

[シンポジウム 2]

抗生剤二重盲検法の実施と反省

(司会) 熊本悦明(札幌医大)

司会のまとめ

熊本悦明

(札幌医大泌尿器科)

はじめに

治療薬の有効性および有用性を評価する方法として、最近、二重盲検法の必要性がかなり強調されるようになり、新しく登場してくる感染治療薬としての抗生剤等(抗菌剤)の評価にも同様な検討が要求されるようになってきている。

しかし、新抗生剤の薬効を二重盲検法で臨床的に検討している第一線の医師の中に、経験的立場から、この方法の必要性に対し、いろいろな疑問や批判が生まれている。それがこのシンポジウムがもたれた大きな理由でもあるといえよう。

その問題点を4人の演者がそれぞれの立場から詳細に分析検討しているが、司会者の立場からそれらの意見をまとめてみたい。

二重盲検比較試験に対する批判を大別すると2つに分けられる。1つは感染症治療薬の薬効判定に二重盲検法が必須かどうかという基本的なものと、もう1つは、その実施内容、言い換えれば、現在一般に行なわれている plan の立て方への批判とがある。

二重盲検法の必要性の問題点

薬効判定法として次の3つのもが考えられる。

- 1) historical control との比較
- 2) 同時試行による open の比較試験
- 3) 二重盲検による同時施行比較試験

この3者のうち推計学的な立場からは3)が良いという結論である。

しかし、一般に感染症の治療を担当する医師の感覚には、治療効果を判定するパラメーターは、主観的判定の加わりやすい自覚症状などと異なり、客観的に判定するものばかりであるから、わざわざ二重盲検にして比較試験を行なう必要はないとする意見がかなりある。

また、それにもまして臨床医にとって問題なのは、2つ比較する薬剤が何であるか分からないということは、医学的立場からも、また患者との関係の倫理的立場からも困惑することが多いことである。人は共に open trial でいちおう有効であることが知られている2剤であるのだから、何も危惧することがないではないかともいう。

しかし、逆にいえば有効であることが明らかな薬剤の僅かな差を、わざわざいろいろな問題があるにもかかわらず、面倒な手続で明らかにする必要が臨床的に意味があるのだろうかという疑問がわく。わずかな差であれば症例の条件で同一の薬でも生ずるものであり、大きな差であれば open trial の段階でとうぜん明らかになっていることである。ただ二重盲検法で検討した data を出すというそれだけの目的のための data におわつてしまい、それ以前の open trial で得た知見を上まわる、臨床医にとって意義のある data が出るであろうかという疑問である。二重盲検法でなければ出ない差、すなわち推計学的にわずかな有意差を出すことが、薬剤の臨床的な有用性を比較することになるのであろうか。

少し極端になるかもしれないが、カラーテレビが最近次から次と新型が出てその差が論じられているが、それぞれに少しずつ特徴があり、わずかな差を多くの人がどれだけ有用と判定するかが問題である。大きな差が出てきて始めてその型が普及してゆくのではあるまいか。

それと同時に open trial でつかみ得ないわずかな差を出すために、第一線の医師に負担をかけた形で行なう必要はないのではないかと考える。どうせ投薬するのだから open でも二重盲検でも同じではないかという論理は、少なくとも感染症という急変してゆく患者に直接接している臨床医にとっては通らないことであり、二重盲検比較試験は感染症には安易な気持では行なえないものであり、行なう時はかなり慎重かつ研究的な態度が必要となる。

そのような考えに立てば、そこから出てくる data は open trial では明らかにし得ない、ひじょうに意義のある結論を得るためのものでなくては、できないことではあるまいか。Open trial でだいたい同じくらい有効ということが分っているものを、あえて二重盲検法で比較試験を行なつて、有意差なし、またはわずかな有意差ありということを証明するために、支払うものが大きすぎる。推計学的に興味があつても医学的に問題がありすぎるように見えてならない。そこに同一薬剤の open trial も行なっている第一線医師からの二重盲検法に対する強い批判、すなわち臨床的に意味のない data を出すための二重盲検法に対する批判が生まれる理由がある訳である。

とすれば、慎重に行なわれている open trial の data で充分であり、何故だめなのかということになる。その場合、第一線の医師以外の人々から指摘されるのは、今までの臨床医の所見のとり方、整理の仕方の曖昧さである。その人々は次のようにいう。「今までの臨床の data は統一がとれていない。施設により check point がまちまちであり、また同一施設でも case によりかなり不

揃いの data しかとつていなかつた。そのような data で正確な判定もできないし、さらに、一歩突込んでいえば、各種感染症の治癒判定基準すらも一定していない。そのようなことで historical control との比較とか open の同時比較を行なつても、薬剤の有効有用性をどのように判断していつて良いか分からないのではないか。ところが、最近行なわれるようになって二重盲検法による比較試験は、関係医師達が少なくともその試験の対象例では効果判定に必要な data をかなり努力して揃えて、しかも同じ基準で判定している。そのようにしてこそ data の比較ができるのではないだろうか。もし、それと同じような check point が理論的かつ同様のものであり、しかも同じ基準で判定された data の集積があるならば、それを対象として、薬剤の効果を比較することが可能となるといえよう。そうならば比較的客観的な data で効果を判定できる種類の感染症では、何も二重盲検法で検討しなくても慎重な open trial で充分であろう。要は、感染症治療経過を追跡する場合、重要な point を必ず check 整理し、同一判定基準で判定できるようになるまでは、現在一般に行なわれている open trial の data の信頼性に問題があり過ぎるのである」と。

その説にも一理ある訳で、やはり現行の二重盲検法を批判する前に少なくとも薬剤の比較試験を行なうための、理論的に整理された check point と判定基準の設定が必要となつてくる。その点、現状はかなり曖昧であるといつて過言ではあるまい。

現行二重盲検法の施行 plan 上の問題点

このように現在の感染症治療の follow の仕方が外から批判されていると同時に、医師側からも別な立場から問題が出ている。

すなわち、はつきりした check point が整理されていないため、二重盲検施行時、あれもこれもとできる限り煩わしさを省みず、data を集めてみようとする傾向がある。何故そのような多くの data が必要があるのかという疑問に対し、もしかしたら差が出るかもしれないからという planner の答が返つてくる。臨床的な有効性を check するには、最小限の必要な data を集めさえすれば良いわけである。ここで問題となるのは、現在、各種症状所見のそれぞれの疾患における重症度への貢献度、いいかえれば重要度係数が明らかでなく、どの所見を追えば正確に、治療効果を判定できるかが決つていないことが少なくない。

その混乱の中で、多々増々弁ず式の data 集めが行なわれる二重盲検法に第一線の医師から批判が出るのも当然といえよう。それほど細かく check しなければ、有意差が出ないならば、たとえ、出たとしてもその結果が

どれほど臨床的に意味があるのだろうかという疑問が出てくる。

そこで、各症例の重症度の算定、ついで各種症状や所見の重要度係数の算定が必須となつてくる。その点の整理がつけば、最少必要数の data の check で十分になり、かつ正確に症状を把握し、かつ治療経過を追求できることになる。各種感染症においてこの問題の整理が急務である。大きく言えば比較のおろそかにされていた感染症治療学の大系作りが迫られているといえよう。その評価法さえ整えば、たとえ二重盲検にしなくても感染症の病像の推移をかなり客観的に把握できる訳で、それによつて出て来た薬剤の効果判定は、同時比較試験の data として、さらには historical control としても充分信頼できるものとなり得る。とするならば、この各感染症の重症度を決める高い重要度係数をもつパラメーターは何か、そして抗菌剤で治癒してゆく経過で、それらのパラメーターがどのように変化してゆくべきものかということを公的に決めてゆくことが焦眉の問題といえよう。

それを、全ての感染症に、現在の段階ですぐできるとは思えないが、可能なものから選び、その治癒 pattern を分析し、check point および判定基準の決定を公的に決め、学会全体で利用できるように、委員会を作つて頂けたらと考えるものである。その点、大越学会理事長も会場で、皆の賛同を得て作つていきたいという発言をされ、この面での前進が期待できるのは嬉しいことである。

次に二重盲検法で問題となるのは、たとえ二重盲検法そのものに意義を認めるとしても、比較する dosis があまりにも“適当”すぎて、その疾患の治療上多からず少なからざる量ではなく、比較する 2 剤とも高 dosis で効果が充分すぎるくらいに出ている量で比較し、両者間に有効度に差がないと論じても、確かにその量で薬効があり臨床的に有用であることはわかつて、二重盲検で比較するまでもなく、両者の最大効果を出している level での比較の結論は、検討以前からすではつきりわかり切つているところである。担当医師からの批判が出るのは当然である。

これはまた薬効を check する投与期間とも関係がある。自然治癒傾向の強い疾患では、量が少なくてもある期間治療すると、自然治癒してゆく。そのため、不完全効果 dosis でも期間が長くなれば、完全効果 dosis との間に差が出ない。

そのため、先の完全効果 dosis での感染症の治癒 pattern の分析決定が望まれるが、その pattern が決つていれば、ある新しい薬剤がどの dosis でその pattern に入る効果を示すかの決定が医学的には興味もたれる。

ところである。すなわち、dosis response curve をかくような形で dosi s をかえての比較試験を行ない、その data から、ある疾患の治療に必要な dosi s を決めてゆく必要がある。

このような検討こそが臨床的に意義のある結論を出すものといえよう。この data は *in vitro* での MIC が *in vivo* でどのような意味をもっているかの分析ともなり得る訳である。

全ての疾患ではこのような試みは不可能としても、可能なものでの data に基づいて、他の疾患での比較試験の dosi s を決定し得る訳である。そして、たとえ実際に臨床的に使用する場合には、ある程度の余裕をみた dosi s の決め方はあるにしても、その薬に臨床的な有効量が決つたところで、他剤の同じ形で決められている有効量間で比較してゆくことこそが臨床的に意味のあることではないだろうか。

