

図6

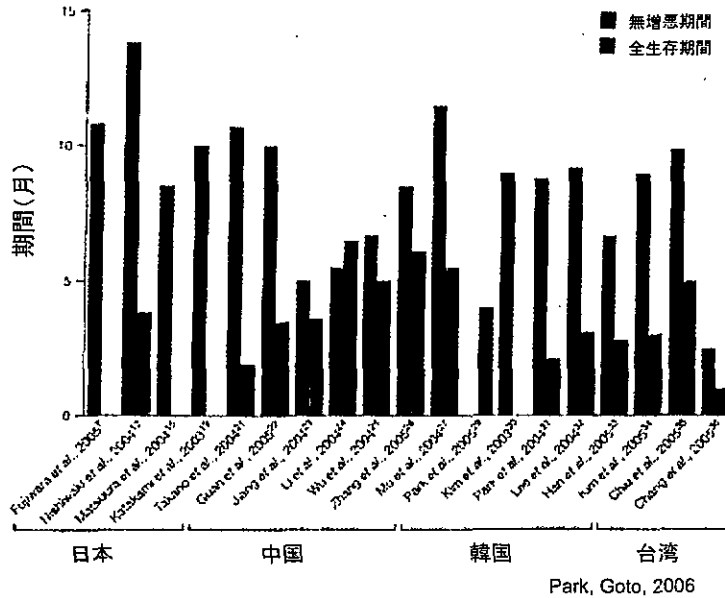


表6: SWOG 0023

二次化学療法+放射線治療に反応性があつたか、癌の進行がなかつた非小細胞肺癌患者でのゲフィチニブの効果。27カ月, 対象患者数=243人

	全生存期間(月)
プラセボ	35
ゲフィチニブ	23

ASCO, 2007

表7

INTACT-1 試験

手術不能 III-VI期の非小細胞肺癌で化学療法歴なしの欧州人患者1093人

プラセボ	* 10.9 mo	● 6.0 mo	▲ 44.8 %
ゲフィチニブ 250 mg	9.9 mo	5.8 mo	50.3 %
ゲフィチニブ 500 mg	9.9 mo	5.5 mo	49.7 %

シスプラチン+ゲムシタピン療法にそれぞれを追加した場合の

- \* 生存期間中央値
- 無増悪期間中央値
- ▲ 反応率

Giaccone et al., 2004

表8

INTACT-2 試験

手術不能 III-VI期の非小細胞肺癌で化学療法歴なしの米国人患者1037人

プラセボ	* 9.9 mo	● 5.0 mo	▲ 28.7
ゲフィチニブ 250 mg	9.8 mo	5.3 mo	30.4
ゲフィチニブ 500 mg	9.7 mo	4.6 mo	30.0

カルボプラチン+パクリタキセル療法にそれぞれを追加した場合の

- \* 生存期間中央値
- 無増悪期間中央値
- ▲ 反応率

Giaccone et al., 2004

た、たとえハード・エンドポイントをパラメーターとして用いたとしても、現実の差がどの程度のものであるかと確認することは重要です。なぜなら、抗がん剤の試験で、生存期間を測定してみると、その効果は期間を0~3.7カ月延長するだけの、比較的わずかな効果に過ぎないからです。これは私たちが抗がん剤に期待しているものとは明らかに異なります(表5)。同様のデータは抗がん剤として用いられる他の化学物質に関してもみられます。これは非小細胞肺癌の治療薬として用いられるイレッサ(一般名:ゲフィチニブ)の作用について収集したデータです<sup>8)</sup>(図6)。ご覧のように、予備データによれば、この薬が全生存に関して有効なようにみえます。しかし、2種の化学療法と放射線治療の後のケースにイレッサを投与するというランダム化比較試験を行ってみると、有意の効果はみられません。実際は、生存期間(月)でみると、プラセボを投与された患者よ

