

平成16年(ワ)第25016号外 薬害イレッサ東日本損害賠償請求事件

原告 近 澤 昭 雄

外

被告 国

外

原告準備書面(30)・(第1分冊)

2009年(平成21年)1月30日

東京地方裁判所民事第24部合議A係 御中

原告ら訴訟代理人

弁護士 白 川 博 清

外

第 1 章 医薬品の有用性評価総論	9
第 1 はじめに	9
第 2 有効性の評価…有効性は確実に	10
1 はじめに	10
2 「くすりとエビデンス」	11
3 「効果と効率」	11
4 光富証人の証言	12
5 「薬事・食品衛生審議会医薬品第 2 部会議事録」 (乙 B 6)	12
第 3 安全性の評価…危険性は鋭敏に	12
1 はじめに	12
2 薬害事件における教訓としての国の認識	13
(1) 「昭和 43 年 5 月 7 日参議院労働委員会議事録」	13
(2) 薬害ヤコブ病確認書	15
3 薬害事件における教訓としての学者の認識	15
4 薬剤疫学の基本的な考え方	16
第 4 医薬品の有用性評価…有効性は確実に, 危険性は鋭敏に	18
1 はじめに	18
2 スモン訴訟の福岡地裁判決	18
3 「医療薬学 I」	18
4 本件における証人の証言	19
5 被告らの主張の破綻	20
第 5 有用性の科学的な「検証」	21
第 2 章 イレッサの有用性評価	25
第 1 節 イレッサの市販前の有効性評価	25
第 1 抗ガン剤の有効性評価の基本的な指標…真のエンドポイント	25
1 はじめに	25

2	抗ガン剤の臨床試験におけるエンドポイント	26
	(1) エンドポイントの種類	26
	(2) 生存期間	27
	(3) 腫瘍縮小効果	28
	(4) QOL, 症状改善等の患者の主観的なエンドポイント	29
	(5) 無増悪生存期間	33
	(6) NCI-PDQの記載	34
3	まとめ	35
4	「抗悪性腫瘍薬の臨床評価方法に関するガイドライン」について	36
	(1) はじめに	36
	(2) 旧ガイドライン注書(4)の記載	37
	(3) 解説文献等の記載	38
	(4) 別府証人の供述	39
	(5) まとめ	40
第2	個別症例による有効性評価について	41
1	はじめに	41
2	症例報告は医薬品の有効性のエビデンスたり得ないこと及びその理由	42
	(1) 出版バイアス, 発表バイアスを回避できない	42
	(2) 選択バイアスを回避できない	42
	(3) 観察バイアスを回避できない	43
	(4) 症例報告のエビデンスレベル	44
	(5) 医薬品の有効性は臨床試験の結果によって評価すべきであること	45
3	光富徹哉証人について	46
	(1) 光富証人の証言内容及び目的	46
	(2) 光富証人の症例報告もやはりバイアスを回避し得ない	47
4	福岡正博証人について	48

5	坪井正博証人について	49
6	まとめ	50
第3	奏効率による延命効果の予測の問題性	51
1	序論	51
2	奏効率が延命につながらない事例がイレッサ承認前に報告されていたこと	53
3	奏効率の判定基準による問題性	55
(1)	奏効率の判定基準に内在する問題	55
(2)	延命の予測精度の低い基準が採用された理由	55
(3)	小括	56
4	相関のみでは、代替エンドポイントの「妥当性」確認には不十分であること	56
(1)	序論	56
(2)	相関のみでは代替指標としての「妥当性」は認められないこと	57
(3)	相関の確認のみでは不十分である理由～タイムバイアス、予後因子バイアスについて	59
(4)	Buyse 論文による相関分析の批判	62
(5)	小括	63
5	被告らが根拠とする研究からは相関の実証すら不十分であること	63
(1)	福岡論文・西丙E 3 4 の 5 = 東丙G 6 0 の 5 について	64
(2)	西條論文（西乙H 3 8 = 東乙F 1 1）について	65
6	Ⅱ相承認の制度設計と、奏効率の位置づけ	66
(1)	「臨床試験計画（プロトコール）の作成と実施，並びに結果の統計解析とその評価について」（西甲D 3 4 = 東甲H 2 2）の記載	67
(2)	新医薬品課審査官（当時）による旧ガイドラインの解説	68
7	小括	69

第4	IDEAL 1, 2の奏効率から有効性を推測することの誤り	69
1	序論	69
2	抗がん剤の奏効率の確認に用いられる一般的基準	70
	(1) 20%の奏効率を必要とする西條証言の内容	70
	(2) 旧ガイドラインにおける期待有効率20%の記載	71
	(3) 旧ガイドラインが, 期待有効率20%を示した趣旨	72
	(4) セカンドライン以降の治療薬としての評価	73
	(5) 小括	74
3	IDEAL各試験におけるイレッサ奏効率の評価	74
	(1) IDEAL-1, 2において見られた奏効率の概観	74
	(2) IDEAL-1の各群全体としての評価	75
	(3) IDEAL-2の各群全体としての評価	75
	(4) プロトコールに照らした評価	76
	(5) 日本人群の結果について	79
	(6) まとめ	79
4	ドセタキセル試験との比較が有する問題点(背景因子の問題)	80
	(1) 被告らの主張	80
	(2) プロトコールには反映されていない議論であること	80
	(3) 歴史的対照の問題性	81
	(4) PS-2患者の割合と奏効率への影響	81
	(5) 西條証言も患者背景の問題を認めていること	83
5	他の既承認薬のセカンドライン患者に対する効果	83
6	小括	84
第5	IDEAL各群の生存期間中央値による有効性の推測について	85
1	被告らの主張の概要	85
2	対照群のない試験における生存期間中央値の評価	85

3	生存期間中央値は副次的評価項目で過大評価してはならないこと	86
4	既存薬における生存期間中央値の報告の概要	87
5	審査報告書の生存期間中央値分析における比較対象	87
	(1) IDEAL-1の評価について	87
	(2) IDEAL-2の評価について	88
6	まとめ	88
第6	承認時点のイレッサの有効性評価についてのまとめ	88
第2節	イレッサ承認前の安全性評価	89
第1	急性肺障害・間質性肺炎について	90
1	抗ガン剤による薬剤性肺障害には死亡例・重篤例も多く見られていたこと	90
	(1) 工藤翔二証人の主尋問及び意見書	90
	(2) 薬剤の安全性評価に関する観察研究の重要性	91
	(3) イレッサ承認前の薬剤性肺障害についての研究	93
	(4) 抗ガン剤一般にあてはめられないとの意見について	95
2	イレッサ承認前の知見では、薬剤性肺障害のうち、臨床経過としてAIP型（病理学的にはDAD）をたどるものがとくに予後不良であることもわかっていたこと	96
	(1) 工藤証人の主尋問及び意見書	96
	(2) AIP/DADについての承認時の知見	96
	(3) 小括	100
3	まとめ	101
第2	ドラッグデザインに見るイレッサの毒性の予見性	101
1	はじめに	101
2	イレッサのドラッグデザインとEGFRの機能	102
	(1) 「ヒト悪性腫瘍における上皮成長因子（EGF）関連ペプチドとそれらの	

レセプター」 David S Salomon ら, Oncology,1995	102
(2) 「上皮増殖因子が新生児ラットのモデルにおいて壊死性小腸結腸炎の進展を減じる」 Bohuslav Dvorak ら, American Journal of Physiology,2002	102
3 EGF R 阻害による肺障害の予見性	103
(1) はじめに	103
(2) 肺は傷つきやすい臓器	103
(3) II 型肺胞細胞の増殖・分化抑制と繊維化	104
(4) II 型肺胞細胞の機能抑制と急性肺障害	111
5 まとめ	119
第 3 非臨床試験に見るイレッサの毒性の予見性	120
1 はじめに	120
2 非臨床試験の意義, 目的	121
3 イレッサ非臨床試験で見られた多くの屠殺例の解釈	124
4 マクロファージ等の肺毒性所見	127
5 イヌ 6 ヶ月試験の肺炎症例等	130
6 ラット 6 ヶ月試験の肺胞浮腫等	133
7 まとめ	134
第 4 東京女子医大永井教授らの実験について	134
第 5 臨床試験, 副作用報告に見るイレッサの安全性の欠如	137
1 はじめに	137
2 臨床試験における有害事象の意味と重要性	138
(1) 有害事象の意味	138
(2) 治験担当医師の判断には限界があること	139
(3) 有害事象か副作用かの最終的な判断	140
(4) 個々の治験担当医師の判断が最終判断でないことは GCP や医薬品承認制度自体が予定している	141

(5) 有害事象の重要性	143
3 副作用報告におけるEAPの重要性	144
(1) EAPによる副作用報告	144
(2) 審査資料としての意味とその重要性	145
(3) EAPのデータは実地臨床で使用される場合に近い情報であること	146
(4) GCPに準拠していないことが副作用情報としての信頼性を低下させるものではないこと	147
(5) EAPによる副作用報告の重要性	149
4 イレッサの臨床試験, 副作用報告に基づく安全性評価 (致死的な急性肺障害・間質性肺炎の副作用発生の予見可能性)	149
(1) はじめに	150
(2) 臨床試験に基づくイレッサの安全性評価	150
(3) 副作用報告に基づくイレッサの安全性評価	164
(4) イレッサによる間質性肺炎の評価のまとめ	190
5 まとめ	192
第3節 イレッサ承認後の有効性評価	193
第1 第Ⅲ相試験に見るイレッサの有効性の欠如	193
1 はじめに	193
2 V1532試験について	194
3 INTACT試験について	200
4 ISEL試験について	201
(1) 試験結果について	201
(2) サブグループ解析について	202
5 SWOG0023試験について	206
6 INTEREST試験について	209
7 IPASS試験について	209

8	まとめ	212
第4節	イレッサの承認後の安全性評価	212
第1	イレッサは他の抗がん剤に比して高度の危険性を有する薬剤であること	212
1	イレッサによる副作用死亡者	213
2	独立行政法人医薬品医療機器総合機構への副作用症例報告	213
3	副作用報告症例の死亡例の数の比較	214
4	肺癌という癌腫に限定した場合の副作用死亡例の数	215
5	抗がん剤の使用者について	216
6	まとめ	217
第2	プロスペクティブ調査について	218
第3	コホート内ケースコントロールスタディについて	219
第4	まとめ	221
第5節	イレッサの有用性結論	221

本書面では、これまでの証拠調べの結果等を踏まえて、被告らの責任総論について述べるものである。第1分冊では、イレッサの有用性の評価について論じ、第2分冊では、被告らの法的責任について論じる。

なお、証拠等の引用については、大阪地裁（「西」と表示）と東京地裁（「東」と表示）の号証番号を並記している。これが同一の場合には、特に東西の区別なく引用している。

第1章 医薬品の有用性評価総論

第1 はじめに

本件では、イレッサの有用性の評価如何が大きな争点の一つとなっており、これは、イレッサの設計上の欠陥による被告会社の製造物責任、あるいは、被告国のイレッサ承認時の違法性等の争点の前提となるものである。

そこで、本章では、医薬品の有用性評価の基本的な考え方について、総論として述べる。

医薬品の有用性は、「有効性」と「安全性」の総合的考慮に基づいて評価される。なお、医薬品は、こうした有用性に加えて「品質」が保たれていることも重要な要素であるが（血液製剤の汚染による薬害エイズ、脳硬膜の汚染による薬害ヤコブ病などは、本来、この品質の問題であったといえる。）、本件においては特に、有効性と安全性に焦点を当てることとする。

そして、有効性と安全性の評価にあたっては、濱証人が述べるように、有効性は科学的に証明されてはじめて有効性の存在が肯定されるのに対し（「有効性は確実に」）、安全性については疑いの段階においても鋭敏に反応して十全の対策が取られる必要がある（「危険性は鋭敏に」西甲E25＝東G31濱意見書p

5)。

そして、このような基本的な有用性評価の視点に立脚した上で、医薬品の有用性が科学的に「検証」されるためには、適正にデザインされた臨床試験、それも比較臨床試験において、その有効性、有用性が「検証」されなければならない。また、こうした医薬品の有用性評価の際における重要な視点は、医薬品承認等の医薬品の適正評価、適正使用を確保するための場面における有用性評価と、承認により一旦有用性を確認された医薬品の実地医療での使用場面における評価を混同させてはならないことである。これら点については、既に西原告準備書面15＝東原告準備書面29において詳論したところであり、ここでは詳細は繰り返さないが、イレッサの有用性評価にあたって極めて重要な視点であることを強調しておきたい。

以下、敷衍して述べる。

第2 有効性の評価…有効性は確実に

1 はじめに

医薬品は、人体にとって異物であるということをその本質とする（西甲F15＝東L91メルクマニュアル第16版日本語版第1版p2494，西福島証人主尋問調書＝東甲L95p5，6）。したがって、医薬品を使用することによって、意図した効果だけでなく、直接目に見える形ではないにしても常に何らかの害作用が生体に生じる可能性がある。

有効性のない物質が市場に出回ることになれば、患者は、そうした物質を使用することによって、有効性のある他の医薬品による治療の機会を奪われたり、あるいは、異物である物質によって何らかの害作用のみを被る可能性が生るなどの不利益を被ることになる。

このようなことから、医薬品の有効性は、その存在が科学的に証明されて初めて肯定され、科学的な証明のない段階においては、有効性は存在しない、すなわち無効であると評価されなければならないのである。

これは、医薬品の有効性評価の基本的な考え方、基本的なルールであり、当然、本件におけるイレッサの有効性評価にあたっても貫かれなければならない原則である。

こうした有効性評価の基本的な考え方は、以下のような文献的な知見においても指摘されているところである。

2 「くすりとエビデンス」 (西甲 F 3 7 = 東 G 6 5)

本書は、東京大学大学院薬学系研究課医療経済学教授の津谷喜一郎氏らの編集にかかる文献であり、2000年1月に創刊された「EBMジャーナル」の第1巻第1号から「くすりとエビデンス」シリーズとして掲載されたものを加筆・修正してまとめたものである (iii 頁)。

同書においては、「臨床的な効果が証明されてはじめて“くすり”になる」の項において、「科学的観点から薬理作用があることだけではくすりとはいえない。臨床的な意味で効果があることが証明されてはじめて“くすり”といえるのである。それが“くすり”の承認のルールである。」 (p 3 2) と指摘されている。

3 「効果と効率」 (西甲 F 1 7 = 東 F 3 3)

本書は、EBM (evidence based medicine : 証拠に基づいた医療) の先駆けとも言える英国の医学者・疫学者であるアーチ・コクラン氏の著述の日本語出版であるが、そこでは、「また、有効という証拠がない限り、それは常に無効だと思っておくべきなのである。」 (p 1 1・8行目末尾) とされている。

4 光富証人の証言

以上のような医薬品の有効性評価の基本的な考え方について、被告側証人である光富証人も、「もちろんそうです。」として一般論としては全く同様の見解であると述べている（西光富反対尋問調書＝東乙L 2 4 p 2 1）。

5 「薬事・食品衛生審議会医薬品第2部会議事録」（乙B6）

イレッサの承認審査にあたっての「薬事・食品衛生審議会医薬品第2部会」において、審査センターの担当事務局も「いきなり本剤を既に標準療法が確立している初回治療の患者さんに単剤で投与して、万一期待されたような効果がない場合は、その患者さんに対しては良からぬこととなりますので」（p 3 2 下の事務局の説明）と述べており、これも、有効性の確認されていない投与方法によって効果がなかった場合には、患者に対して「良からぬこととなる」と説明しており、以上と同様の見解に基づくものと言える。

第3 安全性の評価…危険性は鋭敏に

1 はじめに

医薬品が生体にとって異物であることをその本質とする以上（西甲F 1 5＝東L 9 1メルクマニュアル第16版日本語版第1版p 2 4 9 4，西福島証人主尋問調書＝東甲L 9 5 p 5，6），医薬品の使用は常に何らかの害作用を及ぼす可能性を潜在的にもっていることが本質であると言える。したがって、科学的に証明されて初めて有効性が肯定できるのとは逆に、医薬品の安全性、危険性については、危険性の疑いがある段階で十全な対処がなされなければならない。

医薬品の危険性が科学的に証明されるためには、医薬品の使用過程で生じた

有害事象が医薬品の使用に基づくものであることの証明がなされる必要がある。しかし、そうした医薬品と有害事象との間の因果関係の科学的な証明のためには、当然、多くの時間がかかり、また、その過程で被害を発生させることになり、さらには、その間何らの対策も取られないとすれば、その危険な医薬品の市場における使用により被害が著しく拡大することとなってしまう。

こうした事態がまさに「薬害」なのであり、わが国で繰り返されてきた「薬害」の基本的な構図である。

サリドマイド等の甚大な被害を発生させた薬害事件の教訓としても、医薬品の安全性確保のためには、医薬品と有害事象の間の因果関係の確定を待つのではなく、危険性に疑いが生じた段階で十全な対処をする必要があることを肝に銘じなければならない。

こうした医薬品の安全性確保のための基本的な原則は、まさに過去の薬害事件の教訓であった。このことは、以下のような点に示されている。

2 薬害事件における教訓としての国の認識

- (1) 「昭和43年5月7日参議院労働委員会議事録」（西甲P40＝東L64）

サリドマイド事件にあたって、当時の園田厚生大臣は、医薬品の安全性確保について、昭和43年5月7日の参議院労働委員会において以下のとおり答弁している。

「第一は、これを契機にして薬というものに対する厚生省の基本的状態をはっきりしたい。……それから二番目には、ドイツで昭和三六年にこういう奇形児が出るということが発表になったと同時に、おかしいと思ったら直ちに製造中止を命じて、そして販売中止を命じて、その上で実験をすべきであった。調べてみますと、そういう事実を知り、厚生省の方では大

学に頼んで動物実験をやっておるようでございます。その間しばらく見送っておった。そして、製造中止を命じておる。次に、販売中止をやっておる。この二つの手抜かりがあった。こういう問題は、率直に、製薬会社と、それからこれを許可し販売させた厚生省、及びこういう事件が起こったあとの処置等についての厚生省の責任は私は痛感をしており、これに対する対処をしなければならぬと思います。

なお、また、今後の問題でこれは非常に大きな問題でございます、サリドマイド事件だけではなくて、薬に対する考え方がそのような考え方であるならば、今後新薬がどんどん出てまいりますから、これ以上の問題がどんどん起きてくると思います。したがって、厚生省としては、許可する場合の慎重な態度、あるいは、理論的なあるいは依頼した学者の方の御意見が妥当であると言われてみても、何か事件があった場合には直ちに生命に関することでありますから、製造中止なり販売中止を命じてからその上で検討するという事に考えなければ、薬というものが、一般の薬ではなくて人間の生命につながるものであるというもっと深い精神的な愛情をもって今後処置するということをここで深刻に厚生省並びに担当官は考えなければならぬ。……それで、まずこの際に第一に明確にしなければならぬことは、いままでずいぶん調べてみましたが、正直言って、許可した場合、それからその後の措置について、あいまいな責任のがれのことばを言っておりますが、たとえ訴訟になっておりましたとしても、政府と製薬会社はその責任をとって今後の処置をそれぞれやるべきだ、この点をまず第一に明確にいたしておきたいと思います。」

これは、薬事行政の最高責任者である厚生大臣の公式見解である。ここでも、医薬品の安全性について僅かな疑いが生じた場合には、迅速に対応すべきことが指摘されており、まさにそれが薬害としてのサリドマイド事件の基

本的な、そして重要な教訓であったのである。

- (2) 薬害ヤコブ病確認書（「薬害ヤコブ病の軌跡」西甲P 4 1 = 東L 6 1 p 4 3 以下）

薬害ヤコブ病事件において被告国は、以下のとおり、医薬品の安全性に疑いが生じた時は、その科学的評価を待つことなく、迅速に必要な対応をすることを誓約している。

「万一、医薬品等の安全性、有効性、品質に疑いが生じた場合には、直ちに当該医薬品等について科学的視点に立った総合的な評価を行うとともに、それに止まらず、直ちに必要な危険防止の措置を採るなどして、本件のような悲惨な被害を再び繰り返すことがないよう最善、最大の努力を重ねることを固く確約する。」（p 4 4 第2誓約の2項9行目の末尾部分）

3 薬害事件における教訓としての学者の認識

- (1) 我が国では、これまでサリドマイド事件、スモン事件、薬害エイズ事件、薬害ヤコブ病事件をはじめとする多くの悲惨な薬害事件に対して、薬害根絶のためにその原因を究明し、啓発してきた医学者は少なくなく、福岡スモン訴訟第1審判決でも引用されている元国立療養所東京病院長砂原茂一氏や臨床薬理学の専門家として永年薬害問題を研究してきた現東洋大学教授片平洸彦氏らがあげられる。

- (2) 砂原氏は、その著書「薬その安全性（岩波新書・1976年11月22日、西甲P 3 8 = 東L 6 3）において、医薬品の安全性確保についての基本的な考え方として、サリドマイド事件を受けて、以下のように指摘している。

「真実は何であるかということこそ問題のすべてであり、したがって、根

抛のありそうに思われる問題提起に対して敏感に反応する姿勢が、薬という重要な商品を生産しようとする会社、それを認可する行政当局にとって、もっとも必要なことなのである。」（p 11）

(3) また、片平氏も、その著書ノーモア薬害（桐書房1997年12月10日・西甲P39＝東L57）において、サリドマイド事件の教訓として、次のとおり、医薬品の安全性確保には、「疑わしきを罰す」を原則とすることを指摘する。

「薬害事件の場合、有害性が疑われた医薬品に対する措置は「疑わしきを罰す」を原則とし、迅速に使用中止・回収、その他適切な措置をとるべきです。何故なら、経済被害は後からでも補償が可能ですが、「生命・健康への被害はとりかえしがつかない」からです。」（p 38）

4 薬剤疫学の基本的な考え方

(1) こうしたわが国で繰り返されてきた悲惨な薬害事件の重要な教訓は、医薬品の適正評価、適正使用確保のための薬剤疫学においても基本的な原則として確認されている。

福島雅典証人は、京都大学医学部附属病院等において永年薬剤疫学の研究に携わってきており、薬剤疫学の第一人者ともいえるべき証人である（西福島証人主尋問調書＝東甲L95 p 1以下）。

同証人は、以下のとおり、薬剤疫学における医薬品の安全性評価の基本原則について述べている。

(2) 「医薬品の適正使用と副作用防止の科学」（西甲F16＝東L92）の「薬物療法と因果関係」の項において、

「薬物療法をひとたび始めたら、つまりくすりを飲み始めたら、何が起こっても、極端に言えば電信柱にぶつかっても、くすりのせいではないかと考えるべきである。」（p 43）

と述べており、この点について証言では以下のとおりさらに敷衍して述べている（西福島証人主尋問調書＝東甲L95 p8）

「これは極めて重要なことをごさいますて、通常、臨床試験では、まだ人に対して初めて投与して間もないわけで、十分な経験がございませんから、因果関係について簡単に判断することができません。ですから、その有害事象が起きたときに、それは例えば電信柱にぶつかっても、あるいは転んでも、あるいは交通事故であっても、あるいは自殺であってもということになるわけで、これには、過去に苦い経験と言いますか、非常に重要な事実がございます。かつて、インターフェロンが投与されたときに自殺者が出た。そのときに、初めは、それは自殺だから、個人の事情によるものではないかということを経験も一般に考えたわけですが。しかしながら、その後、そういう例が続いた時点で、これはおかしいということに医師は気づいて、やはりインターフェロンにより、うつ状態が発生して、それが自殺企図につながっているということに最終的に気づきました。ですから、最初の段階で、有害事象というふうに、すべて日常生活を障害するような事象が起きたときは有害事象と定義して、それは全部カウントします。で、ファイルしたうえで、それが本当に薬によるものかどうかを、あと解析する必要がある。ですから、その時点で即断して、それは関係ある、関係ないというふうにしてはならないということをごさいます。」

第4 医薬品の有用性評価…有効性は確実に，危険性は鋭敏に

1 はじめに

以上のような医薬品の有効性，安全性評価についての基本的な考え方は，スモン訴訟の福岡地裁判決が明確に指摘するところであり，また，以下のとおりの文献的な知見によっても確認されている。

2 スモン訴訟の福岡地裁判決（「判例時報910号」西甲P61＝東L80）

「有用性の判断は，以上の諸事情を総合的に判断して，有効性と安全性との比較考量の上に立って行なわれることになる。但し，右の比較考量は，既に述べたところから明らかなように，有効性の認定に際しては厳格に，副作用の発現可能性の認定に際しては緩やかに判断された上でのバランス論でなくてはならない。」（p94最上段の5まとめ欄）

3 「医療薬学I」（西甲F38＝東F60）

本書は，編集当時東京大学大学院薬学系研究課薬効安全性学教室に所属し，その後，2004年時点で国立医薬品食品衛生研究所所長（西甲G2＝東G7表紙及び末尾）になっている長尾拓氏の編集にかかるとの文献である。

「医薬品の評価」「非臨床試験成績の解釈」（p86）においては，非臨床試験における医薬品の有効性の解釈について，

「しかし，動物とヒトの感受性の差や体内動態の差を考慮すれば，これらはあくまで推定の域を出ず，ヒトでの有効性に直接的に置き換えて，薬剤間の効果の比較に使用することは十分な配慮を要する。」

として，有効性評価については，動物実験の結果をヒトに外挿するには十分な

配慮を要するとする一方で、安全性については、

「安全性は毒性試験と一般薬理試験において検討される。基本的に薬効薬理とは逆に発現した毒性や生理作用はヒトでも発現する可能性が潜在的にあると解釈する。」

として、薬効薬理すなわち有効性評価とは逆に安全性評価にあたっては、非臨床試験において観察された危険性はヒトでも発現する可能性があるものとして評価しなければならないことを指摘している。

その上で、医薬品評価の誤りがちが現状として、以下のとおり指摘している。

「一般に非臨床試験の有効性は過大に、安全性は過小に評価される傾向にある。すなわち有効性については、動物に対して臨床投与量の10倍以上を投与したデータで効果があれば有効だと解釈し、安全性については動物に対して臨床投与量の10倍を投与して発現する有害反応を、ヒトに使う量の10倍ですから…と軽んじる。この解釈にいかに問題があるかは既に解説したことから明白である。」（p 87・10行目以下）

4 本件における証人の証言

濱六郎証人は、こうした医薬品評価の基本原則について、「有効性は確実に、危険性は鋭敏に」と表現して以下のとおり指摘している（西甲E25＝東G31濱意見書p5）。

「ある物質を医薬品として一般臨床使用するためには、有効性が、比較臨床試験によって科学的に適切な手続きにしたがって実施され、確実に証明されている必要がある。有効という証拠がないかぎり、それは常に無効

だと思っておくべきなのである[1-1](アーチャー・コクラン, 1971)。

他方, 安全性については, 僅かな危険性を示すシグナルを鋭敏にとらえ, 試験物質や医薬品との関連を疑って対処しなければならない。とくにヒトに初めて用いる臨床試験においては, 「ヒトでは安全性のなお確認されていない物質を, 試みにヒトに用いる」[1-2](砂原茂一, 1979, KICADIS における基調講演より) のであるから, また, 特に第 I 相, I/II 相試験では全く安全性が確認されていない初めての試みであるので, 万が一にも, 危険性を示すシグナルを, 科学的に確立した知見でないなどとして, 安易に試験物質との関連が完全に否定できると判断してはならない。」