現在では動物での治療実験の dose response curve と臨床の dose response curve との正確な相関 curve が明確に分析されずに、MIC や体液濃度等の検討だけから行なわれた dosi s で実際にはその薬剤の最大臨床効果を出す dosi s よりかなり多い dosi s でいきなり二重盲検比較試験が行なわれている。

これでは医師の間に不満が出るのは当然であろう。演者の一人が発言した、現行の二重盲検比較は人がやつて data を出してくれるのは少しは参考になる data もあるから賛成であるが、自分ではやりたくないものであるという現状では困る訳で、自分で意義のあると考える比較試験を行ない、それを検討した結果、必要とあれば、open での data をより正確に証明することができるような、すなわち意味のある結論を出すような plan に従つた二重盲検法を行なうべきである。

今後、この方向で、学会全体としての取組みが必要であることが痛感されてならない。

おわりに

以上のような二重盲検法の問題点をいろいろな観点から討論する機会を得たことは、極めて有意義なことであつた。この討論が感染症治療剤、抗生剤等の有効性・有用性判定の方法の確立への道を拓く1つの斧であつてほしいものと願つている。

1. 統計学的な立場から（比較試験の原理）

佐久間 昭

東京医科歯科大・難研・臨床薬理

薬効評価の1つの形式である比較試験 controlled

trial について、これを支えている論理について述べる。

比較試験とは何かについては、いくつかの見解があるが、さしあたり“効果なり逆効果なりを評価したいと考えている目的の薬のほかに、何らかの基準薬も併行的に用い、これらを患者に割りつけるには無作為化を意識し、比較を基盤にして目的の薬を吟味する方法”と考えておく。

当然のことながら、比較が公平に行なわれなければならない。患者と医師に生じうる心理的な偏りが結論をゆがめ、比較の公平性が失なわれるおそれがあるならば、これに対する歯止めを設ける必要があり、その手段として利用されるものが二重盲検法 double blind technique にほかならない。

この意味からすれば、比較試験の基本的な理解なしに、二重盲検法を活用することは無理である。現在発生している若干の混乱は、この辺のところに原因があるのではないだろうか。二重盲検法とは、データを統計処理することであり、クロスオーバーやマッチド・ペアの利用であり、機器により観測された客観データを扱うことであり、評価基準を申し合わせた薬効評価であり、……とさまざまに語られているが、これらは、二重盲検法と深い関係にはあるものの、本質的な内容ではない。二重盲検法とは、患者と医師に生じうる心理効果による偏りを制御する手だてである。

確率論的評価

多数の系が有機的に組合わされて生体をつくつている。これに薬という刺激を与えたときに反応をみる。薬効評価においては、さしあたり、生体をブラック・ボックス的な変換器とみたらうえ、薬という刺激 x をインプットとして変換器 ϕ に与えたときの反応 y をアウトプットとしてとり出し、反応と刺激との関係を

$$y = \phi(x) + \epsilon$$

というモデルで示す。反応 y は見かけのものである。変換器の個体差、経時的な変化など変換特性にゆらぎがあること、および反応を観測する側でのゆらぎがあることのために、反応 y を (i) 平均的、確定的な部分と (ii) 偶然的、確率的な部分にわけて扱う。ここで偶然的なばらつき ϵ が、本質的な $\phi(x)$ に比べて無視できないというのが生体反応の特徴でもあり、薬効評価にあたつた問題点の1つでもある。 ϕ という変換器の構造機能を知りつくしてしまえば、あるいは ϵ が事実上は無視できるならば、確定論的 deterministic に薬効を論じることができるが、現状においては、ばらつき ϵ を考慮した確率論的 probabilistic な評価が必要である。

ある集団を対象とした場合、確率論的な評価では、残念ながら、個体ごとの議論ができずに、平均的、帰納的

な要約を行なうことになる。日常、臨床家にとって切実な問題は、特定の個体に指向していることが多いため、現在のところ、要約的な結論を個体に演繹するには、十分に注意をしながら、正しい臨床経験に養われた勘に頼るほかはない。

モデル Ω_1

話を簡単にするため、ある標的特性 y の大小が薬効を示すものとし、薬 A を用いたときに

$$\Omega_1: y_i = \alpha + \epsilon_i$$

のモデルがあてはまるとしよう。ある集団での平均的な真の薬効を α とし、 $\alpha > 0$ であることが、定性的な意味で薬効ありということにする。ある 1 例について

$$y_i = \alpha + \epsilon_1 > 0$$

となつたとき、つまり見かけの効果が $y_i > 0$ となつたとき、真の薬効がある、 $\alpha > 0$ と考えることは危険で、たまたま ϵ_1 のいたずらで $y_i > 0$ となる可能性がある。第 2 例では

$$y_i = \alpha + \epsilon_2 < 0$$

かもしれない。

この集団での全個体について薬効を調べつくしたときには、 y_i は平均値 α を中心にばらついている。このばらつきを大きさを示す 1 つの尺度が標準偏差 σ である。

全個体といわないまでも、この集団から n 個体をでために選んで平均値

$$\bar{y} = \frac{1}{n}(y_1 + \dots + y_i + \dots + y_n)$$

をつくってみると、 $y_i = \alpha + \epsilon_i$ であつたから

$$\bar{y} = \frac{1}{n}(n\alpha + \epsilon_1 + \dots + \epsilon_i + \dots + \epsilon_n) \doteq \alpha$$

となるであろう。 $\epsilon_1 + \dots + \epsilon_i + \dots + \epsilon_n$ は、 α から大小への偶然的なゆれを集めたものであるから、ほぼ 0 になるはずである。したがって \bar{y} は、ほぼ真の薬効 α を反映しているはずであるから、 $\bar{y} > 0$ ということになれば、薬効ありといつても、 y_1, y_2 などを用いるときよりも、はるかに失敗は少ないだろう。

しかしながら、別の n 個体について \bar{y} を求めてみると、多少ともくいちがつた値になるであろう。経験的にも理論的にも、 n がある程度大きくなると、こうしてくり返し \bar{y} を求めてみると、 \bar{y} は平均値 α で、ばらつきの尺度が σ/\sqrt{n} である正規分布に近くなる。そこで、1 回の観測セット n 例で求めた \bar{y} については、薬効 α を反映するシグナルと考え、その際のばらつき、標準誤差 σ/\sqrt{n} をノイズと考え、S/N 比

$$t_0 = \frac{\text{シグナル}}{\text{ノイズ}} = \frac{\bar{y}}{\frac{\sigma}{\sqrt{n}}}$$

を計算してみる。分子は、薬効 α という必然的なもので、

分母はばらつき σ/\sqrt{n} という偶然的なものである。統計学によれば、この t_0 という S/N 比が 2 をこえたときに、 $\alpha > 0$ と判断すれば、本当は薬効なしで $\alpha = 0$ であるのに誤っていることは、たかだか 100 回中 5 回程度ということになる。このことが、統計的に水準 5% で有意という内容である。S/N 比が 2 をこえているならば、定性的に薬効ありと判断し、これが誤っていることがたかだか 5% ということになる。有意 significant というのは一種の符丁で、1% で有意ということが、5% で有意ということよりも、強い効果を示しているとうけとめることは危険であり、有意というのは、定性的に薬効ありということの手ごたえを示すにすぎない。

上述したものが検定であり、具体的に α の大きさを知るには推定を行なう。ノイズを用いて幅をつくり

$$\alpha_L \leq \bar{y} \leq \alpha_U$$

とすると、 α_U と α_L は 95% の (平均値 α の) 信頼限界になり、 α_U と α_L の間に真の薬効 α が存在すると宣言すれば、100 回中 95 回は適中している。有意ではあつても、 α_U, α_L からみて、医学的に有意義ではないという場合も少なくない。この辺の意味のとり方を混乱する病気が statistical significantosis である。

ところで、 σ の具体的な値は、集団の全個体を観測しなければわからないというのが普通である。しかし、 \bar{y} が α を反映したように、多数の個体 n についての観測から

$$s^2 = \frac{1}{n-1} [(y_1 - \bar{y})^2 + \dots + (y_i - \bar{y})^2 + \dots + (y_n - \bar{y})^2]$$

を求めると、 s^2 は σ^2 を反映するものになることがわかつていて、 s^2 を σ^2 の代用値として、検定や推定を行なうことができる。

以上のモデル Ω_1 からわかつたことは、多数の個体 n を観測してはじめて、ばらつきの大きさ σ ないし σ/\sqrt{n} を推定することができ、ノイズ σ/\sqrt{n} からわかるように、 n を大きくすれば、ノイズは小さくなり、したがって S/N 比は、 \bar{y} が同じ程度の値でも大きくなり、有意になりやすくなるという 2 点である。

医学的な有意さをみるにしても、ノイズを小さくして、感度のよい検定、精度の高い推定を行なうことが前段階になるため、例数 n (標本の大きさ sample size) は大きいほうが好ましい。

モデル Ω_2

試験に参加する患者にはすべて、医師に接した心理効果、ルーチンの処置の影響……などが影響することが、容易に考えられる。これらをまとめ、平均的に γ で示すと、モデルは

$$\Omega_2: y_i = \alpha + \gamma + \epsilon_i$$

と示せる。この場合には、平均値 \bar{y} は $(\alpha+r)$ を反映することになり、たとえ検定によって有意となつても、シグナルは r で汚染されているから、 $(\alpha+r) > 0$ の結論にとどまり、 $\alpha > 0$ というには、 r について具体的な情報をもつていなければならない。

そこで、薬 A とほかに、薬 B を用いた群について、同時的な試験を行ない

$$\Omega_2' : Z_j = \beta + r + \epsilon_j$$

というモデルを考えてみる。薬 B の真の薬効が β であり、A 群、B 群ともに、患者に共通な効果 r が及んでいる。この平均値は \bar{z} で $(\beta+r)$ を反映している。差によつて 2 群を比べると、 $(\bar{y}-\bar{z})$ となるが、これは $(\alpha+r) - (\beta+r) = (\alpha-\beta)$ を反映し、個別の群においては汚染源であつた r は、平均値の差をとることで消えてしまう。

