

(2) TIP/JIP による評価実例 司会：別府宏園 (TIP)

経口抗アレルギー剤 ———— 林 敬次 はやし けいじ



高槻赤十字病院小児科、医師

1973年より医療問題研究会で主に医療職場の問題を改善する活動をする。その他、1974年より、注射による筋短縮症自主検診団、大阪小児科学会等で喘息薬、乳児突然死問題、BCG問題、経口抗アレルギー剤問題などを通じた薬効評価問題などに取り組む。

経口抗アレルギー剤で喘息治療費用の6割を占める

経口抗アレルギー剤は、アレルギー全般に効く飲み薬ということになっています。ふつう抗ヒスタミン剤は喘息にはよくないと言われていて、ところが、経口抗アレルギー剤は、じんましんやアレルギー性鼻炎だけでなく、喘息に効くと言われます。この点が大きく「抗ヒスタミン剤」と違うところです。経口抗アレルギー剤19種のうち、6種は喘息だけが適応症として承認されています。薬価が従来の「抗ヒスタミン剤」の約20倍にも設定されている理由は「喘息に効く」という点です。ところが、経口抗アレルギー剤の多くの主な作用は抗ヒスタミン作用です。喘息に適応がないのに値段がその他のいわゆる経口抗アレルギー剤と同じになっているもの(2種)もあり、わけがわからなくなっています。しかし、大多数の経口抗アレルギー剤の特徴は、喘息に効くということになっているこ

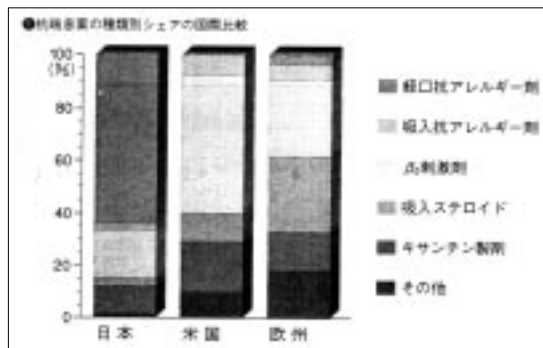


図-1 抗喘息薬の種類別シェアの国際比較 (メディカル朝日 1995, 2月号)

とは変わりはありません。

抗喘息薬の中で経口抗アレルギー剤が占める割合は、欧州では数%、アメリカでは0%、アメリカで経口のロイコトリエン剤が認可されたので少し増えるとは思いますが、これらに対して、日本では実に60%を越えています(図-1)。世界的に異常な状況が日本では起きています。なぜこのような異常なことが起こるのか？

経口抗アレルギー剤が日本においてのみ高く評価されているとしか、考えようがありません。

物を比較する尺度の問題

先ほどのセッション(第4章G, P.200~)では比較試験の意味が問題になりましたが、日本の場合最も問題になるのは、経口抗アレルギー剤にこのような高い評価をしている日本の薬剤の評価方法について考えると、日本特有の効果と安全性を計る尺度の問題がクローズアップされてくるのではないかと、一見どんなに厳密に見える比較試験をやっても、尺度そのものが無茶苦茶では意味がありません。

何でも、物を比較する場合は、尺度が必要です。長さはメートル、重さはキログラムで計ります。「あなたの方が私より重いと思う」というのでは、たとえ高名な科学者や医者が言ったとしても科学的ではないのです。会場の左側の人にはダイエット食を、右側の人には高級レストランのフルコースを10日間食べてもらいどれだ

けの体重の差が出るかを調べようと思うと、できるだけ全員の厳密な体重測定が不可欠です。

ところで、医学の進歩と共に、病気についてもその病態を計る客観的な尺度が症状の分析や検査によって開発されてきました。例えば、細菌性の肺炎なら熱・呼吸数・レントゲン写真・赤沈・CRPなど、肝炎ならGOT・GPT・ウイルス量など、高血圧なら血圧・心筋梗塞の発生・寿命など、ガンなら大きさや寿命です。もちろん何を主な尺度にするかが大切で、ガンは小さくなったが死亡したというのでは何にもなりません。

