

## 教育講演 I

## 検査医学におけるエビデンスとは？

岡田 正彦\*

**【要 旨】** 医療における統計処理の問題点を指摘し、臨床検査のエビデンスがいかにあるべきかを考える。統計処理の基本となっているのは  $t$  検定である。この方法を適用するには、① サンプルが無作為に抽出されている、② データが正規分布している、③ 両グループ間で分散に偏りが無い、という 3 条件をデータが満たしていなければならない。しかし、いずれも臨床研究で取り扱うデータでは満たされていないことが多い。さらに分析法によらず、① 対象が大規模で、② 追跡期間が十分に長く、③ 交絡因子の処理が適切に行われ、④ 総死亡が検討されている、という 4 条件がそろった研究で得られた結論が、本来、エビデンスと呼ばれるべきものである。そのような研究の例は LDL コレステロールに見ることができるが、ほかには見当たらない。検査医学では、正しいエビデンスを得るための努力がなされていない。この点を十分に認識した上で、検査医学の教育に当たる必要があるだろう。

**【キーワード】** 統計処理、交絡因子、総死亡、エビデンス、誤り

## はじめに

1980 年、米国心臓学会発行の *Circulation* 誌に驚くべき論文が掲載された<sup>1)</sup>。ある統計学者が、世界中の主だった臨床系専門誌を対象に、それぞれ統計処理が正しく行われているかどうかを調べた結果を報じたもので、約半数の論文に統計処理の誤りが認められたのだという。

この論文が刊行されてからすでに 30 余年が経過し、臨床医学の研究スタイルも様変わりした。対象者を大規模に設定し、長い年月をかけて追跡・調査するという方法で、データの信頼性も格段に増したはずであった。新しい研究スタイルが普及するに伴い、「エビデンス」という言葉も広く知られるようになり、流行語にもなった。しかし、相変わらず研究デザインの不備、統計処理の誤りが多く認められ、中には特定の商品进行宣传す

るために意図的に結論が歪められた論文さえある<sup>2)</sup>。臨床研究を取り巻く状況は、1980 年当時とあまり変わっていない。

本稿では、医療における統計処理の問題点を具体的に指摘し、臨床検査のエビデンスがいかにあるべきかを考えてみたい。

## I. 統計処理の誤り

統計処理の基本となっているのは  $t$  検定である。2 つのグループ間で測定値に有意な差があるかどうかを調べる方法だが、この方法を適用するには、① サンプルが無作為に抽出されていること、② データが正規分布していること、③ 両グループ間で分散に偏りが無いこと(同じ分布形をしていること)、という 3 つの厳しい条件をデータが満たしていなければならない。

サンプリングに関して、古典的な統計学のテキ

\*新潟大学大学院予防医療学分野(医学部検査診断学) okada1540@ybb.ne.jp

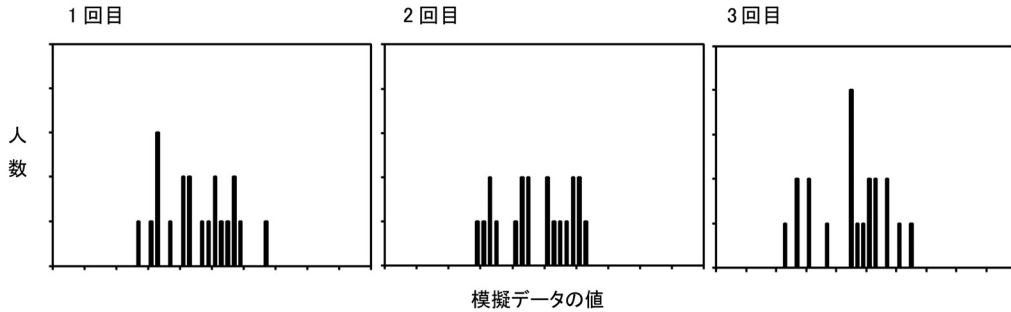


図1 20個ずつの乱数を用いたシミュレーション実験の結果

横軸は検査値に見立てた模擬データ、縦軸は各値を示したサンプル数(人数)。3回の独立した実験の結果をグラフで示したが、少数数のサンプルでは再現性がまったく認められないことが分かる。

ストには「サンプルが無作為に抽出されていれば、その個数は数十程度でよい」などと記述されていた。しかし、この記述が多く誤りを誘発する原因の1つになっていたかもしれない。臨床研究が対象とする個体には、あまりにも変動要因が多く、サンプルの無作為抽出は、いかなる手段を用いても不可能に近い。この問題を回避する唯一の方法は、サンプル数をできるだけ多く取ることなのである。図1は、筆者が行ったシミュレーション実験の結果で、検査値に見立てた20個ずつの乱数をコンピューターで発生させ、ヒストグラムを作成したものである。少数のサンプルでは、グループごとに分布パターンが大きく異なっており、再現性がまったくないことが分かる。再現性の高いグラフを得るためには、少なくとも千例以上のサンプル数が必要となる(データは省略)。