モデル Ω_2 からわかつたことは、かりに不明の影響 r が見かけの薬効に影響していても、2 群を設けて差をとることによつて除くことができるという点である。

この際 A 群で n 例、B 群で m 例を観測するとき、S/N 比は

$$t_0 = \frac{\bar{y} - \bar{z}}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{\bar{y} - \bar{z}}{\sigma \sqrt{\frac{m+n}{nm}}}$$

の形式になり、平均値の差の推定は

$$\delta_L^U = (\bar{y} - \bar{z}) \pm 2 \sigma \sqrt{\frac{m+n}{nm}}$$

の形式になる。与えられた全例数 $k = n + m$ の条件下には、 $n = m$ と両群の例数を等しくしたときに、ノイズの大きさ $\sigma \sqrt{(m+n)/nm}$ が最小になるので、検定の感度も推定の精度も最大になる。

モデル Ω_3

つぎに、薬効の大きさには広い意味の個体差があり、性別、年齢……などの影響が無視できず、これらを ϕ で示したとき

$$\Omega_3 : y_i = \alpha + r + \phi_i + \epsilon_i \\ z_j = \beta + r + \phi_j + \epsilon_j$$

のモデルが考えられるものとしよう。いま、男、女で薬効に差があるとき、試験の便宜上の問題で、薬 A を男、薬 B を女に割りつけたとすれば、これは明らかに好ましくない方法で、 \bar{y} は $(\alpha+r+\phi_0)$ 、 \bar{z} は $(\beta+r+\phi_0)$ を反映するから、差をとつた $(\bar{y}-\bar{z})$ は $(\alpha+\phi_0) - (\beta+\phi_0) = (\alpha-\beta) + (\phi_0 - \phi_0)$ を反映し、性差で汚染されたシグナルになつてしまう。

いま、たまたま性差という形で個体差を指摘できたのであるが、十人十色といつたものが見かけの薬効を修飾しており、これらをひきずり出すことも現実的には無理であるという場合が多い。本来は原因不明の誤差という

偶然的なばらつきではないが、こうした系統的な偏りを、確率的な手段によつて誤差に変身させてしまう手法が無作為割りつけ random allocation, randomization である。

全例数 $k = n + m$ を A が n 例、B が m 例になるように制限をつけたうえ、いずれの個体についても A になるか B になるかを等しい機会にして割りつけを行なう。これによつて、系統的な偏りが偶然的な誤差にこまられるために、ノイズは大きくなるにしても、比較の公平性は確率的に保たれることになる。平たくいえば、全体をよくかきまぜて、2 等分したときには、A, B 群での男女の割合はおおよそ等しくなるであろうということになり、男女といつた直接に目にふれる特性以外についても、およそ平等に分かれるであろうということが期待できる。

モデル Ω_4

無作為割りつけによつて比較の公平性を確率的に確保することはできたのであるが、ノイズが大きくなる点は好ましくない。はつきりとノイズが大きくなるほどの目立つた標識特性であれば、これをあらかじめ指摘できることも多い。見かけの薬効に影響力の大きな特性として、年齢、性別、入院と外来……などがあらかじめマークできるならば、たとえば若年、中年、老年といつた層 strata を設けて、患者を層別し、これらの層ごとに無作為化を行なう。モデルとしては

$$\Omega_4 : y_i^{(s)} = \alpha + r + \phi_i^{(s)} + \epsilon_i \\ z_j^{(s)} = \beta + r + \phi_j^{(s)} + \epsilon_j$$

となり、第 s 層ごとの平均値を比べれば、全体を一括していたとき、つまりモデル Ω_3 では誤差にくり込んでいた $\phi^{(s)}$ を分離することができ、これは r と同じ性質の、この第 s 層に共通に現れる影響であるから、シグナルを汚染することもない。

この層別の手法は、事後的な後層別や徹底した共変数の扱いにも拡張できるが、少なくとも Ω_3 の場合の単純な無作為化よりも、層別の利用では公平性の確保は積極的であり、ノイズも小さくなり、また場合によつては、特定の層では A, B 間に差が認められるということもありうる。

以上に述べたことの要旨は、実験計画における FISHER の 3 原則として知られているもので、それぞれのモデルと対応している。こうした観点からみれば、比較試験とは、薬効評価における FISHER の 3 原則の実践にほかならない。

二重盲検法

以上に述べた比較試験の成立ちをふまえて二重盲検法の意義について考えてみよう。比較の公平性、汚染のないシグナルというものがもつとも重視され、ついでノイ

ズをできるだけ小さくすることで、検定の感度、推定の精度を高める努力が行なわれる。

いま A が実薬、 B が偽薬であるとき、患者ならびに医師に生じうる心理効果が見かけの反応に影響するようであれば

$$y_i = \alpha + \pi_A + \varepsilon_i$$

$$z_j = \beta + \pi_B + \varepsilon_j$$

となり、常識的には $\beta = 0, \pi_A > \pi_B$ であろうから $(\bar{y} - \bar{z})$ は $(\alpha + \pi_A - \pi_B)$ を反映し、 A にとつて有利な方向に、シグナルは汚染されている。医師の場合、 y_i や z_j を評価するとき、実薬、偽薬であることを割引いて判断できるといふ意見もあるが、割引き方が行きすぎて、逆方向に歪めた結論を導びかないだろうか。つまり、薬 A に不利な評価を行なう可能性はないだろうか。

客観的な指標を用いるので、医師については盲検である必要はないという意見もあるが、医師が実薬か偽薬かを知っていることが患者に対する身振り態度に現われて、特殊な効果を生まないだろうか。

患者も医師も人間であり、人間らしさの故に生じる偏りの心配があれば、二重盲検法の採用がすすめられる。もう少し積極的にいえば、常に心理的な偏りの介入のおそれは存在するから、常に二重盲検法の採用が賢明であるということになる。二重盲検法を採用すれば、 π_A, π_B の区別はなくなり、 $\pi_A = \pi_B$ となるから、 $(\bar{y} - \bar{z})$ は α だけを反映するようになる。

二重盲検法の採用によつて、得られる結果の質は、明らかに向上する。薬 A, B の比較にあつて二重盲検法を採用するときには、実際上の手間は、ことさらに面倒になることはあまりないのが普通であろう。ある特定の患者に使用している薬は A か B かのどちらかであることは確実で、 C であつたり、 D であつたりはしない。この際に医師が盲検の状態におかれることで不安を生じたり、責任がもてないと感じたりすること、およびこれらが、どのような形で患者に対してマイナスに働くのかといつた点について、より踏みこんだ研究が必要ではないだろうか。盲検によつて、医師が試されているといつた感じからくる自尊心の問題なども、もつと積極的にさらけ出して吟味をしてみる必要があるだろう。

現状における二重盲検法の是是非非が、科学性とか倫理性の基盤の上で論じられているとは限らず、行政的な強制ないし指導に対する心情的な反ばつなどが中心的な話題の1つになつている。行政的に強制されようがされまいが、二重盲検法のメリットとデメリットは何か、実施にあつての障害は何かなどを具体的にとりあげて吟味をする必要がある。

二重盲検法とからんで、薬効評価の基準を確立すると

いうことも大きな問題としてとりあげられているが、これは二重盲検法とは直接にかかわりあいをもつというよりは、治療学の問題であり、病気がよくなること、薬が有用であることなどを、具体的にどのように考えることが医学的に妥当であるかを問いつめることである。この種の基準が二重盲検であるから、二重盲検でないからといつて、大幅に変化するわけではなからう。

いずれにしても、二重盲検法というものが充分に理解され批判されることなしに、いつしか広く用いられるようになると、二重盲検法で有意と証明されたので、といつた表現をされ、かつての大家のお墨付きに代る、もつとも非科学的な証明に悪用されかねない。少なくとも統計学を逆用し、有意差なしということをして、同等の効果とうけとるような向きも出てきている。ばらつき大きな、例数の少ない比較試験を行なえば、ノイズは大きくなつて、有意ではなくなるし、逆に、例数を増して有意の結果もつくり出せるのである。

統計学を無視することも危険であるが、統計学を無批判に受容することも危険である。もつとも、ここで述べた統計学とは、科学的な道理であつて、妥当な考え方の道標の1つという程度の意味のつもりである。

2. 内科の立場から

三木文雄

大阪市立大学第1内科

抗生剤の比較試験は、ここ数年かなり実施されているが、大半は尿路感染症を対象にしたものである。以下、気道感染症あるいは腸管感染症を対象とした比較試験を例にとり、比較試験実施上の問題について考察を加えたい。

1) 対象疾患の選択

抗菌性物質の治療効果は、あくまで抗菌作用に基づくべきである。したがつて、起炎菌の消長、あるいは感染に直接関係のある症状を正確に把握し得る感染症を対象に選ぶべきである。これらの症状の把握に医師および患者の主観や先入感の入り得ない項目を選べば比較試験において二重盲検とする必要性は少なくなる。いつばう患者の自覚症状や理学的所見などを主にして症状の変化を把握しなければならぬ疾患にあつては、当然二重盲検により、片寄つた判断を防止することが必要とならう。

自然治癒傾向の強い疾患を対象に選んだ場合は、たとえ二重盲検による比較を行なつたとしても、自然治癒にともなう症状の変化に隠蔽されて、薬効を正しく把み得ず、したがつて比較両薬剤の差も把握し難い可能性が多くなる。この際 placebo をおけば、この問題はある程

度解決するが、如何に自然治癒傾向の強い疾患にあつても、なかには更に重症の感染症に移行する可能性も否定し得ないため、感染症の治療に際して placebo を加えての比較を実施することは倫理的に不可能である。

一方において、起炎菌の消長、感染症状の把握がかなり確実に行ない得て、しかも自然治癒傾向がなくても、敗血症や細菌性心内膜炎のように、症例数の少ない感染症においては、比較試験の実施はまず不可能である。

