比較試験」「生存分析」 ⇄ 「生存解析－基礎」

JAMA に掲載された AMATERASU 試験に刷新しました。基礎からサブグループ解析までを含めています。

第10章 「多重代入法」「生存分析」 ⇄ 「生存解析－応用」

JAMA 編集部とのやりとりも含め、多重代入法、競合リスクなどの Stata を用いた解析法も示しました。

第11章 「対象数計算」 ⇄ 「統計解析計画書 (SAP)」

最近、トップジャーナルに投稿する際には研究計画書 (protocol) と統計解析計画書 (SAP) の添付が必須です。今まさに進行中の AMATERASU 2 試験で解説しています。

第12章 「メタ解析」 ⇄ 「メタアナリシス」

メタ解析とメタ分析という2つの呼び名がありますが、中身は同じです。そこでタイトルをメタアナリシスとしました。Stata ver. 17 でメタアナリシスはやりやすくなりました。内容は基本的に同じです。

第13章 「診断学」 同じ

第14章 「時系列分析」 同じ

「プロペンシティ・スコア」は進化を遂げています。しかし、「confounded by indication」という大きな壁はプロペンシティ・スコアを駆使してもランダム化臨床試験を越えることはできないし、かえって読者に誤解を与えかねない」という私の考えもあって、今回の書籍からは外しました。「医師が知りたい医学統計」の Chapter 18. 多変量解析の 223～226 ページに「プロペンシティ・スコア」の実例も交えて解説してあるのでそちらをご参照ください。

読者の皆さんが Stata を使って多くの医学エビデンスを創生し、これが社会実装され、そのことで今よりも多くの人々が健康で幸せな人生を送れることを願ってやみません。