りも劣るというネガティブな成績<sup>9)</sup>でした(表6)。

また、これは他のデータによっても確認されております。例えば、これは腫瘍が大きすぎて手術の対象とはならない約1000人の患者を対象に、化学療法とゲフィチニブの併用療法を行った欧州での臨床試験です。ご覧のように、用量によって示してありますが、生存期間の中央値で測定すると、プラセボと比べても効果がないことがわかります<sup>10)</sup>(表7)。米国で、同様な条件の患者約1000人を対象に行われたINTACT-2試験でも、生存期間の中央値でみると、プラセボ群との間に差がありません<sup>11)</sup>(表8)。これは、ゲフィチニブが無効であることを示す一連のデータです(表9)。しかしながら、この薬剤には、間質性肺炎という非常に重大な毒性があり、これはかなり問題です。効果がないだけでなく、毒性があるというわけですから<sup>12)</sup>(表10)。これは先ほど示した欧州での研究(表4)

とは異なり、米国の食品医薬品局が行った新薬評価の研究です<sup>13)</sup>。これは、医薬品承認に際して行われた評価を13年間にわたって調べた研究ですが、ご覧のように承認に用いられる評価尺度としては反応率が最も多く用いられていることが分かります(表11)。このように、腫瘍の反応率、すなわち腫瘍サイズの縮小という観点だけで有効とされた薬を私たちは用いているわけですが、この腫瘍サイズの縮小が生存に影響するというデータは全くないのです。他にも多くの同様な例をみることができます。これは別の研究ですが、ここでも抗がん剤に関する評価の68%は生存やQOLではなく、腫瘍の反応率で評価されています。定量的な表現が異なっているのです。

したがって、私たちは、医薬品の評価に代理エンドポイントを使用すべきではなく、患者にとって重要なパラメーターを用いて研究すべきだと思います。そして、さらに重要なことは、既にハード・

表9

試験名	対象患者数	ゲフィチニブ 生存期間中央値(月)	プラセボ
INTACT-1	1093	10.9	9.9
INTACT-2	1037	9.9	9.8
TALENT	1172	10.1	9.8
TRIBUTE	1059	10.6	10.8
ISEL	1692	5.1	5.6
IDEAL-1	200	7.9	-
IDEAL-2	200	5.9	-

表10

ゲフィチニブ	
使用患者数	間質性肺疾患
76 (TAIWAN)	4
65 (KOREA)	2
37 (JAPAN)	4
112 (JAPAN)	6
330 (JAPAN)	15
3354 (JAPAN)	19
185000 (WORLD)	1480

Modified from Park and Goto, 2006 2175

表11：抗悪性腫瘍薬の通常承認におけるエンドポイントのまとめ  
1990.01.1～2002.11.1

Total	57
Survival	18
RR	26
RR alone	10
RR + decreased tumor-specific symptoms	9
RR + TTP	7
Decreased tumor-specific symptoms	4
DFS	2
TTP	1
Recurrence of malignant pleural effusion	2
Occurrence of breast cancer	2
Decreased impairment of creatinine clearance	1
Decreased xerostomia	1

Abbreviations: RR, response rate; TTP, time to progression; DFS, disease-free survival.

表12：主要複合エンドポイントを構成する死亡・心筋虚血再発・冠動脈閉塞の個々のエンドポイントがどれだけ全体結果に影響しているか

	Clopidogrel (n=1752)	Placebo (n=1739)	p
死亡	45 (2.6%)	38 (2.2%)	0.49
心筋虚血再発	44 (2.5%)	63 (3.8%)	0.08
冠動脈閉塞	192 (11.7%)	301 (18.4%)	<0.001
主要複合エンドポイント	262 (15.0%)	377 (21.7%)	<0.001

\*TIMI grade 0 or 1 on angiography.