実際に比較試験の実施されている疾患は、肺炎、細菌性赤痢のように客観的に症状を把握しやすい疾患から、急性咽頭炎のように主として患者の自覚症状により症状を把握しなければならぬ疾患もある。これら7報告のうち、No.6だけは薬剤投与を blind とせず、胸部レ線像の読影、重症度および効果の判定時だけ投与薬剤を blind として実施されているが、他は二重盲検法により比較が実施されている。

ここで、二重盲検法による抗菌剤の比較試験実施に際して考えられる別の問題に触れる。感染症の治療に際して、投与薬剤を無効と判断し他の薬剤に変更する場合、それまで投与していた薬剤を、次に投与する抗生剤選択の参考にすることが少なくない。比較試験に際して、効果判定を一カ所に集めて一定の基準で実施しようとする場合、主治医が無効と判断し、その症例に投与されていた薬剤を知るために drop out とすると、無効例が多数 drop out になる危険がある。

なお、比較試験においては、症例のばらつきが比較両薬剤間で均等であることがきわめて重要な条件であるが、この点、投与薬剤はすべて無作為に割付けられており、問題はない。

2) 対照薬剤の選択, 比較両薬剤投与量

前述のとおり、感染症の治療に際して placebo の投与は不可能である。また抗菌力に基づく治療効果を検討するのであるから、何らかの抗菌作用をもつた control との比較ということになる。この場合、対象とした感染症の主な起炎菌が抗菌スペクトラムから外れた薬剤、検討しようとする薬剤に比して対象疾患の起炎菌に対する抗菌力が極端に弱い薬剤、対象疾患の主な起炎菌の耐性頻度の高い薬剤、あるいは、対象疾患の病巣への移行性の悪い薬剤をそれぞれ control に選んだ場合、いずれも新しく評価しようとする薬剤は control に比して良い成績を得るであろうことは、抗菌性物質の場合、基礎的の実験成績から理論的に充分予測が可能である。また、経口投与の場合、消化管からの吸収あるいは副作用の出現に、食事摂取の影響の少ない薬剤と、影響の大きい薬剤の比較の場合、薬剤投与を空腹時にするか食後にするかにより、両薬剤の治療効果あるいは副作用に差の出る可

能性もとうぜん予測されよう。したがって、比較試験のデザインの組み方如何により、人為的に比較両剤の薬効あるいは副作用に差があるようにみせかけることも、とうぜん可能となる。そこで、実際の抗菌性物質の比較試験においては、抗菌力、抗菌スペクトラム、耐性頻度病巣への移行性などの出来るだけ近似した薬剤を control とせざるを得ないことになる。この場合、臨床効果の差はきわめて少ないであろうことは推測に難くない。したがって両剤の差を検出するためには、きわめて多数の症例について比較を行なうと同時に、微妙な差を検出する効果判定法が必要となる。

対照薬剤選択上問題になることに、もう1つ次の場合が考えられる。すなわち、ある特定の菌、たとえばA菌に抗菌力を示すことが、その薬剤の最も大きな特徴とされる薬剤の検討に際して、A菌感染症を対象に選ばず、B菌感染症を対象にして、A菌には抗菌力を示さず、B菌に抗菌力を示す薬剤を control にして比較試験が実施される場合がある。この比較試験において、たとえ両薬剤の効果あるいは副作用に有意差なしの成績が得られたとしても、本来A菌感染症に対しての効果を検討すべき薬剤にとつて、その成績が如何なる価値を有するかははなはだ疑問といわざるを得ない。その薬剤の比較試験はあくまでA菌感染症を対象に実施すべきであり、A菌に対して効果を示す既存の薬剤のない場合は、A菌感染症に対する確実な open trial の成績のほうが、その薬剤の評価にきわめて重要であろうと考えられる。

投与量とも関連があるが、比較試験の実施は、あくまで open trial により、その薬剤の特性がよく検討され、その薬剤のもつとも適切な投与対象疾患、投与方法、投与量を決定した上で、その対象疾患に対して既存の、いちおう正しいと考えられる治療法と比較試験を行ない、両治療法の merit および demerit を比較することが原則となろう。

実際に行なわれた比較薬剤とその投与量をみると、No.1, 2, 5, 7 の4報告は近似した薬剤の等量の比較であり、No.4 は ST 合剤と SMX 単独の薬効差を検討したもの、No.3 は KM に比して抗菌力の強い Aminodeoxy-KM を投与量を減らして KM の効果と比較したもの、No.6 は ABPC に比して吸収性のよい Amoxycillin の効果が ABPC より勝るか、あるいは ABPC の半量で同一効果を得られるかを検討した成績である。

すなわち、目的に応じて、如何なる投与量で比較するかについても、充分考慮を払う必要がある。

3) 重症度判定法, 効果判定法

比較試験においても重症度判定、効果判定の方法が重要であることは、もちろんである。実際に行なわれた効果

判定法は、抗菌剤の効果の最も大切な面の1つである除菌効果のみられているのは細菌性赤痢だけであり、他の呼吸器感染症では、疾患の性格上、細菌学的効果の検討は実際には不可能に終わっている。この際ほとんど全て、主治医により一定の基準を置かず効果判定がされたり、自覚症状の改善日数による効果判定が行なわれている。

起炎菌の消長、レ線所見、臨床検査成績など主観により左右されないデータでの症状の把握の困難な上気道感染症にあつては、これらの自覚症状の改善による総合的な医師のベッドサイドでの効果判定は抗菌性剤の効果判定に際してきわめて重要と考えられる。これらの疾患にあつて、頭痛、咽頭痛、咳などの個々の症状については、化学療法よりむしろ対症療法のほうが速やかに症状の改善をみる場合が多いと考えられ、それら個々の症状の改善だけに把われた場合、感染症の治療剤として、抗菌性物質よりむしろ解熱、鎮痛、鎮咳剤などのほうがすぐれた薬剤であると評価される危険もあり、抗菌性物質としての総合的なベッドサイドでの治療効果の判定が必要になる。しかし、この場合その判断には主観、先入観の入る可能性が多分に存在するので、この意味からも、これらの疾患を対象にして比較試験を実施する時は二重盲検が絶対必要ということになる。

4) 副作用のチェック

臨床検査の上で異常値が現われる副作用は、検査実施者に投与薬剤を blind にすることにより検査成績に先入観念の入り得る余地はなく、二重盲検比較の必要性はない。しかし、自覚症状として出現する副作用の把握に関しては心理的要因や先入観念を完全に除去することは不可能であり、嚴重な副作用の比較は二重盲検によらねばならない。しかし、この場合も副作用チェックの方法に一定の基準を設置すれば、かなり客観性をもたせることが可能となる。

副作用チェックの上でもう1つ重要なことは、生体に致命的な影響を与える副作用であつても、それがきわめて稀にしか出現しない場合には比較試験において有意差なしと片付けられてしまう危険もあり、比較試験を実施しても副作用チェックに万全であるとはいえないことを忘れてはならない。

5) 症例数

前述のとおり、比較試験における対照薬剤には被検薬剤ときわめて類似の薬剤が選択されることとなり、その結果、微妙な差を検出するためには、かなり多数例について比較を実施せねばならないことになるが、実際に行なわれている症例数は、1群 20~30 例の比較により、何%の信頼限界で有意差なしと評価されているのが大多数であり、この成績が実際の临床上に果してどれだけの

有用性があるかはなはだ疑問といわざるを得ない。

抗菌性薬剤の場合、むしろ厳密な効果判定基準を設定して、きめの細かい効果、副作用を検討した成績のほうが、その薬剤の評価にかなり役立つものと考えざるを得ない。しかし、そのためには客観性のある重症度判定方法、薬剤の効果判定基準の設定が先決問題となることはもちろんである。

6) 肺炎における重症度判定基準、効果判定基準の設定について

抗菌性薬剤の open trial についてはもちろんのこと、比較試験においても客観性ある重症度判定法、薬剤判定法が重要であることは前述のとおりである。以下、肺炎について、重症度判定基準、効果判定基準の設定についての検討成績を示す。

対象症例は、昨年の本学会で報告した ABPC と Amoxycillin の比較試験に際して寄せられた肺炎症例のうち、drop out を除いた 70 例である。これらの症例は比較試験実施当初から肺炎の重症度判定、効果判定基準設定のために、症状、検査成績のチェックをできるだけ詳細に、かつ、一定の基準を設けて実施されたものである。また全症例のレ線フィルムは、1カ所に集め経験豊かな小委員会のメンバーにより一定の基準の下に読影、評点を付し、他の臨床症状と併せて、重症、中等症、軽症の3群に分類し、効果判定は将来科学的根拠に立脚した一定の基準を設けることを目的として、特に基準は設けず、小委員会委員の意見の一致をもつて、著効、有効、やや有効、無効、悪化の判定を下された。

以下、小委員会の胸部レ線像の採点、重症度判定、治療効果判定は、絶対に正しいものであるとの前提のもとに、胸部レ線所見、体温、白血球数、赤沈、CRP、血液ガス、チアノーゼ、胸部ラ音、脱水症状、咳嗽、喀痰量、喀痰性状、呼吸困難、胸痛の各数値を基準化して重相関比法により分析し、重症度判定、効果判定が数値でもつてどの程度表わし得るかについて検討した。

この 14 項目のうち、チェックの不充分な血液ガス、CRP、喀痰量、喀痰性状を除いた 10 項目について、重症度判定に関連して各因子の構造ベクトルならびに重みづけを行ない、重症度得点を算出し、小委員会の重症度判定と比較した。その結果、医師判定と得点による重症度判定の相違したのは 70 例中 6 例だけで、肺炎の重症度判定は点数化により、かなり正確に行ない得る。

次に重症度判定に際してのチェック項目を少なくする目的で、10 項目の因子のうち、高相関項目は、いずれか一方で代表させるために各因子の相関を検討したが、いずれの因子間にも高い相関はみられず、止むなく構造ベクトルの高いものから 5 つの因子、すなわち脚部レ線