そして, 同様の点について, 以下のとおり証言している(西濱証人主尋問調書=東甲L102 p3)。

「有効性と, それから, 害がないのだということは, たくさんの人を使って, 統計学的あるいは疫学的あるいは大規模な臨床試験によって, それを使わないよりも差がないということをしちんと証明しない限りは, 1例でこれは関係がないんだということ saying はならない, しかし, 毒性試験とかいろいろな薬剤の性質からこれは起こるだろうなということも1例でもってでも, これは関係あるんだということ saying もいいということです。この点は混同しないように注意しなければいけないと思います。」

他方, 被告側証人である光富証人も, こうした有効性は確実に, 危険性は鋭敏にという医薬品評価の基本原則について, 一般論としては同様の見解であることを認めている(西光富反対尋問調書=東乙L24 p25)。

5 被告らの主張の破綻

以上のような医薬品の有用性評価の基本原則それ自体については、被告側証人である光富証人もまた何らの反論もできなかつたように、被告らとしても肯定するところであろう。

ところが、本件訴訟において現実に被告らが展開している主張、あるいは被告側証人らの証言は、「有効性は確実に、危険性は鋭敏に」という有用性評価の基本原則をないがしろにするものと言わざるを得ず、また、冒頭に述べたとおり、承認場面における医薬品の有用性評価と一旦承認された後の実地医療における使用場面での評価とを意図的に混同させているものと言わざるを得ない。

こうした点について、後に章を改めてより具体的に指摘することとする。

第5 有用性の科学的な「検証」

- 1 医薬品の有効性、有用性が科学的に存在することが確認、「検証」されるためには、最終的には適正にデザインされた臨床試験、それも比較臨床試験において、患者の臨床的な利益（ベネフィット）を確認するための真の指標（エンドポイント）の存在が証明されなければならない。これは、サリドマイド事件を契機に1962年にアメリカで制定されたキーフォーバー・ハリス修正法により明確にされた医薬品評価の科学的原則である（西甲F 22＝東G 38 p 66）。

そして、医薬品の単なる生物学的な活性（抗ガン剤では腫瘍縮小効果や腫瘍マーカー値の低下など）ではなく、患者の臨床的な利益（抗ガン剤では延命効果など）の存在を科学的に証明するためには、対照群をもちいた比較試験を行うほかなく、その比較試験において科学的に「検証」するためには統計学的手法を用いる他ない。そして、統計学的な「検証」においては、統計学の誤用（例えば脱落例を故意にカウントしないなど：西甲F 28＝東L 79「医者が

薬を疑うとき」 p 7 2 以下参照，あるいは「統計的多重性の問題」：西甲 F 3 0 = 東 G 5 6 「臨床試験のエンドポイント」 p 1 2 1 7 参照など）を防ぐ必要があり，そのためには比較臨床試験は適切な真のエンドポイントを用いて，適正にデザインされている必要があるのである。

この点については，別府宏圀証人が以下のように述べているとおりである（西甲 E 3 9 = 東別府証人主尋問調書 p 5， 6）。

「かつては，この薬を飲んだら症状がよくなった，だから効いたんだいう，まあいわゆる昔言われた三た論法というものがございまして，そういう単純なものでは，なかなか薬の評価はできない，なぜならば，病気の自然死というものがあって，独りでに治るものもあれば，あるいはそのほかのいろいろな要素が入って参ります。ですから本当に医薬品の評価をしようと思ったら，やはり2つの均質な，あるいは2つの同じ重みのある患者さんグループに，1つはある治験，実験，被験薬を与える，そしてもう一つは，その対照となる，場合によっては，それらはプラセボという，何の薬効もない，薬理作用のない薬を出す，そして比較するという，そういう慎重な検討が必要です。それに際しては，対象の患者さんの選び方とか，それから実際にその差が出たとき，どこでそれを差があると判断するかというようなことを，数学的な，つまり統計的，確率論的な方法で確認するということが大事で，そのようなきちんとした手続が決まっております，現在では G C P と呼ばれる，要するに臨床試験の実施に関する基準というものがございまして，これはもう今や世界共通のもので運用されているというふうに理解しております。」

こうした点については，西原告準備書面 1 5 = 東原告準備書面 2 9 において，より詳細に主張したところであり，詳細は同準備書面に譲るが，端的に示したものとして，「新医薬品の臨床評価に関する一般指針について」（西乙 D 2 5

＝東乙H28)の以下の部分を指摘しておく。

「臨床試験は客観的に、正確に、首尾一貫した方法で行われなければ科学的評価に耐えうる有意義な試験にはならないが、客観性を保った一定の評価を行うためには統計学的手法以上に優れた方法は今日まだ知られていない。検証目的の試験においてエンドポイントを1つまたはできる限り少数にする理由の1つには、この統計学的手法の誤用による判断の誤りの増大を避けることにある。プライマリー・エンドポイントとしては、一般に、臨床的ないし生物学的に意義があり、客観的測定、観察及び評価が可能で、薬理的にも説明ができる曖昧でないものとする必要がある。」(p689末尾部分以下)

- 2 また、抗ガン剤のように致死的な疾患に対する治療薬の場合には、ともすると医薬品の毒性死を安易に許容するような議論がなされることがあり、事実、本件において被告らも同様の主張をしている。

しかしながら、致死的な疾患であるからといって、その治療薬による毒性死を初めから許容するような議論をすることは誤りであるのは言うまでもない。

この点について、別府証人は、意見書や法廷で以下の様に述べている。

「がんの多くは進行性で致死的であり、予後不良である。抗がん剤では骨髄抑制や脱毛、嘔気・嘔吐などの激しい副作用があるにも拘わらずなお「臨床的有用性」があると評価されるのは、病気自体の重大性に照らして、危険性を容認できると判断されるからであろう。

しかしこのような評価は死を前提とした安易な判断に結びつきやすく、ともすれば危険対益のバランスの客観的・科学的な検証を疎かにするおそれがある。限られた命だからこそ、残された時間は貴重なのであり、副作用で患

者の死を早めたり，苦痛に満ちた時間を患者に強いることがあってはならない。したがって，抗がん剤の「有用性」を判断するにあたっては，通常の薬よりも一層慎重で科学的に厳密な検討が要求される。抗がん剤の「有用性」を示すには，副作用の危険性に見合うだけの十分で確かな「有効性」を証明しなければならないのである。」（西甲E37＝東L78別府意見書p2）

「抗がん剤というのは，もともと，これは死に至る可能性の強い病であるということから，大変判断が難しい場面がございます。しかし，そういうことでですね，どうせこの方は，危ない病気なんだからというような，要するに初めからその危険性，死を前提とした判断をしてはならないと思うんです。これはやはり，例えばその患者さんの余命が半年であるとすれば，その半年はまさに我々の半年よりもはるかに重い意味があるわけです。ですから，そういう意味では，そういう命を短縮することがあってはならないというふうに理解いたします。有用性という判断のときに，そこらをきちんと心にとどめて判断することが大事だろうと考えています。」（西甲E39＝東別府証人主尋問調書p8）

第2章 イレッサの有用性評価

第1節 イレッサの市販前の有効性評価

第1 抗ガン剤の有効性評価の基本的な指標…真のエンドポイント

1 はじめに

第1章で見たとおり，医薬品の有効性は科学的に証明された場合にのみ認められるのが原則であり，科学的な証明がなされない限り，当該医薬品は無効であると評価されなければならない。そして，科学的な証明は，適正にデザインされ，適切な真のエンドポイントを用いた比較臨床試験の結果によらなければならない。

これを抗ガン剤において見れば，抗ガン剤の有効性評価における真のエンドポイントは延命効果であり，全生存期間（overall survival）の延長である。

こうした点について，福島雅典証人は，以下のとおり述べている。

「通常，抗癌剤の有効性と言った場合には，これは延命効果のあり，なしを指します。…最終的に腫瘍が小さくなくてもそれが延命につながるかどうかは別の問題です。…単純に血圧を下げました，ああ，有効ですと，この薬を飲みましようと言ってたら，どういう副作用が出るか分からないし，実際に延命につながるかどうか，つまり，脳卒中や心筋梗塞や，そういうイベントを抑えるかどうか分からない。これを真のエンドポイント，真の評価ポイントと言います。ですから，抗癌剤の評価は，必ず延命できるかどうかで評価しないとイケないんです。単に，その薬を飲んだら，あるいは注射したら，腫瘍が小さくなりましたと。じゃ，1か月経ったら元の木阿弥で，そのあと

はリバウンドで、もっと大きくなってしまいましたでは、全く有効性があるとは言えないわけですから、ですから、真のエンドポイントである延命効果で見ない限りは、有効性については議論できないということです。」（西福島証人主尋問調書＝東甲L95p29,30）

「厳密な比較臨床試験をしない限りは、有効性については議論できない。延命効果については議論できないということになります。」（同p30）

被告らは、腫瘍縮小効果や延命効果、QOL、個別症例報告などの様々な指標をごちゃまぜにして「総合判断」することにより有用性を肯定できるかに主張し、被告側証人である福岡証人も同様の証言をした部分もあるが、こうした被告らの主張が破綻していることは明白である。

なお、イレッサ承認当時の「平成3年『抗悪性腫瘍薬の臨床評価方法に関するガイドライン』について」（いわゆる旧ガイドライン、西乙D7＝東乙H7）についての検討は後に行うが、まず、何よりも抗ガン剤の有効性評価の真のエンドポイントは、本来、延命効果であることを強く認識する必要があり、そのことが旧ガイドラインの評価においても極めて重要な意義を持つのである。

以下、さらに具体的に検討する。

2 抗ガン剤の臨床試験におけるエンドポイント

(1) エンドポイントの種類

「臨床試験のエンドポイント」（「癌の臨床」49巻10号,2003/10 西甲F30＝東G56）は、国立がんセンター研究所がん情報研究部・JCOG（日本臨床腫瘍研究グループ）データセンターに所属している山本精一郎氏が執筆したもので、ガンの臨床試験におけるエンドポイントについて一般的に解説している文献である。

臨床試験におけるエンドポイントとは、研究の目的ないしは指標という意

味で用いられる用語であり（西甲F 3 0 = 東G 5 6 p 1 2 1 1），それぞれの臨床試験における測定尺度（例えば，比較対照群を持たない第Ⅱ相試験における腫瘍反応割合）ないしは証明対象（例えば，比較対照群を持つ第Ⅲ相試験における生存期間の優越性，非劣性等）を指す。

エンドポイントには，「真のエンドポイント」と「サロゲートエンドポイント」があり，「真のエンドポイント」は，患者がどのように感じ，機能し，生存しているかを直接測定する臨床的に意味のあるエンドポイントであり，「サロゲートエンドポイント」は，この真のエンドポイントの代替となるものである（西甲F 3 0 = 東G 5 6 p 1 2 1 2）

また，別の観点からは，「ハードなエンドポイント」と「ソフトなエンドポイント」という種別もあり，「測定されたエンドポイントが客観的で誤差のない場合をハードなエンドポイントと呼び，奏効率や無再発生存期間のように測定者や測定日によって結果が変化するようなものをソフトなエンドポイントと呼ぶ」とされている（甲F 3 0 p 1 2 1 2， 1 2 1 3）

そして，臨床試験においては，「どの段階でもできるだけ真のエンドポイントか効果の証明されたサロゲート・エンドポイントを用いること，できるだけハードで一般的なエンドポイントを用いるべきである。」とされている（西甲F 3 0 = 東G 5 6 p 1 2 1 7）。

以下，個別に検討する。

（２） 生存期間

生存期間は，臨床試験登録日から死亡までの時間として定義されるtime-to-event の一つである（西甲F 3 0 = 東G 5 6 p 1 2 1 3，ここでの生存期間は，「全生存期間— overall survival —と同義である。）。)

この生存期間は真のエンドポイントであり（西甲F 3 0 = 東G 5 6 p 1 2 1 2），また，「研究者の解釈やマスク化していないことによるバイアスの

影響を受けないハードなエンドポイントといえる。」とされている（甲F30＝東G56 p 1213）。

このため、「多くのがん第Ⅲ相臨床試験の真のエンドポイントは生存期間であり」とされ（西甲F30＝東G56 p 1212），また，被告側証人である西條長宏証人ですら，「endpoint としてはもちろん第一に延命効果である。もし生存期間に差が出ない場合は，無病生存期間(disease free survival), QOL, cost benefitなどが検討されることがある。」（西甲F32＝東G58 p 24）としている。

なお，抗ガン剤の第Ⅲ相臨床試験における生存期間の検定は，プラセボもしくはは無治療・緩和療法群（BSC群）に対する優越性試験（イレッサにおけるISEL試験）と，既存の標準治療に対する非劣性試験（イレッサにおけるドセタキセルとの比較国内第Ⅲ相試験）があるところ，非劣性試験においては，非劣性限界，非劣性の許容範囲の設定（イレッサにおけるドセタキセルとの比較国内第Ⅲ相試験においては，非劣性限界としての95%信頼区間の上限値を1.25に設定するデザインをし，実際の試験では95%信頼区間の上限が1.4と上限値を上回ったため，イレッサのドセタキセルに対する非劣性は証明されなかった（西甲C5＝東E10－3））などの試験デザインの妥当性が慎重に吟味される必要がある（西甲F30＝東G56 p 1213, 1214）。

（3） 腫瘍縮小効果

腫瘍縮小効果は，薬剤自体の効果を測る上では真のエンドポイントではあるが，生存を真のエンドポイントとした場合のサロゲート・エンドポイントとして優れているかどうかは疾患の種類に依存するとされる（西甲F30＝東G56 p 1215）。

また，「腫瘍縮小効果の複雑な側面として，評価可能病変の事前特定，評

価のタイミング，画像による評価という方法論的な難しさが挙げられる。同じ奏効率でもCRとPRの割合が異なる場合，奏効期間が異なる場合，腫瘍が縮小した部位の違い，症状改善との関連，転移の程度と大きさをどう扱うかといった問題を含むソフトなエンドポイントである。」とされている（西甲F30＝東G56 p1215）。

福岡証人や西條証人は，腫瘍縮小と延命効果には相関関係があるとするが，腫瘍縮小は延命効果が認められるための必要条件ではあっても，腫瘍縮小が認められたからといって延命効果が認められるという意味での十分条件ではあり得ない。もし，十分条件としての意味も持つとすれば，生存期間をエンドポイントとした第Ⅲ相試験は必要ないこととなるはずであるが，現実には，そのようなものとなっていないことは明らかである。この点は，さらに後述する。

（４） QOL，症状改善等の患者の主観的なエンドポイント

ア QOL，症状改善等は，患者に対するアンケートを主体として調査したものであるため，患者の主観的なものであり，非常にソフトなエンドポイントである（西甲F30＝東G56 p1216，西丙E49の1＝東坪井証人反対尋問調書p79～81）。このため，「正しくそれを測定するためには，妥当性や再現性のある尺度を用いることが必要」である（西甲F30＝東G56 p1216）。

特に，QOL調査は，直接病状について聞く肺癌サブスケールの項目以外にも，「友人を身近に感じる」等の質問項目が多数取り入れられており（西丙H44＝東甲L157），抗ガン剤の治療効果以外の要素に大きな影響を受けることが避けられないものとなっている（西丙E49の1＝東坪井証人反対尋問調書p79。80）。

こうしたことから，西條証人は，「QOLは延命効果で差を検出しがた

い癌腫に対する第Ⅲ相研究での endpoint の1つとして検討されている。QOLの評価には多くの方法が存在しているが、一定の信頼性と妥当性はまだ確立されていない。第Ⅲ相研究の主要評価項目はあくまでも延命効果であり、それを無視したQOLの議論は本末転倒であると思われる。JCOGの研究においても、医療者側が客観的に判断しうるパラメータでないと、経過を追ってみても正確なデータを集めることは困難と示されている。」

(西甲F32=東G58 p27)としており、この文献が出版された2004年10月時点においても、QOL評価における信頼性と妥当性は確立されていないことが示されている。

イ また、QOL評価にあたっては、「バイアスを防ぐためにはマスク化が非常に重要」である(西甲F30=東G56 p1216)。

この点、イレッサのドセタキセルとの比較国内第Ⅲ相試験では、イレッサが経口投与剤であるのに対し、ドセタキセルは点滴による静脈内投与であるため、マスク化、盲検化はできない比較試験であった(西福岡証人反対尋問調書=東丙G58 p21)。

このドセタキセルとの比較国内第Ⅲ相試験では、イレッサ群がドセタキセル群に対して症状改善を見たと言われているが(西甲C5=東E10の3・18のスライド)、一方で「肺癌サブスケールによる評価」では両群で差はなかったと言われている(同スライド)。前述のとおり、QOLの調査項目は、肺癌サブスケールによる肺癌の随伴症状の改善を含む形で、それ以外にも「友人を身近に感じる」というような抗ガン剤の治療効果以外の要素が極めて影響を及ぼすような質問項目で構成されており(西丙H44=東甲L157)、肺癌サブスケールにおいて改善を示さず、QOL評価でイレッサ群に改善が見られたとしても、少なくともマスク化、盲検化されていない同比較試験においては、QOLの改善はイレッサの治療効果でない可能性を否定できないと言わざるを得ない(西丙E49の1=東坪

井証人反対尋問調書 p 8 4 ~ 8 6) 。

ウ さらに、「欠損値の処理も難しい。なぜなら、症状が悪くなるほど欠損値が発生する傾向にあるため、欠損値を無視して解析すると真の状態よりも見かけ上良くなってしまうという誤った結果につながることもあるからである。」(西甲 F 3 0 = 東 G 5 6 p 1 2 1 6) とされる。

欠損、欠測とは、「進行肺癌などの予後不良な疾患においては、終末期に状態が悪くなり Q O L 調査表に答えられなくなる。」(西丙 H 4 5 = 東丙 F 5 6 p 1 4 5) などのために、患者から回答を得られない状態を指す。

こうした欠損、欠測がある場合の解析については、種々の方法が提唱されているが、イレッサにおける I S E L 試験の Q O L 評価においては、単に、測定が可能であった症例のみを取り出して評価されている(西福岡反対尋問調書 = 東丙 G 5 8 p 2 2, 2 3) 。

しかし、こうした完全測定例のみでの評価は、「完全測定例における解析には、大きく 2 つの問題が生じる。1 つは、欠測データが 1 つでもある患者を解析対象から除くため、解析に用いる対象患者数が大きく減少してしまうこと、そして、もう 1 つは、誤った結果を示してしまう危険が高くなることである。欠測データをもつ患者の割合が少なくとも 5 % 未満でない限り、この解析アプローチを推奨できない。」(西甲 F 3 1 = 東 G 5 7 p 2 1 7) とされ、せいぜい 5 % の程度の欠測が許されるに過ぎない。そうでないと、患者背景等を調整してランダム割り付けした試験において、患者背景が調整されていない群間での比較がなされることとなってしまうからである。

この点、イレッサにおける I S E L 試験においては、イレッサ群 1 1 2 9 例中 2 7 1 例が、プラセボ群 5 6 3 例中 1 3 8 例が欠測しており、明らかに 5 % を大きく上回る欠測があったにもかかわらず、なんらの調整もせずに評価されているに過ぎない(西福岡証人反対尋問調書 = 東丙 G 5 8 p

22, 23, ISEL試験の報告論文である西丙E34の8＝東丙G60の8においても、QOL評価にあたって何らかの統計的調整を行った旨の記載はなく、もし調整していれば当然記載されるはずであるところから見ても、ISEL試験のQOL評価は、単に完全測定例だけを取りだした評価に過ぎないことは明らかである)。

このようなQOL評価では、それ自体、信頼性に著しく欠けることは明らかである。

エ このように、QOL、症状改善等の患者の主観的なエンドポイントは、それ自体、主観的なバイアスを排除することが著しく困難なエンドポイントであると言わざるを得ない。

このため被告側証人である西條証人も、「第Ⅲ相研究の主要評価項目はあくまでも延命効果であり、それを無視したQOLの議論は本末転倒であると思われる。」(西甲F32＝東G58 p27)とし、また、「QOLの評価はあくまでも個人的な主観によってなされるため、第三者の価値観による分析は困難である。」(西甲F3＝東F15 p210)、「癌の治療効果判定にQOLの評価を持ち込むことについて、わが国では大きなボタンの掛け違いがあるように思えてならない。まず、QOLという言葉があまりにも安易に用いられていることである。QOLの定義、QOLの測定の方法論と妥当性、成績の評価と患者への還元、いずれをとってもあいまいなまま導入され、言葉が独り歩きしている点を指摘できる。更に危険なのは、QOLという言葉を利用し生物学的治療剤や経口抗癌剤の優位性を強調する動きが見られることである。経口抗癌剤を例にとれば、正確に服用した場合は効果もみられるが、毒性も注射によるものと同以上にみられている。また経口エトポシドのように静脈注射より毒性が強く効果が劣ることが経験されている。」(西甲F3＝東F15 p217)とまで言い切っているのである。

そして、JCOGではQOL調査はあくまで第Ⅲ相比較試験においてのみ調査され、検証的なプライマリーエンドポイントには用いず、探索的なセカンダリーエンドポイント（副次的エンドポイント）と位置づけられているのみである（西甲P125＝東甲L158・2枚目4d）。

平成17年改訂の抗ガン剤の新ガイドライン（西甲D5＝東甲H6）においても、同様にQOLは副次的な評価項目としての位置づけでしかない（西丙E49の1＝東坪井証人反対尋問調書p77）。

（5） 無増悪生存期間

なお、全生存期間をエンドポイントとした臨床試験における、後治療の影響を排除するためなどに、無増悪生存期間がエンドポイントとされることもある。

無増悪生存期間とは、「増悪が観察された」時点か、「あらゆる原因による死亡」の早いほうまでの時間として定義される（西甲F30＝東G56 p1214）。

この無増悪生存期間は、「生存期間のサロゲート・エンドポイントとして、無増悪生存期間が用いられることがある。」（西甲F30＝東G56 p1214）とされているとおり、生存期間という真のエンドポイントに対するサロゲート・エンドポイントに過ぎない。

このように無増悪生存期間は、腫瘍縮小の増悪をもエンドポイントとする点で、結局、腫瘍縮小効果とその継続期間を見ているに過ぎない結果ともなり得ることから、腫瘍縮小効果と同様の問題点を持ち、サロゲート・エンドポイントに過ぎないとされているのである。

しかし、無増悪生存期間をエンドポイントとする臨床試験においては、「マスク化していない試験では群間に偏りが生じる可能性がある。」、「すべての患者を定期的に評価することが必須である。つまり、増悪する可能性

のある部位を全て評価し，ベースラインと追跡の段階ですべての部位を完全に確認すること，それぞれの追跡時に同じ評価方法を用いること，同じ評価スケジュールを用いること，などが必要となる。これらの評価方法・間隔が群間で異なると実際には群間で無増悪生存期間に差がない場合でも見かけ上差が観察されてしまうことがある。」，「試験実施上の問題として，増悪が判断した時期だけを調べるのではなく，無増悪であることを確認した日付も記録しておく必要がある。これがないと増悪が確認されていない対象者に対する打ち切り時期を決定できない。」，「つまり，観察方法によって結果が異なる可能性があり，生存期間よりもソフトなエンドポイントといえる。」

（西甲 F 3 0 = 東 G 5 6 p 1 2 1 4）とされているように，増悪の評価等において，バイアスが生じる可能性が高いという問題点が指摘されているとおり，極めてソフトなエンドポイントに過ぎない点に注意しなければならない。

（6） N C I - P D Q の記載

アメリカの N C I（米国国立癌研究所）では，P D Q（Physician Data Query）として，癌に関する情報を発信している（西甲 F 1 9 = 東甲 F 3 5）。

その中で「癌治療研究に対する証拠レベル（成人）」として，エンドポイントの強さについても記述されており，「強さの降順に挙げる」として，まず，第 1 順位に「全死亡」，第 2 順位に「原因特異的死亡」，第 3 順位に「注意深く評価された生活の質」，そして，第 4 順位として「間接的な代替」つまりサロゲート・エンドポイントが挙げられ，その第 4 順位の中で，「無病生存」，「無増悪生存」，「腫瘍反応割合」の順に挙げられている。

そして，この間接的な代替については，「これらはすべて，研究者の解釈による影響を受けやすい。さらに重要なことに，それらは生存または生活の質などの直接的な患者の利益に自動的に変換されるものではない。」とされ

ている。

このように、抗ガン剤の有効性評価の指標として最も重視されるべきは、全生存期間であることは、国際的なスタンダードなのである。

3 まとめ

以上のとおり、臨床試験のエンドポイントとしては、全生存期間ー overall survival ーが最も信頼に足る指標であることは明らかである。

このため、被告側証人である福岡証人も光富証人も、反対尋問において、全生存期間が最も重視されるべき指標であることを認めざるを得なかったのである（西福岡証人反対尋問調書＝東丙G 5 8 p 2 7，西光富証人反対尋問調書＝東乙L 2 4 p 3 7，3 8）。

もとより、福島証人、濱証人に加え、別府証人も全て、抗ガン剤の有効性は延命効果で見なければならぬことを証言しており、また、シルビオ・ガラティーニ教授もまた、全く同様に指摘している。同教授は、世界有数の研究所であるマリオ・ネグリ薬理学研究所所長であり、ヨーロッパ医薬品評価局委員や世界保健機構（WHO）の顧問なども歴任しており、レジオンドヌール勲章を初めとした国際的な賞も受けている高名な研究者である（西甲E 3 8＝東甲L 8 5 ガラティーニ意見書経歴欄，西甲E 3 9＝東別府証人主尋問調書 p 1 4，1 5）。

ガラティーニ教授もまた、抗ガン剤の有効性評価の指標として、以下のとおり指摘している。

「抗がん剤の有効性評価におけるプライマリー・エンドポイントは、もちろん、生存期間の延長である。既に述べたように、固形がんの種類は雑多な内容であるため、治療への反応か、，，本来の予後によるものかを予測することは難しい。このような条件のもとでは、反応率、無増悪期間、無増悪生

存率といったエンドポイントは、せいぜい抗癌活性の評価指標とみなされるに過ぎず、必ずしも臨床的有用性に関する正当な代用マーカールにはなり得ない。」（西甲E38＝東甲L85 p2）

こうしたガラティニ教授の意見を受けて、別府証人も以下のように供述している。

「先ほど申し上げたように、やはりプライマリーエンドポイントとして一番重視すべきは、やはり延命効果でございます、それ以外のものは非常に変わりやすいということでございます。ですから、例えばQOLを、その延命効果と並列的に並べて、こっちが駄目だけれども、こっちはいいというような、先ほど申し上げた二の矢三の矢としての使い方をするものではないというふうに思っております。」（西甲E39＝東別府証人主尋問調書 p16）。

4 「抗悪性腫瘍薬の臨床評価方法に関するガイドライン」（「旧ガイドライン」西乙D7＝東乙H7）について

（1）はじめに

福岡証人は、「抗悪性腫瘍薬の臨床評価方法に関するガイドライン」の「第Ⅲ相試験成績においては、生存率、生存期間をプライマリーエンドポイントとし、他の適切なエンドポイントとして、QOLなどを求め、これらに対し、何らかの有用性（プライマリーエンドポイントが同等である場合は他の特徴を含めてよい）が示される必要がある。」（西乙D7＝東乙H7・11枚目）とされている部分について、主尋問において、「第Ⅲ相試験によって延命効果が証明されなかった場合、その抗ガン剤には有効性がないという趣旨か」という質問に対し、「何らかのほかの有用性が示されればそれによ