日本特有の尺度 = 「全般改善度」

ともかく、ほとんどの病気は、このような尺度を使って、治療が効果的であるかどうかの評価されなければならない。ところが、日本の場合は、大部分の薬剤の評価に、日本特有の尺度が用いられています。それが、「全般改善度」です。これは、様々な症状や検査成績を参考にして、臨床試験で試験物質を使う前と比較して、どの程度良くなったかを「著明改善」「改善」...「悪化」など、5~7段階の等級に分けるものです。先の食事と体重増のことで考えてみると、医者が体重や皮膚の艶や、患者の印象を元に「著明に肥えた」「肥えた」...「痩せた」などの等級に分けてダイエット食とフルコース食による体重の増加を比較するようなものです。「全般改善度」では、体重に重きを置くか、肌の艶に重きを置くかは決められていません。後に述べるように、体重を計り忘れていた人があっても簡単に等級が決められるのです。

喘息の評価尺度

次に喘息に対する経口抗アレルギー剤を具体例に検討してみました。普通の(世界的な)喘息の評価尺度は、1)肺機能、2)発作の回数や程度、3)試験物以外の喘息用の薬剤の使用量です。これらを数値化して、それぞれについて試験物群と対照群とを比較して統計的検討をします。しかし、日本の方式では、これらの一人一

人のデータを参考にして、一人一人に等級をつけて、試験物群と対照群とを比較する。どの尺度にどれだけの重点を置かなど一切の基準がないので、医師の主観が強くなる可能性があります。(図-2)いかに非科学的かは先ほどの体重のことを考えてみればよく分かります。

(欧米)

肺機能	試験薬群の値	a,b,c.....d,e,f	両群の比較	考察：結論
	対照薬群の値	A,B,C.....D,E,F		
臨床症状	試験薬群の値	h,i,j.....k,l,m	両群の比較	
	対照薬群の値	H,I,J.....K,L,M		
併用薬剤	試験薬群の値	o,p,q.....r,s,t	両群の比較	
	対照薬群の値	O,P,Q.....R,S,T		

(日本)

試験薬群の値	肺機能	a,b,c.....d,e,f	↓	主観	※この3指標から、一人ひとりについて、出発点からの改善の程度を等級をつけてこれらの値を試験薬群と対照薬群で比較する。
	臨床症状	h,i,j.....k,l,m			
	併用薬剤	o,p,q.....r,s,t			
	「全般改善度」	α,β,γ.....δ,ε,ζ		主観	
対照薬群の値	肺機能	A,B,C.....D,E,F	↓	主観	
	臨床症状	H,I,J.....K,L,M			
	併用薬剤	O,P,Q.....R,S,T			
	「全般改善度」	A,B,Γ.....Δ,E,Z		主観	

図-2 日本と欧米の薬効評価の違い

「全般改善度」はごまかしの手段

もう一つの重要な問題は、普通の尺度の一部が欠けていても平気で評価されている点です。例えば、喘息評価で重要な肺機能は、喘息に対する経口抗アレルギー剤の最も厳密であるべき第3相試験でさえ、「改善度」で評価した人数の37%~87%しか実施されていない。13%~63%の人は肺機能の成績なしにも関わらず「全般改善度」によって「有意に良かった」とか「同等だった」と結論しているのが現状です(表-1)。先ほどの、食事と体重の例では、体重を計れない人が居れば比較の厳密さは薄れるが、計れないからといって、肌の艶をみて「その人は著明に肥えた」として計算の中に入れられるようなものです。このようないかげんな「全般改善度」は、ほとんどの経口抗アレルギー剤で、普通の尺度による結果と食い違っていました。いわゆる改善度の評価と、1)肺機能、2)発作の回数や程度、3)試験物以外の併用薬剤があるかどうかの

