(5) 世界的に見た日本の臨床試験のレベルとその推移

司会：谷田憲俊（兵庫医科大学第 4 内科）

編集部：このセッションは、昼休みに読んでおいてくださいということで、英文論文の一部（2 論文）と日本語論文を参加者に配付しました。午後の、論文評価のハダッドスケールで使用するためです。柳氏の話に先立って、浜から論文抽出までの経緯の紹介をしました。

浜：日本の臨床試験を世界的なレベル、一番簡単な基準で一度ざーっと見てみるということをしようと考えたのですが、昨年のオーストラリアのアデレードで開かれたコクラン・コロキウムに参加したときに、カナダの Moher さんという生物統計学者で疫学者でもある方にお会いして相談しましたら、1980 年、1990 年、1995 年の臨床試験論文を 50 ずつくらいを無作為抽出して、ハダッドスケールという比較的確立したものを使って評価してみたらどうかとサジェスチョンを受けました。たまたま、92 年版二重盲検文献資料集という本が手元にありまして、二重盲検の論文数が 6 件以上あった 11 雑誌を抜き出しました。これらの雑誌に載っていた 80 年、90 年、95 年の論文を取り出したわけです。「比較試験」と銘打っているものは取り出しました。第 3 相の臨床試験と銘打っていて比較試験になっていないというものもありました。これはいったい何だと思いました。初めて体験いたしました。第 3 相の臨床試験は当然、比較試験で確認すべきと思いますが。そういうことで第 3 相となっておれば比較試験でなくても含めました。

その後 TIP のメンバーでそれらの雑誌から論文をハンドサーチし、乱数表を使って、50 論文を抽出しました。それを 4 人で、独自に、ハダッドスケールで判定しました。その結果をこれからお話します。

柳（やなぎ） 元和（もとかず） ——— 内科医師

ハダッドスケール

元々は痛みの臨床試験について、本当に意味のあるスクリーニングができるかどうかということで開発されたスケールです。測定項目をたくさん作って、その中で臨床経験のある医師の意見も突き合わせて、意味のある項目を選抜していく。最後まで残った項目で、トレーニングを受けていない人でも、このスケールを使えばある程度はきちんとした評価をできる、そういう作業を経てこのスケールが構成されています。

そのスケールは痛みの研究についてはきちんとしていますが、一般の臨床試験について使えるかどうかは確認されていません。しかしながら、われわれ日本の文献を概括するということを第一の目標にして、これを使って検索することにします。

採点基準

第一には、「無作為」あるいは「無作為化」と書いてあるかどうか。「二重盲検」あるいはこの概念に相当する言葉が書いてあるかどうか。「解析除外例」「脱落例」について詳細に1例も漏らさず記載されているかどうか。それぞれ、「イエス」であれば、「1点」を与える。

付加事項がありまして、質問1．無作為であるということに対して何の記述もなければ、1点を与えておこう。しかし、内容を読んでみて、これはランダマイゼーションではないと考えれば「－1点」として、±ゼロになる。二重盲検についても同様に、プラシーボを使っていないとかダブルダミーにしていない、活性物質が2つある場合にはダミーも2つ必要ですが、そういうことがきちんと記載されていない場合には、不適切と判定できる。つまり、書いていない場合には判定のしようがないが、書いてあって不適切な場合は「－1点」。逆に、手順をきちんとしている場合には「＋1点」加算で、合計2点にする。そういうことで、大まかに論文の点数にバラツキが出てきます。

3点を境にして、4点、5点は、かなり手順としてきちんとしている論文という判定を下せます。1点や2点の論文は問題がある。論文に記載がないか、方法論として認められないようなものだと判定できます。

判定の具体例

具体的には、ランダマイゼイション、どちらの組に入るかが事前にまったく予測できないことを保証する方法については「適切」と認めましょう、しかしそれ以外の方法、誕生日、入院日、交互に割り付ける、こういう方法は適切と言えない。次に来る患者さんはこっちだな、と予測できてはいけない、これは最も基本的なことです。二重盲検についても、飲んでみたらプラシーボと分かるのでは話しにならない。きちんと書いているけれども、実は嘘を書いていた、ということは見破ることは出来ませんが、一応文面については信用せざるを得ない。