いという理解ができる。」との証言をしており（西福岡証人主尋問調書＝東丙G57），被告らも同様の主張をするようである。

しかしながら，このような解釈は，旧ガイドラインの第Ⅲ相試験の主要評価項目（プライマリーエンドポイント）の解釈を誤り，旧ガイドラインの要求した抗ガン剤の評価方法をねじ曲げるものでしかない。

（２） 旧ガイドライン注書（４）の記載

上記の旧ガイドラインの該当部分には，注書（４）が付されており，これによれば，「第Ⅲ相試験で目標としたプライマリーエンドポイントで同等性が証明された場合は，他の特性，例えばQOLの改善（患者の肉体的苦痛の軽減，精神的満足度等）などの有用性が示される必要がある。」（西乙D7＝東乙H7・15枚目）とされている。

すなわち，延命効果が統計学的に示されなかったとしても，QOL等の何らかの有用性が示されれば良いという趣旨ではあり得ず，あくまで，延命効果で同等性が示された場合において，さらに他の有用性が示される必要があるという趣旨であることは明らかである。

なお，ここでの同等性とは，標準治療薬に対する同等性もしくは非劣性の証明を指しており，プラセボ，無治療・緩和療法群との同等性などではないことは，敢えて説明するまでもないであろう。比較対照群をプラセボ，無治療・緩和療法群においた時は，当然，被験薬は，それら比較対照群に対して優越している必要がある。

この点，光富証人は，あたかもプラセボ，無治療・緩和療法群に対する「同等」であっても良いかに述べるが（西光富証人反対尋問調書＝東乙L24 p 137），偽薬であるプラセボや無治療・緩和療法群と生存期間が同等であるということは，被験薬に生存期間のベネフィットすなわち延命効果が無いということに他ならず，以下のような文献等の記載から見ても，旧ガイ

ドラインがそのような趣旨で記載されていると到底考えられない。光富証人が本気でこのようなことを述べているとすれば、それは科学者として失格であると言わざるを得ない。

(3) 解説文献等の記載

ア この旧ガイドライン作成班の班員である下山正徳氏は、旧ガイドラインの解説文献において、「延命効果をエンドポイントにして、安全性、有効性の他に、QOL (quality of life)などを加えて臨床的な有用性について評価する。しかし、QOLだけをエンドポイントにして、延命効果をエンドポイントにしないのは不可である。必ず延命効果があり、QOLもよいというのが最もよい。場合によっては、延命効果はそこそこ同等であるが、QOLが非常によくするというのであれば、これもよい。だから、必ず延命効果をプライマリー・エンドポイントにした試験計画書を出す。」(西甲D15＝東甲H10 p129)としている。

すなわち、作成班の班員である下山氏自身が、あくまで延命効果が存在することを前提として、他の指標を考慮するという趣旨を鮮明にしているのである。

イ 被告側証人である西條証人も、「癌治療における国際化」(西甲H13＝東甲F51)において、「Bridging studyの臨床的考察について」の項、右欄下から6行目末尾以下「薬剤の承認後、薬の survival benefitの確認のため、independentな phase III studyを二つ要求される。これで survival benefitがなければ承認取消となる。」(p211頁)と述べて、旧ガイドラインの承認後の第Ⅲ相試験については、「survival benefit」つまり延命効果を確認するための「independentな phase III studyを二つ」つまり独立した2つの第3相試験が必要であり、これらで延命効果が確認されなければ承認取消となると明確に述べている。

同様に、西條証人は「癌の分子標的と腫瘍マーカーの開発」（西甲H 14＝東甲F 5 2）において、「第Ⅲ相試験においては、over all survival（OS）あるいは TTP を評価するため、これらに代わるサロゲート・マーカーは考えられず、survival benefit の有無を検討することがもっとも重要である。」（p 8 7 右 3 行目）、「臨床試験の早い段階でサロゲート・マーカーをもって有効性が示唆されても、最終的には第Ⅲ相試験で延命効果が証明されなければ臨床応用されないことはいうまでもない。」（p 8 8 左 8 行目）と述べて、抗ガン剤の第Ⅲ相試験においては延命効果の確認が最も重要で、延命効果が確認されない限り「臨床応用」されないと述べている。

西條証人は、イレッサについても市販後第Ⅲ相試験において延命効果が示せなければ承認が取り消されるべきであることを認め（西乙E 2 0＝東西條証人反対尋問調書 p 1 1 3）、また、イレッサが統計学的には有用性を持っていないことを認めている（西乙E 2 0＝東西條証人反対尋問調書 p 1 3 0）。

（４） 別府証人の供述

別府証人も、旧ガイドラインにおける延命効果とQOLの改善等との関係について、以下のように供述している。

「例えば、延命が数箇月延びたといえます。しかし、それが非常に苦痛に満ちた内容であったとしたならば、それは非常にクオリティーオブライフ、生活の質としては非常に劣ることになります。で、そういうことがあってはならない、つまり数字の上だけの延命ではなくて、やはりその人にとって価値ある時間であってほしいということでありまして、QOLをここに加えた理由は、まさに、さらに詳しい、さらに厳しい条件をそこ

に加えているというふうに読み取るべきであると思います。いやしくもこれを、QOLというものを、例えば、延命効果はないけれども、QOLでこれだけ少しよくなったところがあったからというような、言ってみれば、その効果がうまく当たらなかったときの、二の矢、三の矢、というような形で使うべきものではないだろうと思います。」（西甲E39＝東別府証人主尋問調書p10）。

（５） まとめ

このように、旧ガイドラインの趣旨は、あくまで延命効果が統計学的に証明されることが大前提となっていたものである。ただ、その証明が市販後に許されるという構造に過ぎない。

決して、延命効果が統計学的に証明されなくても、他のQOL等の指標によって有用性を判断しても良いなどということではあり得ないのである。

そのため、福岡証人も、反対尋問においては、この事実を認めざるを得なかった（福岡反対尋問調書p29）。

なお、福岡証人は、反対尋問に対して、「延命効果が同等でもいいですよね。」「延命効果が劣れば駄目ですよね。」等と述べている。これは、あるいは、イレッサにおけるドセタキセルとの比較国内第Ⅲ相試験において、「イレッサがドセタキセルに劣っていることが証明されたわけではない」という、独特の議論を前提にした回答であるとも考えられる。

しかしながら、同国内第Ⅲ相試験においては、イレッサのドセタキセルに対する非劣性というプライマリーエンドポイントが証明されなかったのであり、イレッサがドセタキセルに対して同等であるとか、劣っていないなどとは言えない状況であることは明らかである。そして、非劣性試験の性質上、イレッサがドセタキセルに「劣っている」、ないしはドセタキセルがイレッ

サに「優越している」などということは、元々証明対象、対立仮説ではなく、そのようなデザインはされていないのであるから、同試験の結果をもって、「イレッサがドセタキセルに対して劣っていると証明されたわけではない」などという議論は、根本的に誤った極めて非科学的な議論に過ぎない。このような主張こそ、「有効性は確実に」「有効という証明がない限り無効と考
えなければ」という医薬品の有効性評価の大原則を踏みにじるものである。

被告らの主張は、別府証人が述べるように、延命効果に対する副次的評価項目に過ぎないQOL等の指標を、まさに「二の矢、三の矢」として利用しようとするものであり、このような考え方は、旧ガイドラインの趣旨に明らかに反し、到底受け入れられるものではない。

そして、これまで述べてきたように、抗ガン剤の有効性は基本的に延命効果で評価されるべきものである以上、旧ガイドラインにおいて腫瘍縮小効果の確認により承認が許される構造となっているとしても、延命効果の確認のないままでの承認である以上、その際の有効性、安全性の評価は、より厳格になされる必要がある。この点については、さらに後述するとおりである。

第2 個別症例による有効性評価について

1 はじめに

被告らは、西被告会社準備書面（5）p 21以下＝東被告会社準備書面（5）p 20以下をはじめとして、イレッサの投与を受けた個々の患者についての症例報告の存在を列挙する。そして、被告側証人のうち、大阪地裁での西日本訴訟における光富徹哉証人、福岡正博証人、坪井正博証人もまた、いずれも、イレッサ投与例において著効例、スーパーレスポonder例があったなどとしている。

しかしながら、このような個別症例は、医薬品の有効性のエビデンス（証拠）たり得ず、これをことさらに強調することは、むしろ医薬品の有効性評価を見誤ることとなる。

以下、詳述する。

2 症例報告は医薬品の有効性のエビデンスたり得ないこと及びその理由

症例報告は、下記のようなバイアス、すなわち、医薬品の有効性が正しく評価されず歪められてしまうという現象を回避することができず、医薬品の有効性のエビデンスはない。

(1) 出版バイアス、発表バイアスを回避できない

出版バイアスないし発表バイアスとは、結果がうまくいったものについては出版、発表し、そうでないものには、出版、発表しない傾向にあることをいう。

この点、症例報告という手法は、報告者の主観により、結果がうまくいった症例を報告し、うまくいかなかった症例は報告しない、ということが容易に可能である。すなわち、症例報告は、「出版バイアス」ないし「発表バイアス」の影響から逃れることはできない（西光富徹哉証人反対尋問調書＝東乙L24 p56～57）。

(2) 選択バイアスを回避できない

また、ある症例や患者群を選択するにあたって、選択する集団や症例が母集団を正しく代表していないときに、そこで用いられる薬剤や治療法の評価を誤ってしまう事を「選択バイアス」という。例えば、予後因子のよい患者、すなわちもともと諸条件から治りのよい患者を選んでしまい、薬剤や治療法に関係なく疾患が改善した場合に、薬剤や治療法によって改善したのだと誤

ってしまうことがありうる。

この点、症例報告という手法は、集団のなかから個別の症例をまさに選択して取りあげるのもであって、「選択バイアス」から逃れることはできない（西光富徹哉証人反対尋問調書＝東乙L 2 4 p 5 7～5 8）。また、同一症例においても、そのうちの一部の情報のみ切り取り、例えば結果や背景因子等を選択的に情報として除外するというも行われがちであるが、これもまた「選択バイアス」によるものである。

（3）観察バイアスを回避できない

また、ある症例の観察者が、薬剤や治療法とその結果の関係について予断を有している場合に生じるバイアスのことを「観察バイアス」という。例えば、当該薬剤投与例に効果が出れば、それは当該薬剤ないし治療法によるものである、と考えがちである、という予断が考えられる。

これについて、元国立療養所院長の砂原茂一氏は、その著書において

「患者というものは、1人1人特別な条件をもっていて、2人として同じ患者というものは存在せず、1人の患者にある治療をほどこしたあとで病状が良くなることが観察されたとしても、同じ病気の他の患者に同じクスリを与えたとき同じような効果が期待できるとはかぎらないこと、さらに病気には自然治癒というものがあって、くすりをつかわず、手術をしなくとも、自然によくなることがある」「よくきくはずのくすりをつかったのだから、また、一生懸命治療したのだから、病気のよくなったのはそのくすりのおかげに違いないと、自分自身いいきかせているにすぎません。」（西甲G 4＝東甲L 4 9 p 9 0～9 1）

と述べ、てるてる坊主や千人針の例を引いた後、

「つまり、時間の前後関係を因果関係ととりちがえるのです。こういう考え方の誤りはいうまでもありませんが、治療の場合には、しろうとの

みでなく医者もしばしばこのような誤った判断に導かれます。」

と述べている。

すなわち、個別の症例を観察して、投与した医薬品に有効性があったと判断するのは、「飲んだ」「治った」「効いた」の「3た論法」として、医薬品の有効性を判断するにあたって、最も初歩的基本的な誤りなのである。

症例報告という手法は、観察者＝担当臨床医であり、当然、ある薬剤や治療法を使用したことを知っているのであり、「観察バイアス」に極めて影響を受けやすい（西光富徹哉証人反対尋問調書＝東乙L 2 4 p 5 9）。

（4）症例報告のエビデンスレベル

N C I（米国国立癌研究所）－P D Qのサイト（西甲F 1 9＝東甲F 3 5）には、ガン治療のエビデンスレベルについて、研究デザインの観点からその高低を順位づけた記載がある。これが、高い順から「1. ランダム化対照臨床試験」，「2. 非ランダム化対照臨床試験」，「3. ケースシリーズ」となっており、3の中でも、「i 集団ベースの連続シリーズ，ii 連続したケース（集団ベースでない），iii 非連続のケース」とさらに細かく高低が定められている。

この基準に照らして考えると、被告側証人が主尋問で行った症例報告は、それぞれ学会や論文等で別々に発表されたもので、連続したケースではないから、エビデンスレベルとしては最も低い3 iiiの「非連続のケース」である。なお、同基準には、「臨床経験は最も弱い形態の研究デザインである」と記載されている。

さらにいえば、さきほどのN C I－P D Qの基準によれば、「研究または臨床経験はデザインの強さとエンドポイントの重要性の両方によって順位付けされる」とされている。被告側証人らは、単に症例報告をしたのみであって、集団ベースの腫瘍反応割合についてのデータを提供しているわけではな

いので、上記基準のD iiiにもあてはまらない。

これら、被告側証人らが本法廷で行った症例報告は、イレッサの医薬品の有効性のエビデンスとしては「なし」というのが端的な結論である。

(5) 医薬品の有効性は臨床試験の結果によって評価すべきであること

このような症例報告等の観察研究が回避できない上記のようなバイアスを出来る限り除去し、医薬品の有効性を正しく評価するためになされるのが、比較臨床試験、とくにランダム化二重盲検比較臨床試験なのである（西甲F 40「Evidenceと臨床試験」＝東甲F 61）。

抗ガン剤の旧ガイドライン等をみれば明らかなおり、我が国において、抗ガン剤が、臨床試験を行わず、症例報告やその集積を根拠として承認されることはあり得ない。これは、症例報告では、医薬品の有効性が正しく評価されず歪められてしまうというバイアスを回避することができないからにほかならない。

この点について、工藤翔二証人は、東日本訴訟における主尋問において、医薬品の有効性評価のあり方として、

「本当の全体としての有効性というのは、そういう患者さん全体をマスで取り扱って、統計的に処理して生存期間がどれぐらい伸びるかとか、そういうことで判断していくものです。」（東工藤翔二証人主尋問調書＝西では未提出p 112）

と証言する。これは、医薬品の承認の根拠となる有効性の評価は、統計的手法、つまり臨床試験によって行うべきであるとする趣旨である（東工藤翔二証人主尋問調書＝西では未提出p 48～49, p 51～52）。すなわち、1つの症例報告の結果から、他の症例でどの程度類似の結果が得られるかというのは、症例報告そのものからは分からない。そうすると、どれだけの症例でどれだけの効果が出るのかというのを見るためには、全症例を見なければ

ばならず、結局、臨床試験によらざるを得ない（東工藤翔二証人主尋問調書＝西では未提出 p 49）。全く正当な指摘である。

また、個別症例が医薬品の有効性評価の根拠となり得ないということは、東日本訴訟の西條長宏証人の文献（東甲 F 59＝西甲 F 58 『癌診療と EBM Part 2』巻頭言）においても、下記のとおり、強調されているところである。

「学会に出席して頻繁に目にする抄録のタイトルとして『当院における〇〇の治療成績』あるいは『当院における〇〇診療の現状』がある。これらの大半は治療計画書（プロトコール）なしに、当然 IRB（Institutional Review Board）の審査もなく IC（Informed Consent）form もなく、clinical practice として行われたきわめて多岐にわたる診療（治療）による奏効率や、生存期間を retrospective にまとめたものが多い。これらは当然エビデンスレベルとしてはきわめて低いものであり、学問の進歩に寄与するものではない。また、これらの発表の最後に演者は必ず『今後さらに症例を重ねて検討したい』としめくくることが多いが、症例を重ねても何ら得ることがないことは自明である。」

これもまた、正当な指摘である。

以下、これに対して、医薬品の有効性の根拠として本法廷で症例報告を紹介している光富徹哉証人、福岡正博証人、及び坪井正博証人の個別の証言内容を確認する。

3 光富徹哉証人について

(1) 光富証人の証言内容及び目的

光富証人は、その証言の相当部分を自ら医師として経験した症例報告の説明に費やしている。光富証人が、別症例報告をもって医薬品の有効性の根拠としようとする意図のもとに証言をしていることは、光富証人自らが「示唆

に富む症例を紹介することにより，医療現場にたずさわらない方々にゲフィチニブによる肺がん薬物療法の実地臨床における実態と特徴を理解してもらうことに加え，ゲフィチニブに関する一専門家の医学的評価を提供するためである」（西乙E 1 2＝東乙L 1 4 p 1）と述べているとおりである。すなわち，光富証人は，イレッサがよく効く（と思われる）個別症例を選択して報告し，イレッサの「医学的評価」の根拠としようとしているのである。

（2）光富証人の症例報告もやはりバイアスを回避し得ない

しかしながら，光富証人の症例報告も，やはり，上記述べたような各種のバイアスを回避し得ていない。

光富証人は，「症例報告をするに当たっては，診療行為中に見られた事象のすべてを取り上げる必要はなく，自ら提供しようとする医学的評価を形成する上で必要かつ十分と考える範囲で，その一部を取り出し，整理することになる。」（西乙E 1 2＝東乙L 1 4 p 2）と，「選択バイアス」があることについて自ら認めている。法廷において紹介した6症例のうち，症例1・4以外の4例について死亡したことすらも当初は触れず，投与後いつ亡くなったかなどは意図的に情報として除外している。また，他方で，光富証人は，イレッサの毒性については，ほとんど全くと行っていいほど触れていない。わずかに，意見書において「一方で，副作用として間質性肺炎を発症した患者は，当診療科では100症例強中3症例を経験しており，そのうちの2症例は間質性肺炎によりお亡くなりになられた。」とあるのみで，その症例経過も，患者背景因子も全く分からないし，証言にあたってこれを確認検討したという形跡すら見受けられない。

光富証人自身も，自己の文献において，「専門誌に掲載される論文には，ポジティブデータが選択される傾向となる（出版バイアス）。また，研究者としては，ポジティブデータを発表したいために，ある場合は特殊な条件下

のポジティブデータを売りとして発表しがちとなる」（西甲H15＝東甲F62）と述べているところであるが、上記のとおり、光富証人は、本法廷における証言においても、自ら「示唆に富む症例」をセレクトした、としており、出版バイアスの影響にあることを自ら認めている。光富証人自身も、「エビデンスレベルからいえば臨床試験にそれに比べて低いものであることは言われるまでもなく十分承知している」（西乙E12＝東乙L14p5）と述べているところである。

4 福岡正博証人について

福岡正博証人は、「第I相試験の対象になるような非常に進行した非小細胞肺癌患者について、イレッサが著効を示したことは、正に驚くべき経験であった」（西丙E33＝西丙G59p9）と述べ、また、主尋問においても、スーパーレスポonderという言い方をして、イレッサの有効性の根拠の一つとして、こうした患者さんが存在していることを挙げている（西福岡正博証人主尋問調書＝東丙G57p35）。

しかしながら、このスーパーレスポonderという趣旨不明のことばを用いて、個々の著効例が医薬品の有効性の根拠となるという考え方自体、医薬品の評価をゆがめるものである。すなわち、既に上記2（3）で述べたとおり、個々の症例のみをみて、スーパーレスポonderがある、これをもってイレッサが効いていると判断することは、医師の観察バイアスに歪められた判断でしかないのである。

なお、未承認薬であるが、癌患者について著効例の症例報告がいまも後を絶たないものとしては、丸山ワクチンが有名である（西甲P79＝甲L143）。福岡証人は、丸山ワクチンについての著効例の存在をもってその有効性を評価することができないとした上で、イレッサの著効例なるものについては、「丸山ワクチンと一緒にしてもらっては困る」などと述べながらも、自らもまたイ

レッサについて全く同じことを述べていることにはたと気づいたのか、イレッサの著効例は、イレッサにとって有効性の根拠とはならず、「一つの情報」（西福岡正博証人反対尋問調書＝東丙G 5 7 p 4 7～4 8）に過ぎず、検証されていない、ということ認めざるを得なかった。

このように、福岡証人のいうスーパーレスポonderなるものは、イレッサの有効性の根拠となり得ない。そればかりか、仰々しい言葉でもって、イレッサの医薬品評価を見誤らせる有害な「情報」として悪用され得るものといわざるを得ない。

5 坪井正博証人について

そして、坪井正博証人も、また、東日本訴訟における証人尋問において、実臨床での症例を取り上げて、イレッサの有効性評価を行おうとしている。

例えば、坪井証人は、イレッサについてのV-15-32試験の中の1症例を取り上げ、「延命効果があるであろうと強く推察しています」とする。

しかしながら、坪井証人は、反対尋問においては、主尋問で取り上げた経験症例3例は、いずれも学会や論文などで公表したものですらく（西丙49の1＝東坪井正博証人反対尋問調書p 5 5）、臨床医の目から見て個々の患者にとってベネフィットがあったと思われる症例を紹介したものにすぎず、この3症例をもって、「医薬品の承認の根拠としての有効性」の根拠とはならないことを自認している（西丙49の1＝東坪井正博証人反対尋問調書p 5 6）。

また、同時に、坪井証人は、V-15-32試験において取りあげた1症例では確かに効いているようにみえても、全体の結果として、延命効果（非劣性）は証明されていないということは、他方では短期で死亡されてしまう症例があるためであると考えられるからであることを認め（西丙49の1＝東坪井正博証人反対尋問調書p 5 7～5 8）ている。

このように、結局のところ、坪井証人は、医薬品の有効性評価にあたっては、臨床試験の対象患者総体において延命効果があるかどうかが重要なのであり、その一部の個別症例を取り上げてイレッサの有効性評価を行うのは適切ではなかったことを認めざるを得なかった。

他方で、坪井証人は、V-15-32試験での間質性肺炎発症例3例の経過については何ら答えなかった（西丙49の1＝東坪井正博証人反対尋問調書p58）だけでなく、イレッサ投与群の中から、短期間で死亡された症例や、重篤な副作用が発生した症例を取り上げて、「イレッサは延命効果どころか余命短縮効果がある」と結論づけることは妥当ではない。それは仮説である旨述べた上で、原告代理人の「1例だけ取り上げて、治療効果があるというのも仮説なんじゃないですか。」という質問には事実上答えず、ごまかす（西丙49の1＝東坪井正博証人反対尋問調書p59）など、イレッサについては、なりふりかまわず、公平性を欠く証言態度で本件訴訟に臨んでいることが明らかである。

6 まとめ

以上述べたとおり、個別症例の存在を医薬品の有効性の根拠としようとする被告会社の主張、及び光富証人、福岡証人、坪井証人の各証言は、いずれも症例報告である以上バイアスを免れ得ないものであり、失敗に終わっている。エビデンスのないものをいくら集積してもエビデンスにはならない。医薬品の有効性は、上記第1で述べたとおり、比較臨床試験の結果によってはじめて証明されるものであり、臨床試験結果が出ているにもかかわらず、個別症例をもって医薬品の有効性を論じようとする事自体、医薬品の適正な有効性評価を妨げるものでしかない。

なお、これら被告側証人のなした症例報告は、いずれもイレッサ承認後のものであり、イレッサの承認時までには審査資料とされていたものではないことに

についても念のため付言しておく。

第3 奏効率による延命効果の予測の問題性

1 序論

(1) 被告らは、承認時におけるイレッサの有効性の最大の根拠として、I DEAL-1の結果、特に日本人群における奏効率（正確には「反応率」）の存在を挙げている。

しかし、既に述べたように、奏効率は、延命効果のような真の臨床的ベネフィットを計測するものではなく、延命効果の代替エンドポイントに過ぎない。更に言えば、一定程度の奏効率が確認されたとしても、これをもって薬剤の延命効果が直ちに推測できるというものでは全くなく、有効性の見込み判断においては極めて慎重な態度が求められる。

被告側申請の西條証人は、自身が執筆者として名を連ねる論文において、「奏効率を根拠に加速承認を行うことが妥当であるかどうかについて検証されていない」と指摘し、「医薬品行政には、…リスクやハームを上回るだけのベネフィットがない新薬が出回ることを防ぐことが求められているわけだが、現在の日本の制度がそれにも合ったものであるとは言い難い。」として、批判的な見解を述べている（西甲H18＝東甲G55 p586）。ここに述べられるような、奏効率を根拠とした第Ⅱ相承認の問題性については、論文執筆時（2004年）のみならず、イレッサ承認当時においても同様であり、奏効率のみに依拠して安易に有効性を見込みを肯定するようなことは許されない状況にあったものである。

(2) 本項においては、以下の観点から、奏効率による抗がん剤評価の問題

性について整理し、奏効率の高さをもって直ちに延命効果が予測されることが誤りであることを明らかにする。

- ① まず、高い奏効率を示しても延命効果に繋がらないという問題性はイレッサの承認、販売開始以前に明らかとなっており、このことは承認時のイレッサの評価にあたって大前提とされなければならないということである（後述「2」参照）。
- ② 次に、このような問題性は、「奏効率」の判定方法が第Ⅱ相試験段階におけるスクリーニングの目的に沿うように定義されていたことによる当然の帰結であるということである。この点からも、抗がん剤の第Ⅱ相試験結果について、第Ⅲ相試験へと進むスクリーニングの意義を超え、市販承認の前提として延命効果を予測しようとする際には慎重さが求められる（後述「3」参照）。
- ③ この点について、被告側申請の証人らは、奏効率と生存期間中央値に統計学的な相関関係があることをもって、延命予測の論拠とする。しかし、そもそも、このような相関すらなければ奏効率は代替エンドポイントたり得ないのであって、相関があるからと言って、奏効率が延命効果を十分に予測させるかどうかを合理的に明らかにしたことはない（後述「4」参照）。
- ④ 前記③について更に敷衍すれば、被告らが根拠として挙げる福岡論文（西丙E 3 4の5＝東丙G 6 0の5）、西條論文（西乙H 3 8の2＝東乙F 1 1の2）を見てみると、そもそも、相関関係の存在についてすら、極めて不十分な確認にとどまっているものであるから（後述「5」参照）、奏効率と生存期間中央値との相関に関する被告らの主張は、それ自体が根拠を欠くものである。
- ⑤ そして、奏効率が精度の低い延命予測しか提供できないことは、抗がん剤承認の全体の制度の枠組みにおいても、当然の前提とされている（後述

「6」参照)。

2 奏効率が延命につながらない実例がイレッサ承認前に報告されていたこと

(1) そもそも、第Ⅱ相試験において一定の奏効率が観測されていたのに、第Ⅲ相試験で延命効果の検証に失敗した例は枚挙にいとまがない。奏効率が延命につながらない点については、既にイレッサ承認以前の知見となっていた。

