

平成25年12月9日

日本計量生物学会
会長 大橋靖雄 様

厚生労働省保険局総務課

「産科医療補償制度」の見直し案に係る統計学的解釈について（依頼）

平素より厚生労働行政にご協力を賜り厚く御礼申し上げます。

さて、当課が事務局として運営をしております「社会保障審議会医療保険部会」におきまして、現在、分娩に係る医療事故の救済を行い紛争の早期解決を図るとともに、事故原因を分析し産科医療の質の向上を目的とした「産科医療補償制度」において、「補償対象となる在胎週数の見直し等」について議論を行っております。

この度、本制度の運営組織である日本医療機能評価機構事務局より、当部会に対し、「補償対象となる在胎週数の見直し」について提案がなされたところですが、その際、提案根拠とされた統計学手法及び解釈等の妥当性について、「第三者の生物統計学専門家に見解を求め、それらを踏まえてあらためて議論してはどうか」との意見を頂いております。

つきましては、ご多忙のところ、誠に恐縮ですが、生物学・医学などの諸分野の研究について、計量的・数学的・統計的手法の普及の推進等を目的に活動されている貴学会よりご意見を賜りたくお願い存じます。

以上、ご高配賜りますよう、宜しくお願い申し上げます。

日本計量生物学会 (Biometric Society of Japan)

設 立

1980年9月

現会長 大橋靖雄 先生

東京大学大学院医学系研究科公共健康医学専攻疫学保健学講座
生物統計学 教授

設立趣旨

生物学・医学・農林水産学・生態学・環境科学などの諸分野における研究を計量・数学的・統計的手法を用いて推進するとともに、研究手法の普及などを目的として設立されました。

このような手法は従来、生物統計学、生物測定学などと呼ばれていましたが、最近の計量的・数学的手法の発展に鑑み、「計量生物学」と呼称することとしました。国際計量生物学会 (IBS) の日本支部としての機能を有し、関連研究者相互の交流を行っています。

発 信

- ・臨床研究に関する日本計量生物学会声明 (平成 25 年 9 月)
- ・統計化の行動基準 (平成 25 年 4 月提出 同年 11 月改訂)

平成 26 年 1 月 20 日

「在胎週数別脳性麻痺発生率の統計的分析」に対する生物統計家としての意見

日本計量生物学会産科医療保障制度ワーキンググループ

東京大学大学院医学系研究科教授 大橋靖雄
(2013-学会会長)

東京理科大学名誉教授 吉村 功
(1997-2000 学会会長)

九州大学名誉教授 柳川 堯
(2001-2004 学会会長)

1. 経緯

厚生労働省保険局より日本計量生物学会に対して、産科医療保障制度の制度見直しに関する（公財）日本医療機能評価機構提出資料（平成 25 年 12 月 5 日第 12 回社会保障審議会医療保険部会資料 1）中の統計的記載について、その妥当性を専門家の立場から評価してほしい旨の依頼があった。2013 年 12 月 26 日に開催された日本計量生物学会理事会においてこの件について審議がなされ、現会長・大橋靖雄がワーキンググループを組織し、当ワーキンググループの責任において報告をするべき旨の決議がなされた。本報告書はワーキンググループ（大橋、吉村、柳川）による結論である。なお、本報告書では産科医療保障制度の背景とその社会的医療的意味、見直しをもたらす社会的医療的意味については触れず、以下の統計的判断

在胎週数の基準

在胎週数 33 週における脳性麻痺の発生率と各週における脳性麻痺の発生率との間の統計学的な有意差を検定したところ、在胎週数 30 週以下については統計学的な有意差があるものの、31 週、32 週においては 33 週との有意差は認められないとの結果であった。

したがって、在胎週数の基準について現行の「在胎週数 33 週以上」から「在胎週数 31 週以上」へ見直すことが適当と考えられる。

にのみ検討結果を述べる。

2. 総体的結論

統計学的検定（仮説検定）で有意差がないことを根拠にしてこのような判断を下す

ことは、統計学的に妥当でない。すなわち 30 週以下と 31 週の脳性麻痺発生頻度間に統計的有意差が存在しないことは事実であるが、これをもって判断の根拠とすることは妥当ではない。

3. 理由その 1

仮説検定は、差のあることが確率的に明瞭であるかどうかを判断する手法である。この手法で差があるとは言えないという結論が出たことをもって、差がないことが示されたと判断してはならない。これは仮説検定という手法の最も基本的な特徴である。