脱落についても理由が分からないのは、厳密には「脱落」とは言わない。後追いの調査怠慢である。具体的には、一人ひとりの患者さんを追いかけて行って、どういうときにどうなったときちんと記載していないと、きちんとしたトライアルとは言えない、というのがコクランなどではそういう考え方です。もしも「来院せず」とあった場合にどうするかは（今回の4人の採点者間で）議論できていないのですが。私は、「来院せず」とあるだけでは「0点」だと思うのですが。何の情報も得られないわけで、その人に対してどういうフォローアップであったかが分からない。

論文の読み方——エパルレスタット論文を例に

無作為化の論文を読むときに、通常、われわれはサマリーを読みます。メドラインでサマリーだけなら手に入ります。これを読んで、だいたいの当たりを付けます。みなさんにお配りしている資料の英文論文では、サマリーに、「Duble blind：ダブルブラインド」という言葉が出てきます。しかし「Randomization：ランダマイゼイション」はどこにもない。どうもよくわからないな、となりますと、中身を読んでいかなくてはいけない。中身を読むとなると、小見出しが大事です。「Introduction：イントロダクション」には書いていないです。次の小見出しに Patients and Methods 方法論が出てくる。どういう人を対象としているかがわかる。ところがここを読んでも Randomization のことは書いていない。

この論文は、エパルレスタットが日本で承認されるに際して参考にされた日本語文献を英語に置き換えたのです。実は、日本語の論文には、「無作為割り付けをした」と書いてあります。何症例かをひと組にして、それをランダマイズドして、各施設に配りましたという表現です。これは従来の論文でランダマイゼイションというときに、コクランでいうところのランダマイゼイションには当たらない。おそらくこの英文を作成するときに、Randomization として送ったのではないか、しかし、削られたのではないかと私は思います。この論文は、アルドース還元酵素剤のレビューでは、採用されませんでした。おそらく査読した人が、Randomization と認めないと言ったのではないかと想像しております。

論文の読み方——トルレスタット

もう一つの英文論文（資料参照）はトルレス

タットです。サマリーの中に、Randomizationとあります。サマリーを読んだだけで分かります。あとは中身を読んで確認をしていく。小見出しの「Drug administration」にも randomisedという言葉があります。しかしこれは、ずっと全体を読んだ限りでは問題があると私は思います。Randomizationしているからといって、すべて良いとは限らない（このことの議論は今回は省きますが）。結論だけを申しますと、この論文はトルレスタットをかなり有望だと言っているのですが、昨日もお話しましたように（H. 糖尿病用剤）、トルレスタットは開発が中止されました。こういうような論文が一つあるからといって、信用するわけにはいかない。

ハダッドスケール評価例―その1

アトピー性皮膚炎に対する臨床効果という論文は、サマリーは英文です。これを読んだだけでは分かりません。本文を読みますと、英語論文と同じで、小見出しは、「はじめに」「対象および方法」…とあり、最後に「全般改善度」というのがあります。「試験方法」というのがあります。これには、プラシーボを使ったかとか、ランダマイゼイションしたかとかが、まったく書かれていません。これで第3相長期試験です。

認可にかかわる非常に重要な、しかもアトピーですから長期に服用しなくてはいけないだろう、そういうものでランダマイゼイションの記載がない、プラシーボの記載がない。どういうことかとずっと読んでいきますと、20mg、30mg、40mg、症例の構成という記載があります。やはり比較試験のようです。なぜ比較出来るのか理解できなかったのですが、主治医が投与量を決めるらしいのです。方法論に「治験薬は症状によって適宜増減してもよい」と書いてあります。投与量は主治医が決めて、その各ミリ数間で比較するという、比較試験です。この評価を4人がどうくだしたかは後で紹介します。

ハダッドスケール評価例―その2

塩酸イプサピロン、不安神経症に対する臨床評価の論文です。これは前期第2相、用量設定などを決めていく試験です。これも、小見出しは、「はじめに」「対象および方法」…とあって、どういう対象を選ぶか、インフォームド・コンセント、薬剤、スケデュールなどが書いてあります。形式は整っているように思います。試験薬剤を見てみますと、「外観上、識別不能な2.5mg錠、5mg錠、プラシーボの3剤を使った」と記載があります。識別不能とはっきり書いてありますから、一応文章上は信用せざるを得ない。投与量および投与方法については、「シングル・ブラインド」と書いてあります。ダブル・ブラインドではないことがわかります。