(2) この点に関しては、被告側証人である西條医師自ら、「従来の抗がん剤で、第Ⅱ相試験での奏効率の大小では第Ⅲ相試験での生存のベネフィットを予測できないことが指摘されてい」たことを認める（西乙H38の2＝東乙F11の2 p582（西乙E20＝西條証人反対尋問調書 p65で確認））。福岡証人も、多くの臨床試験の報告では、奏効率は新抗がん剤群で高かったが、生存がないという結果が存しており、非小細胞肺癌に対する治療評価は奏効率では不適切で、生存率、QOLのような真のエンドポイントの改善が重要と断じている。（西甲H10＝東甲F56（西福岡証人反対尋問調書＝東丙G58 p36で確認））。この文献においては、表4（p86）、表5（p87）において、この当時報告されていた試験結果について、一般的に紹介されている。ここに挙げられた合計12の比較試験のうち、実に5つの試験において、より高い奏効率が見られた群において、生存期間中央値（被験群の半数が死亡するまでの期間を指す。「MST」との略称も用いられる。）が短くなるという「ねじれ現象」が見られている。下の表は、「ねじれ現象」を示した5試験を抜粋したものである。このうち、Giacconeの試験においては、18%もの差で高い奏効率を示した群において、延命とは正反対の結果となる、生存期間中央値の短縮が見られた点で注目される。

表	報告者	レジメン	症例数	奏効率	奏効率の差	生存期間中央値
4	Gatzemeier	Cisplatin/Cisplatin+Paclitaxel	206/202	17/26	9 %	8.6/8.1m
5	Giaccone	CDDP+PCT/CDDP+VM16	166/166	47/29	18 %	9.4/9.7m
5	Belani	CBDCA+PCT/CDDP+VP16	190/179	22/14	8 %	7.7/8.2m
5	Crino	CDDP+GEM/CDDP+MMC+IFO	154/152	40/28	12 %	35/38wk
5	Niho	CDDP+CPT/CDDP+VDS	100/103	29/22	7 %	45.4/49.6wk

この外にも、2000年の小細胞肺癌の研究において、既に、「第Ⅱ相試験で有望と思われた試験レジメンが、生存期間につながることはまれにしかないことが明らかとなった」ことが指摘されていた（西甲H27＝東甲F53，チェン論文（西乙E20＝西條証人反対尋問調書p65で確認））。

また、西條証人も、この文献を引用して、「奏効率が生存のベネフィットを示唆する適切な代理エンドポイントであるかという点には議論の余地がある。」と述べているのである（西甲H18＝東甲G55 p582左段）。

なお、西條証人は、このチェン論文は小細胞肺癌を対照とするもので、非小細胞肺癌には当てはまらないかのように証言するが（西乙E20＝西條証人反対尋問調書p65以下），自身の論文においてそのような限定を施していないことは明らかで（西甲H18＝東甲G55），かかる反論は、自らの立場を場当たりの変遷させた供述に過ぎない。

(3) このように、既に報告されていた実例から、奏効率が抗がん剤の有効性である延命効果を予測させる代替エンドポイントとして必ずしも妥当ではないということは、イレッサ承認以前において既に明らかになっていたことであった。

3 奏効率の判定基準による問題性

(1) 奏効率の判定基準に内在する問題

一定の奏効率が確認されても、それが延命につながらないという問題の理由としては、予後因子バイアスやタイムバイアスといったバイアスの存在も指摘されるが、この点は後述することとして、ここでは、まずもって、奏効率の判定基準そのものに内在する問題があることを指摘する。

I D E A L 試験で採用された腫瘍縮小効果判定基準は、修正 WHO 基準と呼ばれるものであって、部分反応（P R）に該当するには、計測腫瘍の 5 0 %以上縮小が 4 週間以上継続することとされていた（西丙 C 1 = 東丙 D 1 申請概要 p 4 6 0, 西乙 E 1 8 = 東乙 L 1 0 西條意見書 p 1 0 ~ 1 1）。なお、現在は RECIST 基準（西甲 G 7 = 東甲 L 4 7）が広く用いられているところ、その実質において WHO 基準と大きく変わるものではない。

しかし、この基準は、延命効果を予測するための代替エンドポイントの基準としては緩やかに過ぎるのであって、全く十分なものではないのである。

この点は、被告側申請の西條証人自身が、自らの論文で、「腫瘍の大きさが 5 0 %以上減少することは腫瘍細胞の最小殺傷のみを意味しており、全身腫瘍組織量において生物学的に意味のある減少を必ずしも示唆するものではないことである。加えて、相対的に短い反応期間（1 ヶ月以上）は、生存期間を有意に延長するには不十分であると思われる。」と指摘している（西乙 H 3 8 = 東乙 F 1 1 訳文 5）。

(2) 延命の予測精度の低い基準が採用された理由

抗がん剤の第 II 相試験において、このような緩やかな基準が採用されたのは、第 II 相試験は、単に、効果のない薬剤を早期にスクリーニングするという「探索」目的の試験として位置づけられるからである（西原告第 1 5 準備

書面 p 14 以下，西甲 F 46 = 東甲 F 80 訳 p 2 右段参照）。

この点，通常の承認過程においては，第Ⅱ相の後，市販承認される前の段階において，第Ⅲ相における延命効果の「検証」が行われることが前提となる。このため，第Ⅱ相段階では，第Ⅲ相に進むか否かのスクリーニングができれば十分である。この予備的な段階で延命予測の精度を上げることにのみ捕らわれて過大な労力や時間を割いてしまつては，制度全体として非効率・不経済となつてしまう。そこで，第Ⅱ相で用いられる「奏効率」という指標についても，延命予測の精度をある程度犠牲にしても，より効率よく第Ⅲ相試験へのスクリーニングを行えるように設定すべきとされ，専門家の議論を踏まえて，上述のような基準が策定されたものである。

このように，現在，一般に用いられている「奏効率」という指標は，延命効果は後の第Ⅲ相で確認するという大前提のもと，延命予測の精度を一定犠牲にしながら設定されたエンドポイントといふことができる。

(3) 小括

以上のような「奏効率」の性質からして，第Ⅱ相試験の奏効率の結果のみをもって真のベネフィットの証明を代替できたと誤信し，安易に有効性を肯定することは許されない。腫瘍縮小効果の判定に用いる基準そのものに内在する限界がある以上，ある試験で一定の奏効率が認められたとしても，その結果から延命効果を予測するには慎重な検討が必要なのである。

4 相関のみでは，代替エンドポイントの「妥当性」確認には不十分であること

(1) 序論

ところが，本訴訟において，被告らは，奏効率と生存期間中央値に単純な相関があることをもって，奏効率が，延命効果を予測するのに有益な代替エンドポイントであることの論拠とし，被告側申請にかかる西條・福岡各証人

もまたそれに沿う証言をしている。

しかし、かかる主張及び証言は、以下の点において完全な誤りである。

- ① まず、被告らが言う統計学的な相関の点をもって、奏効率が延命効果の代替エンドポイントして成立することには全くなならない。そのような統計学的相関すら認められなければ、もはや代替エンドポイントたりえないという意味では必要条件ではあっても、相関さえあれば十分に代替エンドポイントたりうるという十分条件ではないのである。被告らの主張は、あえてこの点を混同している。
- ② また、被告側証人らが指摘する統計学的相関の根拠とする論文には、相関の確認の限度においてさえも問題があり、そもそも、かかる相関があることの根拠自体に問題があるといえる。

本項においては、上記①の点について述べ、上記②については項を変えて後記5項において整理することとする。

(2) 相関のみでは代替指標としての「妥当性」は認められないこと

ア 代替指標としての「妥当性」(バリデーション)について

一般に代替エンドポイントの候補となる指標について、真のベネフィット予測の精度の存否を客観的に確認する統計的手法は、この分野では一般に「バリデーション」、「妥当性研究」などと呼ばれている。先に述べたように、真のエンドポイントを推測させる代替指標として成立するには、最低限の必要条件を満たしていることは当然として、更に十分条件も満たさねばならない。この点については、山本精一郎氏も「サロゲートエンドポイントを用いるためには、それが上に挙げたような統計的な要件、臨床的な要件を満たしていることを証明してから用いる必要がある。これを妥当性研究と呼ぶ。」とする(西甲F30=東甲G56 p1212右段)。本書面においては、上記した点から、かかる両条件を合わせて、代替指標

としての「妥当性」と表記して、以下論ずる。

代替指標の「妥当性」の確認に関しては、いくつかの手法が提案されている。しかし、そのいずれにおいても相関関係の確認のみで足りる、とするものはない。例えば、代替指標の「妥当性」確認の基準として、最もよく知られて繰り返し引用されるものが Prentice (プレнтаイス) 氏が提案するものである。旧ガイドラインにも引用される生物統計学者 Fleming (フレミング) は、この基準について、

「… (1) 生物マーカーが、臨床エンドポイントと相関関係を有すること、(2) そのマーカーが、介入の臨床効能エンドポイントに対する最終的効果を十分に捉えられることである。条件 (1) だけで代替エンドポイントの正当性を実証するのに十分だと勘違いしている者も多いが、条件 (2) のほうが充足しにくく、ずっと実証が困難なのである。(西甲F47=東甲G98訳文p4。なお、この訳文においては、原文中の「validation (バリデーション)」について「正当性の裏付け」という訳語を当てている。)。」

と整理し、別の文献においても、「代替エンドポイントの検証にあたって、この相関関係は、必要条件ではあるが、十分条件ではない。」と繰り返し強調している(西甲H60=東甲G106訳文p1右段)。

この点に関しては、RECIST 基準の筆頭執筆者である Therasse (テラッセ) 氏による文献(西甲F46=東甲F80訳文p5右段)を初めとして、各種学術文献においても繰り返し指摘が行われている。また、外国文献のみならず、上述の山本精一郎氏による日本文献でも、「単に相関があるだけでは、この(挿入: サロゲートエンドポイントを用いる必要)条件を満たさない。」と明記されている(西甲F30=東甲G56p1212右段)。

このように、相関関係は、「妥当性」の必要条件とはなるものの、相関

のみでは十分条件たりえない点については、明らかに、各種専門文献に一致する一般的な知見となっていたものである。

イ 西條，福岡証人らに共通して見られる混同

他方，西條証人も，意見書において「実証 (validate)」という，同じ言葉を用いているものの（西乙E 1 8 = 東乙L 1 0 p 1 0），その引用する文献（西乙H 3 8 の 2 = 東乙F 1 1 の 2）の内容を見ると，奏効率と生存期間中央値との間の単なる相関の検討以上のものを含むものではない。これは，代替エンドポイントの「妥当性」の確認という文脈においては，完全な誤用と評価しうるものである。（福岡証人も，相関のみを検討する西丙E 3 4 の 5 = 東丙G 6 0 の 5 に依拠して，奏効率からの延命予測を行っている点で，同様の誤りを犯しているといえることができる。）。

なお，西條証人が意見書において，

「なお，近年は，…腫瘍縮小効果と延命効果が必ずしも相関するとは限らないという考え方も見られる。…イレッサの承認当時においては，がんの専門医のコンセンサスは…腫瘍縮小効果と延命効果は相関するというものであり，…承認当時のかかるコンセンサスが合理的な根拠に基づいていたことを否定するのは，明らかに適切ではない。（西乙E18 = 東乙L10p13）」

と述べている箇所に関しては，反対尋問において，西條証人が根拠として挙げる自らの論文で確認したのは，奏効率と延命効果の相関などではなく，単に，奏効率と生存期間中央値の相関に過ぎないことを認めた（西乙E 2 0 = 西條証人反対尋問調書 p 6 4）。この証言からも，西條証人において概念の整理が不十分なまま混同していることが垣間見える。

(3) 相関の確認のみでは不十分である理由～タイムバイアス，予後因子バ

イアスについて

ア はじめに

このように、奏効率と生存期間中央値の間に相関関係が存在したとしても、奏効率が必ずしも延命効果の代替とはならない、即ち、相関関係の存在のみで代替指標としての「妥当性」を有することにならないということは、既にイレッサ承認以前に確立した知見というべきであった。

その実質的な理由として、奏効率判定における予後因子バイアス、タイムバイアスというバイアスの影響が挙げられる。この点は、西條・福岡両証人も認めるとおりである（西乙E20＝西條証人反対尋問調書p62，西福岡証人反対尋問調書＝東丙G58p34）。

イ 予後因子バイアス

ここで、予後因子バイアスとは、一般には、各患者の持つ予後因子（全身状態（PS）、喫煙歴、治療組入期間等）を予め均質に各被験群に振り分けること（無作為化）をしないで、群を単純比較すると、各群の予後因子の差異の影響により見かけ上の違いがもたらされてしまうことを指す。そして、奏効群と非奏効群を比較する場合などには、予後因子と密接に関連する結果から群の振り分けを行うことにより、各群の予後因子に偏りが生じてしまい、見かけ上の薬剤の効果があるかのように観察されてしまう。さらに、複数の試験結果を対象とした相関の確認についても、上にも引用した Buyse の文献にも説明されるように、

「ある研究で、ほとんどの被験者の疾病が非常に進行して全身状態が不良であれば、低い奏効率と短い生存が観察されると予想される。反対に、ほとんどの被験者の限局期にあり全身状態が良好であった試験では、高い奏効率と長い生存が観察されると予想できる…（西甲H61＝東甲F90p3～4）」

ため、生存と奏効の双方に作用する予後因子の影響により、薬剤の効果の有無に関係なく生存と奏効の相関関係が得られてしまうことになるが、これも予後因子バイアスの影響として整理することができる。

この点に関して、西條証人も、反対尋問において、ある特定の被験者群における生存期間中央値は、薬剤の活性とは無関係に、状態のいい患者が集まれば長くなり、状態の悪い患者が集まれば短くなる可能性を認める。そして、予後的に良好な患者のみのデータ蓄積により、その薬剤活性とは無関係に長期の生存期間中央値となる場合がある、と明言している。（西乙H38の2＝東乙F11の2（西乙E20＝西條証人反対尋問調書p63において確認））。

ウ タイムバイアス

次に、タイムバイアスとは、「奏効者に関しては、“奏効（response）”が観察されるまで生存していることが必要なので、治療による延命効果のあるなしに関係なく、非奏効者よりも生存時間は長くなるというバイアス」のことをさす（西甲F33＝東甲F54 p16（西福岡証人反対尋問調書＝東丙G58 p34））。この場合も、予後因子バイアスと同様、薬剤の効果に関係なく生存と奏効の相関関係が得られてしまう（なお、福岡証人は、多数の試験を対象とした解析についてはあたかもバイアスの影響を受けないかのように証言しているが（西福岡証人反対尋問調書＝東丙G58 p35以下）、その根拠が明示されることなく、西條証人の見解とも相違している点で、信用性は皆無である。）。

当然のことながら、このような相関をどれだけ精密に確認したところで、薬剤の効果から独立したものである可能性がある以上、奏効率の代替エンドポイントとしての「妥当性」の根拠にはならない。すなわち、ある薬剤投与群において一定の奏効率が得られても、論理的に、延命効果が予測さ

れることにはならないのである。

(4) Buyse 論文による相関分析の批判

ア 相関のみに依拠して代替エンドポイントを評価する手法の問題点

生物統計学者 Buyse の論文（西甲 H 6 1 = 東甲 F 9 0）は、西條氏の分析（西乙 H 3 8 の 2 = 東乙 F 1 1 の 2）、福岡氏の分析（西丙 E 3 4 の 5 = 東丙 G 6 0 の 5）のような、奏効率と生存期間中央値の相関についての分析を正面から批判する内容となっている。

ここでは、

「複数の試験がある場合、腫瘍縮小と生存の関係の評価方法として直観的に訴えやすい（が、重大な誤解を招きやすい）方法は、各研究の奏効率を、（例えば生存期間中央値といった）生存の要約データと対比させてプロットし、これらのポイントを通じて回帰線に適合させるやり方である。…回帰分析は、これらの疾病の中で行なわれた臨床試験で報告された奏効率と生存期間中央値に基づいて行われた。有意な回帰スロープは、生存の延長は奏効率の向上と相関していることを意味すると解釈された。」

と福岡・西條各証人が行ったものと同種の解析の存在を概観したうえで、

「このアプローチの背後にはいくつかの重大な統計学的問題がある。」

と断じている（西甲 H 6 1 = 東甲 F 9 0 訳文 p 2）。その理由として、この後の記述において、上述した予後因子による問題の外、群としてのデータを用いた解析において、当該群を構成する個別患者における個体差が捨象されてしまう問題、といった点を指摘する。

イ Buyse 論文の位置づけと、証人らの証言の信用性が低いこと

この Buyse 論文（西甲 H 6 1 = 東甲 F 9 0）は、実は、西條証人の上記文献（西乙 H 3 8 の 1 = 東乙 F 1 1 の 1）を、同一の掲載誌上において評釈する Gelmon 氏の論説（西乙 H 3 8 の 2 の p 1 の目次の左上参照）「主要評価項目の明確なポイント： 肺がんにおける第Ⅱ相臨床試験」において、「Buyse および Piedbois は、ある極めて重要な論文において、…このような調査の限界を注意深く概説した」ものとして引用されているものである（西甲 H 6 2 = 東甲 F 9 1 訳文 p 2 左段）。したがって、この分野の研究としては良く知られた内容であり、専門家証人である以上、西條・福岡両証人も、一般的な知見として、当然知っていなくてはならないものである。特に、Gelmon 氏に自らの論述を論評されている西條証人においては、この文献を確認すべきは当然である。

にも関わらず、本件訴訟における両証人は、Buyse論文が問題提起する上記の観点を全く無視するものであり、その信用性は低いと言わざるを得ない。

（５） 小括

以上より、福岡、西條両証人が、奏効率による延命予測の根拠として、相関の存在のみを挙げている点については、この主張の当否の検討を待つまでもなく、主張自体において失当で、不完全なものであることに留意されるべきである。

5 被告らが根拠とする研究からは相関の実証すら不十分であること

以上より、西條・福岡各証人は、奏効率による延命予測の根拠として相関のみを挙げる点において、主張自体失当との評価を受けても仕方のないものである。さらに、両証人が挙げる２つの文献の内容を精査してみると（西條証人の東甲 F 1 1 の 2 = 西乙 H 3 8 の 2，福岡証人の西丙 E 3 4 の 5 = 東丙 G 6 0 の

5) , 単なる相関の確認の限度においてすら, 実証ありとは到底言えないようなものであった。上述のとおり, 相関は, 「妥当性」を論ずるにあたって必要条件として当然にその確認が必要となるものであるが, 被告らが論拠とする統計解析については, この限度においてさえも十分な根拠がない, ということになる。以下, 詳説する。

(1) 福岡論文・西丙E 3 4 の 5 =東丙G 6 0 の 5 について

まず, 福岡証人が執筆者の一人となっている丙E 3 4 の 5 =東丙G 6 0 の 5 について検討する。

この文献における統計解析の対象とされた各数値は, 各文献中に報告された数値(要約データ)をそのまま用いたものである。各試験, 各患者について, 背景因子の調整, 症例登録数による重み付けはされていない。例えば, 表 1, 2 (丙E 3 4 の 5 の 2, p 8 ~ 9) の症例数を見ても,

- ・わずか 1 7 症例の群 (表 1 ・下から 1 7 行目) ,
- ・ 2 8 8 症例の群 (表 1 ・下から 5 行目) ,
- ・ 4 8 8 症例もの群 (表 2 ・下から 4 行目) が,

それぞれ, 対等な 1 つの群としてしか評価されていない点だけを見ても, 果たしてまともな結論が導かれるものなのか, 大きな疑問を抱かずにはいられない。

さらに, この解析については,

「今回の調査には, 多くの不均質な臨床第 II 相試験が含まれている。これらの試験では, 症例登録数に大きな差があり, また患者背景の多くは不均質で, 試験前に受けた治療法の数にも大きな差がある。奏効の評価方法にも差がある。それによって誤った結論が導かれる可能性がある。」

と, 執筆者自身認めているものである (西丙E 3 4 の 5 =東丙G 6 0 の 5

p 28)。

このような精度の低いデータを対象として、相関に関する統計解析を行ったところで、到底、相関関係の存否を検証したなどとはいえないことは明らかである（西福岡証人反対尋問調書＝東丙G 58 p 32参照）。

(2) 西條論文（西乙H 38＝東乙F 11）について

次に、西條証人自身も執筆者であり、相関の論拠とするこの解析において、まず第1に着目すべきは、無作為化第Ⅱ相臨床試験では奏効率（訳文における「RR値」）が生存期間中央値値（生存期間中央値の値）と相関していないという事実である。これらの治験では、生存期間中央値は、奏効率（RR値）とは無関係に一定していることが示されている（西乙H 38の2＝東乙F 11の2 p 4左段）。この論文において、わざわざ、無作為化第Ⅱ相試験を、他の単アーム試験と区別して解析したのは、無作為化試験は各群間の各種背景因子の調整なども要することから、結果の信頼性がより高いと考えられたためである（p 5左段参照）。にもかかわらず、相関が示されなかったということは、仮に他の信頼性の低い試験の解析において相関が示されていたとしても、これが疑わしいということを示唆するものである。

第2に、本件論文中に「ある特定のRR値（奏効率）で広範囲の生存期間中央値を示すことは今後の課題を含んでいる。」（p 5左段）と明記されていることからわかるとおり、この解析において確認された相関の精度は極めて低いものである。これでは、ある一つの群でRR値の増加が見られたとしても、生存期間中央値の延長を合理的に予測することなどできない。例えば、この分布図（p 4・図1）を見ればわかるとおり、奏効率が10%の群と、20%の群があったとしても、10%の群の生存期間中央値が9ヶ月ほどあり、20%の群の生存期間中央値が6ヶ月ほどしかなく、実際には奏効率の増加が余命の短縮をもたらすといった極端な事態が生じる可能性ですら

も排斥できない。

当然のことであるが、「妥当性」の必要条件としての相関関係の確認にあたっては、単に弱い相関の存在があるだけ（一般にはp値0.05以下か否かで判定する。）では足りない。ばらつきも少なく、また、奏効率の増加に伴って大きな生存期間中央値の増加が見込めること（相関直線の傾きの大きさを判定）が必須となる。したがって、ここに示されるような相関に過ぎなければ、「妥当性」の前提となるべき相関の証明としては、極めて不十分なものとなる。

西乙H38の2＝東甲F11の2の記述を見ても、「RR値と生存期間中央値間に有意な相関が認められた」（p4左段，p5左段）としながらも、「RR値（奏効率）と生存期間中央値間に乖離が見られる。」として、相関について否定的な評価を前提とした論述を行っている。さらに、この論文の「考察」においても同様に、相関に関する否定的な評価を前提に、その原因について詳細に論述する（p5左段以下）。具体的には、各試験における測定の不正確さ、選択バイアスの問題、上述した奏効率測定の限界、などといった点が指摘されているもので、これらの点からしても、この論文において確かめられたとされる相関を過大評価できないことは、同論文自体が自認しているものである。

以上のとおり、この論文においては奏効率と生存期間中央値との相関が確かめられたとされるものの、奏効率が延命効果に対する代替指標としての「妥当性」を有することなど何ら認められず、そればかりか、相関についてすら過大評価できないものである。これらのことは、この論文自体が認めるものである。

6 II相承認の制度設計と、奏効率の位置づけ

抗がん剤開発、承認制度の枠組みについても、奏効率による延命予測に限界

があることを当然の前提としており，安易に有効性を肯定することを未然に防ぐように考えられていたものである。仮に，西條，福岡両証人の述べるような「延命の蓋然性」が高い精度で認められるというのであれば，第Ⅱ相承認制度を設計するにあたり，莫大な費用，時間，医療機関の負担を要する第Ⅲ相試験を省略できる制度設計もあり得たはずである。しかし，実際にはそのように考えられたことは全くなかった。

(1) 「臨床試験計画（プロトコール）の作成と実施，並びに結果の統計解析とその評価について」（西甲D34＝東甲H22）の記載

旧ガイドラインが作成される2年前，厚生省がん研究助成金指定研究「固形癌の集学的治療との研究」班により，「臨床試験計画（プロトコール）の作成と実施，並びに結果の統計解析とその評価について」と題される論文（西甲D34＝東甲H22）がまとめられた。執筆者を見ると，西條証人や下山氏を初め，後に旧ガイドライン作成メンバーとなった者も数名含まれている。

この文書を見ると，との記載で，抗がん剤の臨床試験においても，第Ⅱ相試験が「探索」に位置づけられることを明確にしたうえで（p2511右段），

「小規模な第Ⅱ相試験的な臨床試験で，既存の治療法よりずば抜けて良い治療効果を示すことが明らかになった場合は，中規模の第Ⅱ相試験は必ずしも必要ではない。むしろ，生存効果が得られるかどうかを目的にした第Ⅲ相試験的な無作為化比較試験に進む方がよい。」

とする。この記載から，奏効率からすれば「ずば抜けて良い治療効果」で示した薬剤でさえも，なお，第Ⅲ相試験の省略はできないことがわかる。

このように，抗がん剤承認においては，いかに高い奏効率が見られたとしてもそれだけで市販承認の根拠としては足りず，単に第Ⅲ相試験に進めてよ

いというだけ、という位置づけとされてきたことがわかる。したがって、イレッサの I D E A L 試験程度の奏効率の存在をもって、有効性が確実視できたなどという被告らの主張は失当である。

(2) 新医薬品課審査官（当時）による旧ガイドラインの解説

当時の厚生省薬務局新医薬品課審査官である川原章氏は、平成3年3月の「新医薬品研究開発フォーラム」という、旧ガイドラインの内容の解説（西甲D35=東甲H19）において、

「第Ⅲ相に関して特に申せば、これについては、かならず延命効果を見る、ということである。腫瘍は小さくなったけれども、本当に命が延びたのかどうかという問題である。…外国の論文等の中には、腫瘍は縮小したけれども、延命については疑問な結果が出たといった報告も見られている（p47）。」

という腫瘍縮小効果による延命予測の限界を確認する。このため、旧ガイドラインによっても、原則的には

「第Ⅲ相は申請の時点までにできていれば一番よろしい。（p48）」と述べる。このように、旧ガイドラインにおいても、他の医薬品と同様、最終的には、第Ⅲ相試験において真のエンドポイントである延命効果で検証したうえで、承認が行われるべきという原則が維持されていたことは明らかである。

そのうえで、例外的に、Ⅱ相承認が要求される例として、

「腫瘍縮小効果が非常に高い薬、たとえば従来PRまで含めて20%程度までしかなかった癌腫に70%といったものが出てきたような場合に…そういった効果の高い薬もあるということで第Ⅱ相までで申請は認める。（p48）」

と述べている。この記載から、「非常に」高い奏効率がなければⅡ相承認を

与えない趣旨であることも十分に読み取れるものである。そうでなくとも、少なくとも既存の薬剤と同程度の奏効率しか示せていないような場面においてまで、あえてⅡ相承認を与えるようなことは許容しない趣旨であることが明らかである。

7 小括

以上、本項においては、高い奏効率が生存期間中央値の延長には繋がらない実例が、承認当時、既に数多く存在していたこと（前述「2」）、そもそも「奏効率」の判定方法は、第Ⅱ相段階におけるスクリーニング目的に沿うように定義されていて、延命予測の精度を最重視するような設定とはなっていないこと（前述「3」）、被告ら申請証人の論拠は、科学的な根拠を欠くものであること（前述「4」「5」）について論述を行った。これらの点からも明らかとなり、イレッサの評価にあたり、有効性の根拠として奏効率を用いるときに、これを過大評価してはならないものである。