所見、赤沈、チアノーゼ、ラ音、体温について、上記同様構造ベクトルならびに因子の重みづけから重症度得点の算出を試みたが、この5因子だけによつてもかなり正しく重症度判定が可能である。

次に数量的な効果判定の基準化について検討した。感染症の治療に伴う症状の軽快に際して、どの因子がどれだけの速さで改善していくか、それに応じて、どの因子について治療開始後のどの時期の動きをチェックすれば、抗菌剤の効果を最も正しく把み得るかは、きわめて重要な問題である。

模型図(略)での D, E, d, e を無効あるいは悪化と判定するのに異論を唱える人はまず無いと考えられる。問題は X 点での状態を治療前の状態と比較すると、治療の判定は可能であるが、薬効判定はまず不可能である。また治療前の症状の軽重により、同じ速さで症状の改善をみた A, a の薬効をどう比較評価するかも問題になる。

実際の肺炎症例について、前述の各因子の治療前値と治療後のどの時期の数値の変動がもつとも効果判定に関連が深いかを検討しようとしたが、治療開始後の各時期とも、個々の項目のチェックの脱落症例が多く、効果判定に重要な因子として、どの項目のどの時期の成績を最も重要視すべきかの解析結果を得ることは不可能に終わった。止むなく、前述の重症度得点の算出を治療前、治療開始3日後、7日後、10~14日後の4時期について行ない、その得点差をもつて、治療効果の算定を実施したところ、小委員会による効果判定が、著効、有効、やや有効、無効になるに従がい、得点差は小さくなるが、明確に分けることは不可能になり、いつばう治療前との比較の日を治療開始7日後、さらに10~14日後と、後にするに従がつて、治療前後の点数差は大きくなるが、小委員会による効果判定結果の差による点数差の差異も小さくなり、薬効判定よりむしろ治癒判定を示すこととなり、これらの方法により、治療効果を正確に表示することは不可能に終わった。

しかしながら、治療ともなる症状軽快の速さは、各因子によりかなりの差を示すことは明らかであるので、今後さらに同一基準により観察、効果判定を行なつた症例を加えて、それぞれの因子について、最も差の大きく現われる時点の数値を基にすれば、効果判定をかなり正確に、かつ数量的に行ない得るものと信じる。

7) まとめ

1つの感染症に対して、臨床的に有効な薬剤が幾種類も存在することは、同一疾患であつても、その治療に当つて、患者の条件、疾患の性状などにより、薬剤選択上の自由のあることで、臨床的にきわめて有意義である一方、同一疾患に対して有効性を示す複数の薬剤につい

て、効果、副作用の両面での merit および demerit を客観的に評価した成績をもつことも治療上重要である。

化学療法剤の評価は、あくまで抗菌作用によりもたらされる治療効果について判断するべきである。自然治癒傾向の少ない疾患を対象に選び、起炎菌の消長や感染による炎症に直接関係があり、しかも客観的に把み得る症状所見を基にして判断すれば、効果判定に際して心理的要因の入り得る余地はほとんどないであろう。従がつて、これらの客観的に把み得る症状、所見のデータを基にした重症度判定基準および効果判定基準を設定すれば、薬剤の絶対的評価を正しく行なうことが可能であり、それによつて、他の薬剤の効果と比較することも可能になる。もちろんこの場合の比較は、同時比較ではなく、historical control になるが、同一の感染症にあつても、その病型や起炎菌の変遷、耐性菌の増加などの問題により、同時比較の成績といえども、あくまで比較を実施した時期における薬効の比較を表わすものであり、理論的には常に同一薬剤について比較試験を実施し続けねば、その成績を実際の臨床に役立たせることは不可能となり、結果的には絶対評価を反覆実施することと同一になり、同時比較でなければ、その成績を信頼できないとは言えないことになろう。

二重盲検法は、医師、患者双方の先入観念や心理的要因を除外し、公平な比較をし得る点で、すぐれた比較の方法であることは論をまたない。一般に二重盲検の成績は、推計学的手法によりデータの解析が行なわれ、有意差云々という表現で結果が示される関係上、その成績は open trial の成績より価値の大きいものとしてうけとられがちであるが、あくまでも一定の条件における比較の成績であることを念頭に置く必要がある。薬効の絶対評価の困難な薬剤の場合にはとも角として、客観性のある薬効評価の可能な抗菌性薬剤の場合には、二重盲検法による比較試験の必要性より、むしろ抗菌作用による薬効の評価を正しく把む方法、さらには微妙な薬効の差を把む方法を確立して、それによる絶対評価をできるだけ多数の症例について実施することのほうがより重要ではないかと考える。

文 献

- 1) 中川圭一、他：日伝会誌 46：210, 1970
- 2) 勝 正孝、他：日伝会誌 46：332, 1970
- 3) 杉山茂彦、他：日伝会誌 47：77, 1971
- 4) 三木文雄、他：第20回日本化学療法学会総会、1972
- 5) 水谷 明、他：Chemotherapy 21：1221, 1973
- 6) 塩田憲三、他：Chemotherapy 21：1535, 1973
- 7) 勝 正孝：他：日伝会誌 (投稿予定)

3. 小児科の立場から

紺野昌俊

帝京大小児科

臨床における抗生物質の有効性

臨床における抗生物質の有効性とは、他の生体に直接作用させて薬効を期待する薬剤とは異なり、本来は、1) 試験管内抗菌力が良好であり、2) 生体内、ことに病巣内への移行が充分であり、3) 起炎菌が消失し、4) 臨床反応が良好である、という4原則が認められれば、それで充分であると、私は考える。

しかしながら、比較試験が必要となってくる理由は、ないわけではない。以下にそれを列挙すると、

1. 他剤より優れていると主張する場合

その内容としては

- a 抗菌力が強い
- b 吸収がよい、或いは排泄がよい
- c 副作用が少ない
- d 生体内代謝或いは蛋白結合率が異なる

というようなことが、臨床使用上、どのような利点や欠点となつて現われてくるのかという、いわば明らかに他の薬剤に目標を決めて、その薬剤との優劣を論ずるような場合。次には

2. 剤型・投与方法による臨床反応の相違を知るためのもの

これは、同一薬剤の中でも、生体内移行の良否が、臨床の上でどのような利点、欠点をもたらすかというもので、これは、同一薬剤内での比較ということになる。

3. その他として

- a 上記の4原則から少し外れて、有効性を実証する根拠に乏しいような場合
- b 起炎菌の決定が不可能であり、臨床反応および臨床検査成績の上からも、抗生物質の有効性との相関がなかなか得られないような場合

等があり、この際は、既存の有効と考えられる抗生物質あるいは、対症療法または疾患の自然経過との比較が必要となってくる。

現在行なわれている比較試験とその問題点

以上のような原点に立つて、現在行なわれている比較試験法を眺めてみると、ほとんどが二重盲検法で行なわれているが、その理由をしらべると、

1) 他剤より優れているとか、2) 他剤とほぼ同等の効力を有するか? とか、3) 有効性の実証、とかいうような目的を持つており、それ自体は、医学上の問題として、重要なことであるとしても、そのような医学上の必要性から、医師が自主的に行なっているものは、極めて

少なく、製薬側の謳う効能効果の実証、或いは行政側の指導という形で行なわれているのが実情である。

このような目的で行なわれている二重盲検法は、必然的に他剤より優れている、あるいは他剤と同等か? という目的を負わされる結果となるが、このような現状における二重盲検法には、種々の問題点を内蔵している。

第1には“有意差なし”ということから、直ちに同等の効力を有すると判定しているものが、かなり散見されるが、推計学上有意差がないということは、あくまでも有意差がなかったということであつて、同等ということではなく、“ほぼ同等であろう”と言うためには、極めて多くの症例を必要とするし、それに該当する臨床比較試験はほとんど不可能に近いことである。

第2に効果判定基準が、二重盲検法に限らなくても、比較試験を行なうのには、合致しないものが、かなり見られる。比較試験を行なうためには、どのような判定基準が必要かということについては、すでに何回も述べているので略するが、その一部は後述する予定である。

第3に、現在、抗生物質に用いられている常用投与量は、たとえば、有効率80%を示すために必要な最低量ではなく、もつと余裕のある量である。

第4に、このように効果判定基準が曖昧で、投与量も十分な効果を有する量で比較試験を行なつて、果してここから何が得られるのだろうかということがある。割り付けの平等さを主張する人もいるが、単に割り付けが平等であるということが、その薬剤の有意性を立証することにはならない。

第5に、現在までに行なわれている二重盲検法の中には、ただ二重盲検法を行なつたというだけで、retrospectiveな解析、或いは今後の二重盲検法に対する反省としての問題点の指摘が不充分であるものがあり、bed sideのDoctorの意見が反映していないものもかなりある。

第6に倫理上の問題があげられる。本来、医療は医師と患者との信頼関係の上に成り立っているものであり、薬は患者の苦痛を和げるものでありながら、現在の医療不信感の中で、患者から明確な承諾を得て二重盲検法を行なうことは、本当に可能なのであろうか。多かれ少なかれ、投薬の内容には触れずに実験しているのが現況である。これらの実験は、本来行政管理のもとになされるべきものであるにもかかわらず、今は医師だけの責任に負わされているのである。

抗生物質の比較試験に必要な資料

本来、感染症と抗生物質との関係は、抗生物質を致命的な肺炎等に使用して、救命し得たことから出発した。それは全く画期的なことであつたとしても、この臨床的な医師の感覚は、その後の多くの医師に、治れば有効だ

という錯誤力？ を招いたことも確かである。

ついで、抗生物質は自然治癒傾向があるような感染症にも使用されるようになり、これらの疾患に対する抗生物質使用の意義は、疾患の改善度の短縮、あるいは遷延、悪化の防止にあつたにもかかわらず、依然として、治れば有効であるといった考え方に支配されて来たことは否めないことである。