Lapostolle et al., 2007 2183

エンドポイントで承認された既存薬があるにも拘わらず、しばしば代理エンドポイントを用いた評価で新薬が承認されていることです。これは非倫理的なことだと私たちは思います。

複合エンドポイントの利用

私たちが非常に注意しなければならないもう一つのポイントは、複合エンドポイントが使われている時です。ご存じのように、多くの場合、ただ一つのエンド

ポイントで評価する場合には事象の発現も比較的少ないものです。そのため、比較する2群間に差を見いだすためには、多数の対象患者を必要とすることになります。しかしながら、最近では、事象の発現がもっと多くなるように、いくつかのハード・エンドポイントをまとめた複合エンドポイントを使用する傾向がでてきました。そうすれば、対象患者数がなくても利用できるからです。しかし、複合エンドポイントの使用には問題もあ

ります。

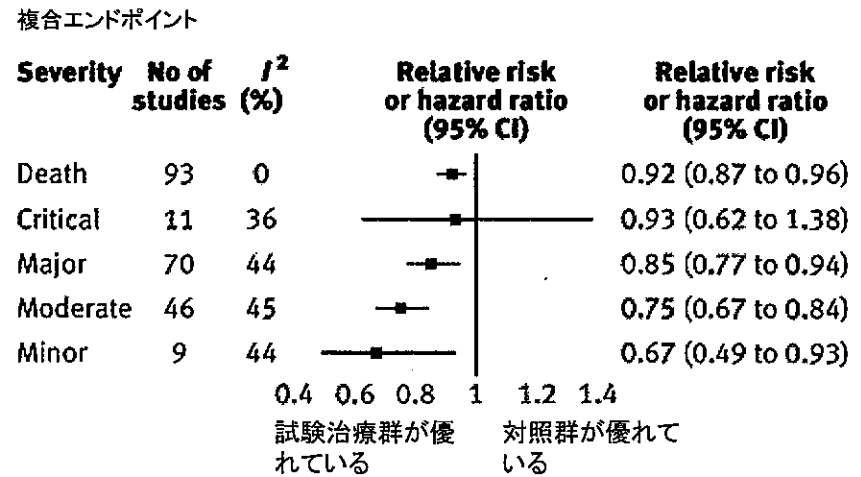
私たちが複合エンドポイントについて語るとき、実際に行われることは、たとえば心血管疾患、心血管薬、心血管死、心筋梗塞、血行再建術、入院など、全ての事象をまとめて数え、事象の発現を比較するわけです。疾患、心臓脈管薬、心臓血管死亡率、心筋梗塞、脈管再生、入院、およびすべてが合計されます、そして、比較はこのように実行されます。例えば、表12でいえば<sup>10)</sup>、最下段にある複合エンドポイントを見るとクロピドグレルとプラセボ間に有意差がみられます。この例では、クロピドグレル群のほうがプラセボ群に比べて有意にエンドポイントの減少があるので、有効だということになるわけです。しかし、死亡に関しては差がなく、虚血発作の再発も関係ありません。しかし、冠動脈閉塞に関しては有意差があり、関係がみられます。これらはいずれも重要なパラメーターですが、個々のパラメーターをみると、重要なパラメーターが重要性の少ないパラメーターによって影響される例があることが分かります。

これは多くの研究でいえることです。図7はその1例です<sup>10)</sup>。多くの臨床試験で複合エンドポイントが用いられているのが分かります。図に示すように、介入の効果は患者にとって非常に重要なエンドポイントよりもむしろあまり重要でないエンドポイントのほうに顕著にみられます。したがって、複合エンドポイントを用いた論文を読む場合は、複合エンドポイントを構成する個々のエンドポイントの重要度がどれくらいかを自分に問いかけてみる必要があります。結果の意味を判断するのは、そうした問いかけがあってはじめてできることだからです。(以下、次月号に続く)