第4 IDEAL 1, 2の奏効率から有効性を推測することの誤り

1 序論

イレッサ承認にあたっては、IDEAL-1, 2各試験で確認された奏効率を根拠に有効性が肯定された。しかし、一般に抗がん剤の有効性を推定するためには、最低でも、20%の奏効率の確認は必要と考えられているところ（後述「2」）、IDEAL-1, 2各試験を全体として見れば、このような最低限度の奏効率ですら確認されていなかったものである（後述「3」）。

被告らはドセタキセル試験と比較することにより、イレッサ奏効率を肯定的に評価すべきと主張するが、このような評価手法には歴史的対照による問題が

あり、実際にも単純比較を許さないような背景因子の違いが明らかとなっていたから（後述「4」）、イレッサ有効性を根拠づけることはできない。

また、他の既承認薬によっても、イレッサ承認当時、既に、IDEAL-1, 2各試験に見られたものを上回る奏効率が確認されていたものであるから（後述「5」）、被告らの前提は、イレッサの有効性を過大に評価するもので失当である。

2 抗がん剤の奏効率の確認に用いられる一般的基準

(1) 20%の奏効率を必要とする西條証言の内容

主尋問で、西條証人は、セカンドラインの抗がん剤の奏効率について、「20%以上であれば」ようやく「ある程度効く抗癌剤というところに分類」に達すると述べる（西乙E19＝西條証人主尋問調書p15～16、西乙E20＝西條証人主尋問調書p68）。さらに、平成14年当時に抗がん剤として市場が求めていた薬剤を問われ、

「先ほど、奏効率が30%と言いましたけれども、これは不十分ですから、これを50%くらいに持って行けるようなサイトトキシク（細胞毒性）ドラッグです。（西乙E19＝西條証人主尋問調書p16）」

と述べている。ここから、まず、30%もの奏効率であっても不十分と断じており、さらにこの奏効率を20%も引き上げて50%に達することを目標としていることが読み取れる。したがって、西條証人も、20%という数値は効果の存在を予測しうる最低水準と認識していることが窺える。

この点に関しては、西條証人自身の論文で、統計解析におけるカットオフポイントを奏効率20%に設定した根拠の説明において、

「20%以上の奏効率を有する薬剤が非小細胞肺癌に対して活性あり、と考えられているからである。（西乙H38の2＝東乙F11の2訳文p3左段）」

と述べている。ここで、「活性」とは「生物学的活性」とも呼ばれるもので、一般に、臨床的な真のベネフィットの確認の前段階として、客観的に抗腫瘍効果など、薬剤の作用が外形的に確認しうるかどうかを問題とする概念である。一般的には、活性すら見られない被験薬であれば、真のベネフィットは到底期待し得ないものと評価され、開発を中止すべきとされていた。

なお、西條証人は反対尋問になると、上記主尋問内容の確認の質問にも関わらず、この点を曖昧にした。その上で、IDEAL 1 結果である15.5%でも十分などとも反論もしている（西乙E 20＝西條証人反対尋問調書p 69）。しかし、これらの証言は、上記の証言や文献において自ら設定した前提を、都合良く場当たりの的に変遷させるもので、信用できない。

(2) 旧ガイドラインにおける期待有効率20%の記載

西條証人の上記評価は、旧ガイドライン（西乙D 7＝東乙H 7）における期待有効率20%の記載とも一貫する（西乙E 19＝東西條主尋問調書p 24でも確認）。

具体的には、旧ガイドラインにおいては、前期第Ⅱ相試験における期待有効率の設定に関して、一般的に20%以上を目標とすることとし、「この期待有効率以上の効果がなければ有用な抗悪性腫瘍薬としては認められないことになる。」との前提に立つ（西乙D 7＝東乙H 7・7枚目）。

旧ガイドライン作成班の一員でもある下山正徳氏による旧ガイドライン解説においても、「一般的に有効率は20%以上を目標にする。どのくらい効く薬を開発するかが念頭になれば、新薬の開発はできないので、その目標を設定しているわけである。期待する有効率を期待有効率といい、固形癌では一般的に20%以上が目標となる…。この期待有効率とその rejection error（すなわちβエラー）を設定することにより期待有効率を示さない薬、すなわち無効な薬を早期に判別する。」とされるべきことが明記されている。

(西甲D 1 5 = 東甲H 1 0 p 1 2 6 左段以下)

(3) 旧ガイドラインが、期待有効率20%を示した趣旨

もともと、旧ガイドラインが、一般的には、期待有効率を20%に設定すべきとしたのは、20%の奏効率が确实視される薬剤には当然にⅡ相承認を与えて良いとする趣旨などと解すべきではない。

なぜなら、上述のとおり旧ガイドラインは、原則形態としては第Ⅲ相試験まで終えて承認を与えるべきことを想定していて、20%の奏効率は、この原則形態において、第Ⅲ相に進むための最低基準を示したものに過ぎない

(この点については、第Ⅲ相終了後の承認を前提とした通常形態の治験の進め方を論述した厚生省作業班による文献(西甲D 3 4 = 東甲H 2 2)においても、同じ期待有効率が設定されていることからわかる。)。そして、単に第Ⅲ相に進めても良いかという場面においては、未だ市販の前段階にとどまるため、有効性の判定をやや緩和することも許容されるのに対し、第Ⅱ相承認を与える場面においては、市販により多数の一般患者に流通する以上、より厳格な基準を用いるべきだからである。

この点、第Ⅱ相のスクリーニング判定の際には、なるべく、有効なものを誤って振り落としたいくないという要請(βエラーの最小化をより重視する。)があり、これに沿うように期待有効率も想定される。20%という奏効率は、確実に効果が期待できるようなものではなく、西條証人も「まあ効くと判定するようなパーセンテージ」(西乙E 1 9 = 西條証人主尋問調書25)といった程度の、あいまいで低水準なものではあるが、第Ⅲ相への移行の判定という目的の限りにおいては、これで十分と考えられたものである。

他方、Ⅱ相承認に求められる水準を考えてみると、上述の川原氏の論述(西甲D 3 5 = 東甲H 1 9)を見る限り、かろうじて第Ⅲ相に進ませてやってもいいか、といった程度の被験薬全てにつき、医薬品承認を与える趣旨と

は到底解されない。したがって、旧ガイドラインにおける期待有効率20%という設定は、単に開発中止とならないために最低限クリアすべき水準である、と理解すべきものである。

(4) セカンドライン以降の治療薬としての評価

この外、西條証人は、イレッサの低い奏効率でも許容すべき論拠として、非小細胞肺癌、特にセカンドライン以降の治療全般において、総じて、抗がん剤の効果が小さく奏効率も低かったことを挙げる（西乙E19＝西條証人主尋問調書p25参照）。

しかし、この立論は、まずもって前提に誤りがある。この当時、他剤による非小細胞肺癌のセカンドライン患者を対象とした試験における奏効率の結果を見ると、イレッサ承認当時であっても、併用で、33%（西甲H48＝東甲F82）、40%（西甲H49＝東甲F83）、単剤でも、19%（西甲H50＝東甲84）、21%（西甲H51＝東甲F85）、5.2%CRを含む34%（甲H52＝東甲F86）などが報告されていた。

IDEAL1, 2の試験の計画にあたって、このような実態を反映してか、期待有効率を15～20%、閾値有効率5%として被験者数が設定されている。IDEAL-2が期待有効率を若干低めの15%に取っていることを除けば、旧ガイドラインに示された一般的な基準を正にそのまま採用しているものである。したがって、試験計画時においては、非小細胞肺癌のセカンドライン治療についても、通常の一般的な固形癌に求められる有効率の基準に沿うべきことが想定されていたことがわかる。

また、そもそも他剤の奏効率が軒並み低かったとしても、Ⅱ相承認のハードルを下げる論拠とはならない。なぜなら、奏効率は、真のベネフィットである延命効果を予測して初めて意味があるところ、他剤の奏効率が小さいからといって、論理的に、低い奏効率の被験薬によっても延命を期待できる確

率が高まることにはならないからである。

(5) 小括

以上より、一般に、抗がん剤に求められる最低限の生物学的活性として、約20%の奏効率が必要と考えられていたことがわかる。この点については、西條証人自身も認め、旧ガイドラインも同様の前提に立つものである。しかも、この水準は、第Ⅲ相試験に進むために最低限必要とされるものに過ぎず、市販承認に要求されるような、より確実な延命予測の水準には達しないものとして理解すべきものとなる。

このような理解を前提として、以下、イレッサのIDEAL試験において見られた奏効率について検討する。

3 IDEAL各試験におけるイレッサ奏効率の評価

(1) IDEAL-1, 2において見られた奏効率の概観

これらの試験で見られた奏効率をまとめたものが、下の表である（西丙C1＝東丙D1 p 470, 502）。IDEAL-1に関しては、申請資料概要に複数の表が掲載されているが、このうち、最も信頼性が高い数値は審査センター判定による表ト-73（西丙C1＝東丙D1 p 470）であり、下の表や、後の論述においても、これらの数字を前提とすることとする。

試験	用量	人種	奏効率	信頼区間
IDEAL-1	250mg	日本人	25.5%	14.3～39.6%
		日本人以外	5.8%	1.2～15.9%
		合計	15.5%	

	500mg	日本人	25.5%	14.3～39.6%
		日本人以外	9.3%	3.1～20.3%
		合計	16.3%	
IDEAL-2	250mg	日本人以外	11.8%	6.2～19.7%
	500mg		8.8%	4.3～15.5%

(2) IDEAL-1の各群全体としての評価

上述のとおり、一般に、抗がん剤に最低限求められる奏効率は20%とされているところ、イレッサのIDEAL-1, 2のうち、平均的に高い数値を示したIDEAL-1の奏効率でさえ、承認用量の250ミリ群で全体として見れば15.5%にとどまっていたものである。しかも、この15.5%という数値は日本人群の奏効率によって大幅に引き上げられていたものであるが、西條証人も認めるとおり、日本人群の奏効率は、単に背景因子の偏りを調整すれば、外国人群の低い数値並みとなっていた可能性もあるという（西乙E20＝西條証人反対尋問調書p71～72）。にもかかわらず、最低水準の20%もクリアできないというのであるから、この試験をもって、イレッサの奏効率を高いなどと評価できないことは明らかというべきである。

(3) IDEAL-2の各群全体としての評価

IDEAL-2についても、いずれの群においても奏効率は、わずか11.8%（250mg）、8.8%（500mg群）であり、一般に求められる20%奏効率の水準を大きく割り込む結果となっていた。したがって、この試験を前提としても、イレッサの奏効率を高いとは、到底評価しえないものであった。

(4) プロトコールに照らした評価

ア プロトコールの「解析方法」により有効性判定を行う必要性

以上、一般的な抗がん剤について設定される期待有効率20%を前提に、イレッサの奏効率の評価を行ってきた。もっとも、本来、イレッサの承認根拠となったIDEAL-1, 2の各試験の評価にあたり、まずもって、第一義的な評価基準として検討されるべきは、事前に、試験計画書（プロトコール）において設定されていた有効性の検定方法である。このように、プロトコールにおいて、事前に評価方法を明記しておくのは、治験の報告者が、後付けの、医学的にもっともらしい理由を付けて過大に有効性を報告することを防止するためである。つまり、ICH-E9にも示されており、事前の予想に反して芳しい結果が得られなかった場合に、バイアスの影響や、不当な動機からごまかしてしまうことのないように、客観的に妥当な基準を事前に明示しておくものがプロトコールなのである（西甲P15＝東甲H3 p23のV5.1, p28の5.5参照）。

そこで、原告代理人らにおいて、改めてイレッサのIDEAL1, 2の試験結果を検討したところ、実際には、各試験のプロトコールに照らして見れば、複数の群において有効性の検定に失敗し、イレッサの奏効率が高いなどとは、到底評価しえないものであった。

イ プロトコールの「解析方法」における有効性検定の方法

IDEAL各試験のプロトコールにおいては、以下のような方法で、有効性検定を行うことが、事前に明記されていた。

IDEAL1については、申請資料概要で、「奏効率の95%信頼区間の下限が5%を上回っていた場合、真の奏効率は5%以上であると結論づける。」（丙C1 p462・表ト-66）と記載されている。

同様に、IDEAL2についても、申請資料概要において、奏効率「5

%は他に有効な治療がない場合の実薬の許容される最小率として選択される。」ことを前提として、試験結果から得られた奏効率の信頼区間下限がその5%を下回るようであれば、有効性はないものと評価すべきとされていた（西丙C1＝東丙D1 p498「解析方法」）。

ウ 旧ガイドラインにおける閾値有効率との関係

このように、IDEAL各試験において、一般に求められる期待有効率20%を大幅に下回る、5%奏効率を基準としているのは、旧ガイドラインにも記載された一般的な閾値有効率の水準をもって、無効な薬剤の判定を行おうとしたためと考えられる。

閾値有効率とは、有効率の判定にあたって帰無仮説として設定される奏効率の値であるところ、下山正徳氏による旧ガイドライン解説によれば、

「近代的な方法では、期待有効率の他に、さらに無効な抗がん剤の判定に使う閾値有効率を適切に決める。この閾値有効率は5%以上とするのが通例である。」

とされる（西甲D15＝東甲H10 p126左段）。

この点、当然のことながら、あまりに奏効率が低い抗がん剤については、到底、延命効果など期待できないもので、有効性は否定される。そこで、第Ⅱ相試験の症例数の算出にあたっては、一定の奏効率、例えば5%という目安を設定し、これ以下であれば開発中止としてよい、という前提で行われる。ここでの目安となる奏効率（上述の5%）が、「閾値有効率」と呼ばれるものなのである。

これを更に説明すれば、次のとおりとなる。例えば、ある被験者群で得られた抗がん剤の奏効率について、その信頼区間の下限が5%を割ってしまったとすれば、確率論として、真実の奏効率が5%以下である可能性、すなわち効果のない抗がん剤である可能性が一定程度以上に残ることとな

る。統計学的に言えば、当該抗がん剤が無効な薬剤である可能性（帰無仮説）を排除（棄却）することをもって、当該抗がん剤の効果を証明しようとしたところ、その証明に失敗したということである。そのような証明ができない場合には、期待有効率以上の有効性を得る可能性はないとみなすことにする、という意味になる。この場合、市販など許されないのはもちろんのこと、その時点で開発を中止してよい、という結論になる。

なお、旧ガイドラインの閾値有効率の記述箇所において想定された場面は、より探索的な色合いの強い前期第Ⅱ相試験のものであり、その閾値は極めて緩やかな水準として設定されている。つまり、この水準は検定に失敗すれば開発中止と判定する、という意味での最低限度の基準なのであり、この検定に成功したとしても、延命効果の予測ができて市販承認が許容されるような帰結とはならないことについては十分に留意すべきである。

エ IDEAL各試験の結果

まず、IDEAL 1を見ると、250mg海外群における奏効率の信頼区間の下限（1.2%）においても、500mg海外群の同下限（3.1%）ともに、この5%を下回ってしまっている（申請資料概要にも同様の否定的な評価の記載がある（西丙C1＝東丙D1 p 487～488））。）。。

IDEAL 2についても、500mg群の奏効率の信頼区間下限（4.3%）がこの5%を下回ってしまっている。250mg群の信頼区間下限は6.2%であるが、これは、プロトコールで最低ラインとされた11例をわずかに1例上回ってかろうじてクリアしたという結果であった。

このように、IDEAL各試験においては、複数の群において、開発中止の結論を導くべきとされるような、低い水準における有効性の検定にすら失敗したという有様であった。後述のとおり、検定に成功した日本人群についても、背景因子の問題があった以上、全体としてイレッサの奏効率

が高いなどと評価することは許されないものである。当然のことながら、延命効果が精度高く推認できるかのような結論を導くことはできない。

(5) 日本人群の結果について

この点、被告らは、IDEAL 1の日本人群の奏効率によりイレッサの効果を強調するような主張をしている。

しかしながら、IDEAL 1試験においては、日本人群と日本人以外群との背景因子に著しい違いがあった。特に、試験結果に大きく影響する患者の全身状態（PS）については、日本人群において、全身状態が不良の患者（PS 2）の割合が著しく少ないという偏りがあった。被告会社自身、IDEAL 1での国内外の結果の差について最大の原因がこのPSの偏りにあったことを説明していた。また、この患者群の背景因子の偏りを調整すれば、日本人群の結果も、外国人群の低い数値に近づいていた可能性もあることは、被告側申請にかかる西條証人ですら認めていたのである（西乙E 20＝西條証人反対尋問調書p 71～72）。

このようなことを考えれば、IDEAL 1試験のうち、日本人群の結果のみを取り出し、患者背景の著しい偏りを無視して高い有効性が期待できるなどとすることは全くの誤りである。

(6) まとめ

以上のとおり、IDEAL 1, 2試験ともに、通常、抗がん剤に最低限要求される奏効率20%の水準を達成することができず、また、プロトコールに予め明記された検定にすら失敗してしまった群も続出したのであった。ここで留意すべきは、プロトコールに規定された検定の水準は、これに失敗すれば開発中止としてもよいような、極めて緩やかな有効性の検定であったということである。イレッサに関しては複数の群で、この緩やかな検定にすら

失敗したということであり，その有効性を過大評価することが許されないことは明らかである。

I D E A L 試験結果については，イレッサには臨床的に意味がありうる程度の腫瘍縮小効果は認められなかったもので，当然の帰結として，延命効果など期待すらできないものと判断すべきものであった。

4 ドセタキセル試験との比較が有する問題点（背景因子の問題）

(1) 被告らの主張

申請資料概要（西丙C 1 = 東丙D 1）において，被告会社は，約10%のI D E A L - 1の海外奏効率も含め「いずれの民族群においても単独両方で臨床的に意味のある奏効率が得られ」とし（西丙C 1 = 東丙D 1 p 5 0 9），審査報告書（西乙B 4 の1 = 東乙B 4 の1）でも約11.8%のI D E A L - 2の奏効率について同様の評価をしている。

このような評価の前提として，まず，セカンドライン非小細胞肺癌患者に対する延命効果を検証した Shepherd（シェファード）氏によるドセタキセル試験があり，I D E A Lの各群の奏効率がこの試験のドセタキセル奏効率7.1%（西丙H 2 2 の2 = 東甲F 2 0）を上回ったことが主要な根拠とされている（西丙C 1 = 東丙D 1 p 5 0 9，西乙E 2 0 = 西條証人反対尋問調書 p 7 2）。

(2) プロトコールには反映されていない議論であること

しかし，そもそも，真にこのような評価を行うべきであったならば，上述のとおり，当初のプロトコール作成段階において目標症例数の算出にあたり，期待有効率，閾値有効率を低く設定すべきものであるし，これに応じた有効性検定の基準が事前に設定されるべきである（西甲P 1 5 = 東甲H 3 p 2 3 V 5. 1，p 2 8 の5. 5。結果報告について p 3 3 参照。）。

にもかかわらず、IDEAL各試験の症例数設定の際には、何らこれに類する検討をしないまま、通常どおりの期待有効率（IDEAL1で20%、2で15%）と閾値有効率（IDEAL1, 2ともに5%）が採用されていた。そして、有効性の検定においても、5%の水準が採用された点については、上述のとおりである。これらの試験計画を見ると、イレッサの奏効率が低いとの試験結果が出てから、慌ててドセタキセル比較を持ち出した議論がなされるようになったという経緯が読み取れる。

このため、ドセタキセルとの比較の議論は、その信用性について、極めて慎重に吟味すべきこととなる。

(3) 歴史的対照の問題性

このような後付けのドセタキセル比較を持ち出した経緯に目をつぶったとしても、ドセタキセル試験との比較においてイレッサの有効性を基礎づける論法は、歴史的対照を用いる問題性、つまり過去の試験結果との単純比較は誤った結論を導く可能性があるという問題をはらんでいる（西甲G4=東甲L49p211, 西乙D7=東乙H7・8枚目（西福岡証人反対尋問調書=東丙G58p37で確認））。

西條証人も執筆者として名を連ねる厚生省研究班による文献（西甲D34）においては、端的に、「RCTを敬遠して、歴史的対照などの非無作為化臨床試験だけに依存すれば、臨床医学の科学性が崩れ去るのは自明である。」とまで述べられている。この点に関連して、福岡証人も、再主尋問において「ヒストリカルな比較」（歴史対照による比較と同義）については問題があり、最終的には、第Ⅲ相試験による確認が必要と認めた（西福岡証人反対尋問調書=東丙G58p80-81）。

(4) PS-2患者の割合と奏効率への影響

この外、ドセタキセル試験との比較論は、以下のようなP S（パフォーマンス・ステータス）による患者背景の相違の問題点もある。

まず前提として確認しておくべき事項として、I D E A L - 1 の国内群と海外群の間で見られた約20%もの奏効率の違いについてさえも、統計解析（多変量解析）の結果、民族差などは考慮することなく、単に、各群の間のP S、腺がんなどの患者背景因子の違いのみで説明しうるものとされている（申請資料概要（西丙C 1 = 東丙D 1 p 4 9 3 - 4 9 4）、審査報告書（西乙B 4 の1 = 東乙B 4 の1 p 3 9））。このうち、全身状態が不良なP Sが2の患者割合の違いを見ると、国内群は海外群の約2分の1（割合の差としては約8%）という違いがあった。

そこで、比較対象となったドセタキセル試験（西丙H 2 2 の2 = 東甲F 2 0）の患者の背景因子を検討してみると、I D E A L - 1 の日本人250ミリ群では、状態の悪いP S - 2の患者割合がシェファード試験の僅か4分の1程度しかなく、患者背景に大きな違いがあることが判明した（西丙H 2 2 の2 = 東甲F 2 0 p 5の表1、西丙C 1 = 東丙D 1 p 4 6 8の表ト - 7 0）。西條証人も認めるとおり、このような患者背景の違いを無視して、単純に結果を比較して論じることは妥当ではない（西乙E 2 0 = 西條証人反対尋問調書p 7 2）。

そもそも、Shepherd 試験については、原著論文中に、セカンドライン、すなわち

「プラチナ製剤で治療した非小細胞肺癌患者を対象にドセタキセル単剤療法を検討した7つの第II相試験が報告されている。これらの7試験では、…全体の奏効率は14~24%であった。（2頁）」

「我々が実施した試験の全体の奏効率7%は、どのドセタキセル第II相試験で報告された値よりも小さいが、…。このいくらか残念な結果は、より進行した状態の悪い癌患者を選択したことが原因かもしれな

い。(12 頁)」

と記載があるとおりに、患者背景の偏りによって、通常期待しうるよりも低い奏効率となっていた可能性が明記されている。したがって、一般的な奏効率の比較対象として適切であったのか、疑問が残るものである。

さらに、IDEAL-1 の海外群については、PS-2 割合でいけば 18.9%で、上記ドセタキセル試験のPS-2 割合25.5%を下回り、これだけを見れば本来、より高い奏効率が期待できたはずであった。にもかかわらず、承認用量250mg群における奏効率は5.8%であって、ドセタキセル試験の7.1%すらも下回る結果であった。(西丙C1=東丙D1 p 468 「表ト-70」)。

(5) 西條証言も患者背景の問題を認めていること

患者背景に関連して、IDEAL-1の日本人群の奏効率が高いとの評価について再確認すると、IDEAL-1の日本人以外の250mg群でのPS-2は18.9%(10例)、日本人は5.9%(3例)と3倍以上の違いがある(西丙C1=東丙D1 p 468)。したがって、西條証人も認めるとおりに、海外群並みの比率でPSの悪い患者をグループに入れた場合には、日本人の群でも、IDEAL-1の海外群のような結果に近づく可能性がある(西乙E20=西條証人反対尋問調書 p 71~72)。IDEAL-1の日本人群の奏効率は、偏った状態の良い患者群での結果といえ、イレッサが日本人に対しては高い腫瘍縮小効果があるなどと即断することはできないものであった。

5 他の既承認薬のセカンドライン患者に対する効果

このように、Shepherdによるドセタキセル試験(西丙H22の2=東甲F20)との対比においても、イレッサの奏効率は必ずしも高い奏効率を示せてい

ない。それでは、IDEAL-1試験の全体としての平均奏効率15.5%は、他の既存の承認薬との対比において、ずば抜けた奏効率を示すものであったといえるのであろうか。

他剤による非小細胞肺癌のセカンドライン患者を対象とした試験における奏効率の結果を見ると、イレッサ承認当時であっても、併用で、33%（西甲H48=東甲82）、40%（西甲H49=東甲83）、単剤でも、19%（西甲H50=東甲84）、21%（西甲H51=東甲85）、5.2%CRを含む34%（甲H52=東甲86）などが報告されていた。したがって、日本人群の背景因子の偏りで膨らんだ15.5%の数値を前提としてもなお、イレッサIDEAL-1の奏効率は、既に患者の選択肢となっていた既承認の薬剤（もしくはその組み合わせ）以上の効果を示すものとは評価しえないものであるし、非常に効果が高いとか、ずば抜けているなどとは到底言い得ないような、お粗末なものであった。

以上より、イレッサについて、非小細胞肺癌のセカンドライン患者に限定したとしても、他の既承認薬を超えるような奏効率があったとは評価しえない。これは、イレッサを第II相で承認したところで、奏効率で見ても、患者に対して既承認薬を超える利益をもたらすものではなかった。したがって、延命効果の検証はもちろんのこと、副作用の確認も不十分なまま、あえて承認を急ぐべき必要性があったとは評価できない。

6 小括

以上より、IDEAL-1, 2各試験で確認された奏効率は、全体として見れば、一般に、抗がん剤の開発継続に必要な最低基準とされる20%の奏効率（前述「2」）ですら下回るものであり（前述「3」）、また、ドセタキセルのShepherd試験との比較も背景因子の偏りにより適切な分析とはいえないもので（前述「4」）、イレッサ有効性を根拠づけることはできない。さらに、他

の既承認薬との比較においてもイレッサの奏効率が殊更、優れているともいえない（後述「5」）。

ゆえに、奏効率を基準としても、イレッサの有効性は高いとはいえない状況にあったものである。

第5 IDEAL各群の生存期間中央値による有効性の推測について

1 被告らの主張の概要

被告らは、以下のIDEAL-1, 2試験の生存期間中央値（生存期間中央値）が臨床的に有意義なものであると評価できるもので（西乙B4の3=東乙B4の3 p 54）、これがイレッサ承認の根拠の一つとなりうると主張する。

IDEAL-1 250mg群 7.6ヶ月

IDEAL-1 500mg群 7.9ヶ月

IDEAL-2 250mg群 6.5ヶ月

IDEAL-2 500mg群 5.9ヶ月

2 対照群のない試験における生存期間中央値の評価

しかし、このような対照群のない試験における生存期間中央値というものは、比較試験による延命効果の検証とは全く性質が違うものである点に注意すべきである。つまり、延命効果というものは、適切な患者割り付けによって患者背景因子を均等化したうえで、対照群との比較において、はじめて確認しうるものである。この対照群との比較、という性質は、「延命効果」というエンドポイントを論ずるにあたり、欠くことのできない要素なのである。

これに対し、適切な比較対照のない試験の被験者群における生存期間中央値というものは、その群の患者背景因子によって大きく変動しうるものであり、仮に、一見良好な数値が得られたとしても、これが、被験薬の効果によって得

られたものか、単に患者群の背景因子の偏りによってもたらされたものなのかを見分けることは不可能である。したがって、ある試験における生存期間中央値を、別個の試験や、一般的な臨床成績などと対比したところで、有効性の根拠としては全く信頼性がないと言っても過言ではない。