論文評価の第一関門――せめて形式を

つまり、形式だけでだいたいその試験の持つ臨床的な意味の重みが推定できる、ということです。しかし、これは第一段階です。論文の最終的な質とは関係ありません。たとえ形式を整えていても質的に悪い論文はあります。しかし、逆はないのではないでしょうか。形式が整っていないのに、良い質の論文が出るだろうか。そこにわれわれは着目したのです。ですからこれは論文評価の第一関門です。

こういうチェック項目を作って論文を読んでいけば、書いているかどうかは分かる、つまり医学に素人の方々でも分かる、チェックできる。例えば薬を病院でもらった。その薬についての論文を取り寄せた。全部を読むのは大変ですが、少なくとも、ランダマイゼイションしているか、プラシーボを使っているかはチェックできる。もしも使っていないとなると、この薬はちょっとおかしいのじゃないか、と疑問を持つきっかけになるだろう、というのがわれわれの狙いです。

慢性心不全におけるエナラプリルマレエイト

これは2つの試験を一つの論文にまとめています。1回投与における運動能の改善と、長期投与における「全般改善度」が出ています。プラシーボを使い、識別不能としています。問題は、9症例分を1組とする割り付けを行ったとあります。これが本当にランダマイズド・コントロール・スタディかどうかは議論の対象になると思います。

患者背景に、NYHA分類で2度、3度の患者ばかりです。4度つまり重症の人はいないということです。

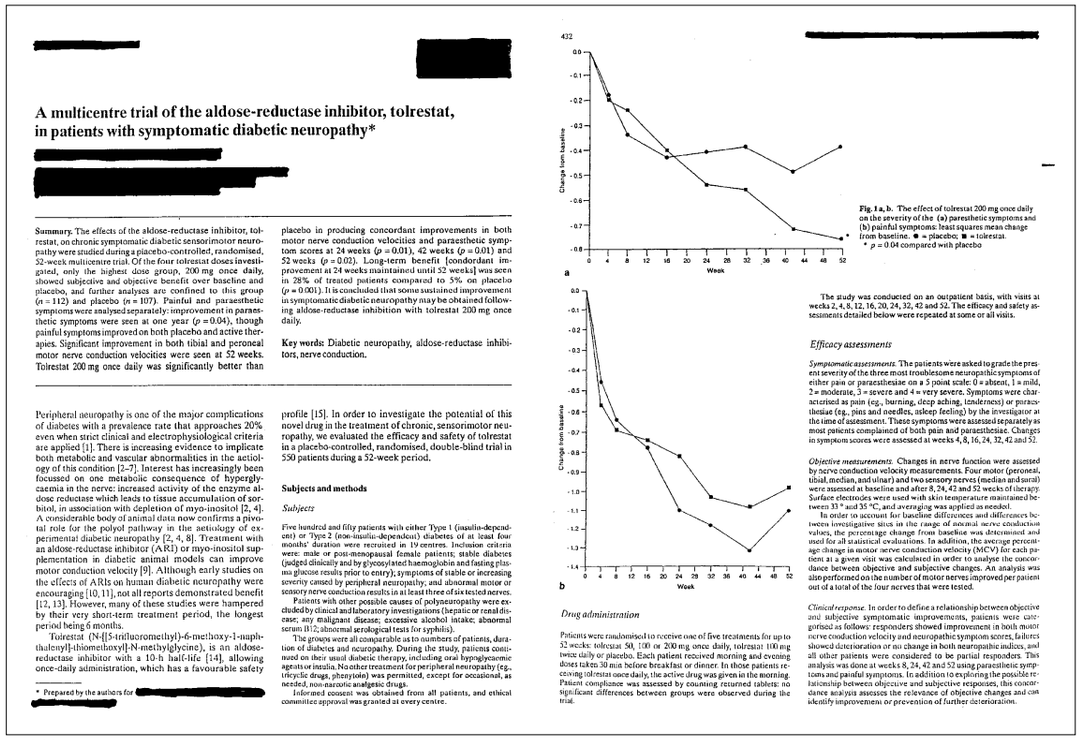
結論は、試験物質とプラシーボとで差がなかったんです。しかし、エナラプリルは国際的に認められている良い薬です。国際的に認められている薬剤が、「良い」という結論を出せなかった論文です。これは日本の文献で、良い薬を良いと言えなかった例です。有意差が出せなかったんです。この論文を読むと、日本の臨床家は、国際的に認められているエナラプリルに対して、あんまり効いていないと思うかもしれない。そういうおかしなことになる例として紹介しました。