いつぼう、抗生物質の効果判定は、起炎菌の消失を原則とすべきだという批判もあり、そのような姿に変換しつつあることも事実であるが、そこにも複雑な問題が生じてくる。

たとえば、起炎菌が消失するまでに至る時間の短縮が、抗生物質の優劣を決めることであつたとしても、bed side の Doctor が感ずる臨床的な手応え、あるいは患者の感ずる苦痛からの解除というものは、起炎菌の消失に至る時間の短縮よりも、それと直接関係を有しているかどうかは必ずしも定かではない。いわゆる随伴症状の時間的な改善度の良否が、医師の側にも患者の側にも、この薬は効いたという優劣性を感じさせるのである。いつたい、このような優劣性と抗生物質の作用とは、どこでどのように関係しているものであろうか。

二重盲検法に限らず、比較試験の際には、第1に、感染症の自然経過にみられる臨床症状ならびに臨床検査成績の経時的、経日的な集積が先ずなされなければならない。

第2には、既知抗生物質を使用した際の臨床症状ならびに臨床検査成績の経時的、経日的な資料の集積が必要である。

そして、これらの集積の中から

- a 起炎菌の消長だけで効果判定可能なもの
- b 理化学的検査所見だけで効果判定可能なもの
- c 臨床症状を加える必要のあるもの
- d 臨床症状だけで効果判定をして差し支えないもの
- e 患者の訴えで判定可能なもの
- f これらの組合せが必要なもの

等が、疾患によつて分類される筈であり、しかも、それらのチェックポイントはいつがいちばん重要であるか、そして優劣を論ずるためには、どの時点で行なえばよいかということが出てくる筈である。そして、最後にいちばん重要なことは

- g 組み合わせで判定を行なわなければならない時には、それぞれの判定項目に対する重要度が考慮されなければならない。

小児の急性気道感染症における比較試験

以下に、私が関与した比較試験における問題点を、主

として効果判定法に関連して、2~3 の問題点を述べる。私達が行なつた Clindamycin-palmitate と Chloramphenicol-palmitate の比較試験の成績をみよう。これは各年令における臨床症状の相違を考慮して、ほぼ同一年令あるいは月令であること、発熱して 24 時間以内であること、白血球増多があり細菌の 2 次感染が関与していると考えられる症例であること、発病が季節的に同時期であること、という 4 つの関門をもうけ、これらの 4 つに適合する症例を対として選ぶことの困難さから、先ず Clindamycin-palmitate を投与する症例を選んだ時には、それと対になるに適合する症例を、精力的に目的意識的に症例を探して、それには Chloramphenicol-palmitate を使用するというやり方で行なつたもので、有効率は Clindamycin-palmitate が 76.5%、Chloramphenicol-palmitate が 64.7% で、もとより両者間に有意の差はない。しかし、この結果は bed side の Doctor には極めて不満なことであつた。Bed side の Doctor は Clindamycin-palmitate のほうが臨床効果がよかつたと主張するのである。

そこで臨床評価項目の相関性を分析してみると、投薬開始後 72 時間目の下熱が、臨床効果判定の上で、57.68% の比重を占めており、中検からの検査結果が不明な時点では、bed side の Doctor としても、患児の母親としても、臨床の手応えとしては、それが最も確実な手掛りであるに違いない。臨床検査成績との相関をみても、下熱ということが、臨床反応の上で、密接な関係を有していることは確かめられたが、bed side の Doctor が両側間に差があると感じた理由は、実は 72 時間目の下熱ではなく、72 時間までに至る経過の相違にあることに気付いた。

両剤を使用した群では、病日、発熱ともに全く同じような状態の小児を選んだにもかかわらず、投薬開始後の経日的な最高体温の経過を関数で求めると、両剤間のパターンには何となく差が認められ、bed side の Doctor だけでなく、患児の母親としても、この下熱傾向の相違から、印象としては、Clindamycin-palmitate のほうがよく効くと感じるとに違いない。そして、下熱傾向の優劣が、果して、この薬剤間の臨床における抗菌力の優劣を直ちに物語つていないものではないかも知れないが、この両剤の比較試験を行なつた Doctor 達は、以後の小児の急性気道感染症には、Clindamycin-palmitate を先ず使用するに違いないし、母親もそれを支持するであろう。医師と患者との信頼関係というものは、そのようなものであり、これが、もし二重盲検法で、bed side の Doctor が、その薬剤の内容を知らなかつた時には、このような実感が湧いてこないことも確かである。二重盲

検法で bed side の Doctor にこのような apprehensibility を求めることも決して不可能ではないが、その時には薬剤の素姓も判らないまま、どちらを貴方なら使用するかという問いかけに答えた結果が、統計上判明するだけで、実感とは程遠いものであろう。

肺炎における比較試験

次に私が第2コントローラとして参加した MDI-PC と AB-PC の合剤と AB-PC の二重盲検法による肺炎の臨床比較試験についての retrospective な解析によって得られた問題点について述べる。この試験は、北大の村尾教授および関連病院の先生方によってなされた仕事であるが、効果判定基準は、東大医科研・深谷博士が考案された肺化膿症に対する判定法を modify して用いている。すなわち、自覚症、他覚症、X-P 等をほぼ同格として、点数加算し、改善率で臨床評価を表わし、総合判定した。

この点数による薬剤投与開始後、1週間目の分布をみると、ほぼ正規分布を示し、この効果判定法は、一見極めて理想的に見えるが、たとえば、連続した値を示す正規分布の中から、0.3 以下が著明改善で、0.31 は改善だという区別は、何を根拠に分けられるのであろうか。本来、これらの点数による分け方が可能となる条件は、著明改善、改善というような集団は連続した分布の上にあるのでなく、別個な集団として明確に分けられるようなものでなければならない。このような点数の集団にするためには、効果判定の中で、何がいちばん重要であるかということをしらべ、それには、重要度といった定数を乗ずることにより可能な筈で、これは、この後の解析の上に種々の問題を投げかけて行くのである。たとえば、咳嗽が2点から1点になつた場合と、X-P 陰影が2点から1点になつた場合、数値としては同格であるが、臨床上の重要性は同義であるとは解釈しにくく、その重みづけに妥当性を欠くように思えるのである。

このようなことから、この資料の解析には試験前の観察項目および治療状況の影響を考慮する多変量解析法を試みる必要が生じてきた。その手順は省略するが、結論づけをすると、自覚症状の咳嗽、喀痰および胸痛というようなものは、投薬開始後1週後には改善していることを示し、他覚所見の中でも、体温、ラ音、白血球等も1週以内に改善し始める。つまり、これらの項目は、1週以内の判定には、極めて重要な役割りを演じているのであり、もしこのような症状により薬剤間の優劣を論ずるためには、もう少し投与開始直後から数日間の短期観察が必要であつたかも知れず、bed side の Doctor や患者が、この薬がきいたという手応えは、もつと短い期間の中に存在しているのかも知れないのである。

Bed side Doctor が独自に行なつた効果判定をみると、初期要因のうち、入院例と自覚症状の高いものほど、5段階評価法では改善度が低いことを示しており、男性は女性よりも改善され易く、発病からこの薬剤による治療に至る経過日数の長いものほど改善度がよく、これらは当初の層別割り付けの際に、もつと明確にしなければならぬ問題を残している。

薬剤の欄をみると、1週目の効果判定では、主治医による判定は合剤のほうがよいという判定が下されているが、点数による判定では合剤のほうがよいという判定は出て来なかつた。これは、点数による5段階判定をくずし、著明改善だけを有効、改善から悪化までを無効とした判定法を使用して統計処理した時にだけ、主治医の判定と合致する成績を得た。

つまり、主治医は何に重要性をおいて判定しているのか、点数による判定法は、それを何処に反映させているのかという gap を物語っている。2週以後についても述べたいことがあるが省略する。

さて、以上述べたような、主治医の判定と点数による判定の素質は、主治医は主として何を基準として効果を判定しているかということ、単相関係数から検討しなければならない。

主治医の判定は、1週では X-P 所見、血沈、白血球という或る程度客観性を持つた他覚所見と密接な関連を有しており、主治医が重要視した他覚所見のほかに、咳嗽、喀痰との相関も高く、ラ音とも相関を有する。いっぽう、咳嗽と喀痰の改善は正の相関を示し、このことは喀痰の改善に伴つて咳嗽も改善されるという同じ現象をダブって裏表からみている可能性を生じ、本来は1つの判定項目にまとめるか、咳嗽は効果判定項目から削除するのが妥当であるのかも知れない。

このようなことから、自覚症状と他覚症状をそれぞれまとめて、効果判定との関連をしらべると、主治医の効果判定のほうが他覚症状に重点をおいており、点数による判定は他覚症状とも自覚症状とも関連を有するという成績であり、一般的に主治医による効果判定は客観性に欠けると言われていることとは反対に、主治医のほうが、より客観的なデータに基づいて判定をしているという極めて皮肉な結果も出てきた。

この解析は、愛知県がんセンター、大谷元彦博士の御指導によるものであり、私は第2コントローラとして、勝手な注文をつけたのであり、多変量解析を行なつたのは抗生物質関係の仕事では、おそらく始めてと思われ、膨大な労作に取り組み、主治医としての客観性のある姿勢をくずさなかつた、北大・村尾教授始め、一門の方々に改めて深い敬意を表するものである。

いずれにしても、二重盲検法を実施するにあたっては、これらの retrospective な解析を通じて、問題点を指摘して是正し、よりやり易い方向に、より正確な方法にと、一步一步改良していくことが、目下の急務であると考える。

4. 泌尿器科の立場から

名出 頼 男

名古屋保健衛生大泌尿器科

二重盲検法の本質論については、他の論者によつて詳しく論じられたとおり、薬効比較試験の1方法であり、方法論自体とくに欠陥を云々されるべきものではない。比較試験において、客観的パラメータが数多く（定量的なものであればなお充分であるが）、これらが常に主観的パラメータと対応しているという証拠が揃えられれば、少なくとも一時点においては、二重盲検法を採用しないで比較評価が可能な筈である。二重盲検法が、新薬の評価に当つて、学会内における十分な論議を経ないで登場して来た背景には、従来の各種薬剤（抗菌物質と限定しない）の薬効評価法が、必ずしも万人を納得させ得ないばかりか、検者自身でさえ、retrospective に見て疑問を懐くことであり、数々の疑点が近年になつて露わにされ、その疑問——時として極めて primitive なもの——に医学者ないし医師が答えを持ち得ない状況さえ起り得る事態になつて来た故であろう。むろん、二重盲検法の広範な採用に、為政者側の業界界に対する要請のあつたことが大きく作用してはいるものの、このようなものが要求として出てくること自体、評価に当つてきた研究者としても反省の必要な点があると考えられる。