よって、上記の I D E A L 各試験の生存期間中央値をもって、イレッサの有効性を根拠づけることはできない。

3 生存期間中央値は副次的評価項目で過大評価してはならないこと

申請資料概要の試験計画から、I D E A L 各試験において、主要評価項目はあくまで奏効率であり、生存期間は、他の複数の副次的評価項目のうちの一つでしかないことが明らかである（西丙 C 1 = 東丙 D 1 p 4 6 2, p 4 9 8）。

したがって、平山証人も認めるとおり、承認当時、イレッサの有効性の主たる評価対象となるものはあくまで奏効率であり、生存期間中央値の数値は副次的なものとして捉えるべきものである。これは、I D E A L 各試験プロトコールにおいて、生存期間中央値は主要評価項目とされておらず、副次的な位置づけしか与えられていないことの当然の帰結である。

このように、複数の副次的評価項目がある場合に、そのうちの一つの指標で良好な結果を示したとしても、それは、偶然の要素が入り込みやすい外、単に「いいとこ取り」をしているだけの可能性を否定することができず、信頼できる有効性の根拠を示すものではない（西甲 D 3 4 = 東甲 H 2 2 p 2 5 1 1 右段「エンドポイントが多くなれば、本来差がないものでも、…もっとも大きな治療差が認められたエンドポイントを事後的に選択することで、判断を誤ることになる。」との記載参照）。

I D E A L 各試験の主要評価項目（プライマリーエンドポイント）は、あくまで奏効率なのであるから、その結果を最重視すべきは当然である。生存期間中央値を含め、他の指標の肯定的結果を過大評価してしまうことで、誤った結

論に至る可能性がある点には注意が必要である。

4 既存薬における生存期間中央値の報告の概要

西甲H47～52＝東甲F82～87は、イレッサのIDEAL試験と同種の、非小細胞肺癌のセカンドライン第Ⅱ相試験である。これらの試験における各群の生存期間中央値を見ていくと、それぞれ9ヶ月（西甲H47＝東甲87のArm A）、8ヶ月（西甲H47＝東甲82のArm B）、8.5ヶ月（甲H48＝東甲82）、42週間（甲H51＝東甲85）、40週間（甲H52＝東甲86）と、いずれもIDEAL各試験の結果を上回る数値を示している（甲H49、50＝東甲83、84は、報告なし）。

5 審査報告書の生存期間中央値分析における比較対象

(1) IDEAL-1の評価について

なお、上述の「審査報告書(3)」(西乙B4の3＝東乙B4の3 p 54)のIDEAL-1の生存期間中央値の評価にあたり、比較対象として用いられている「J Clin Oncol 18:2095-2103, 2000」という研究は、実は、奏効率の比較対照ともなった上述のShepherdのドセタキセル試験である(西丙H22の2＝東甲F20)。上の奏効率の確認で見たとおり、歴史対照の問題、患者の背景因子がIDEAL-1のものはPSが良好で単純な比較が許されない点については、奏効率の評価と同様である。特に、生存期間中央値については、患者背景因子によって、極めて大きな変動を見せるもので、異なる試験の間での比較はほとんど意味を有しない。ましてや、Shepherd試験とIDEAL各試験の比較においては、予後に大きく関わるPSや腺がん割合といった患者背景因子に大きな違いがあることは、一見して明らかである。にも関わらず、これを無視して単純比較しようとする被告らの態度は、科学性の放棄とさえも評価しうるものである。

特に、生存期間中央値の評価との関連で着目すべきは、腺がん患者が、一般に長期生存期間中央値を示すことである。例えば、あるセカンドライン非小細胞肺癌試験（西甲H 5 1 = 東甲F 8 5）においては、腺がん、非腺がんの生存期間中央値はそれぞれ5 1 週、2 2 週と、倍以上の差を示している。そこで、Shepherd 試験の腺がん患者の割合を検討すべきこととなるが、この重要な患者背景について、論文中には明記されておらず（西丙H 2 2 の2 = 東甲F 2 0）、生存期間中央値比較をするための前提情報を欠いているというべきである。

以上より、IDEAL-1の結果から、安易にイレッサの延命効果の存在について推認することは慎むべきである。

（2） IDEAL-2の評価について

IDEAL-2の生存期間中央値評価にあたっては、「プラチナ系抗癌剤及びタキサン系抗がん剤治療後」すなわちサードライン患者群において、「平均的な予後は4ヶ月程度と推測される」という前提事実に依拠しているのであるが（西乙B 4 の3 = 東乙B 4 の3 p 5 4）、その根拠はここに明示されておらず、全く不明確であり、そもそも歴史対照の評価にあたって不可欠となる背景因子や治療水準の推移等の影響について検討すらできない。したがって、IDEAL-2の生存期間中央値評価が「臨床的に有意義である」と断じるには、全くもって科学的根拠を有さないものである。

6 まとめ

以上より、IDEAL-1, 2の生存期間中央値をもって、イレッサの有効性の根拠とするのは、完全な誤りである。

第6 承認時点のイレッサの有効性評価についてのまとめ

これまで述べてきたように、そもそも、抗がん剤の第Ⅱ相試験での奏効率をもって延命効果を推測すること自体に大きな問題性があることは、イレッサ承認時点では明らかとなっていた。

そして、イレッサについて I D E A L 試験結果を具体的に見ても、その奏効率を積極評価することには様々な点において問題があった。

更に、対照群を置かない I D E A L 試験での生存期間中央値などをもって、イレッサの効果を評価する根拠とすることは全く認められないことであった。

このような点を総合して考えれば、承認時点において、イレッサについて相当の腫瘍縮小効果があると評価すること自体に大きな問題があった。むしろ、I D E A L 1 試験の日本人群には、外国人群や別のドセタキセル試験の患者群と比較しても、際だって全身状態の良好な患者に偏っていたという患者背景の偏りなどもふまえて試験結果を考えれば、イレッサが日本人の非小細胞肺癌患者の治療において有効性、すなわち延命効果を有しない薬剤である可能性を念頭に置くべきような状況にあった。

第2節 イレッサ承認前の安全性評価

以下では、イレッサが承認された平成20年7月5日以前におけるイレッサの安全性の評価について述べる。第1章において述べたとおり、医薬品の安全性は、「危険性は鋭敏に」の予防原則のもとに評価されなければならないところ、イレッサが致死的な急性肺障害・間質性肺炎を惹起するということについては、既に承認前の段階において明らかであった。

以下、薬剤性肺障害についての当時の知見、イレッサのドラッグデザイン及び作

用機序，非臨床試験の結果，永井教授らによる動物実験の結果，臨床試験の結果を踏まえて，イレッサの安全性の欠如を明らかにする。

第1 急性肺障害・間質性肺炎について

1 抗ガン剤による薬剤性肺障害には死亡例・重篤例も多く見られていたこと

(1) 工藤翔二証人の主尋問及び意見書

工藤証人は，大阪地裁の西日本訴訟における主尋問及び同証人の意見書において，イレッサ承認前の段階，すなわち平成14年7月当時は，急性肺障害・間質性肺炎についての知見は未熟であった（西工藤翔二証人主尋問調書＝東乙L16 p1～8，p14～21，西乙E17＝東乙L18 p6～7）とし，その理由として，間質性肺炎の頻度や患者数が少なく，研究といえは症例報告しかなかった（西工藤翔二証人主尋問調書＝東乙L16 p14～21，乙E17＝東乙L18 p7～8）ことを挙げている。

そして，工藤証人は，①これらの研究のエビデンスレベルは「低い」（西工藤翔二証人主尋問調書 p20）として知見として重要視できないとし，かつ，②この当時の薬剤性肺障害の予後についての知見として，「圧倒的にはやっぱりステロイドをやればよくなる」（西工藤翔二証人主尋問調書＝東乙L16 p27）と述べ，さらに，中川の報告（西乙H34の4＝東乙F13の4）を挙げて「薬剤性肺障害はほとんど治るといのような，そういうメッセージですね」（西工藤翔二証人主尋問調書＝東乙L16 p27）とまとめている。

しかしながら，下記に述べるとおり，反対尋問においては，工藤証人の意見は維持されなかった。

(2) 薬剤の安全性評価に関する観察研究の重要性

まず、これらイレッサ承認前に行われていた、急性肺障害・間質性肺炎についての症例報告等の研究は、例えば光富証人、坪井証人らが強調するようなイレッサの有効性についての症例報告等とは異なり、薬剤の毒性についてのエビデンスとしては非常に貴重なものであった。

薬剤の「有効性」について、一般に、症例報告などの観察研究のエビデンスレベルが介入研究に劣るとされている理由のひとつは、起こった事象と薬剤との因果関係が完全には確定できないからである。すなわち、有効性は「慎重に」検証されなければならない、いわゆる「3た論法」，「飲んだ、治った、効いた」とはいえない。そのため、医薬品の有効性を検証するためには、薬剤の暴露／非暴露に研究者自身が介入する研究、すなわち無作為化盲検比較臨床試験に代表されるような介入試験を行わなければならないのである。この点については、既に、本章第1節第2において述べたとおりである。

他方、薬剤の安全性は「鋭敏に」評価されなければならない。薬剤の毒性すなわち副作用は、その定義上も、「病気の予防、診断もしくは治療、または生理機能を変える目的で投与された（投与量にかかわらない）医薬品に対する反応のうち、有害で意図しないもの」であり、その「医薬品に対する反応」とは、「有害事象のうち当該医薬品との因果関係が否定できないものを言う。」とされている（西丙D 3＝東丙H 3）のであり、因果関係が確立されないイベントであっても、副作用として十分に注意しなければならないのである。

さらに、薬剤の毒性については、その性質上、倫理的に介入研究を行うことは困難とされており、研究としては症例報告等しか行うことができない（西工藤翔二証人反対尋問調書＝東乙L 1 7 p 5 1～5 2）。

このように、「因果関係が否定できない」有害事象が副作用とされ、かつ、

薬剤の副作用について介入研究を行うことが困難である以上、薬剤の安全性に関しては、症例報告等の観察研究を、貴重なものとして重視しなければならないということは当然である。このことについて、反対尋問において工藤証人も認めざるを得なかった（西工藤翔二証人反対尋問調書＝東乙L 1 7 p 5 2～5 4）。

実際に、「薬剤性肺障害の評価、治療についてのガイドライン」の「はじめに」という欄（西丙H 4 6＝東丙G 7 2 p II）においても、「扱う対象の性質上、無作為試験を組むこと自体が不可能であり、レベル4あるいはレベル5の内容がほとんどである。」とされているが、これは、すなわち、同ガイドラインが、主として症例報告等の観察研究に基づいて作成されていることに他ならない。

また、医薬品情報の収集や評価のしかたについて書かれた教科書「医薬品情報学」にも、有効性についてのエビデンスとしては介入研究がエビデンスレベルは高いとされているが、安全性については介入研究が困難な場合が多く、観察研究がエビデンスとしては高く評価される。」（西甲F 3 9＝東甲G 6 6 p 1 0 2）とされており、「まれな病気や副作用などは、症例報告や観察研究が診断と治療に関する唯一の情報源となる」（西甲F 3 9＝東甲G 6 6 p 1 3 8）とされているところであり、患者や頻度の少ない副作用である急性肺障害・間質性肺炎という毒性をみるにあたっては、症例報告等の観察研究が極めて重要であることが繰り返し指摘されている。

臨床試験においても、医薬品の安全性評価の場面においては、症例報告は、安全性の評価には用いることが予定されている（東工藤翔二証人主尋問調書 p 5 3）。すなわち、安全性の面については、臨床試験は、（参加適格に厳しい制限があったり、併用薬も制限されるといった）限られた条件の下で行われることや、比較的少ない症例数で行われることなどから、副作用を十分に把握するには限界があり、安全性については、市販後の副作用報告（症例

報告の一種)が義務付けられ、承認申請段階でも、例えば先行して承認された外国での市販後の副作用症例やEAP症例のように、治験外で発生した副作用も報告が義務付けられている(例えば、西丙D3=東丙H3「治験中に得られる安全性情報の取り扱いについて 1934 頁一番下の段落「3. 緊急報告のための基準1 報告すべきもの (1)重篤で予測できない副作用は全て緊急報告の対象となる。」1935 頁第二段落「治験依頼者または企業は、重篤で予測できない副作用の報告を受けた場合、それが緊急報告の必要条件に当てはまる内容の場合は、情報源が何であれ該当する規制当局に迅速に報告しなければならない。」)。

(3) イレッサ承認前の薬剤性肺障害についての研究

工藤証人によれば、イレッサ承認前の段階においては、急性肺障害・間質性肺炎の分野では、「厚生省特定疾患びまん性肺疾患研究班」に代表的な研究者が集っており(西工藤翔二証人主尋問調書=東乙L16 p5~6)、薬剤性肺障害の研究の第一人者は、近藤有好、中川和子であったという(西工藤翔二証人主尋問調書=東乙L16 p17)。

このうち、中川の報告については、工藤証人は、主尋問においては中川の論文を「薬剤性肺障害はほとんど治るといような、そういうメッセージですね」と評価して、自らの「圧倒的にはやっぱりステロイドをやればよくなる」(西工藤翔二証人主尋問調書=東乙L16 p27)という意見を支持するものとして挙げておきながら、反対尋問においては、中川の報告をエビデンスが弱いとしている(西工藤翔二証人反対尋問調書=東乙L16 p72~73)。無論、工藤証人に指摘されるまでもなく、中川の報告は、例えば、論文中に3つ列挙されている研究のうちどの研究について報告したものかすら不明であり、さらにいえば、患者の母集団が何人か、そのうち抗ガン剤を用いた者は何人か、なども全く分からないのであって、少なくとも、抗ガン

剤による薬剤性肺障害の予後が良好であるとの結論を導出することは到底出来ない。

また、他方、近藤のいくつかの研究については、工藤証人は、「当時としては非常に重要な臨床研究であった」（西工藤翔二証人主尋問調書 p 26）として評価している。このうち、昭和47年に発表された「薬剤による肺障害」（西甲H43＝東甲G80）と題する近藤らの論文は、ブレオマイシンを使用した「新潟大学医学部，歯学部，脳研，長岡赤十字病院，水原郷病院，西新潟病院において昭和46年12月末日までに使用した282例」についてのケースシリーズであるが，同論文の「BLM-pneumoniaの転帰」という項には，BLM肺傷害の胸部X線所見について「全肺野びまん性に出現して増悪するもの，並びに肺感染症合併例に死亡例が多く見られた」「全肺野びまん性に出現するものは概して急速に悪化して呼吸困難に，あるいは心不全に至るので充分注意する必要がある」（同 p 409）との記載があり，全肺野びまん性に出現する病型については予後不良であるという病理学的な分析も行っている。さらに，昭和55年に発表された「薬物による肺炎」（西乙H34の1＝東乙F13の1）においては，少なくとも10種類以上の抗がん剤ないし免疫抑制剤投与例を含む306例もの症例レビューを行っている。このように，近藤の研究は，いずれも，とりわけ抗ガン剤による薬剤性肺障害に死亡例ないし重篤例が多数あることを示している。

また，各種解説書も近藤の研究等に基づき，平成8年に改訂第1版が出された「がん化学療法の副作用対策」には，「抗がん剤投与に起因する肺毒性も致命的な転帰をたどりうる」（西丙H11＝東丙F4 p 344），「抗がん剤による肺毒性はそのほとんどが間質性肺炎の型をとる。・・・間質性肺炎から肺線維症へ移行すると考えられており予後不良である。」「同じような障害を，頻度の差はあるものの，すべての抗がん剤で起こしうる」（同 p 109）との記載がある。また，平成11年に第1刷が出版された「ステロ

イド薬の選び方と使い方」においては、「a. 薬剤による間質性肺炎」という項目に、「抗悪性腫瘍薬によるものの予後は不良で、50%以上の死亡率が報告されているが、・・・」「抗悪性腫瘍薬によるものは治療が困難である。」との記載がある（西丙H33＝東丙F24 p107）。さらに、平成14年4月に第1版が発行された「呼吸器腫瘍学ハンドブック」には、「間質性肺炎を認めたなら、薬剤の中止が対応の第一段階である。中止により病勢の進行を抑えられることもあるが、中止にもかかわらず進行することも多い。・・・抗癌剤による肺の障害は頻度が高いとはいえないが、ひとたび発生すると致命的になることも多く、抗癌剤使用時のみでなく、その後も引き続いて注意深く経過観察する必要がある。」（西丙H32＝東丙F23 p146）との記載があるなど、抗ガン剤による薬剤性肺障害の予後が不良であることが示されている。

このように、とりわけ抗ガン剤による薬剤性肺障害については、予後不良の経過をたどるものもあったという知見は、イレッサ承認前の段階には既に存在していたものである。

（4）抗ガン剤一般にあてはめられないとの意見について

この点について、工藤証人は、ブレオマイシン、メトトレキサート、小柴胡湯等の少数の薬剤についてはある程度は傾向が分かっていたものの、それを抗がん剤一般にあてはめることはできず、多数例の集積をみなければ、その特徴はわからないなどと述べる（西乙E17＝東乙L18 p12，西工藤翔二証人主尋問調書＝東乙L16 p33，乙H32 p2）。

しかしながら、ブレオマイシンやメトトレキサート（抗がん剤としても用いられていた）のみならず、ブスルファンや、ゲムシタビンやイリノテカンなど、他の抗がん剤においてもやはり同様に急性肺障害・間質性肺炎による死亡例重篤例が出ていたのである。ほかならぬ工藤証人自身も、平成14年

4月に発表した「抗癌剤の副作用対策」において、「抗がん剤による肺障害は、1度発症すると治療が中断され、時には死に至ることもあり、また、GEM, CPT-11など多くの新薬にも肺毒性の報告があるため、現在でも臨床上重要な問題とされている。」「ほとんどの抗がん剤が急性・亜急性の肺傷害を呈する可能性がある」と述べている（西甲H35＝東甲F71 p315）とおり、少なくとも、抗ガン剤のごく一部についてしか急性肺障害・間質性肺炎が起きず、抗ガン剤の開発や販売においてさほど留意する必要はないという状況ではなかった。

2 イレッサ承認前の知見では、薬剤性肺障害のうち、臨床経過としてAIP型（病理学的にはDAD）をたどるものがとくに予後不良であることもわかってきたこと

(1) 工藤証人の主尋問及び意見書

工藤証人は、薬剤性肺障害について病理型と予後の関連について述べた西甲H6「抗癌剤による肺傷害」について、症例数が少なく、単なる問題提起に過ぎず、そのために原著論文化を差し控えたと述べる（西工藤翔二証人主尋問調書＝東乙L16 p22）。

しかし、工藤証人の上記論文は、薬剤性肺障害のうち、臨床経過としてAIP型（病理学的にはDAD）をたどるものがとくに予後不良であるとの当時の知見に合致し、これを支持するものであり、問題提起という位置づけにとどまるものではなかった。

(2) AIP/DADについての承認時の知見

工藤証人は、薬剤性間質性肺炎は、病理組織像がDAD（びまん性肺胞傷害）パターンをとるものは予後が不良であると述べている（西工藤翔二証人反対尋問調書＝東乙L17 p59）が、イレッサ承認時もまた、同様の知見

が存在していたのである。

ア ATS/ERSのガイドライン

また、病理組織としてDADパターンをとる間質性肺炎は、IIP特発性間質性肺炎の臨床経過でいえば、急性間質性肺炎、AIPといわれているところ（西工藤翔二証人反対尋問調書＝東乙L17p60）、IIP特発性間質性肺炎については、既に2002年にATS（米国胸部学会）／ERS（欧州呼吸器学会）による国際分類がなされ、また診断方針のガイドラインも整備されている（西工藤翔二証人反対尋問調書＝東乙L17p60）。

西甲H36＝東甲G78「間質性肺疾患診療ガイドライン」は、その「日本語版への序」（p vii～viii）を見ればわかるとおり、そのATS/ERS等が主導して作成された3つのガイドライン、すなわち1999年に公表されたサルコイドーシスのガイドライン、2000年2月に公表されたIPF特発性肺線維症のガイドライン、そして、2002年6月に公表されたIIP特発性間質性肺炎のガイドラインを日本語訳したものであるところ、2002年6月のIIPのガイドライン、日本語では「特発性間質性肺炎の診断指針」には、AIPの「臨床経過」として、「死亡率は高い（50%以上。多くは発症から1～2か月で死亡する）」と記載されている（西甲H36＝東甲G78p69）。

なお、この、ATSのガイドラインの作成経緯は、

「8. IIPの国際分類（2002）」の項、「IPF以外のIIPに含まれる疾患群の位置づけを明確にすることを目的として、1998年5月から、ATS、ERSの合同作業として、わが国からの研究者も加わって、IIP国際分類委員会が発足」（西甲H36＝東甲G78p133）

とされているように、IIPの診療ガイドラインや国際分類も、2002

年に突如できたものではなく，1998年5月から国際分類委員会が4年をかけてそれまでの知見の整理をしてきたものである。

I I Pのガイドラインの「要約」においても，「3. 文書作成の根拠」の項に，「MedLineで検索（1966年～1998年12月の期間）された英語で書かれた論文・・・」とされており，ガイドライン作成の根拠となった文献は，1998年までに出そろっていた知見が中心となっていた（西甲H36＝東甲G78 p43）。

さらに，ATSにおいては，間質性肺疾患に関して，1998年に国際分類委員会が発足する以前より，1990年代から，米国だけでもシンポジウムが数多く行われていた（西甲H36＝東甲G78 p136～145）。とくに，同p136には，「96年のATS学会において「I I P：現状と治療」と題した卒後教育コースが開催され，この時点においてI I Pガイドラインの骨格ともいふべき事項は示されていた」との記載がある。

イ 国内の文献

また，国内の文献でいえば，2000年8月に第1版が出した西甲H37「難病の最新情報」においては，この表によれば，特発性間質性肺炎の中の一つの病型である急性間質性肺炎（A I P）については，死亡率は62%，生存期間は1～2か月とされている（西甲H37＝東甲F72 p258）。

なお，この西甲H37の記載がそのよりどころとしているのは，末尾の文献（8），間質性肺炎の病理組織分類の第一人者である Katzenstein（カッツェンシュタイン）の1998年の文献（西甲H38 Katzenstein"Idiopathic Pulmonary Fibrosis"p1303）によれば，AIPで，死亡率は62パーセントとなっている（西工藤翔二証人反対尋問調書＝東乙L17 p62）。

さらに，西甲H39＝東甲F74「特発性間質性肺炎（I I P）の急性

増悪について」は、前掲近藤らの研究報告で、厚生省の「1993年度びまん性肺疾患調査研究」報告書に掲載されているものである（工藤証人は、西日本訴訟における主尋問で、間質性肺炎については、この厚生省のびまん性肺疾患調査研究班の班長が日本を代表する研究者である、としていたが、研究班の報告もまた当時の研究としては一番レベルが高かったとしている。西工藤翔二証人反対尋問調書＝東乙L17p63）。この研究報告においては、IIPの急性増悪例4例が紹介されている。うち3例はステロイド剤のパルス療法が無効で死亡、最終的には全例死亡しており、病理像としては、「DADの像がみられた」とされ、これらの所見は「急性増悪の特徴的所見」とされている（西工藤翔二証人反対尋問調書＝東乙L17p63）。

ウ 工藤証人らの研究

また、工藤証人らによる1998年の研究（西甲H6＝東甲F47，西甲H40＝東甲F75）も、抗癌剤による薬剤性間質性肺炎のうち、病理像としてDAD，臨床経過としてAIPの経過をたどるものは致死的であるとの知見と整合し、またそれを支持するものである。西甲H40＝東甲F75「抗がん剤による肺障害と特発性間質性肺炎の急性増悪」は、厚生省の「1998年度びまん性肺疾患調査研究」，すなわち、工藤証人の表現を借りれば当時国内では「一番高い」レベルの研究として報告されたものあり、西甲H6＝東甲F47の発表を論文化したものである。

この報告のサマリー部分には、「AIPパターンは5例に認められ、急激な経過をとり短期間に呼吸不全に陥った。ステロイド薬に対する反応性は不良で全例呼吸不全により死亡した」，「抗がん剤による肺傷害は急性経過が主体であり，IIP合併は致死的な肺傷害を発症する危険因子と推測された」 「臨床経過は急激で発症後約1週間で呼吸不全に陥った。ステロイド薬に対する反応は不良で2例では一時的な改善が認められたが、最

最終的に全例呼吸不全により死亡したこの中には、3例のIIPあるいはIIP合併例が含まれていた」と記述されている（西甲H40＝東甲F75 p 61）。具体的な「結果」として、AIPパターンの肺傷害は、「最終的に全例呼吸不全により死亡した」「case 1では剖検によりびまん性肺胞障害（DAD）が確認された」（西甲H40 p 62～63）。また、「従来より抗癌剤、放射線療法によりIIPが急性増悪することが報告されている」「肺癌治療における抗がん剤による肺傷害は、その主体が急性型に変化する可能性が示唆される」「今後、AIPパターンの病像をとりうる抗がん剤の使用にあたっては充分留意する」と記述されている（西甲H40＝東甲F75 p 63～64）。

また、証人工藤の別の文献には、「急性型のうち、病理学的にDADを示すものはステロイド薬に対する反応が悪く、予後は不良とされている。」との記載がある（西甲H35＝東甲F71 p 316）。

工藤らの研究は、抗癌剤による副作用としての肺障害についても、特発性間質性肺炎の病型分類にしたがって考えることを当然の前提として、このうち、AIP/DAD型をたどるものは予後不良であることを示したものであり、当時の医学的知見に合致するものである。

（3）小括

これらの国内外の文献、及びATSのガイドラインの作成状況をみるに、イレッサ承認当時には、特発性間質性肺炎のうち、AIPあるいはIIPの急性増悪症例、すなわち病理像としてDADをとるものについては予後が不良であるということは、医学的知見として存在していた（西工藤翔二証人反対尋問調書＝東乙L17 p 63～64）。よって、2002年6月以前、遅くとも1998年ころにはすでに、AIP/DADが予後不良であるというガイドラインの記載根拠となる知見は十分にあった。

3 まとめ

以上見てきたとおり，イレッサ承認時の段階において，抗ガン剤による薬剤性肺障害は基本的に予後が不良な重大な副作用であり，とりわけ，A I P / D A D型をたどるものは予後がとくに不良であって，抗ガン剤の開発や販売に際しては極めて重大な注意を払うべき副作用であるとの知見が存在していた。

第2 ドラッグデザインに見るイレッサの毒性の予見性

1 はじめに

本件ではイレッサによる急性肺障害・間質性肺炎の毒性被害が問題となっているところ，後述のとおり，イレッサによって致死的な急性肺障害・間質性肺炎が相当な頻度で発症することは，少なくとも臨床試験等の段階で十分に判明していた。しかしながら，イレッサは日本での承認が世界最初であったにもかかわらず，こうした急性肺障害・間質性肺炎という致死的な毒性が発症することについて，承認時点で十分な注意が払われたとは言い難い状況にあり，そのようなイレッサの毒性の軽視が本件のような多大な被害を生ぜしめている。

イレッサによって致死的な急性肺障害・間質性肺炎が発症することは，既にイレッサのドラッグデザインそのものからも当然に予期されたものであり，その意味で，イレッサによる急性肺障害・間質性肺炎は，まさにイレッサの主作用であるE G F R阻害に必然的に付随する副作用であって，予測不能な副作用ではなかった。