4人の評価結果

4人とも評価が一致していたのが、1980年は17件、1990年は15件、1995年は24件です。95年になると論文の半分くらいは一致している。4人のうち1人は全く医学の素人です。それでも4人一致するのが24件も出てくるというのは、スケールが単純であることで比較的判定しやすい。1点や2点の論文は、このスケールでは判定しがたい要素があるということです。このスケールには限界がある。その分析はまだできていません。しかし、何のトレーニングも受けていない4人がこれだけ一致しているというのは、大まかなところでは有望ではないか、スケールとしてかなり使えるのではないかと思います。

みなさんに資料をお配りした論文の45番（エパルレスタット）は、1点、0点、1点、1点でした。ひどいです。97番（トルレスタット）は3点、1点、1点、1点です。一人は評価が甘いですね。122番（塩酸イプサピロン）は、3点、2点、2点、2点。きょうご紹介した3論文は、点数が低いところで一致している。

5点は不思議と4人が一致しますね。1980年に



A multicentre trial of the aldose-reductase inhibitor, tolrestat, in patients with symptomatic diabetic neuropathy*

[redacted]

Summary. The effects of the aldose-reductase inhibitor, tolrestat, on chronic symptomatic diabetic sensorimotor neuropathy were studied during a placebo-controlled, randomised, 52-week multicentre trial. Of the four tolrestat doses investigated, only the highest dose group, 200 mg once daily, showed subjective and objective benefit over baseline and placebo, and further analyses are confined to this group ($n = 112$) and placebo ($n = 107$). Painful and paraesthetic symptoms were analysed separately: improvement in paraesthetic symptoms were seen at one year ($p = 0.04$), though painful symptoms improved on both placebo and active therapies. Significant improvement in both tibial and peroneal motor nerve conduction velocities were seen at 52 weeks. Tolrestat 200 mg once daily was significantly better than placebo in producing concordant improvements in both motor nerve conduction velocities and paraesthetic symptom scores at 24 weeks ($p = 0.01$), 42 weeks ($p = 0.01$) and 52 weeks ($p = 0.02$). Long-term benefit [coordardant improvement at 24 weeks maintained until 52 weeks] was seen in 28% of treated patients compared to 5% on placebo ($p = 0.001$). It is concluded that some sustained improvement in symptomatic diabetic neuropathy may be obtained following aldose-reductase inhibition with tolrestat 200 mg once daily.

Key words: Diabetic neuropathy, aldose-reductase inhibitors, nerve conduction.

Peripheral neuropathy is one of the major complications of diabetes with a prevalence rate that approaches 20% even when strict clinical and electrophysiological criteria are applied [1]. There is increasing evidence to implicate both metabolic and vascular abnormalities in the aetiology of this condition [2–7]. Interest has increasingly been focussed on one metabolic consequence of hyperglycaemia in the nerve: increased activity of the enzyme aldose reductase which leads to tissue accumulation of sorbitol, in association with depletion of myo-inositol [2, 4]. A considerable body of animal data now confirms a pivotal role for the polyol pathway in the aetiology of experimental diabetic neuropathy [2, 4, 8]. Treatment with an aldose-reductase inhibitor (ARI) or myo-inositol supplementation in diabetic animal models can improve motor conduction velocity [9]. Although early studies on the effects of ARIs on human diabetic neuropathy were encouraging [10, 11], not all reports demonstrated benefit [12, 13]. However, many of these studies were hampered by their very short-term treatment period, the longest period being 6 months.