文 献

- 1) Zinman B, Hoogwerf BJ, Duran Garcia S et al. : The effect of adding exenatide to a thiazolidinedione in suboptimally controlled type 2 diabetes: a randomized trial. *Ann Intern Med.* 2007 Apr 3;146(7):477-85.
- 2) Moher D, Dulberg CS, Wells GA: Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA.* 1994; 272(2):122-4.
- 3) Rochon PA, Fortin PR, Dear KB et al.: Reporting of age data in clinical trials of arthritis. Deficiencies and solutions. *Arch Intern Med.* 1993; 153(2):243-8.
- 4) Geddes J, Freemantle N, Harrison P et al. : Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. *BMJ* 2000; 321 (7273):1371-6.
- 5) Psaty BM, Weiss NS: NSAID trials and the choice of comparators--questions of public health importance. *N Engl J Med.* 2007; 356(4):328-30.
- 6) Barbui C, Cipriani A, Brambilla P et al. : "Wish bias" in antidepressant drug trials? *J Clin Psychopharmacol.* 2004 Apr;24(2):126-30.
- 7) Apolone G, Joppi R, Bertele V et al. : Ten years of marketing approvals of anticancer drugs in Europe: regulatory policy and guidance documents need to find a balance between different pressures. *Br.J. Cancer* 2005; 93(5):504-9.
- 8) Park K, Goto K: A review of the benefit-risk profile of gefitinib in Asian patients with advanced non-small-cell lung cancer. *Curr Med Res Opin.* 2006; 22(3):561-73.
- 9) Kelly K, Gasper LE, Chansky K et al. : Low incidence of pneumonitis on SWOG0023 : A preliminary analysis of an ongoing phase III trial of concurrent chemoradiotherapy followed by consolidation docetaxel and Iressa/placebo maintenance in patients with inoperable stage III non-small cell lung cancer. *Proc. Am. Soc. Clin. Oncol.* 2005 (abstract#7058)
- 10) Giaccone G, et al. : Gefitinib in combination with gemcitabine and cisplatin in advanced non-small-cell lung cancer: a phase III trial--INTACT 1. *J Clin Oncol.* 2004; 22(5):777-84.
- 11) Giaccone G et al.: Gefitinib in combination with paclitaxel and carboplatin in advanced non-small-cell lung cancer: a phase III trial--INTACT 2. *J Clin Oncol.* 2004; 22(5):785-94.
- 12) Park K, Goto K: A review of the benefit-risk profile of gefitinib in Asian patients with advanced non-small-cell lung cancer. *Curr Med Res Opin.* 2006; 22(3):561-73.
- 13) Johnson JR, Williams G, Pazdur R. : End points and United States Food and Drug Administration approval of oncology drugs. *J Clin Oncol.* 2003 Apr 1;21(7):1404-11
- 14) Lapostolle F, Cateineau J, Lapandry C et al. : Endpoints in studies on myocardial infarction. *Lancet* 2007 ; 369(9571) : 1430.
- 15) Ferreira-Gonzalez I, Permyer-Miranda G, Busse JW et al.: Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns. *J Clin Epidemiol* 2007; 60(7):651-7.

図7：エンドポイントを患者にとっての重要度別にみた場合の介入効果の大きさのばらつき



Ferreira-González et al., 2007 2187

未熟児への初期のステロイド剤治療：神経系の後遺症

Early treatment of premature infants with steroids: neurological sequelae  
Prescrire International (2007): 16(89): 108-109

ステロイドを呼吸窮迫症候群の未熟児に使用する方法は日本でも行われている。しかし、極めて質のよいランダム化比較試験の結果、その長期予後が悪いという明瞭な結果が得られた。短期使用でも同様の傾向が認められていることから、短期でも長期でもいずれも、ステロイドは使用すべきでない。

また、この結果は合計146人のランダム化比較試験で有意の差が得られたほど明瞭である。小児に対するステロイド全般の治療目的での使用も極力控えるべきことを示唆している重要な結果であると思われる (HR)。

まとめ

- 過去30年以上にわたり、呼吸器障害を有する未熟児には、出生直後からある程度の期間、ルーチンにステロイド剤が使用されてきた。
- 短期間で発症する害作用が多く報告されており、未熟児が3歳になるまでを調べた調査では、神経系の後遺症の問題が指摘されている。
- 出生直後から4週間にわたる全身ステロイド剤治療のプラセボ対照ランダム化試験に参加した小児146人が8歳まで追跡された。