そうである以上，非臨床試験，臨床試験等の開発段階において，より注意深くイレッサの肺毒性が検討されなければならなかったのであり，そのような当たり前の注意が払われていれば，承認時点において肺毒性が軽視されることも

あり得ず、本件のような被害を惹起させることもなかったのである。

以下、イレッサのドラッグデザイン自体から肺毒性が予見されたことについて詳述する。

2 イレッサのドラッグデザインとEGFRの機能

イレッサは、ガン細胞の上皮細胞成長因子受容体（EGFR）に作用してチロシンキナーゼの磷酸化を阻害することによってガン細胞の増殖を抑制することをコンセプトとしてデザインされている（西丙C1＝東丙D1他）。

しかし、EGFRは、ガン細胞に過剰発現するとされていたものの、必ずしもガン細胞のみに特異的に存在するのではなく、正常細胞にも存在して、正常細胞の再生、分化等に極めて重要な役割を果たしていることは、イレッサ承認当時においても十分に判明していた知見である（西濱証人主尋問調書＝東甲L102p5以下）。

したがって、こうしたEGFRを阻害することによって、正常細胞にも何らかの悪影響を及ぼしかねないことは当初から十分に予見された事柄であった。

こうしたことは、以下のような知見に示されている。

(1) 「ヒト悪性腫瘍における上皮成長因子（EGF）関連ペプチドとそれらのレセプター」 David S Salomon ら， *Oncology*,1995，西甲E4＝東甲F4

「EGFRはまた正常肺と肺がんの両方に発現している。特に、EGFRの陽性染色が気管支上皮の刷子縁（ブラシ縁）に沿って見つかった」（訳文p15の下から4行目）

少なくとも1995年時点において、EGFRは、正常肺にも存在していたことが判明している。

(2) 「上皮増殖因子が新生児ラットのモデルにおいて壊死性小腸結腸炎の

進展を減じる」 Bohuslav Dvorak ら， American Journal of Physiology,2002,西甲 E 6 = 東甲 F 6

この論文では，壊死性腸炎を誘発させたラットを，EGF 欠乏ミルクでの飼育群，これに EGF を加えた飼育群，母乳での飼育群の 3 群に分けて観察したところ，EGF 欠乏ミルク群は出血が多く，腸管の異常も多かったのに対し，EGF を加えた群は，壊死性腸炎の発現を 50% 程度抑えたこと，あるいは，EGF 欠乏ミルク群が EGF を加えた群よりも EGF R の発現が多かったことなどが示されている。

こうした知見からは，EGF，EGF R は，傷害された組織の修復に重要な役割を果たし，生命の維持に重要な役割を担っていることが示されている（西濱証人主尋問調書 = 東甲 L 1 0 2 p 8）。

3 EGF R 阻害による肺障害の予見性

(1) はじめに

このように，EGF R は正常細胞にも存在して，その分化，再生等に重要な役割を果たしていることからすれば，EGF R 阻害薬を開発するにあたっては，それ自体として，生体に対する悪影響の可能性を慎重に吟味しなければならなかった（西濱証人主尋問調書 = 東甲 L 1 0 2 p 8 以下）。

そして，それにとどまらず，EGF R を阻害することによって，致命的な肺障害を発生させる可能性があることの予見可能性も，イレッサ承認時点において十分に存していたのである。

(2) 肺は傷つきやすい臓器

肺は人体の中でも広く大気に接し，また，全身の血流が肺を巡ることなどから，他臓器に比較しても傷つきやすい臓器である（西工藤証人反対尋問調書 = 東乙 L 1 7 p 2 1，2 2）。

この点については、被告側証人である工藤証人がその意見書で以下のとおり述べている。

「肺は、人体の中で最も広く大気に接する臓器となっており、絶えず微生物、粉塵・刺激性ガスなど外的要因の直接作用に晒されている。また、肺は、他の臓器と異なって血流が心臓に直列に配されているため、全身の血流が肺を巡ることによって、膠原病などの全身性疾患や薬剤副作用など本来全身的な影響が肺に疾患を引き起こす。また、このような吸気中・血液中の外的及び内的要因に対する、肺の有する防御機能の過剰な反応が様々な免疫学的機序による多彩な疾患を形成することもある。こうした肺の疾患のうち、外的要因及び内的要因によって肺に起こる炎症を肺炎という。」（西乙E 1 7 工藤意見書（国）＝東乙L 1 5 p 3）

（３） II型肺胞細胞の増殖・分化抑制と繊維化

ア 間質性肺炎と肺胞腔内の繊維化

肺胞が傷ついた場合、II型肺胞細胞が増殖してI型肺胞細胞に置き換わることによって、肺胞の修復が行われる。この点について被告側証人である工藤証人は、その意見書で「肺胞では、I型肺胞上皮細胞が傷害を受けると、この過程を経てII型肺胞上皮細胞がI型肺胞上皮細胞に分化して修復される。」（西乙E 1 7 工藤意見書（国）＝東乙L 1 5 p 3）と述べている。

そして、間質性肺炎については、その繊維化の機序として、「そのため、間質性肺炎では、I型肺胞上皮細胞が損傷する一方で、II型肺胞上皮細胞は修復を目的として増殖しI型肺胞上皮細胞に分化するが、慢性に経過すると炎症は器質化し、繊維芽細胞や膠原繊維が過剰に産制されることによ

って繊維化を起こす。」（西乙E 1 7 工藤意見書（国）＝東乙L 1 5 p 3）と述べている。

つまり、間質性肺炎の基本的な病態としては、工藤証人の説明によれば、肺胞腔内の繊維化、肺胞構造の改変（リモデリング）であるということになる（西工藤証人主尋問調書＝東乙L 1 6 p 8～1 4）。

イ II型肺胞細胞と繊維芽細胞

間質性肺炎の基本的な病態である肺胞腔内の繊維化については、II型肺胞細胞と繊維芽細胞との間での「陣取り」が行われ、繊維芽細胞がII型肺胞細胞の増殖に勝った場合に、肺胞腔内の繊維化が生じるとされている。この点については、工藤証人も引用している福田氏の文献において以下の通り説明されている。

「断裂基底膜周囲の繊維芽細胞は活性化され、本来接着していた細胞外基質からは遊離して、基底膜断裂部より肺胞腔内へ侵入し、増殖する。肺胞腔内に残存する再生性II型上皮と、腔内繊維化巣は陣取りを行う（図2－c）。上皮傷害の程度が軽く、再生もよい場合（図2－d 白矢印部）には腔内繊維化は形成されず、再生性II型上皮により被われ、正常細胞への再構築がなされるが、腔内繊維化が勝った部位（図2－d 青矢印部）では、繊維芽細胞の筋繊維芽細胞化とともに、盛んな細胞外基質の産生、沈着が起こる（図2－e）。わずかに残存したII型上皮の再生により不完全な小肺胞構造が作られるが、周囲は完全な繊維化に陥り、本来の肺胞構造は改築される。」（西乙H 3 6 の1＝東乙F 1 4 の1 p 2 6 右欄）

そして、こうしたII型肺胞細胞と繊維芽細胞の「陣取り」については、

生体の本来の反応としては、繊維芽細胞はⅡ型肺胞細胞の増殖を促進する一方で、Ⅱ型肺胞細胞は繊維芽細胞の増殖を抑制することが、以下の知見のとおり知られており、生体はより正常な修復を行う方向に反応していることが示されている。

「ラットⅡ型肺胞細胞は、肺繊維芽細胞の増殖を阻害する(in vitro 実験で)」 Tianli Pan, Robert J. Mason, Jay Y. Westcott, and John M. Shannon, 2001 (西甲H 4 6 = 東甲F 7 7)

「繊維芽細胞は、試験管内実験でも、肺の発達段階でも、Ⅱ型肺胞細胞の分化と増殖を促進する。」(訳文冒頭のアブストラクト)、「われわれは以下のように結論する。この同時培養システムにおいて、Ⅱ型肺胞細胞は繊維芽細胞の増殖を抑制した。」(訳文末尾部分)

このように、Ⅱ型肺胞細胞による正常な修復がなされるならば、肺胞腔内の繊維化が生じることはなく、間質性肺炎へと進展することが避けられると考えられる。

このことは、後述の東京女子医大永井教授らの実験論文でも、「EGFRのリン酸化の阻害が再生上皮の分芽増殖を抑制することによりブレオマイシンにより誘発された肺線維症を増強することを示唆している。」(西甲E 8 = 東甲G 6 訳文 p 1 下から3行目)と同様の認識が示されている。

また、この永井教授らの実験とは異なった結論を導き出したとする石井教授らの論文においても、「上皮の損傷および修復の遅延が、繊維発生に重要な役割を果たすのではないかと示唆されている。」(西乙H 3 4 の 8 = 東乙G 4 4 の 3, 4 訳文 p 6 上から5行目)と、この点については同様の認識が示されている。

ウ II型肺胞細胞の増殖・分化にEGFRが関与

以上のとおり、本来、II型肺胞細胞は繊維芽細胞の増殖を抑制し、他方、繊維芽細胞はII型肺胞細胞の増殖を促進する中で、傷ついた肺胞の修復が行われていくのであるから、II型肺胞細胞の増殖、再生機能が損なわれた場合には、こうした正常な修復過程が阻害されてしまうこととなる。

そして、以下の知見に示されるとおり、II型肺胞細胞の増殖、分化にはEGFRが関与しているのであり、イレッサはそのEGFRを阻害する以上、傷ついた肺胞のII型肺胞細胞による正常な修復を阻害することにより、繊維芽細胞の増殖、肺胞腔内の繊維化へと進展させてしまうことは十分に予見されたことである。

- a 「EGF 投与により、胎児アカゲザルの肺での肺胞 II 型細胞の分化が加速」 Plopper CG ら Am J Physiol. 1992 Mar;262(3 Pt 1):L313-21 (西甲 E 5 4 = 東甲 F 6 5)

この論文は、アカゲザルの胎児にEGFを投与した時のII型肺胞細胞の変化を観察したものであり、以下のとおり、EGF、EGFRがII型肺胞細胞の分化を促進し、また、SP-A（サーファクタント、これについては後述）の合成も活性化することが確認されている。

「妊娠第三トリメスター（妊娠後期）に外来性にEGFを投与すると、霊長類胎児での肺胞 II 型細胞の構造的機能的細胞分化を促進すると結論づける。これらの成熟変化は、肺の総合的な成長やガス交換領域の形態形成に有意な変化をもたらさずに生じる。」（訳文P 1 最初の段落の下から4行目以下）

「霊長類の妊娠後期にEGFを投与すると、肺胞II型上皮細胞の細胞分化を加速させる。」（訳文p 6 「考察」の項1行目）

「肺胞Ⅱ型細胞の分化をEGFがどのような機序で変化させるにせよ、われわれの研究からは、妊娠の最終トリメスター期の胎児にEGFを投与すると、肺胞Ⅱ型細胞の細胞分化を加速させるだけでなく、SP-Aの合成も賦活化することが明確である。」（訳文p9の本文末尾下から3行目）

- b 「ラット肺における上皮成長因子」 Raaberg L ら.Histochemistry. 1991;95(5):471-5. (西甲E55=東甲F66)

この論文は、ラットの胎児、新生児ラット、成体ラットの肺におけるEGFの存在について研究したものであるが、「出生の2日前から、その後の生存中を通じて、Ⅱ型肺胞上皮内にEGF免疫反応性が存在していることを報告する。」（訳文p1の冒頭部分の第2段落）とされ、成体ラットのⅡ型肺胞上皮内にもEGF、EGFRが存在していることが示されている。

- c 「上皮成長因子受容体を欠いたマウスにおける上皮の未発達と多臓器の欠陥について」 Paivi J Miettinen ら, Nature,1995, (西甲E3=東甲F3)

この論文では、遺伝子操作によってEGFRを持たないマウスを作り出した結果、EGFRを持つマウスに比較して、様々な点で生理学的な発達不全・機能障害が見られ、生後8日間しか生存できなかったことが記載されている。

そして、「呼吸困難が生じたのは、こうした肺胞が広範囲にわたって破壊されていたことに原因がある」（訳文p4・3行目以下）とされ、また、「これらのマウスの肺は比較対照となる正常なマウスのそれに比べ、劇的に異なっていた。これらのマウスの肺は胸膜の表面のそばにあ

るつぶれた肺胞に填塞され、あるいは胸膜のそばに膨張した終末気管支が存在した。つぶれた肺胞の中のわずかだけが、肺胞の界面活性物質である SP-C あるいは SP-A の着色が認められた。」（訳文 p 3 下から 4 行目以下）とされているように、肺の機能障害が直接の死因となっており、肺胞の界面活性物質（サーファクタント）である SP-C あるいは SP-A がわずかしか検出されずに、肺胞がつぶれてしまったことが示されている。

すなわち、EGFR を阻害することによって、肺胞が破壊されて虚脱し、また、肺胞をふくらませておくためのサーファクタント産生が阻害され、肺胞がつぶれて機能障害を起こして死に至ったことが示されている（西濱証人主尋問調書＝東甲 L 1 0 2 p 6～8）。

以上のとおり、II 型肺胞細胞の増殖、分化、あるいは、II 型肺胞細胞によるサーファクタント産生には EGFR が関与することが示されており、イレッサがこの EGFR を阻害することとなれば、傷ついた肺胞の修復過程において、II 型肺胞細胞の増殖、分化が阻害され、結果として繊維芽細胞の増殖が勝って、肺胞腔内の繊維化、間質性肺炎へと進展してしまう可能性のあることは、イレッサ承認当時、既に知られた知見だったのである。

なお、工藤証人は、その主尋問において、上記 c の Paivi J Miettinen（ミエチネン）らの Nature の論文（西甲 E 3＝東甲 F 3）について、胎生期の問題を指摘したものに過ぎず、成体にはあてはまらないなどと述べていた。しかしながら、上記 b の Reaberg（リーベルグ）らの文献（西甲 E 5 5＝東甲 F 6 6）では成体ラットの II 型肺胞細胞にも EGFR が存在していることが示されており、また、胎生期の肺の分化機序と同様の肺の修復機序が成体においても観察されることは、以下の文献で指摘されているのであって、上記 a の Plopper（プロッパー）ら（西甲 E 5 4＝東甲 F 6

5) や上記 c の Miettinen ら (西甲 E 3 = 東甲 F 3) の実験結果を, 単に胎生期のものであり成体にはあてはまらないなどと切り捨てて解釈することなど到底許されなかったことは明らかである。

d 人体組織学 1996 (西甲 H 4 4 = 東甲 F 7 6)

同文献は, 1996年に第2版が出版された組織学のごく基本的な文献であるが, p 138以下で胎生期の肺の分化機序が記載され, 「ヒトの扁平肺胞細胞 (I型細胞) と大肺胞細胞 (II型細胞) が内胚葉由来の未分化な立法上皮から分化することは以上のごとく明らかである。小動物の発生では, さらに大肺胞細胞 (II型細胞) から扁平肺胞細胞 (I型細胞) への分化が観察され, 実験的条件においても確認されているので, ヒトの場合も同様な分化が起こっている可能性が大である。」 (p 140左欄の第2段落部分) とされて, 未分化の立法上皮から I, II型肺胞細胞が分化し, また, II型肺胞細胞が I型肺胞細胞に置き換わって肺が形成されていくとされている。

他方, p 142の4) では, 「実験的肺障害後の再生と分化」, a) 「大肺胞細胞 (II型細胞) から扁平肺胞細胞 (I型細胞) への分化」の項で, 傷ついた肺の修復過程において, II型肺胞細胞が I型肺胞細胞に分化して修復されることが記載されている。そして, p 143のb) 「大肺胞細胞 (II型細胞) の起源」では, 「肺胞の障害が局所的に高度の時は呼吸細気管支細胞から肺胞上皮が再生される。」 (6行目以下) とされ, また, 「障害を受けた肺胞へ続く呼吸気管支の無線毛立法上皮細胞の造成が, 投与3日目に始まる。」 (10行目以下), 「9日目には分化が始まり, 12日目には2種類の肺胞上皮細胞に分化し, 肺胞が再生される。このように一時期肺胞に腺様構造が形成され, 胎児肺の腺様期の構造に類似するため, 肺胞の胎児化ともいわれる」 (下から7行

目以下)とされており、成体の肺胞の修復過程でも、胎生期と同様の機序によって修復が図られていることが示されているのである。

結局、工藤証人が Miettinen らの実験結果を胎生期のものに過ぎないとして切り捨てる発想は、第1章に指摘した「危険性は鋭敏に」の原則に背を向けるものに他ならないというべきであり、こうした安全性軽視の姿勢そのものが本件の薬害を引き起こしたと言っても決して過言ではない。

エ まとめ

以上のとおり、Ⅱ型肺胞細胞の増殖、分化を阻害することによって肺胞腔内の繊維化が促され、間質性肺炎へと進展する可能性があるところ、Ⅱ型肺胞細胞の増殖、再生にはEGFRが深く関与しているため、イレッサがEGFRを阻害することによって、間質性肺炎へと進展してしまう可能性があることは、イレッサのドラッグデザインそのものからも十分に予見可能なことであった。

この点については、被告側証人である工藤証人も、傷ついた肺の修復過程においては、EGFRを阻害することによって異常な修復（リモデリング）してしまう可能性があったことを認めざるを得なかったのである（西工藤証人反対尋問調書＝東乙L17p21～30）。

(4) Ⅱ型肺胞細胞の機能抑制と急性肺障害

ア 急性肺障害の発症機序

以上は、肺胞腔内の繊維化の観点から見た場合の間質性肺炎の発症機序として、イレッサによるEGFR阻害が間質性肺炎へと進展させてしまう可能性があったことを検討したものであるが、間質性肺炎の中でもより重篤な急性間質性肺炎（AIP/DAD）の発症機序においても、同様にE

GFRを阻害することによって、急性間質性肺炎を発症させてしまうことを予見することが十分に可能であった。

急性間質性肺炎（AIP/DAD）と成人呼吸窮迫症候群（ARDS）は、急性間質性肺炎が原因不明なものとして分類され、成人呼吸窮迫症候群が原因の判明したものとして分類されているが、両者の基本的な病態は同一であるとされる。この点は、工藤証人も証言しているとおりでであり（西工藤証人反対尋問調書＝東乙L17p34）、また、「ALI/ARDS（急性肺障害／急性呼吸促迫症候群）とDAD（びまん性肺胞障害）：病態と治療」日胸：63巻1号、2004年1月」（西甲H31＝東甲G105）においても、「DADはARDSの代表的な病理所見として知られ、敗血症とか肺感染症など様々な原因で生じますが、臨床的にも病理学でも種々病因を検索しても原因不明な間質性肺炎がある。それをAIPと命名したのです。」（p3右8行目）、「いまお話しいただいたように、ALI/ARDSがゼプシス等諸々、原因があって起こってくる病態名称として使われますが、病態は基本的に同じだけれど、原因がはっきりわからないものをAIPと呼んでいるということですね。」（p4左）とされているとおりである。

この急性間質性肺炎、成人呼吸窮迫症候群の基本的な病態は、Ⅱ型肺胞細胞のポンプ作用（肺胞腔内に浸潤してきた浸潤液等の吸収、除去機能）が阻害されたり、あるいは、サーファクタントの機能が阻害されることによって、肺胞が虚脱等する状態である。

この点については、「間質性肺炎一びまん性肺疾患」（西甲H32＝東甲G76）において、「ARDS（急性呼吸窮迫症候群）の病態」として、以下のとおり記載されている。

「Ⅱ型肺胞上皮細胞には肺胞腔内の水分を吸収・除去する働きがある

が、細胞損傷に伴ってその機能が障害される。肺胞腔内に漏出したタンパク質を豊富に含んだ水分は肺サーファクタントの機能を障害して肺胞虚脱を助長する。これにより、肺コンプライアンスの低下とガス交換障害が招来され、低酸素血症が生ずる。」（p 12・18行目以下）

したがって、Ⅱ型肺胞細胞のポンプ機能が阻害されたり、サーファクタントの機能が阻害されることによって、急性間質性肺炎と同様の病態に進展する可能性があることになる。

イ Ⅱ型肺胞細胞のポンプ機能とEGFR

Ⅱ型肺胞細胞のポンプ機能がEGFRに関連していることは、以下の知見により明らかである。

「上皮成長因子がラットの肺の液体クリアランスを増加させる」
Epidermal growth factor increases lungliquid clearance in rat lungs the American Physiological Society 85:1004-1010, 1998（西甲E 56＝東甲F 67）

同論文では、「上皮成長因子(EGF)は上皮細胞の増殖を賦活化し、肺胞上皮細胞モノレイヤーでの Na^+ 流量や $\text{Na}^+\text{-K}^+\text{-ATPase}$ の機能を高めることが報告されている。肺胞Ⅱ型細胞(AT2)での $\text{Na}^+\text{-K}^+\text{-ATPase}$ レベルの上昇は、増殖性肺損傷のモデルで能動的 Na^+ イオンの輸送やラットの肺胞上皮を通じた肺水腫のクリアランスの増加が伴うとされてきた。そこで、エアロゾル化したEGFをラットの肺に投与すると、能動的 Na^+ イオンの輸送と肺の液体クリアランスを増加させるかどうか調べた。」

（訳文 p 1 冒頭部分）とされ、EGFとⅡ型肺胞細胞でのポンプ機能の関連を調べたことが示されており、結論として、「以上の結果は、EGF

エアロゾルを in vivo 投与することで、ラットの肺胞上皮の $\text{Na}^+\text{-K}^+\text{-ATPase}$ 活性をアップレギュレートさせ、肺の液体クリアランスを増加させるという仮説を支持するものである。」（p 1 第 1 段落の下から 3 行目以下）とされ、EGF の投与により II 型肺胞細胞のポンプ機能が上昇することが示されている。

したがって、EGFR を阻害することによって、II 型肺胞細胞のポンプ機能が阻害されることは十分に予見されることであった。

工藤証人も、イレッサによって EGFR が阻害されることによって、II 型肺胞細胞のポンプ機能が阻害され、急性間質性肺炎と同様の病態を助長してしまう理論的可能性は否定できないことを認めている（西工藤証人反対尋問調書＝東乙 L 1 7 p 3 5～3 7）。

ウ サーファクタントの機能と EGFR

(ア) サーファクタント（肺表面活性物質）は、II 型肺胞細胞から産生され（西工藤証人主尋問調書＝東乙 L 1 6 p 1 1 「表面活性物質を作るとか」）、肺胞表面を覆って肺胞の表面張力を低下させて肺胞をふくらませるために極めて重要な物質である。この点については、組織学の基本的な文献である「ジュンケイラ組織学」（西甲 H 3 3＝東甲 G 7 7）において、以下のとおり記載されているとおりである。

「組織切片では、II 型肺胞上皮細胞に特徴的な、顆粒状または泡沫状の細胞質がみられる。この顆粒は…明確な層板小体として見出される。…（右欄 2 行目）また層板小体は絶え間なく形成され、細胞の頂上部表面で放出されることも示されている。層板小体は、肺胞表面上に広がって肺胞表面を被覆する物質のもとになっている。この物質は

肺胞表面の張力を低下させる肺表面活性物質 *pulmonary surfactant* である。…（下から10行目）肺表面活性物質は、肺の有機的な営みにいくつかが際だった特徴をもっているが、もっとも大切なことは肺胞細胞の表面張力を低下させることである。表面張力を低下させることにより、より少ない吸気で肺胞を膨らませ、呼吸運動の仕事を軽減させることができる。もし表面活性物質がなければ、肺胞は息を吐く時につぶれてしまうだろう。」（p 369）

「新生児の呼吸窮迫症候群は、生命を脅かす肺疾患で、肺表面活性物質の欠如によって起こる。…（p 370・2行目）未熟な肺は、肺表面活性物質の量も組織も不十分である。…（8行目）呼吸窮迫症候群では、肺胞はつぶれ、呼吸細気管支と肺胞管は拡張し、浮腫を生じる大量の間質液を含んでいる。」（p 369～370）

したがって、サーファクタントが不足することによって、肺胞がつぶれる、すなわち虚脱した状態となるのであり、新生児の呼吸窮迫症候群は、まさに、こうしたサーファクタント産生不良によって引き起こされる肺胞虚脱であることが分かっているのである。

また、「表面活性物質の層は安定しているものではなく、常につくりかえられている。」（p 370右欄の下の段落部分）とされ、サーファクタントは、常に作り替えられているものであることも示されている。

(イ) 次に、サーファクタントの機能として、肺で生じた炎症を防御する機能があることが知られていた。

「肺サーファクタントプロテインAが *in vivo* でのLPS誘導性サイトカインならびに一酸化窒素の産生を阻害する」Borron Pら, *Am J Physiol Lung Cell Mol Physiol* 278;L840-L847,2000（西甲E57＝東甲F68）では、サーファクタントを欠損させたマウスを用いて、リポ多糖

(炎症性サイトカイン分泌を促進する作用を持ち炎症を誘導する。) による肺の炎症の機序についての実験の結果が記載されているが、ここでは、「われわれの得たデータおよび他の研究者の得たデータは、SP-Aが免疫細胞に直接作用して、LPSにより誘導される炎症を抑えることを示唆している。以上の結果は、内因性あるいは外因性のSP-Aが、肺のLPSにより誘導されるサイトカインならびに一酸化窒素の産生をin vivoで阻害することを示すものである。」(訳文 p 1 冒頭第1段落下から5行目以下)、SP-AすなわちサーファクタントプロテインAがLPS(リポ多糖)により誘導された炎症を抑制することが示されたとしている。

(ウ) このように、II型肺胞細胞により産生されるサーファクタントは、肺胞をふくらませて虚脱させないために極めて重要な役割を果たしていると共に、炎症防御機能をも有している物質である。

そして、こうしたII型肺胞細胞により産生されるサーファクタントは、EGFRに関連していることが、以下のとおり、イレッサ承認当時、既に知られた知見となっていたのである。

すなわち、「胎児の肺でのMr = 35,000の肺サーファクタントプロテインの合成に対する上皮成長因子と腫瘍増殖因子-βの異なる効果」Whitsett JA. J Biol Chem 262;2908-7913, 1987 (西甲E 5 7 = 東甲F 6 9)は、ヒト胎児の肺組織を用いて、EGF刺激によってサーファクタント産生が促されるか否かを調べた実験であるが、「上皮成長因子(EGF)による刺激効果は、早くも2日の時点で検出され、5日後まで続いた。EGFに対する応答は用量依存的であり(0.01-10 ng/ml), [³⁵S]メチオニンの免疫沈降SAP-35への取り込み量の増加が伴っていた。」(訳文 p 1 の5行目以下)とされ、EGFの用量依存的にサーファクタント産生が増加したことが示されており、結論として、「EGFに応答して生じる

SAP-35 合成の増加は高度に選択的なものである。」（訳文 p 6 下から 1 4 行目以下）とされ、サーファクタント産生が EGF, EGF R に強く関連していることが示されたとしているのである。

また、上記のとおり、Plopper らの実験においては EGF 投与によってサーファクタント産生が活性化することが示されており（西甲 E 5 4 = 東甲 F 6 5）、逆に、Miettinen らの実験においては EGF R 欠損マウスにおいて、サーファクタント産生が僅かで肺が虚脱した状態となっていることが示されていたのである（西甲 E 3 = 東甲 F 3）。