他の2つの条件(正規分布であること、2グループの分布形が同じであること)についても、臨床研究で取り扱うデータでは満たされていないことが多い。 $t$ 検定のほかにも相関係数、回帰直線などの統計処理法がよく用いられるが、やはり同じ問題がある。

## II. 最強の統計処理法

$t$ 検定、相関係数、回帰直線など基本的な統計処理法には、もう1つ共通する弱点(誤りを犯す要因)がある。以前、「家庭内暴力は砂糖の取り過ぎが原因」という記事が新聞に載っていたことがある。戦後、国内で砂糖の消費量が増えるとともに

に家庭内暴力の件数も増加している、というデータがあり、誰かがそのような研究発表をしたらしい。どう考えてもありえない話だが、その原因は、戦後の日本社会で起こっていたであろう数々の変化を無視し、単純に砂糖の消費量と家庭内暴力の件数の相関係数を求めてしまったことにある。正しい関係を求めるには、両者に影響を与える要因のデータをすべて集め、それによる変動成分を取り除くという操作がある。

このように、2つの出来事の間には介入する要因は「交絡因子」と呼ばれる。それらの影響を取り除いて真の関係を求める統計処理法もいろいろあり、総称して「多変量解析」という。正規分布をしていないデータに対しても適用できる便利な方法が「多重ロジスティック解析」で、これが最近の臨床研究でもっぱら使われている。最強の統計処理法と言ってよいだろう。

## III. 臨床検査のエビデンス

ところで臨床検査には、たとえば肝機能検査のように現在の異常を判定するためのものと、将来の発病の原因(リスク因子)を測るためのものの2通りがある。後者の代表は、血清LDLコレステロール値、中性脂肪値、空腹時血糖値、血圧などであろう。これらの検査法には、健常参考値(またはカットオフ値)の決め方が難しいという共通した特徴がある。健常参考値は、通常、「健常と思われる人の95%が含まれる範囲」と定義されるが、健常かどうかを判定するには検査を行うし

かなく、ジレンマに行き当たる。これを回避するには、特定の集団(コホート)で当該検査値を求めておき、長い年月をかけて発病の有無を追跡調査するしかない。

このように、①対象が大規模で、②追跡期間が十分に長く、③交絡因子の処理が適切に行われた、という3条件がそろった研究で得られた結論が、本来、エビデンスと呼ばれるべきものである。対象者の人数は、数千名以上、追跡年数も5~10年をかけたものが多い。検査医学の分野における、そのようなエビデンスの代表は、血清LDLコレステロール値と血糖値に見ることができる。

図2は、イタリアで行われた大規模調査の結果で、血清LDLコレステロール値と11年後の総死亡との関係を示したグラフである<sup>3)</sup>。男女で死亡率が2倍も違うことや、どちらも滑らかな曲線になっていてカットオフ値が定め難いことなどが分かる。一方、図3は、空腹時血糖値と数年後の糖尿病性網膜症発症率との関係を示したグラフである<sup>4)</sup>。LDLコレステロールの場合と異なり、110mg/dl前後を境に発症率が急上昇している、これまで伝えられてきた同検査のカットオフ値がほぼ納得のいくものであることを示している。この2つの研究では、総死亡、あるいは糖尿病性網膜症発症が検査を行うことの目的とみなされているが、研究デザインの立場から「エンドポイン

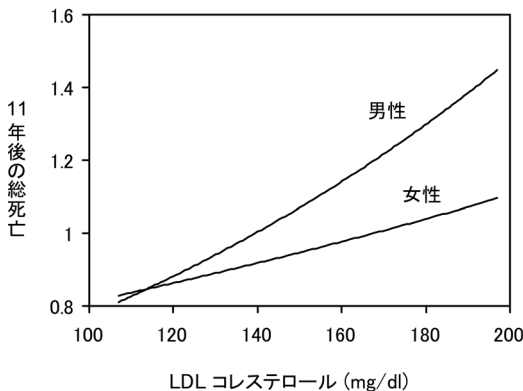


図2 血清LDLコレステロール値と11年後の総死亡(ハザード比)との関係  
(文献<sup>3)</sup>のデータより筆者が作図したもの)

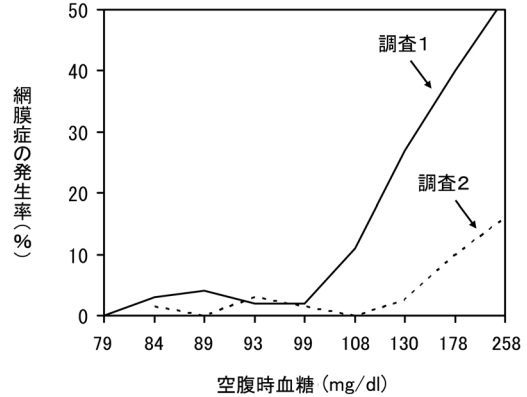


図3 空腹時血糖値と将来の糖尿病性網膜症発症率との関係  
(文献<sup>4)</sup>のデータより筆者が作図したもの)