改善度の人数に対する肺機能、臨床症状・併用剤の記載人数の割合

改善度の人数に対する 肺機能、臨床症状、併用剤の記載人数の割合					
文献番号	改善度	肺機能	臨床症状	併用薬剤	
小児	1	239人	?%	86%	90%
	2	277	?	97	86
	3	152	36	?	?
	4	199	30	89	90
	5	208	?	?	?
	6	223	15	99	99
成人	1	226	?	99	96
	2	189	37	92	92
	3	221	87	94	94
	4	236	87	103	104
	5	260	77	94	95
	6	166	79	90	94
	7	213	84	73	?
	8	262	73	88	88
	9	218	50	79	79
	10	179	?	90	90
	11	212	65	78	80
	12	229	74	93	?
	13	203	73	97	100
	14	209	64	88	89
	15	187	59	85	86
	16	264	67	89	89

注：臨床症状では発作点数を、併用薬剤では治療点数を優先して採用した。

表-1

改善度による評価と「3指標」による評価の関係

試験薬剤名	対照薬名	改善度による評価	3指標による評価
小児喘息			
トラニラスト(T)	フラーシ-キ(P)	T>P	NS
ケトチフェン(K)	フラーシ-キ(P)	T>P	T>P?1
レビロラスト(R)	トラニラスト(T)	R>T	R>T?2
ハミロラスト(P)	トラニラスト(T)	NS	P>T?3
オキサトミド(O)	トラニラスト(T)	NS	NS
オサケレド(O)	トラニラスト(T)	NS	NS
成人喘息			
タニラスト(W)	トラニラスト(T)	W>T	NS
チルフェナジン(T)	ケトチフェン(K)	T>K	NS
エピナスチン(W)	ケトチフェン(K)	W>K	NS
ケトチフェン(K)	フラーシ-キ(P)	NS	K>P
トラニラスト(T)	フラーシ-キ(P)	T>P	T>P?4
イソララスト(K)	トラニラスト(T)	K>T	K>T?5
スプラタスト(I)	トラニラスト(T)	I>T	I>T?6
ハミロラスト(P)	トラニラスト(T)	P>T	P>T?7
メキサチン(M)	ケトチフェン(K)	NS	M>K?8
トラニラスト(T)	クロモリシタム-フラーシ-キ(P)	T>P	T>P
オキサトミド(O)	クロモリシタム-フラーシ-キ(P)	K>P	K>P
オサケレド(O)	フラーシ-キ(A)	O>A	O>A
トラキサクス(T)	クロモリシタム-フラーシ-キ(P)	NS	NS
レビロラスト(R)	トラニラスト(T)	NS	NS
アピラスチン(E)	ケトチフェン(K)	NS	NS
アムレキサクス(A)	トラニラスト(T)	NS	NS

?1,?7:極めて不自然な経過を示す
 ?2,?3:観察項目の一部の一時的有意差のみ
 ?4:一日量の少ない方のみ有意差が出るという不自然なもの
 ?5:6週のうち3-4週のみK>T
 ?6:重症度のみ、発作・発作・治療点数その他は逆なし
 ?8:観察項目9分の2のみM>K

表-2

3指標を使った評価と、どのように異なるかを具体的に示します(表-2)。トラニラストとケトチフェンを見て下さい。改善度では2剤とも効いたという評価になっている。ところが3指標によると、トラニラストはまるで有意差がありません。ケトチフェンも昨日浜さんがスライドで紹介したように何らかの操作が入った有意差ではない。

普通の尺度では有意差が証明されていないのに、「改善度」では有意差があったというものがほとんどでした。逆に、ケトチフェンのみでした。

危険性の煙幕＝「安全度」

効果の評価ばかりではありません。安全性についても「安全度」という医師の主観の評価が優先されています。安全性評価に欠かせない血液・尿検査がきちんと実施されていなく、必須の検査もなしに「安全度」が評価されています。また、安全性については、医者や薬剤師にもその情報がまともに伝えられていません。

このようなことを経て、海外では認められないような物が日本ではいかにも効果があるかのように作られていくのです。

また、プロニカの長期試験では喘息死と交通事故死が出ているのに「安全度」を「問題なし」として、どこにも考察されていない。まさに「安全度」が副作用を隠す働きをしているのです。