- デキサメタゾン群の平均身長はプラセボ群よりも低く(女児121.3cm対124.7cm:p=0.01, 男児122.8cm対126.4cm:p=0.03), 平均頭囲も小さかった(49.8cm対50.6cm:p=0.04)。
- 運動や認知能に関する種々のテストのスコア、学校の試験の結果は、デキサメタゾン群の小児のほうが低かった。
- デキサメタゾン群に障害がより多く見られた(28/72人対16/74人, P<0.05)。
- これらのデータは、未熟児が出生後初

期から長期(4週間)にわたり全身ステロイド剤治療を受けることが、神経系に害作用を引き起こすことを裏付けている。呼吸器障害を有する未熟児に対するこのようなステロイド剤の使用法は中止すべきである。

ステロイド剤はいまだに、未熟児の出生直前と直後の両方に広く使われている。危険-益比のバランスが徐々に明らかになってきており、使用される状況(出生前か後か)によって異なることも分かってきている。

妊娠34週目以前の場合、分娩前の数日間、母親へのステロイド剤を使用することが乳児の罹病や死亡を減らし、害作用は許容範囲内であることが分かっている<sup>1)</sup>。

過去30年以上もの間、重症呼吸器疾患(呼吸窮迫症候群)のある未熟児にステロイド剤を処方することは、標準的な治療とされてきた<sup>2-4)</sup>。出生後4日間のステロイド使用を調べたデータは、適切な方法で行なわれた試験が20件あり、約3000人の未熟児が対象となっていた<sup>2)</sup>。

これら試験のメタ分析の結果では、ステロイドは気管支-肺の形成異常を減じ、抜管を早めるという短期的な利益があると結論された<sup>2)</sup>。

しかし、この治療法は、これらの試験で短期間の害作用が確認されており、危険-益比のバランスでは、危険(害)のほうが大きいことがわかった。そして、このたび、長期間の追跡調査でも、神経系の発達に害を与えることが示されたのである。

#### 短期間で多くの害作用

短期間の臨床試験を分析すると、ステロイド治療を受けた未熟児はプラセボ群に比べて、高血圧や糖代謝障害だけでなく重篤な胃腸障害(胃腸からの大出血、腸壁の穿孔)を起こしやすいことが示されている。成長の遅れ(身長および頭囲)や、肥大型心筋症も報告されている<sup>25-27)</sup>。

ランダム化比較試験(少人数の乳児が対象であることが多い)では、幼児を1歳から4歳まで追跡したところ、神経系の異常が起こる可能性が示されている<sup>2,8-12)</sup>。そして、長期追跡した新たなデータで、この危険性が確認されたのである(a,b)。

#### プラセボ対照ランダム化比較試験の参加者を8年間追跡

1992年から1995年の間に、アジアの6病院でランダム化比較試験が行なわれた<sup>13)</sup>。この二重遮蔽試験では、出生時体重が500gから1999g(平均約1400g)の未熟児262人を当初の対象とした。新生児は全員、出生後6時間以内に人工呼吸器を装着された児である<sup>13,14)</sup>。そして12時間以内にプラセボまたはデキサメタゾン0.25mg/kgを1日2回(静注)が開始され、これが当初7日間続けられた。その後、漸減され、合計28日間使用された。

死亡率(当初の試験期間)は両群とも同程度であり(33%対30%)(訳註1)、両群で159人が学童期まで生存しただけであった(c)<sup>14)</sup>。

生存した幼児の92%にあたる146人(デキサメタゾン群72人、プラセボ群74人)が平均8歳まで追跡された<sup>13)</sup>。評価項目は、理学療法士による機能的な検査、「特別に訓練を受けた」教師による学校の成績の評価、神経小児科医による神経系の検査、そして母親へのインタビューである(d)。評価の担当者は、デキサメタゾン群あるいはプラセボ群のどちら

かに割り付けられた児であるかは知らされていない。

#### 成長遅滞

デキサメタゾン群の小児の平均身長は有意に低かった:男子は122.8cm対126.4cm( $p=0.03$ ),女子は121.3cm対124.7cm( $P=0.01$ )。平均頭囲も有意に小さかった(49.8cm対50.6cm; $P<0.05$ )。