Tolrestat (N-[[5-trifluoromethyl)-6-methoxy-1-naphthalenyl]-thiomethoxyl]-N-methylglycine), is an aldose-reductase inhibitor with a 10-h half-life [14], allowing once-daily administration, which has a favourable safety profile [15]. In order to investigate the potential of this novel drug in the treatment of chronic, sensorimotor neuropathy, we evaluated the efficacy and safety of tolrestat in a placebo-controlled, randomised, double-blind trial in 550 patients during a 52-week period.

[redacted]

* Prepared by the authors for [redacted]

Subjects and methods

Subjects

Five hundred and fifty patients with either Type 1 (insulin-dependent) or Type 2 (non-insulin-dependent) diabetes of at least four months' duration were recruited in 19 centres. Inclusion criteria were: male or post-menopausal female patients; stable diabetes (judged clinically and by glycosylated haemoglobin and fasting plasma glucose results prior to entry); symptoms of stable or increasing severity caused by peripheral neuropathy; and abnormal motor or sensory nerve conduction results in at least three of six tested nerves.

Patients with other possible causes of polyneuropathy were excluded by clinical and laboratory investigations (hepatic or renal disease; any malignant disease; excessive alcohol intake; abnormal serum B12; abnormal serological tests for syphilis).

The groups were all comparable as to numbers of patients, duration of diabetes and neuropathy. During the study, patients continued on their usual diabetic therapy, including oral hypoglycaemic agents or insulin. No other treatment for peripheral neuropathy (eg., tricyclic drugs, phenytoin) was permitted, except for occasional, as needed, non-narcotic analgesic drugs.

Informed consent was obtained from all patients, and ethical committee approval was granted at every centre.

Fig. 1 a, b. The effect of tolrestat 200 mg once daily on the severity of the (a) paraesthetic symptoms and (b) painful symptoms: least squares mean change from baseline. ● = placebo; ■ = tolrestat. * $p = 0.04$ compared with placebo

The study was conducted on an outpatient basis, with visits at weeks 2, 4, 8, 12, 16, 20, 24, 32, 42 and 52. The efficacy and safety assessments detailed below were repeated at some or all visits.

Efficacy assessments

Symptomatic assessments. The patients were asked to grade the present severity of the three most troublesome neuropathic symptoms of either pain or paraesthesiae on a 5 point scale: 0 = absent, 1 = mild, 2 = moderate, 3 = severe and 4 = very severe. Symptoms were characterised as pain (eg., burning, deep aching, tenderness) or paraesthesiae (eg., pins and needles, asleep feeling) by the investigator at the time of assessment. These symptoms were assessed separately as most patients complained of both pain and paraesthesiae. Changes in symptom scores were assessed at weeks 4, 8, 16, 24, 32, 42 and 52.

Objective measurements. Changes in nerve function were assessed by nerve conduction velocity measurements. Four motor (peroneal, tibial, median, and ulnar) and two sensory nerves (median and sural) were assessed at baseline and after 8, 24, 42 and 52 weeks of therapy. Surface electrodes were used with skin temperature maintained between 33 ° and 35 °C, and averaging was applied as needed.

In order to account for baseline differences and differences between investigative sites in the range of normal nerve conduction values, the percentage change from baseline was determined and used for all statistical evaluations. In addition, the average percentage change in motor nerve conduction velocity (MCV) for each patient at a given visit was calculated in order to analyse the concordance between objective and subjective changes. An analysis was also performed on the number of motor nerves improved per patient out of a total of the four nerves that were tested.

Clinical response. In order to define a relationship between objective and subjective symptomatic improvements, patients were categorised as follows: responders showed improvement in both motor nerve conduction velocity and neuropathic symptom scores, failures showed deterioration or no change in both neuropathic indices, and all other patients were considered to be partial responders. This analysis was done at weeks 8, 24, 42 and 52 using paraesthetic symptoms and painful symptoms. In addition to exploring the possible relationship between objective and subjective responses, this concordance analysis assesses the relevance of objective changes and can identify improvement or prevention of further deterioration.