抗菌製剤の場合を考えると、薬効試験判定のための客観的パラメータは、他種薬剤に比し多い筈である。感染が、常識的には寄生体の侵入に始まり、それが symbiosis として定着するものでない限り、宿主側にこれを reject するための機転が働き、その多くは一連の炎症反応として見られる。われわれが客観的パラメータとしてとらえ得るものは、従がつて1つは寄生体の消長であり、1つはこれに伴う炎症反応の強さの変化である。しかしながら宿主つまり患者が、最も影響を受けるのはこの炎症反応のうち、痛み、発熱の2つである。いつぼろ、この炎症による自覚症状の極めて弱い——というより炎症反応自体が、古典的4大症候を揃つて示さないというべきか——慢性症が、宿主自体の Life span に影響を及ぼすほどの強さで持続することの多いのもよく知られた事実であり、事実、感染症治療に当つて医師を最も悩ませるのはこの種の感染症である。

対象を尿路感染症にしばつて考えると、前者、つまり典型的感染症としては急性（単純性）膀胱炎、および腎盂腎炎が（急性尿道炎、急性前立腺炎、急性副睾丸炎等、生殖器にまたがる感染もこれに入るが）ある。他方の慢性症は、種々の基礎疾患（そのあるものは医原性疾患であるが）の上に成立するもので、基礎疾患自体の経過が、感染症の経過を左右する種のものである。

以上、いわば周知の事実に立つて考え、またわれわれが、抗菌剤評価に当つて採つてきた方法を省みて見ると、多くの場合起炎菌の数と、尿中白血球数という炎症反応を最も端的に反映する客観的パラメータの動きを主に追ひ、さらに、発熱ある場合はこれも重要なパラメータとして見ている。症状等の主観的パラメータは、多く低い評価しか与えられていない。これは研究者側から見て、症状の変化がかならずしも、感染の経過と併せ見て、信頼すべきものと考えられないという経験からくる結論めいたものと、抗菌剤は本来細菌の生体内発育を抑えるもので、細菌の消長だけを見ればよいとする論法とが偶々相似た評価方法論を導くに至つたためと考えられる。

しかし、さらに省みると、この主観的パラメータの変化が詳細に追求され、客観的パラメータの動きとどのような相関にあるかを調べられたことが少なかつたのも事実である。ここに、二重盲検を行なつて見る余地が僅かに残されていることを否定はできない。

以下に実例について検討を加えて見ることとする。

第1の例は、神戸大石神教授が主催された piromidic acid と nalidixic acid との女子急性膀胱炎を対象とした二重盲検試験である。患者の年齢構成、初診時の症状、尿所見等、両群それぞれ約 60 例、併せて約 120 例について有意差のある分布を示さない。その試験結果をみると、piromidic acid 53 例中著効 34 例 (64%)、有効 9 例 (17%)、無効 10 例 (19%)、nalidixic acid 61 例中、著効 35 例 (57%)、有効 10 例 (16%)、無効 16 例 (26%) で、推計学的に有意差なしの結果であつた。この時、尿所見（細菌数、白血球）や症状の変化を通覧し、それぞれが矛盾した動きを示したものと判定基準を変えることにより、有効率が変つてくることになるが、客観的指標である、尿中細菌の減少または消失と、尿中白血球の減少に矛盾を見たものは、piromidic acid で 2 例に、白血球減少だけで、細菌減少度不満のものがあり、nalidixic acid では、上と同様のものが 1 例、減の傾向を示したものが 1 例と両群ともに、2 つの主要な客観的パラメータの動きが併行したものが大部分であつた。また、主観的パラメータ、つまり排尿痛とか頻尿といったものの好転率も、両群ともに、客観的パ

ラメータの動きを裏切るほどの矛盾を示さなかつた。さらに附言すると、現在進行なわれた、急性膀胱炎を対象とするいくつかの類似した試みは、耐性菌の率の多い薬剤を用いたものを除いて、いずれも同様の結果を示している。

この結果から、直ちにこの種の二重盲検試験の価値を云々することはできないが、逆にこのような試験結果から両者が全く同等の力を持つているかどうかということになると疑問を示す人が少なくない。周知のとおり、この両薬剤の主な差異は、グラム陽性桿菌に対し、piromidic acid が、nalidixic acid より遙かに強い試験管内抗菌力を示すことである。しかし、尿路感染症においては、グラム陽性球菌（ほとんどが *Staphylococcus epidermidis* または *aureus*）によるものは数少なく、両群ともそれぞれ 10 例に満たぬ症例数しかなく、これに起因する膀胱炎だけについては piromidic acid のほうが、高い有効率を示しながら、推計学的には有意差なしとなり、また、むろん全体に対しても有効率に差の出るほどの影響を与えていない。また、このような active placebo を用いるやり方では直ちに表面には出ないが、急性膀胱炎では自然治癒率が時によつて異なるがかなり高率で、1 週以内に 10~20% くらいにも達し、従がつて僅かな薬剤の力が、極めて急速な治癒をもたらす可能性が高いと考えられることである。この観点からすると、急性膀胱炎は薬剤試験の対象としては問題ありということになり兼ねないが、尿路感染症の中で、比較的内容の揃つた人数の多く得られるものとして極めて用い易いことと、自然治癒傾向のあることを逆用し、この疾患において治癒率の低いものは臨床的に無価値と考えられることから、試験対象として頻用されて来ている。従がつてこの疾患を二重盲検試験対象として価値あるものにするためには、試験方法論に工夫を加えて、果して、薬剤の力の微妙な差をこの疾患を通して出し得るかどうかの検討が必要となる。

ここでこの方法論改良の問題を論ずる前に、このような自然治癒傾向を示すものとは別な疾患を対象として行なわれたもの（このような試験の数は後述する理由により極めて少なからざるを得ない）の代表として、市川篤二名誉教授を moderator として行なわれた DKB-gentamicin の二重盲検試験結果を示す。この両薬剤ともに抗菌 spectrum が広く、耐性菌が極めて少ないことから、慢性症を主に、それぞれ 70~80 例に用いた試験である。結果は、gentamicin は 79 例に用い、3 例の不明例を除く 76 例中、著効有効併せて 39 例、やや有効 18 例、悪化および不変（無効）例 19 例であり、DKB は 69 例中（不明なし）著効有効合計 30 例、やや有効

19 例、無効 20 例であつた。症例内容は両者とも有意差のない分布であり、また判定基準は、慢性症に無症状のものが少なくないことや、症状内容が不揃いな故もあつて、細菌および白血球の消長が主なるパラメータとなつたようである。有効率に有意差のないのは抗菌 spectrum、体内分布動態等から考え、当然とも言える結果と常識的には受け取れる。しかしこの両者とても、類似薬剤ではあるが、抗菌力、血中濃度等に、いくぶんの差があり、これが臨床的に有効性の差を生む可能性はこのような一律の試験結果からは、必ずしも否定できない。

このような代表的な従来の試験結果から引出されることは、ともかく、試験管内抗菌力、体内分布動態等の差が、どこまで臨床的に薬剤の力として反映するであろうかという問題である。従がつてこれを解くための方法論の案出が先ず行なわれないう限り、比較試験そのものも不完全であれば、また二重盲検を同じ方法論を用いて行なうことも無意味である。

しかし、この方法論の案出は必ずしも容易でない。動物試験であれば ED₅₀ を求めるという確立されたものがあるが（ただし、感染モデルが必ずしも、各種疾患に対応したものとして確立されていない難点があり、多くのばあい敗血症型の全身感染モデルによる薬効比較に終り、尿路感染症等の局所疾患に対する有効性とこれらに対するものとの良い対応がとれるという確証も得られていない）、臨床的にこれを求めることは、患者の中から協力的な人を volunteer として多く得られない限り極めて困難な作業であり、とくに重症感染症においては全く許さるべきものではない。このジレンマを如何に乗越えて行くかが、これからの感染治療学にとって不可欠の問題である。故に、ここで、薬剤投与による治癒経過が、どのような形をとるか仮想的な time course を考え、これに基づく臨床例の観察法を作り上げる試みを行なつて、これから患者負担を少なくする方法が見出せないであろうかと考えてみた。先ず考えられる治癒経過をみると、ED₁₀₀ があると仮定した疾患で治癒率と薬剤量の対応（投与法は問わない。いずれかの一定したものとすればよい）は、sigmoid curve として表わされ、ED₅₀、MD₅₀ 等が求められる。いつぼう、この単位投与量が変化したばあい、治癒の time course も変つてくるとすれば、多重曲線が得られる。また、これを一定の治癒率に達するまでの時間という見方をすれば、曲線群が得られる筈である。実際にこのような曲線群が得られるかどうか、臨床例を用いて証明しようとするれば、多数の揃つた症例が必要となるので、やはり女子急性膀胱炎を対象として行なうしかないこととなる。このばあい問題と

るのは外来患者を対象として行なわれるこの種の試験が、time course をとること自体をかなり困難にするだけでなく、Nurse, Laboratory technician 等 paramedical staff の緊密な協力を必要条件とすることである。また、time course をとるばあい、個人差が強く影響する時間中の限界を知る必要があり、またこの個人差を小さくする工夫も試みるべきである。この preliminary study として、前日、斎藤博士(虎の門、泌尿器科)の演説された大越会長を長とする amoxicillin 最小必要量推定の試みが行われたわけである。その結果は、amoxicillin 1回投与方法での治癒限界は 50 mg で、それ以下の投与量では、治癒率が急激に下ることが推定された。ただこの study の初期に、投与後2時間まで、time course を追ったものは尿中細菌消失率が24時間目の消失率と相関せぬように見え、早期に個人差が大きく、この時間帯の作業は無理を伴うとの中間結論があつたが、これは投与時の消化管内の状態が一定せず、吸収の遅れが人によつて大きく異なるためとも考えられ、投与方法を変え現在検討中である。