註と補足：差が実際に存在する場合に仮説検定によって「有意差がある」という結論が導かれる確率を「検出力 (power)」という。上記の説明は、検出力が確保されていない状態で「有意差がないから差がない」「有意差がないから同等である」と結論することはおかしい、判断できるだけの情報量が確保されているとは限らないから、ということ述べている。たとえば 1 例ずつの 2 群間検定は絶対に有意にならない。実は、この「有意差はないから同等」という論理により、1980 年代後半まで厚生省 (当時) は多くの (後に無効と判明したものも含め) 薬剤を認可して非難された過去がある。1998 年の国際ハーモニゼーション (ICH) E9「統計ガイドライン」で、この問題は統計的に妥当な「非劣性試験」を質保証のもとに行うべし、として解決された。この明確化には多くの日本人生物統計家の努力があった。

4. 理由その 2

「見直しの判断」は第 33 週の脳性麻痺発生率を軸にした仮説検定で各週の脳性麻痺発生率が第 33 週と同じかどうかを判断しようとしている。

ところが第 33 週の発生数はわずか 1 例である。わずか 1 例という数値は統計学的判断を下すのにあまりにも少なすぎる。仮にこの週にもう 1 例が発生していたら、第 33 週の発生率は 2 倍ということになり結論が全く変わってしまう。そのデータ不足を示すのが、信頼区間の幅の広さである。たとえば第 32 週のオッズ比の信頼区間を別添 2 の図から読み取ると次の値になる。

下限 約 0.02、 上限 約 100

これは第 32 週の発生率が第 33 週と比べて、50 分の 1 かもしれないし、100 倍かもしれないということである。このデータにはその程度の情報しか含まれていないのである。

信頼できる判断を下すには、もう少し多くのデータ・情報が必要である。

註と補足：疫学の大家ロスマンの教科書 (Rothman K (著) 矢野・橋本訳：ロスマンの疫学、篠原出版、2004.) でも、検定の結果である p 値は効果の大きさと精密さを一緒にした指標であり、疫学研究で過度に尊重することを避けるよう指摘がある。彼は「p 値関数」の利用をすすめているが、これはさまざまなレベルで (通常の 95% だけでなく) 信頼区間を計算する考え方である。信頼区間が大きい場合には、上記のように「情報量が少なく明確な結論がだせない」ことが分かる。ちなみに「見直しの判断」において 33 週 (脳性麻痺は 290 件中 1) と有意差がないとき

れている 31 週（151 件中 2）の間の p 値は両側で 0.2709（Fisher 直接確率検定）と確かに有意でないが、31 週の 33 週に対する相対危険は点推定値で 3.841、95%信頼区間は対数変換後の正規近似で 0.35 から 42.0、正確な無条件区間で 0.13 から 24154 となる。つまり 31 週は 33 週に比べ発生率が 10 倍以上という可能性も否定できないほど信頼区間は広く、信頼性の高い結論は出せないことが分かる。

5. より望ましい統計的分析

仮にこの程度の少量のデータから判断を下さざるを得ないとしても、仮説検定という論理的に不適切な手法を用いるのではなく、より適切な手法を用いて判断を下すべきである。

より適切な手法を選ぶには、工業の品質管理において「変化点の推定」として検討されている問題としてとらえるのが妥当であろう。変化点というのは、経時的変化の中で発生率に急な変化が認められる時点のことで、品質管理の分野では、不良率の急増あるいは急減の原因を探る際に、推定されるものである。本件においては、変化点が脳性麻痺が増える週数にあたる。あるいは毒性試験の分野において「最大無影響量の推定」として検討されている問題を応用することも考えられる。これは毒物を与えない対照群に対して投与用量を複数変化させた群の毒性発現を観察し、どこまで毒物の影響がないと判断できるかを推定する方法である。本件においては、無影響量が脳性麻痺発生が正常分娩と同様と考えられる週数にあたる。

本件に関しては、変化点あるいは無影響量を必要な精度で推定できるようにデータを収集し、然るべき専門家の助言を受けて、上記「見直し判断」を再検討することが必要である。その際には、精度のみならず、一般化可能性が保証されるデータであるのか、換言すれば本件の沖縄のように特定地域のみでデータで十分か検討する必要もあろう。

註と補足：最大無影響量の推定問題については、柳川堯編：統計科学の最前線、九州大学出版会、2003. の第 4 章に詳細な記述があり、その中でも仮説検定の適用が不適切であることが述べられている。これに替わる方法として推奨されているのが、モデル選択のための統計量 AIC による方法 Yanagawa-Kikuchi-Brown 法であり、これを 28-36 週に適用すると、31 週と 32 週の間に脳性麻痺発生率にジャンプがあるモデルの当てはまりがよいことが示される。

以上