Drug administration

Patients were randomised to receive one of five treatments for up to 52 weeks: tolrestat 50, 100 or 200 mg once daily, tolrestat 100 mg twice daily or placebo. Each patient received morning and evening doses taken 30 min before breakfast or dinner. In those patients receiving tolrestat once daily, the active drug was given in the morning. Patient compliance was assessed by counting returned tablets: no significant differences between groups were observed during the trial.

▲資料

5点が一つありました。平均値はこういう採点に意味があるかどうかは別として、目安として報告します。1980年が4人の平均の平均が2点くらい。1990年は落ちまして、1.5点くらいです。1995年が挽回して1.7点くらいに回復してきました。1990年というのは、臨床試験として（質が）落ち込む時期だったのではないかと想像しております。昔のほうがむしろ良かった。90年代に入って、臨床試験の質が非常に低下している可能性がある。それがいろいろ批判を受けて、今、回復期にあるということではないでしょうか。

浜：柳さんの推測は、あながち間違いではないのです。薬事法が改正されたのが1979年です。それ以前から行われていた臨床試験が発表されるのが1980年で、大きな薬害が表沙汰にならないできた1990年以降に承認される薬剤がかなりルーズになってきたことと、多少関係するかと思います。

柳：1990年代に新薬として認可されたものは危ない、と思ったほうがいい。

浜：1995年頃に承認された薬剤というのは、1990年頃に臨床試験の報告がなされたものが多いですから。承認すべきでなかったような問題の新薬は94年、95年よりは96年のほうが少なかったです。多少、今後はましになるのかなと考えます。

谷田：最後のプレゼンテーションは非常に分かりやすいものでしたから、みなさん、消化不良にはならずに済んだかと思いますが、却って暗澹たる思いになるかもしれません。製薬会社が一方的に責められているように思われるかもしれませんが、実際に責められるのは日本の医者、研究者です。いかに今まで、いい加減なことをしてきたかを認識していただくものだったと思います。医療関係者だけでなく他の方々も含めて、自らのレベルを上げていかなくてはいけないということをこのデータは示していると思います。

■質疑／討論■

近藤：（医師、放射線科）質問です。イプサピロンの不安神経症に対する論文は前期第2相ですから、もともと一種の増量試験だと思うのです。ですからダブルブラインドや無作為化を要求するのは無理なのではないかと思いましたが、どうなのでしょうか。僕の知識が足りないのかもしれませんが。

柳：完全な用量設定試験であれば、ダブルブラインドにする必要はむしろありません。だからプラシーボを使う必要もないと思うのです。

近藤：プラシーボというのは、錠剤を…。

柳：こういうデザインで行われていること自体がおかしいのです。完全な用量設定試験であればプラシーボと比較するということは…。

近藤：はい、わかりました。第2点は、著者名や施設名をブラインド（編集部：参加者に資料として配った文献の著者や施設名を黒塗りしたこと。資料参照）にしたのはどういう理由からですか。

柳：施設とか著者の名前によって、採点者がこれは良い論文かどうかと先入観で見ないようにです。（会場爆笑）

近藤：さっきの浜さんの話で少し奇異に感じたのは、評価したのがどの薬なのか、名前が出て来ないことですが。

浜：最後のほうでちらっとお見せしました。

近藤：はい、少しね。どこかで発表されておれば別ですが。

浜：いえ、まだです。

近藤：匿名の批判は批判足りえないはずですが、批判する相手をきちんと出していかないのも批判にならないのではないかと思うのです。（柳：これは批判のためではなく評価の仕方のセッションですが。）ええ、そのことは分かります。ただ、評価するにも匿名の文献というのどうかなと思ったものですから。データがブラインドである理由は先ほどの説明でわかりましたが、最初にそういう説明がほしかったと思います。

柳：はい。雑誌名とか著者の名前で論文に対する評価が、かなり引きずられるという研究があるんです。ですから敢えてブラインドにさせていただきました。97番の論文について補足するのを忘れました。「脱落」「中止」の内訳をちゃんと書いています。しかし、53症例中31例が脱落とはどういうことか。（会場笑い）けしからんじゃないか、というのが私の評価です。