下部単純性炎における作業は、このようなやり方で、患者に与える苦痛も少なく(24時間後で見ると、尿中細菌は消失せぬまでも、penicillin 剤投与特有の菌隔壁形成だけ障害された elongation が見られ、これとともに症状改善はかなりの程度に見られる)、協力も得られ易いが、上部尿路感染は、発熱腹膜刺激症状等、全身状態に影響を与えることが少なくないため、症例およびそれぞれの例における投与量の選択には極めて慎重とならざるを得ない。ところが、たまたま latent chronic infection とでもいうべき症例を経験し、この例で適量決定が比較的容易であつたので、上部尿路感染の診断法を含めて例示する。

29才、主婦、急性腎盂腎炎の既往あり。右軽度の脊椎肋骨角部の鈍痛で来院、静脈性腎盂撮影で、右中腎杯に極微の変形(?)を認めるだけ。尿中 *E. coli* 2×10^5 /ml、白血球は 5~10/HPF、膀胱炎症状も軽度。上部尿路感染はこれからは確定できず。次に Bladder wash out technique (FAIRLEY 法)と、Differential urine collection (STAMEY 法、分腎採尿法)を連続して行ない、利尿下に右腎尿だけに、 5.5×10^4 /ml の菌を証明した。治療として、cephalexin 1日 250 mg を分2投与し、症状軽快、尿中菌減少を見たので維持量として1日 50 mg に減量したところ、症状やや増悪し、尿中菌 2.2×10^5 /ml、尿中白血球計となつたので、1日量 250 mg と増量したが、今回はやや有効程度しか反応せず、さらに1日量 750 mg として始めて症状軽快、菌陰性化、尿中白血球消失を見、以後同量を継続し経過観察中である。

全経過を通じ発熱を見ない。この例は、典型的急性腎盂腎炎とは異なるが、症状軽く、上部尿路感染の確定診断にやや手間取つたものの、適正投与量の決定まで患者自身も医師も余裕を持つて経過を追えた利点があつた。従がつて、いわゆる適正投与量を決めるには、このような例において preliminary に1つの標準量を想定し、症例の緊急度の軽いものから漸次対象を拡げて行くのが、時間は要するものの医の倫理に背馳しない方法と考えられる。ただ各種薬剤において、それぞれ同様の方法をとるのは極めて困難である。そのため、前述の急性膀胱炎を対象とした Dose-response, Time course をさらにそれぞれの薬剤の pharmacodynamics と対応して観察することで、下部尿路感染における *in vivo-in vitro* の関連づけがどこまで可能であるかを検討した上で、これに基づいて上部尿路感染における必要量・観察法を類推するといった作業が主となつてくると思われる。

慢性症における同様の作業は、基礎疾患の軽重によつて治癒経過が左右されることとなるので、基礎疾患の程度を揃える作業が先ず行なわれるべきで(疾患層別)、これと薬剤の試験管内抗菌力、pharmacodynamics, etc. をからませて行くことになるが、何としても、基礎疾患の層別が最大の作業となり、このばあい感染準備状態のパラメータを何にとるか、層別可能か否かが決つてくることになる。この点に関しては泌尿器科学的尿路機能測定方法論の未熟さもあつて、かなり将来の問題となるかも知れず1つの努力目標とすべきものであろう。

追加発言(1) 桑原章吾(東邦大学微生物)

抗菌薬の絶対の有効性を知るには、とくに二重盲検の必要はない。しかし、現在多数の類薬が存在する環境条件ではそれらの相互の優劣を知るためには、判定基準が完全に整備されていない状態では、二重盲検が必須と考えられる。

追加発言(2) 清水喜八郎(東大第一内科)

比較試験をするのには有効性をみるより有用性をみるためのものであると考える。たしかに、桑原先生がいわれるように効果判定不十分なものもあり、有用性の判定は現時点では無理であり、基礎から臨床への橋渡しをするために、逆に基礎的なデータは不十分な面もあり、必ずしも類似薬としての価値評価ができない場合もあり、現時点では過渡期であると私は考えてみたい。

追加発言(3) 大越正秋(慶大泌尿器科)

二重盲検法そのものは薬剤比較の方法として全く正しく、反省の余地はないが、抗生剤の比較試験としての二重盲検法には多くの問題点がある。そのうち最も重要なことは、パラメーターとして何を選ぶか、その判定時期を投与後何日にするか、投与量は各々の位を用いるか

など、治療効果判定の方法の問題である。

抗生剤が細菌を消滅させることを目的としている以上、パラメーターとしては菌の陰転化の有無を最も重要視すべきであることはいうまでもない。この点で急性単純性膀胱炎は実験対象疾患として最も適当している。

なお、急性単純性膀胱炎は他疾患に比べれば、backgroundがよく揃っており、またその症例数が多いという特長もある。

本疾患における起炎菌の薬剤投与後の消失の有無を治療効果の判定のパラメーターとするならば、これは客観的事実であつて、してこれを blind のもとに比較する必要はないものと思われる。

副作用その他については別であつて、盲検を必要とするであろう。

[新薬シンポジウム 1]

経口カルベニシリン (Carbenicillin Indanyl Sodium) について

(司会) 真下啓明 (東大医科研)

I 抗菌力・動物実験

五島 瑳 智子
東邦大学微生物

I-CB-PC の抗菌力について各研究機関の成績をとりまとめて報告する。

1. *In vitro* 抗菌作用

1) 標準株における抗菌スペクトラム

I-CB-PC の抗菌力を CB-PC と比較すると、グラム陽性菌においては一般に I-CB-PC の抗菌作用が強く、とくに *Streptococcus* の MIC が CB-PC より低い値を示している。グラム陰性桿菌においては両者の間に大差はないが、*Proteus* の中には CB-PC の抗菌力がやや優るものがみとめられた。

2) 臨床分離株の感受性

Staphylococcus aureus 168 株の感受性は、CB-PC と I-CB-PC との差は少なく、ほぼ同程度の分布を示した。

Streptococcus pyogenes 85 株では CB-PC の MIC のピークは 0.19 $\mu\text{g/ml}$ 、I-CB-PC は 0.01~0.02 $\mu\text{g/ml}$ で I-CB-PC は CB-PC の 10~20 倍の抗菌力を示した。

Streptococcus faecalis 2 株についても株数は少ないが *S. pyogenes* の場合と同様、I-CB-PC 0.78 $\mu\text{g/ml}$ 、CB-PC 6.25 $\mu\text{g/ml}$ の MIC を示し、I-CB-PC に対する感受性値が低い。

Diplococcus pneumoniae 5 株は I-CB-PC 0.2 $\mu\text{g/ml}$

以下、CB-PC は 0.39 $\mu\text{g/ml}$ の MIC である。

Corynebacterium diphtheriae 8 株では、I-CB-PC 1.56 $\mu\text{g/ml}$ 、CB-PC 6.25 $\mu\text{g/ml}$ にピークがある。

グラム陰性菌のうち *E. coli* 187 株、*Shigella* 38 株、*Salmonella* 10 株では I-CB-PC と CB-PC はほとんど差が認められないが、*Proteus mirabilis* 69 株、*Morganella* 21 株では、I-CB-PC が 0.78 $\mu\text{g/ml}$ 、CB-PC が 1.56 $\mu\text{g/ml}$ にピークがあり、僅かの違いがみられた。*Proteus vulgaris* 9 株では両剤ともほとんど無効である。

Pseudomonas aeruginosa 209 株、*Klebsiella* 73 株、*Serratia* 46 株では両剤の差はほとんどない。

3) 殺菌作用

S. aureus 209-P に対し、CB-PC は 0.39 $\mu\text{g/ml}$ の濃度で殺菌効果をあらわすが、8 時間以後は菌数の上昇がみとめられる。1.56 $\mu\text{g/ml}$ 以上では 8 時間以後も完全に菌の発育がおさえられる。

I-CB-PC は 0.045 $\mu\text{g/ml}$ で殺菌作用を示すが、8 時間以後も菌の発育がみられないのは 0.19 $\mu\text{g/ml}$ 以上の濃度である。*S. aureus* における両剤の差は、MIC の差と平行している。

4) I-CB-PC と CB-PC の感受性相関

上記の成績からグラム陽性菌については、I-CB-PC の抗菌力が CB-PC より強いことが確認されたが、これは両剤の構造上の差に由来する脂溶性の違いがグラム陽性菌の菌体内取込みに影響し、I-CB-PC の MIC が低くなるものと考えられる。

I-CB-PC は腸管から吸収されると Indanol がはなれて CB-PC となるので、生体内ではこの差はほとんどないと考えてよい。事実、*in vivo* 実験で、両剤の MIC の差の最も大きい *S. pyogenes* の株を用いてマウスの感染治療実験を行なつても、両剤の ED₅₀ に差は認められない。

またグラム陽性菌には PC-G の含量の影響が考えられるが、I-CB-PC の PC-G 含量は CB-PC に比べて多くはないので、この点も否定されてよい。

2. β -lactamase に対する安定性

S. aureus、*E. coli* の β -lactamase に対する安定性は、*E. coli* ではほとんど差がなく、*S. aureus* では CB-PC のほうが安定である。

3. *In vivo* 抗菌作用

マウスを用いた実験感染治療実験では、*P. aeruginosa* 5 株について、I-CB-PC を経口、CB-PC を皮下に、感染後 1, 3, 5 時間後に 3 回分割投与したが、両剤の効果はほぼ同じであつた。

しかし、*E. coli*、*P. mirabilis*、*P. aeruginosa* 各 1 株