

はじめに

医学領域の研究では、個人差の大きいヒトを対象に、治療法や薬剤の効果を客観的に評価すると共に、一見しただけでは見逃されてしまう潜在的な関連性を見つけることが目標です。近年は evidence-based medicine という言葉に表されるように、誰が見ても明らかな証拠となる研究、すなわち「この薬を投与すると生存率が改善する」といった仮説を検証する研究（検証的手法）がもてはやされています。このような研究により様々な誤解や迷信が否定されたという素晴らしい功績がもたらされたものの、膨大な費用をかけた大規模研究でありながら不適切な研究モデルにより誤った結論に導かれた研究も少なからず存在します。では、このような不適切なモデルが、なぜ研究実施前に指摘されなかったのでしょうか？ 多くの場合、データ解析を適切に実施する手順を知らないまま、興味関心のある仮説だけを検討してしまったことにあるようです。

私たち臨床研究者は、日常臨床現場から得られる情報に対して常に疑問を持ち、今まで気付かれなかった関係性を見い出して研究に結びつける探索的手法をとることが多く、最初から研究テーマや仮説が存在するわけではありません。実際には、手もとのデータを様々な角度から検討し、分布形式や関連を把握し、このような地道な作業で抽出された様々な情報から証明すべき仮説を練り上げ、大規模研究を計画・実施するのです。では、どうすればデータを多角的にとらえ、潜在的な関係を抽出することができるのでしょうか？ このようなノウハウについて合理的な手順で解説することは、実は非常に難しいのです。

本書は、医学領域の研究で用いられる統計手法を統計ソフト JMP® で簡単に計算させたい方が対象です。私自身、JMP がバージョン 5 の時代からのユーザーですが、それまで利用していた統計ソフト StatView® とのコンセプトの違いに戸惑ったものです。その後、JMP はバージョンアップを重ねて着実に機能が強化されてきましたが、ソフトウェアが複雑化すればするほど希望する統計手法を手早く実施するために簡潔なマニュアルが求められます。一方、臨床医学では多変量を扱う研究が増加し、単なる統計手法の選択だけでなく、多変数間の関連性に配慮した解析手順が求められるようになりました。このような要望に応えられるよう、JMP の操作手順を解説した単純なマニュアルにとどまらず、前述した問題に対して私が実際に行っている手順を交えて、統計学ではなく統計学的解析の手順を理解できる書籍を目標としています。例題として多変数の研究を想定し、段階を追ってデータを読み解き、最終的な仮説を練り上げて検証するまでを通じて、医学研究で頻用される統計手法と私自身が目にした研究論文の問題点を解決するための工夫を盛り込み、JMP で計算する手順を網羅しました。研究を実施する際にも、また査読者として統計手法を検討する場合にも、JMP の潜在能力を引き出すよう本書を活用頂ければ幸いです。

最後に、本書の構想から出版まで長年にわたり支援してくださったオーエムエス出版の皆様に、この場を借りて心から感謝いたします。

平成 30 年 1 月

長田 理

JMP 医学統計マニュアル

目次

はじめに	003
第1章 統計解析の基礎知識	007
① 医学研究での統計解析	008
② データの属性と統計解析	009
③ パラメトリック分析とノンパラメトリック分析	010
④ 多重比較検定法の選択	013
⑤ 検出力と標本サイズ	014
⑥ 臨床疫学的指標とカットオフ値の決定	015
⑦ 交互作用を検討する	017
第2章 医療分野における統計解析の定石	019
① データの特徴を把握する	022
② 集計表を用いて出現状況を検討する	026
③ 複数の変数間に潜む関係を調べる	029
④ 2群の数値データを比較する	039
⑤ 3群以上の数値データを比較する	045
⑥ 説明変数を用いて目的変数を予測する	055
⑦ 寿命（生存率）に及ぼす影響を検討する	063
⑧ 疫学的評価指標を検討する	071
⑨ 研究の信頼性を高める	073

CONTENTS

第3章 JMPでの解析手順	075
① JMPの基本的操作	076
② 記述統計 – データの分布を把握する	100
③ 分割表分析 – 集計表で検討する	104
④ 多変数の相関分析 – 複数の変数間に潜む関係を調べる	108
⑤ 主成分分析・因子分析 – 複数の変数間に潜む因子を調べる	112
⑥ 対応のあるペアの解析 – 変化量を調べる	116
⑦ 一元配置分析の解析 – 2群及び3群以上の複数群を比較する	118
⑧ 分散分析・共分散分析 – モデルのあてはめによる解析(1)	124
⑨ 反復測定分散分析 – モデルのあてはめによる解析(2)	130
⑩ 重回帰分析 – モデルのあてはめによる解析(3)	134
⑪ ロジスティック回帰分析 – モデルのあてはめによる解析(4)	138
⑫ 判別分析	146
⑬ ノンパラメトリック生存時間分析	150
⑭ Coxの比例ハザードモデルによる生存時間分析	154
⑮ 関数モデルによる生存時間分析	158
⑯ 再生モデルによる生存時間分析	162
⑰ 標本サイズ/検出力の解析	166
⑱ 傾向スコアを用いた背景因子の調整	174
索引	180
参考文献	182