浜：もう一度、JIP/TIPが評価した95年～96年承認の新薬リストをお見せします。この中にはヒスマナールもあります。これは抗アレルギー剤でトリルダンと非常によく似た薬剤です。これ自身ももちろん（副作用として）QT延長から不整脈を起こすんですが、抗アレルギー剤としてどれか一つ残すとすれば、まあこれかな、と。しかし非常に注意して使わなくてはいけないということで、5ではなく4の評価にしました。95年には、結構、6つまり駄目だというのがあります。

今日はみなさんに一応お見せするために間に合わせようと非常に粗っぽい段階のものしか準備できませんでした。実際の評価しようとすると、非常なエネルギーが要ります。もしもみなさんの中に一緒に評価しようという方がおられましたら、是非手伝っていただきたいと思います。

佐藤：（東京医科歯科大学）非常に重要なことですので発言を。97番と45番の論文は、そもそもコントロールを置いた比較試験ではない論文なのです。それをもってきて批判するのは全くナンセンスでして…。

谷田：ちょっと待ってください。これは批判ではなくてトレーニング練習問題でして、評価項目を見ていただいたら分かると思うの

ですが、トレーニングなのです。

佐藤：だから、トレーニングであるにしてもあまりにも不適切な教材を取り上げたと思います。私の専門は主に市販後調査です。臨床試験のことは必ずしも専門ではない。先ほど近藤さんもおっしゃいましたが、そういう私にさえ、教材として不適切ということがわかるというのは非常に残念な発表だと思います。

例えば97番について。そもそもコントロールを置いた比較試験ではなくて、3群に分けていますが、あくまでも用量関係をみるためにプラシーボを置いているわけで。

谷田：これはコントロール・スタディではない、ということが分かればいいのですが。その背景はまた別のことでして。

佐藤：いや、論文の評価ということではなくて、本来そうではないスタディに対してダブルブラインドかどうかと判断させるのはどうかと思うのです。つまり二重盲検なり比較試験でない論文を取り上げて…。

谷田：ですから、この論文を読んで二重盲検でないということが分かっていただけたらいいのです。それから先は別の機会に話していただけますか。いろいろ消化不良の点はあると思いますが、時間が迫っておりますので。（会場で挙手あり）あ、まだ（会場笑い）、はい。

津谷：122番の論文ですが、実施方法のところで「割り付けは4症例分を1組とし、割り付け表はコントローラーが保管した」。この「割り付けは4症例分を1組」というのが何のことか分からないとおっしゃっていましたが、これはいわゆる「ブロック・ランダマイゼイション」というものです。割り付けの基本中の基本です。私の講義を受けた人は、学生も消費者グループの人もちゃんと知っています。つまり各施設に、この場合には2例、2例が行くようにという、施設をブロックとして割り付けたということです。

柳：それ、世界的な方法と考えられていますか。

津谷：はい、世界的な方法です。

柳：何十という施設に対して。

津谷：ええ、そうです。

柳：この点についてはまた議論したいと思います。

谷田：TIPの評価というのは、「物質」から「薬」と名前を変えたものが、果して本当に「薬」として承認してよかったのかどうか、またその「薬」が実際に臨床で本当に役に立つかどうかを判断するものです。

新しいGCPが今年4月から導入されて、今まで指摘されていたような問題は起こらないはずなのですが、日本の医療状況や研究者、医師、医薬品業界を含めて、果して新GCPに対応できるのだろうかと考えると、疑問の部分がかなりあると思うのです。そういう部分は、市民や患者さんが監視していかなくてはならないことだと思います。情報の公開に関して、次のセッション(P.460)で話がありますのでみなさんに十分にお考えいただきたいと思います。

本当に申し訳ないのですが、時間がないのでこのセッションはこのへんで終わりにさせていただきます。

編集部：ブロックランダマイゼーションは世界的に行われている方法であるが、1998年11月に厚生省医薬安全局から示された新しい「統計原則」ではブロックを十分長くし、2種以上の長さのブロックを用意し、ブロックの長さが知られないようにするなど、中央で管理すべき内容について詳しく規定している。2例、2例、合計4例を各施設に送り付けた後、中央は関知しないという従来の方法は明瞭に否定されている。
TIP Vol.14,No.2,P.15-19, 1999参照

第7章

最近の課題、医療事故 —(K)