#### 運動能力の検査で「劣る」との結果

デキサメタゾン群の小児は運動能力が劣っていた。デキサメタゾン群ではプラセボ群の2倍の小児が、年齢の5パーセント以下以下のスコアであった。(デキサメタゾン群40%,プラセボ群20%)( $P=0.01$ )<sup>13)</sup>。協調運動能も劣り、視覚や手と目の協調運動能も劣っていた。

視力と聴覚の障害は両群で差がなかった(訳註2)。

#### 認知能力で低いスコア

知能指数(IQ)は、総合解析(78.2点対84.4点, $p=0.008$ )でも、言語能や非言語能で分けたサブ解析でも、デキサメタゾン群のほうがプラセボ群よりも低かった<sup>13)</sup>。素早さや集中力、系統立てて考える能力の点数(同じスケールで試験された)もデキサメタゾン群で低かった。

特別なニーズのための学校に通っていた子は両群で同人数いた。デキサメタゾン群の小児は、計算や音声の書き写し、文法の点数が低かった。しかし、話し言葉の理解、聴覚的理解、読解、語義の区別、単語の分析、適応力の点数には差がなかった(訳註2)。

#### 全体的に障害が大きい

このランダム化比較試験の筆者らは、脳性麻痺、認知能力が年齢の5パーセント以下、補聴器が必要そして/または視力が20/60以下であることを障害と定義した<sup>13)</sup>。これら一つ以上の障害を持っていた子は、デキサメタゾン群28人(39%)で、プラセボ群16人(22%)よりも多かった( $P<0.05$ )(e)。

脳性麻痺の小児はデキサメタゾン群のほうが多かった(17/72対9/74)が、統計学的に有意の差ではなかった( $p=0.07$ )<sup>13)</sup>。

#### 未熟児の出生後初期から長期にわたる全身ステロイド剤治療は中止すべき

未熟児の治療に関するランダム化比較試験で、これほど長期の追跡を行なった

例はまれである。これらの結果は重要であり、レベルの高いエビデンスを提供している。これらは、出生後初期から長期(4週間)にわたる未熟児への全身ステロイド剤治療は、神経系に障害を起こす危険があることを裏付けている。このようなステロイド剤処方法は、もはや行なわれるべきではない(訳註3)。

a- 新生児期の後期(出生後2週間以降)に、デキサメタゾン全身療法のプロセボ対照ランダム化比較試験に1週間参加した小児を13~17歳まで追跡したデータもある。このうちの多くが追跡不能であった(ランダム割付された287人中142人だけが追跡された)。また、プラセボ群の小児の39%にもステロイド剤が使用されたため、結果の信頼性はその分低い<sup>15,16)</sup>。

b- ガイドラインでは使用しないよう推奨しているにもかかわらず、2000年代になっても、ステロイド剤はいまだ頻繁に使用されている<sup>15,20)</sup>。

c- これらの小児が2歳になるまでの追跡では、デキサメタゾン群で神経発達の異常が多く、成長がより遅いことが示されたが、有意の差はなかった<sup>21)</sup>。

d- 両親の社会経済的状況は両群で類似しており、母親の学歴や新生児予防接種の頻度なども同様であった<sup>13)</sup>。

e- プラセボ群の小児のうち7人(デキサメタゾン群は6人)が(試験期間終了後に)、呼吸器に問題が生じたため、3日から5日間ステロイド剤が少量処方された。しかし、彼らは分析に含められた。彼らが除外されたとしても、障害はデキサメタゾン群のほうが多いままである。この試験で分析された他の母集団特性値は、詳細不明<sup>13)</sup>。

訳註1: 当初の試験期間までの死亡は33%対30%、学童期までの死亡例は53人(40%)対50人(38%)であった。有意の差はないとはいえ、むしろデキサメタゾン群の方が死亡者が多い傾向があり、菌血症や臨床的敗血症はデキサメタゾン群が有意に多かった。その上長期予後も悪かったので、良いことは何もなかったということである。

訳註2: 神経機能10指標中7指標で有意にデキサメタゾン群が劣り、そのうち