ト」、統計処理上は「目的変数」「従属変数」などと呼ばれる。

#### IV. 究極の目的変数

近年、医療のさまざまな分野で行われている大規模な追跡調査、あるいは臨床試験の多くは、「総死亡」を究極の目的変数とみなしている。原因を問わず、あらゆる死亡を合計したものが、とくに生命予後に関わるような重大な疾患の場合、何らかの医療行為を施す以上、その結果が死亡率の減少につながるようなものであってほしいからである。しかし検査医学の分野では、総死亡を目的変数とした研究は今のところ図2で示したもののしかない。ほとんどの研究で、「特定の疾患の発病」、あるいは「特定の疾患による死亡」と設定されているにすぎないのである。

ある降圧剤の効果を確認するために行われた研究で、意外な事実が分かった。血圧を下げる一番の目的は、脳卒中を予防するためだが、その薬は確かに脳卒中死亡を減少させる効果が確認できた。しかし、心筋梗塞による死亡が逆に増えてしまい、結局、総死亡は改善していなかったのである(一部の降圧剤に脂質代謝を攪乱する副作用があるため)。これでは薬を服用する意味がない。それにもかかわらず、降圧剤について行われたほとんど大規模臨床試験では、「血圧が下がること」「脳

卒中の発症が減少すること」「脳卒中死亡が減少すること」などが目的変数となっていて、総死亡が無視されてしまっている。中には、意図的に隠されていたりする論文もあるらしい。同じ過ちを検査医学の分野でも犯している可能性がある。図3で示した空腹時血糖のデータにも同じ状況があるが、大規模な追跡調査が行われているという点では、大いに評価されるべきものであろう。翻って、大部分の検査データでは、追跡調査さえ行われていないのである。

数年前に始まった特定健診・特定保健指導(いわゆるメタボ健診)では、血清 HDL コレステロール、中性脂肪なども必須項目として位置づけられている。しかし、どちらの検査法についても、総死亡との関係を調べた研究データは今のところ存在しない。

HDL コレステロールの低値は、冠動脈疾患のリスク因子としてよく知られており、その重要性については論を俟たない。しかし、総死亡との関係を示すエビデンスは、やはり存在しない。

一方、中性脂肪は、主にカイロミクロン、VLDL、IDL の3種類のリポ蛋白によって血液を搬送されているが、脂肪自体が血管壁を傷害することも、また血管壁にたまることもなく、病気の原因とはなりえない。問題は中性脂肪の担送体であるリポ蛋白にある。まずカイロミクロンは、食事由来の脂肪酸搬送を担っているものであり、冠動脈疾患のリスク因子にはなりえず、一方、VLDL と IDL は、複数のコホート調査から、将来の心血管障害発症のもっとも重要なリスク因子であることが示されている<sup>5)</sup>。とくに VLDL は、血管内皮細胞に直接作用して血液凝固を促進する<sup>6)</sup>。つまり、血液中に存在する中性脂肪の総量(善玉と悪玉を混ぜたもの)を測っても意味がないのである。実際、数々の疫学調査によれば、中性脂肪値だけが単独で高くとも、ほかに異常がなければ、将来の心血管障害の発病には結びつかない。中性脂肪は、LDL コレステロールのように悪玉成分に限定した測定法が必要になるだろう<sup>7)</sup>。(睥炎で著しい高中性脂肪血症をきたすことがあるのは、

中性脂肪がリスク因子として作用しているのではなく、結果である可能性が高い。インスリン分泌が低下するとリポ蛋白リパーゼ(LPL)が作用しなくなるからである。)

結局、メタボ健診では、エビデンスのない検査を国民に強いていることになる。

## V. ま と め

紙面の制約で、心血管障害のリスク因子となる検査データに特化して統計処理上の間違いについて述べたが、検査医学全体に関わる重要な問題でもある。現在の検査医学は、あまりにも多くの問題を抱えている。この点を十分に認識した上で、検査医学の教育に当たる必要があるように思われる。

## 文 献

- 1) Granz S. Biostatistics: how to detect, correct and prevent errors in medical literature. *Circulation* 1980; 61: 1-7.
- 2) Editorial, Is GSK guilty of fraud? *Lancet* 2004; 363: 1919.
- 3) Tikhonoff V, et al. Low-density lipoprotein cholesterol and mortality in older people. *J Am Geriatr Soc* 2005; 53: 2159-64.
- 4) WHO, Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia. Report of a WHO/IDF Consultation 2006.
- 5) Malmberg K, Båvenholm P, Hamsten A. Clinical and biochemical factors associated with prognosis after myocardial infarction at a young age. *J Am Coll Cardiol* 1994; 24: 592-9.
- 6) Nilsson L, Gäfvels M, Musakka L, Enslin K, Strickland D.K, Angelin B, et al. VLDL activation of plasminogen inhibitor-1 (PAI-1) expression: involvement of the VLDL receptor. *J Lipid Res* 1999; 40: 913-9.
- 7) Okada M, Saito T, Yoshimura H, Noguchi Y, Ito T, Sasaki H, et al. Surfactant-based homogeneous assay for the measurement of triglyceride concentrations in VLDL and intermediate-density lipoprotein. *Clin Chem* 2005; 51: 1804-10.