このようなことをTIP誌や、日本臨床薬理学会などいろんな場で発表してきました。一つだけ反論がありました。聖マリアンナ医科大学の方ですが、(スライド省略)上から発作点数、治療点数、喘息点数とありますが、これをもって、「効いた」と、あるアレルギー専門雑誌に書いておられました。中からエピナスチンだけを引っ張り出して、効いたことにしている。コントロールがなくなっているのです。こういうことをアレルギー専門雑誌に平気で書くこと自体が、日本の医学雑誌のレベルを示していると思います。

以上が、経口抗アレルギー剤の評価尺度の問題点です。こういういい加減な評価尺度で承認

された物であるにも拘わらず、日本では60数%の非常に大きなシェアを持っているのが現状です。「改善度」を主体にしてどうにでも誤魔化せるような臨床試験をして、製薬企業に都合のよい結果にして売り出していくという、製薬会社丸抱えの学会、アレルギー学会や小児アレルギー学会などの問題点に辿り着くのではないかと思います。

「全般改善度」は、経口抗アレルギー剤に限

■質疑/討論■

日高：(医師、耳鼻科)一つだけ質問いたします。喘息の病因は何でしょうか。必ずしもアレルギーだけではないですよね。

林：そのことについては後で議論を。

日高：国際的な治療と日本の治療の違いがあると思います。国際的な治療は一般的な気道の慢性に対する治療でしょう。日本と海外とで喘息に対する認識が違うのだと思います。だから治療方法も違って来た。薬剤の使用量があれだけ違うというのはちょっと信じられない。やはりそれなりの理由があるからではないかと思えます。逆に、外国で抗アレルギー剤を使わないのは喘息が必ずしもアレルギー疾患ではない、という認識があるからだと思います。

林：先ほどからお見せしました「全般改善度」の評価方法の問題点について、理解していただけていないような。僕は、その点を問題にしているのです。喘息の認識が違うかどうか、ではなくて。

浜：議論の焦点を、喘息に対する認識の問題ではなく、全般改善度という評価方法についてしていただきたいのですが。

日高：耳鼻科だけでの話はよくあるのですが、小児科のドクターとこういう風に共通の話題を持てる機会があまりなくて、つい質問がずれてしまいました。後で、またよろしくお願いたします。

らず、大部分の薬剤の評価に使われている尺度です。多くの学会・大学医局などを通じ、このような非科学的な評価方法で、効かない単なる化学物質が「効く」とする試験をし、厚生省・薬事審議会もこのような試験方法をガイドラインで強要する一方で「薬」として認可する。その結果、製薬企業は効かない薬で大儲けをし、国民は莫大な浪費をし、患者は不要な副作用と出費を押し付けられているのではないのでしょうか。

林：はい、是非後でじっくりと。

浜：アレルギーであるということの認識は同じだろうと思います。

大橋：(統計学者)一言コメントを。現在、ICH(国際ハーモナイゼーション)がヨーロッパ、アメリカ、日本の3極で議論されていまして、全般改善度というのは多分、今後使われなくなると思っています。

林：全てにですか。精神科でもですか。

大橋：精神科とか一部の領域では許されるでしょうが、それは信頼性・妥当性という問題が確保されればという括弧つきです。

林：世界的に言われているいわゆる総合的評価法と日本で使っている全般的改善度あるいは有用度、安全度というシステムと混同されておられるのではないのでしょうか。

大橋：いいえ、主観が入った評価方法は今後許されなくなるということです。ガイドラインによって使えなくなる。しかし、精神科領域など、どうしても主治医の主観の評価を交えなくてはならないものは認めるけれども、基本的には今までの総合改善度というものは極めて使われにくくなる状況が現出します。

近藤：では大橋さんに質問したいのですが、今後使われなくなるのはいいとして、今まで使って評価してきた薬はどうなるのですか。

大橋：そこはコメントを控えます。どういう形で再評価されるかは知りません。