

編集後記

1980年代初頭に、日本の製薬メーカーからFDAへの申請が、多重性問題を理由に次々と却下されるというセンセーショナルな出来事があった。当時日本では、後知恵解析やNS同等が公然と認められ、多重性問題に関してもほとんど未知であったため大混乱が生じた。そこで、統計的多重比較法を研究テーマの一つに掲げていた筆者にコントローラー委員会から要請があり、本誌に多重性問題を解説する(1984; 12(2): 309-19, 12(3): 589-610)と共に、同等性(非劣性)検証のための新しい方法の提案(1986; 14(3): 467-75)等を行った。その流れで、1992年には厚生省から「臨床試験の統計解析に関するガイドライン」が刊行され、さらに1998年にはそれもベースの一つとして、「国際ガイドラインICHE9」が発刊され、「臨床試験のための統計的方法」は一段落したかに見える。しかし、臨床試験の現場は多様であり、ガイドラインで全て事足りるはずはないので、これに甘んじて落ち着くことには警鐘を鳴らしたい。

臨床試験ではないが、2007年12月25日に、タミフル投薬は非投薬に対し、異常行動発症を半減せしめるという驚くべき報道が一瞬流れ、その後各種メディアが一斉に沈黙を守ったことを記憶されているだろうか。これは当日のデータで、タミフル被投与者の発症率 $911/7,392 = 0.123$ と非投与者の発症率 $546/2,477 = 0.220$ とを比較した疫学班の「明瞭な推測結果」に基づく報道であり、疑いの余地なく見える。しかしながらこの主張は、タミフル投与群と非投与群が予め設定された2群比較の場合に正当化されるものであり、この場合は事情が異なる。この場合は、当初9,869例全体が非投与群であり、それが病状の進行に伴い最終的に被投与7,392例、非投与2,477例に変遷した。従って例えば初期の発症者については、タミフル投与例数(分母)は7,392より遙かに小さく、タミフル非投与例数(分母)は2,477より遙かに大きい。つまり、上記の推定値は被投与者については過小評価、非投与者については過大評価になっている。この実質分母に当たるものはリスクセットと呼ばれ、発症例ごとにその時点のリスクセットを考慮して発症率を評価すべきところを、最終集計のみで短絡的に解析した結果が誤報道の基であるが、専門家以外では、いや、疫学の専門家ですら実際に陥った誤りであり、改めてレディーメードでは済まされない統計解析の難しさを思い知らされる。当日の審議会ではこの誤りを指摘し、長時間の議論の末疫学班の解析は白紙に戻されたのだが、事前に用意されたプレスリリースが一瞬だけ流れたというのが事の真相である。

臨床試験の最近の話題として、アダプティブデザインと国際共同治験がある。背景にdrug lagがあり、これらの効率化の方向は必須とも言える。しかしながら、前者については例数変更等の柔軟性と引き換えに支払う代償について十分認識する必要がある。また、後者については保証される日本人に対する有効性・安全性の範囲を従来方式と十分比較し、認可後に情報を補うプログラムを検討する等のことが必須である。これらについて、かつてほど賑やかな論争が見られないまま定着しつつあるのがやや心配な昨今である。1993年の新GCP以来進行していると言われる国内臨床試験の空洞化もその一因なのだろうか。

(広津千尋)