【助成 39－57】

評価表現分析で読み解く医学論文査読者の嗜好と思考：査読自動化へ向けた基盤研究

代表研究者　福島県立医科大学附属病院臨床研究教育推進部　副部長・特任准教授　　大前　憲史

〔研究の概要〕

論文の質を複数人の専門家で評価する査読の仕組みは、医学エビデンス創出の基盤を成す。しかし、日本では、臨床医が忙しい臨床の傍ら研究に従事することも多く、査読を学ぶ機会は劇的に不足する。また、査読そのものも、性質上閉鎖的で、実態が不透明な部分も多く、査読教育を行う上での大きな障壁となってきた。本研究は、実証的エビデンスに基づき現行の査読の実態を明らかにし、効果的な査読教育の開発に繋げることを目的とする。そのため、まず、公開査読制を採用する医学研究雑誌の査読レポートを集積し、査読に特化した膨大な言語表現データベース（コーパス）を構築する。さらに、コーパスを利用して、医学研究のデザインや方法論的枠組みに着目した疫学的分析と査読特有の構造や評価表現に着目した言語学的分析から、査読者の思考や嗜好を質と量の両面から明確化する。

〔研究経過および成果〕

本研究の最大の目的は、医学論文査読に特化した、膨大なデータから成るコーパスの構築である。査読対象となる医学論文には、一般的に IMRAD（背景・方法・結果・考察）と呼ばれる共通の枠組み的構成要素があり、その構造を体系的に捉えることはそう難しくない。実際、著者側の視点で作成された医学論文コーパスは既に構築されており、コーパスを用いた論文執筆サポートツールもいくつか存在する。他方、査読レポートは、雑誌ごとにある程度形式が定められてはいるものの、論文ほど定型ではないことが多い。そのため、最初のステップとして、査読レポートの構成要素を明らかにし、意味的・構造的まとまりに分類する作業が必要となる。言語教育の分野では、このまとまりを Move と呼んでいる（Move の下位区分を Step と呼ぶ）。Move を分析することで、効果的に論文の構造を明らかにしたり、説得力のある文章を構成したりできるようになる（図参照）。

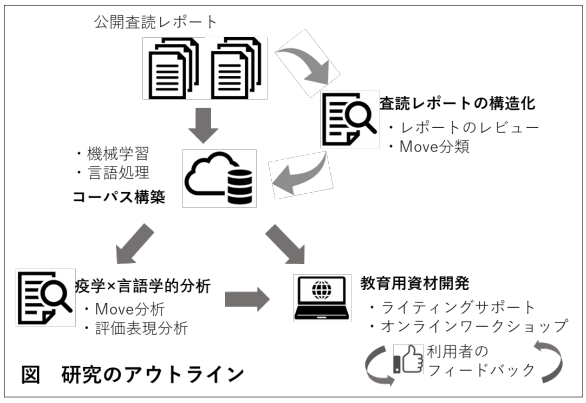

図　研究のアウトライン

我々はまず、研究の第一段階として、公開査読制を採用する医学研究雑誌のうち、4 大トップジャーナルの1つとしても知られる The British Medical Journal (BMJ)に絞って、BMJ に掲載されたランダム化比較試験（RCT）論文に対する査読レポートを解析することにした。トップジャーナルでは一流の査読者が査読を行うため、質の高い査読レポートの入手が期待できる。また、観察研究と比較して、RCT では研究の方法論がある程度決められており、査読レポートも構造化しやすいことが想定される。さらに、RCT は医学以外の領域でも広く用いられる研究デザインであり、得られ

た知見の汎用性も期待できる。

そこで、BMJ で出版された RCT 論文のうち、出版時期が最近のものから適当に 10 本抽出した(うち、2 本は最初の査読結果で Reject と判断されたものを含めるようにした)。査読を受ける前の原稿とそれに対する初回の査読レポートを BMJ のウェブサイトから PDF ファイルで入手し、Python を用いて査読者のコメント部分のみ自動的に抽出できるようにした。その上で、2 本の RCT 論文に対する計 10 名の査読者コメントを意味的なまとまりに分類し、各パートに意味づけするアノテーションを 2 名の臨床疫学専門家が独立して作業した。その後、互いの作業結果を照らし合わせたが想定以上に両者の見解の相違が多く、言語学および教育学の専門家も加わって両者の合意形成がなされた。このプロセスでの気づきとして、同一の雑誌であっても査読者による査読レポートの形式上の異質性が非常に高いということと、その一方で、査読者が言及する項目や内容にはある程度の集約が見込めることがわかった。そこで、作業の効率化を図るため、引き続いての 4 本の RCT 論文に対する計 22 名の査読者コメントへのアノテーションは、まず 1 人の臨床疫学専門家で行い、それをもう 1 人の臨床疫学専門家が検証するようにした。この作業では、マイナーな相違点はいくつかあったものの、メジャーな相違点は 1 か所のみで、それも両者の議論を通して容易に合意形成が得られた。

これまでの解析の中で、査読者が最も言及することの多かった項目は元論文の方法に関連する事項であった。特に、①アウトカムの測定や定義、②サンプルサイズ設計や解析手法、③研究対象者の選定、④当初のプロトコールとの相違について、より明確化するための追加情報を要求する内容のコメントが多かった。次に多かった、元論文の考察に対するコメントでは、①選択バイアスや一般化可能性に関連する事項に加え、②結果についてのメカニズム的な解釈を求める内容がよく言及されていた。これらの過程で得られた知見は、下記の論文や学会にて報告を行った。

今後は、さらに解析する査読レポートの数や種類を増やし、査読者コメントの意味的・構造的分類(Move や Step)について理論的な飽和を目指す。その上で、コーパスに取り込んだデータを利用して、各パートでどのような言語表現が使用されることが多いか、評価表現辞書を用いた言語学的解析を加えていく予定である。

〔発表論文〕

1. 大前憲史. 効果的な論文査読を行うために知っておきたい重要な視点. 老年看護学(日本老年看護学会) 第 27 巻第 2 号. 2023 年 1 月.

〔発表学会〕

1. 大前憲史. 量的研究論文、私はこう書く. PCR Connect 2022. 2022 年 12 月. オンライン.
2. 大前憲史. 研究論文のどのパートが最も重要なのか? Introduction, Methods, Results, Discussion. 日本臨床疫学会第5回年次学術大会 2022 年 11 月. 東京.
3. 大前憲史. 効果的な論文査読を行うために知っておきたい重要な視点. 日本老年看護学会第 27 回学術集会. 2022 年 6 月. オンライン.