このように、II 型肺胞細胞から産生され、「常に作り替えられている」サーファクタントは、EGF R に高度に関連していることは明らかであり、イレッサが EGF R を阻害することによって、サーファクタント産生が阻害され、肺虚脱を招いたり、炎症を防御できなくなってしまう、急性間質性肺炎と同様の病態を招いてしまう可能性があることが、イレッサ承認当時の知見においても明らかだったのである。

工藤証人も、少なくとも可能性として、イレッサがサーファクタント産生を阻害してしまうことを否定できなかった（西工藤証人反対尋問調書 = 東乙 L 1 7 p 4 1）。

(エ) なお、サーファクタントの炎症防御機能については、イレッサ検討会の委員でもある東北大学の貫和氏が同様の機序を考えていたことが「EGF R チロシンキナーゼ阻害薬は II 型肺胞上皮機能を低下させる」分子呼吸器病 Vol.10 No. 3, 2006 井上彰他（貫和敏博）（西甲 E 5 9 = 東甲 G 7 4）で示されており、「呼吸器感染症をはじめとするさまざまな肺での炎症において、II 型肺胞上皮細胞から産生される肺サーファクタント蛋白（とくに SP-A）は、防御因子としてその修復に大きく関与しているとされ、SP-A ノックアウトマウスではリポ多糖（LPS）による肺での炎症が増悪することや、急性呼吸急迫症候群（ARD

S) や放射線肺臓炎などの患者においては、血清SP-Aの増加を認めることなどが既に報告されている(6～9)」(p57左第2段落)、
「もとより胎生期の肺組織ではEGF刺激によりII型肺胞上皮細胞からのSP-A産制が促されることが知られていることから(12)、今回、われわれは、EGFRを抑制するゲフィチニブがII型肺胞上皮細胞からの肺サーファクタント蛋白産生を減少させることにより、炎症に対する肺での防御能力が低下しILDへ進展するとの仮説をたて、それを検証する目的で本研究を行った。」(p57右2行目)とされている。

(オ) さらに、Bryan Corrinらは、Pathology of The Lungにおいて、「肺胞の損傷と修復」(西甲H55＝東甲G91)で、「急性呼吸窮迫症候群と乳児呼吸促迫症候群の双方とも病理的变化が類似していることから、臨床的特徴および放射線学的特徴も相互に非常に似ている。」とし、また、「双方とも共通した事象サイクルが始まるため(図4.1)、原因の如何にはかかわりなく最終結果は同じになる」として図4.1を示している。図4.1では、「上皮と内皮損傷→サーファクタント欠乏→虚脱と浮腫→剪断力」の4項目が円を描いたサイクルとして描かれており、成人の急性呼吸窮迫症候群が「上皮と内皮損傷」から始まるのに対し、新生児呼吸窮迫症候群が「サーファクタント欠乏」から始まるという違いはあるものの、いずれもサイクルとして「上皮、内皮損傷」や「サーファクタント欠乏」を引き起こす要因と結果になっていることを示している。したがって、イレッサによってサーファクタント産生が阻害された場合には、それが引き金、原因となって「上皮と内皮の損傷」にも連なり得ることが示されているのであって、イレッサによる肺傷害は、単に傷ついた肺の修復阻害だけでなく、サーファクタント欠乏が肺虚脱等の肺傷害を引き起こす可能性があることが示されているのである。

5 まとめ

以上のとおりであり、イレッサの開発コンセプトであるEGFR阻害というドラッグデザインにおいては、そもそもEGFRはガンに特異的に発現しているものではなく正常細胞にも存在し、かつ、上皮細胞の再生、増殖に極めて重要な役割を果たしていることは、イレッサ承認当時において、既に知られた知見であった。

それだけでなく、EGFRは、肺胞上皮の正常な修復のために欠くことのできないⅡ型肺胞細胞の再生、増殖に強く関連しており、これを阻害することによって繊維芽細胞の増殖を促し肺胞腔内の繊維化、間質性肺炎へと進展してしまう可能性のあったことや、同様に、EGFRは、Ⅱ型肺胞細胞のポンプ機能やサーファクタント産生にも強く関連しており、これを阻害することによって、急性間質性肺炎（AIP/DAD）と同様の病態を招いてしまう可能性があったことは、イレッサ承認当時においては、既に知られた知見だったのである。

濱証人は、原告ら代理人が作成した模式図（西甲P71＝東甲L106）に基づいて、イレッサにおける肺傷害の発症機序を説明している（西甲E4＝東甲F43濱証人主尋問調書p6以下）。この濱証人の同模式図における説明は、まさに以上のような知見に裏付けられたものに他ならない。

このように見ると、本件で問題となっている急性肺障害・間質性肺炎は、イレッサのドラッグデザイン自体に内在し、それに由来する副作用、毒性だったのであり、医薬品としての主作用に必然的に付随する副作用、毒性であって、そもそも予測不可能な副作用、毒性ではなかったことは明らかである。

こうしたことから、神戸大学医学部附属病院教授・薬剤部長の奥村勝彦氏は、「みんなで考えようくすりのリスクとベネフィット」ーくすりの適正使用協議会・第13回日本医療薬学会年会共済シンポジウムーのディスカッション（西甲E60＝東甲G81）において、「イレッサはEGF（上皮成長因子）レセプターをターゲットとします。EGFは癌細胞に多いが、体中、全部にあるわ

けですから、そんなにターゲティングできるわけがないと思います。私どもの大学の大学院の人たちにはスタートから間違っていると講義していました。それで私のところではあまり使っていません。」（p 25左欄最終行以下）と述べている。

したがって、イレッサのドラッグデザインがこのようなものである以上、少なくとも、非臨床試験や臨床試験においては、イレッサのEGFR阻害による毒性、とりわけ致命的になりやすい肺毒性については、慎重な配慮をもって検討されなければならなかったというべきであり、そのようなごく当たり前の慎重な検討が加えられていれば、後述のとおり、臨床試験段階等における肺についての有害事象を軽視するようなことはあり得なかったことは明白である。

第3 非臨床試験に見るイレッサの毒性の予見性

1 はじめに

前項で検討したとおり、イレッサは、その開発コンセプトであるEGFR阻害というドラッグデザイン自体から見ても、正常細胞にも作用して、種々の毒性を発現させたり、とりわけ致命的な肺障害を生ずるかもしれないという可能性のある医薬品であった。すなわち、イレッサは、その本来予定する主作用に必然的に不随する副作用として、種々の毒性、とりわけ肺障害を生ずる可能性を内在していたのである。

したがって、非臨床試験における種々の毒性状況を検討するにあたっても、こうしたイレッサの本来的な毒性、とりわけ肺毒性について十分に慎重な吟味がなされなければならなかった。

しかるに、以下にのべるとおり、被告会社における非臨床試験は、イレッサの毒性を殊更に軽視し、あるいは隠蔽しようとしたのではないかとさえ疑われる

ようなものでしかなく、被告国もまた、イレッサのドラッグデザインから当然に予期される毒性，とりわけ肺毒性についての非臨床試験における十分な吟味を怠ったと言わざるを得ないのである。

2 非臨床試験の意義，目的

非臨床試験（前臨床試験）は，文字通りヒトへの使用前に動物によって医薬品の毒性等を確認するための試験であり，期待する効果発現量と毒性量，無毒性量を確認して，臨床試験段階への移行の可否の判定をし，最初にヒトに使用する場合の安全量を決定し，さらに，臨床試験段階で特に注意すべき毒性を把握するなどの目的を持っている（西甲E 2 5＝東G 3 1 濱意見書 p 9 以下，西濱証人主尋問調書＝東甲L 1 0 2 p 9 以下）。

特に，非臨床試験では，ヒトへの臨床用量よりも高用量が用いられる場合が多いが，これは，ヒトに使用した場合に少ない頻度で発現するかもしれない毒性をも把握するために行われるものである。したがって，そうした高用量で観察された毒性についても，当然，ヒトで発現する可能性を十分念頭において検討されなければならないのは当然である（西甲E 2 5＝東G 3 1 濱意見書 p 1 0，西濱証人主尋問調書＝東甲L 1 0 2 p 1 0）。

また，ヒトの臨床試験あるいは市販後の使用数は莫大な数に上るが，一般に非臨床試験における被験動物数はせいぜい十数頭から数十頭に過ぎないことからすれば，非臨床試験において発現した毒性が頻度としては低かったとしても決して軽視してはならない（西甲E 2 5＝東G 3 1 濱意見書 p 1 0 以下，西濱証人主尋問調書＝東甲L 1 0 2 p 1 0 以下）

これは，第1章で指摘した長尾氏らの指摘において，「一般に非臨床試験の有効性は過大に，安全性は過小に評価される傾向にある。すなわち有効性については，動物に対して臨床投与量の10倍以上を投与したデータで効果があれば有効だと解釈し，安全性については動物に対して臨床投与量の10

倍を投与して発現する有害反応を、ヒトに使う量の10倍ですから…と軽んじる。この解釈にいかに問題があるかは既に解説したことから明白である。」（西甲F38＝東F60「医療薬学I」p87・10行目以下）とされているところと全く同様の指摘である。

こうした点について、前述の砂原茂一氏は、医薬品の非臨床試験—動物実験における安全性の考え方につき以下のとおり指摘する（西甲P38＝東L63）。

「しかし、動物実験にはヒトにおける臨床試験からは引き出すことのできない貴重な情報を引き出すことができるという利点もあるのである。ヒトには到底与えることのできない大量の薬を与えて、そのような極限状況ではじめてとらえることのできる、いわば潜在的な毒性を発見することができる。そのような副作用は普通にはヒトには起こらないかもしれない。しかしきわめて稀には特殊な条件の下で、あるいはその薬に特に敏感な人には同じようなことが起こらないとはいえないから、動物実験からの情報に基づいて警戒心を強めることが望ましいということになる。また動物の場合は実験条件の設定が自由だから、ある薬のいろいろな量を、いろいろな期間与え続けて適当な時期にその動物を殺し、臓器を取り出し薬の影響をくわしく研究することもできる。したがってヒトの試験では到底得られないような貴重な情報を前臨床試験が与えてくれるはずである（p115）。」

「そのうえ人間の場合とはちがって動物実験は条件の設定が自由であり、何回でも繰り返すことができるし、偶然的な機会に依存するしかない臨床観察では、しばしば見逃しやすい事実を速やかにとらえることのできる鋭い切れ味をもっている。サリドマイドにしても市販前に妊娠中の動物に対する安全性試験—催奇形性試験が行われていればあのような歴史的な悲劇が回避されたにちがいないし、キノホルムの場合もスモンとの因果関係が気づかれた

後に行われた程度の動物実験がもっと早い時期に行われていれば、因果関係が少なくとももう少し早くとらえられたにちがいない。四頭筋拘縮症にしても筋肉注射の局所筋肉への影響を個々の注射薬について動物実験で確かめておくことが新薬許可の申請のさい要求されていたとしたら、あのような事件が頻発しなかったであろう。したがって前臨床試験—動物実験の必要性、重要性をどれほど強調してもなお足りないであろう（p 120）。」

また、医薬品副作用被害救済・研究振興調査機構（現在の医薬品医療機器総合機構）の藤森観之助氏は、「市販前の安全性について—非臨床試験のデータをどう生かすか—」（臨床評価29巻1号p 7以下：2001年）において、医薬品の安全性確保についての非臨床試験の意義を以下のとおり指摘する（西甲G6＝東F89）。

「すべての段階の臨床試験あるいは臨床段階において、事前に有害事象の発現を最小にする、あるいは対応する方策を予測する手段を取ることが重要である。臨床試験で現れる有害事象に関しては1. 重篤度、2. 有害事象の理由、3. 用量／血中濃度相関、4. 可逆性（回復性）、6. 類薬有害事象との類似性、7. 用法による回避について事前に検討する必要がある（6, 7はママ：代理人注）。本来、非臨床試験はそのために最大限利用できるように計画されているのであり、その結果に関しては、臨床試験とも密接に関連させ、事前に、あるいは事後であっても評価を行うべきである（p 12右側）。」

「開発過程における非臨床試験は臨床試験を補完するものである。……従って、非臨床試験の意義は臨床試験に入る前にヒトでの有効性〔臨床薬理〕および安全性を予測することと共に、臨床試験結果の評価を理論的に補完することにある。非臨床試験のすべてのデータは臨床での有効性と安全性の評価

に利用するためにあり，薬剤学的試験データ，非臨床毒性試験データ，非臨床薬理試験）ICH-C TDによる分類では効力を裏付ける薬力学的試験，副次的薬力学的試験，安全性薬理試験および薬力学的相互作用）データおよび非臨床薬物動態試験のすべてのデータが臨床データの臨床薬理効果の評価に総合的に利用される（p 13左側）。」

このように，非臨床試験は，単にヒトへの投与段階となる臨床試験段階に進むことの可否，あるいは臨床試験における用量を定めるだけでなく，当該医薬品の安全性確保のために，臨床試験結果とも合わせて，総合的に考慮されなければならないという，これもまた当然の事理が指摘されている。

3 イレッサ非臨床試験で見られた多くの屠殺例の解釈

(1) イレッサの非臨床試験においては，以下のような屠殺例が生じている（西丙C 1 = 東丙D 1 申請資料概要 p 196 以下）。

単回投与試験でラット雌 5 匹中 4 匹

イヌ 1 ヶ月試験で 1 頭

ラット 6 ヶ月試験で 4 頭

イヌ 6 ヶ月試験で 2 頭

非臨床試験において被験動物が屠殺処分されるのは，「死にかかった場合，徒に死を待つより屠殺処分を行う方が多くの知見が得られる」ため（西甲D 1 = 東甲H 1 「医薬品毒性試験法ガイドライン（H1.9.11 薬審 24 号，H5.8.10 新薬審第 88 号）」）であり，これらの屠殺処分された被験動物は，いずれも「死にかかった」と判断されたものに他ならない。

これだけの数の被験動物を屠殺処分しなけりばならなかつたということ自体，イレッサの毒性の強さを物語るものである。

イレッサの非臨床試験，特にイヌ 6 ヶ月継続投与試験において用いられた

最高用量である 25 mg/kg/day は、AUC（時間血中濃度曲線下面積、被験物質の総曝露量の算定において示される単位）で比較すれば、ヒト臨床用量のわずか 1.8 倍に過ぎない（西甲 E 25 = 東 G 3 1 濱意見書 p 2 1 以下、西濱証人主尋問調書 = 東甲 L 1 0 2 p 1 2 ~ 1 4, 西丙 C 1 = 東丙 D 1 p 2 1 3, 3 2 2）。

このような用量で、これだけの被験動物を屠殺処分しなければならなかったことはイレッサの毒性を強く物語るものであり、非臨床試験、さらには臨床試験を通じて、少なくとも市販前にはイレッサの毒性の慎重な吟味が必要だったことは明らかである。

(2) また、それにとどまらず重大な問題は、イレッサの非臨床試験において、これだけの被験動物を屠殺処分したにもかかわらず、その死因に繋がりうる病変が不明なまま放置されていたことである。

これらの屠殺処分された被験動物の死因に繋がる病変が不明であることは、濱証人の意見書（西甲 E 2 5 = 東甲 G 3 1 p 1 3 以下, p 2 0 以下）で詳述されており、詳細は同意見書の記載に譲るが、本来、非臨床試験の最大の目的が被験物質の毒性の把握にある以上、屠殺処分した被験動物からは、被験物質の毒性が明らかになるような所見が得られないと何の意味もない。

特に、イヌ 6 ヶ月試験やラット 6 ヶ月試験では、途中で最高用量群の用量を 25 mg から 15 mg に減量している（西丙 C 1 = 東丙 D 1 p 2 0 8, 2 1 2）。こうした減量自体、通常ではあり得ない事態であるにもかかわらず、試験実施者は、試験を継続するために敢えて減量しているのであって、これは、最高用量群においては毒性が強く出過ぎると判断したからに他ならない（西濱証人主尋問調書 = 東甲 L 1 0 2 p 1 4 以下）。

それにもかかわらず、屠殺された被験動物の死因に繋がる所見が判然としないというのは奇妙としか言いようが無く、これは、屠殺処分が早すぎたた

め死因に繋がる所見が把握できなかつたものとする（西濱証人主尋問調書＝東甲L102p12）。

非臨床試験の目的が被験物質の毒性の把握にあり、被験動物数が決して多くないことからすれば、屠殺処分に至る動物から得られる情報は極めて貴重であるはずであるが、屠殺が早すぎて情報を得られなかつたということになれば、欠陥のある非臨床試験であつたと言わざるを得ないのである（西濱証人主尋問調書＝東甲L102p20以下）。

さらに言えば、これだけの被験動物を屠殺処分したにもかかわらず、屠殺が早すぎて死因に繋がる所見を把握できなかつたケースが決して少なくないことからすれば、さらに穿ってみれば、敢えて、イレッサの毒性を隠蔽すべく屠殺処分を早めたのではないかとまで疑われるところである。

こうした点について、特にラット6ヶ月試験において、高用量群のラットが8週目に屠殺され、用量が減量されているにもかかわらず、当該屠殺ラットの死因が不明とされている点などを受けて、濱証人は以下のように証言している（西甲E43＝東濱証人主尋問調書p23以下）。

「これはもうほんとに信じ難いことなんですけれども、イヌの場合には、かろうじて10日目に屠殺した症例のイヌの解剖所見として、死因につながる病変が、慢性肺炎というのはちょっとおかしいんですけれども、肺虚脱というふうな格好で、一応、肺虚脱というふうなことがうかがえるような所見が記載されておるわけなんですけれども、ラット8週目と11週目という、比較的早期に死亡したあるいは屠殺せざるを得なかつた、そういう症例について、死因が全く記載されていないということは非常に不可解なことです。といいますのは、毒性試験というのは、死亡した症例が生じた場合には、その死亡につながる病変がどこに現れるのか、どういう毒性を生じて動物が死ぬのかということを知ることが一番大事なことであります。

それは当然ヒトでどういう毒性で、ヒトが死亡するようなことが現れる場合には、どこが障害されて死亡するのかということを知るために、極めて重要なことでもありますから、その所見が、8週目で屠殺した例、それから11週で屠殺した例について、全く記載がないと。死因が分からないということはあり得ないことでもあります。こういう実験は、実験として非常に価値が低いのではないかと。またもちろん、EGFR阻害によって起きた病変が現れたということはいかえまでするので、その意味ではいいんですけども、安全性を極めて低く見積もるような、そういう形の報告書になっているというふうに考えます。」

(3) このように、イレッサの非臨床試験においては、屠殺処分された被験動物が決して少なくなく、それ自体イレッサの毒性の強さを物語っているものであり、また、そうした屠殺処分から十分な毒性情報が得られていないことは、イレッサの非臨床試験の不十分性を端的に示しているというべきである。

4 マクロファージ等の肺毒性所見

(1) これまでの準備書面でも指摘したが、イレッサの非臨床試験では、イレッサ群の被験動物に肺胞マクロファージ浸潤が観察されている。

特に、ラット6ヶ月試験ではイレッサ高用量群40頭中10頭に認められ対照群には1頭も観察されず、イヌ6ヶ月試験ではイレッサ高用量群8頭中3頭に認められ、同様に対照群では1頭も観察されていない（西甲B5, 6 = 東甲C3, 4）。とりわけラット6ヶ月試験における出現率は、統計的に見ても明らかに有意な差を示している（西甲E25 = 東G31 濱意見書p17）。

ドラッグデザインの項で述べたとおり、イレッサはEGFRを阻害することでII型肺胞細胞の機能を阻害し、肺胞虚脱を招く可能性があったのであり、

そのことによりマクロファージが増加することは十分に考えられることである（西甲E25＝東G31 濱意見書p17，西濱証人主尋問調書＝東甲L102 p15以下）。少なくとも，マクロファージは，「旺盛な貪食能を有し，大食細胞とも呼ばれる。炎症局所で壊死組織や病原体などを貪食し処理する働きがある」（西甲G3＝東甲G61 p104）とされるものであり，マクロファージが認められたことは，肺に炎症が起きたことを示唆するものである。

いずれから見ても，このように肺虚脱や炎症性変化など何らかの異常所見を示すマクロファージの増加が対照群では1頭も観察されていないにもかかわらず，イレッサ群にのみ観察されたことは，イレッサの肺毒性を示唆するものとして検討される必要があったことは明らかである。

(2) これに対し，被告らは，これらのマクロファージの発現は，自然発生的に観察されるものと変わらないなどとする（西丙C4＝東丙G74 工藤意見書）。しかしながら，上記のとおり，肺泡マクロファージの増加はイレッサ投与群のみにみられ，溶媒対照群にはただの1例も見られておらず，しかも，ラット6ヶ月試験ではイレッサ群の発現率が有意差をもって観察されたという状況なのであって，こうした結果を目の前にしてもなお，「自然発生的」などとしてかたづけられる発想は，まさに「危険性は鋭敏に」の基本的な原則を無視するものでしかない（西濱証人主尋問調書＝東甲L102 p16）。

そもそもこれまでの準備書面でも指摘したとおり，毒性試験における被験動物は，被告会社も詳細に主張するGLPに基づき厳格に管理されていたはずであり，そうであるからこそ，溶媒対照群には自然発生的な肺の炎症性変化が見られなかったと評価するのが最も素直且つ合理的な評価であることは疑いを入れる余地がない。イレッサ投与群のみの管理が杜撰であったはずもなく（もしそうであれば，毒性試験としては不適合であり，承認審査資料と

しては使用に耐えないものに他ならない。) , 溶媒対照群にみられなかった所見がイレッサ投与群のみに見られ, それも高用量群に特徴的に認められる以上, こうした所見について, 安易にイレッサとの関連を否定する評価こそ, 医薬品の安全性確保についての原則を踏みにじっているという他ない。なお, 被告会社が前提とする自然発生的なイヌのマクロファージ等の報告は, 被告会社が提出した証拠上からは必ずしも十分に明らかではないが, そこで報告されている動物群の飼育・管理が, G L Pに則ったほどの厳格な管理がなされていたとの根拠はない(西丙G 1 1 ~ 1 3 = 東丙F 3 9 ~ 4 1)。そして, ビーグル犬における泡沫細胞浸潤は, せいぜい2 1例中1例, 4 2例中2例程度の頻度であるのに比べれば(丙G 1 2 = 東丙F 4 0), イレッサの反復投与試験にみられた高用量群のマクロファージ所見の比率は格段に高い。

以上のとおり, 少なくともイレッサにおける肺胞マクロファージの所見については, これをイレッサとの関連を簡単に否定し切ることができるようなものであり得ないことは余りに明らかだったというべきである。

なお, 濱証人がイレッサの肺虚脱に伴って間質に存在した肺胞マクロファージが肺胞に出てきたと説明する機序について, 工藤証人は, 肺胞マクロファージは間質には存在しないかに述べていたが, 「ジュンケイラ組織学」(西甲H 3 3 = 東甲G 7 7) p 3 6 8 図1 7 - 1 9においては, 明確に「肺胞中隔内のマクロファージ」との記載があり, また, 「肺胞のマクロファージは, 肺胞中隔中に見い出されるが, しばしば肺胞表面にも見られる。」(西甲H 3 3 = 東甲G 7 7 p 3 7 1) とされており, マクロファージは, 間質中に存するのが原則的な位置であるとされているのである。

さらに, 工藤証人は, 現在でもイレッサによる間質性肺炎によって肺胞マクロファージが増加するという所見はないなどと述べていたが, 日本におけるイレッサによる間質性肺炎発症例の症例報告によれば, 「肺胞腔へのマクロファージの集積」が認められているのであって, この点の工藤証人の認識

も正確とは言い難い（西甲H58＝東甲G94 訳文 p 2・13 行目以下）。

5 イヌ6ヶ月試験の肺炎症例等

(1) イヌ6ヶ月試験では、高用量群のメス1頭が体重減少、摂餌減少により10日目に切迫屠殺され、この屠殺を契機として、高用量群の用量が25mgから15mgに減少されており（西丙C1＝東丙D1 p 212）、実験者は、この屠殺例の一般状態の悪化等がイレッサの毒性であり、且つ、試験継続のためには減量しなければならないと考えたことが示されている。

そして、同屠殺例は、ケースカードによれば、肉眼所見では左肺上葉が小さく蒼白化しており、剖検所見では「慢性肺炎」が見られたとされている（西甲E17＝東甲G17）。申請資料概要では、この屠殺例については腎乳頭壊死が見られたとの記載がされているのみであるが（西丙C1＝東丙D1 p 212）、ケースカードによれば、この腎乳頭壊死は軽微（minimal）とされており死因に繋がる所見とは考えられない（西甲E17＝東甲G17）。

したがって、この屠殺例の死因に繋がりうる病変は、「慢性肺炎」とされた肺病変だったことは明らかである（西甲E25＝東G31 濱意見書 p 18 以下、西濱証人主尋問調書＝東甲L102 p 17 以下）。

そして、本来健康であるはずの被験動物が僅か10日で「慢性」の肺炎となるとは考え難く、また、前項で詳論したとおり、イレッサのEGFR阻害作用によって、肺サーファクタント産生が阻害されて肺虚脱に至る可能性があることなどを前提にすれば、この「慢性肺炎」とされた病変は、急性の肺障害、肺虚脱だった可能性を否定しきれないというべきである（西甲E24 濱意見書 p 19、西濱証人主尋問調書＝東甲L102 p 18 以下）。

したがって、このイヌ6ヶ月試験での屠殺例は、まさに、ヒトにおける急性肺障害の毒性を示していたものであり、どれだけ譲っても、臨床試験段階

において、ヒトの肺毒性について慎重な吟味が必要だったというべきである。

また、同じくイヌ6ヶ月試験では、高用量群のオス1頭に、限局性の肺胞中隔化生が認められている。化生は、正常の組織から、正常には存在しない組織に置き換わることであり、肺胞中隔（間質）の肺胞細胞が減少ないしは消失し、間質優位ないしは、間質のみに置き換わったことを意味する。「interstitial proliferation：肺間質増殖」に相当し、間質性肺炎に近い所見である（西甲E25＝東G31濱意見書p20）。

(2) これに対し、被告らは、この「慢性肺炎」の所見は、実験以前から存在した肺炎が慢性化したものであり、特にイレッサとの関連を検討する必要はないなどとする。

しかしながら、薬事法14条3項等に基づき非臨床試験の適正確保を定めたGLP（医薬品の安全性に関する非臨床試験の実施の基準に関する省令・西丙D1＝東丙H1）によれば、「動物を用いた試験を行う試験施設は、動物を適切に飼育し、又は管理するため、飼育施設、飼料、補給品等を保管する動物用品供給施設その他必要な施設設備を有しなければならない」（9条2項）、「試験に従事する者は、前項の観察又は試験中に試験の実施に影響を及ぼすような疾病又は状況が見られる動物を、他の動物から隔離するとともに、試験に使用してはならない。」（12条2項）、「試験に従事する者は、飼育施設、動物用品等を衛生的に管理しなければならない。」（12条5項）とされている。

このように、医薬品の承認申請目的の非臨床試験はGLPによって厳格に管理された被験動物を使用するのであって、実験以前から慢性の肺炎を呈していたとは考え難い。工藤証人は、慢性肺炎は肺胞性肺炎が器質化したものであり、炎症が器質化した時点では外見所見から判断しにくいなどと述べているが、そうであれば、かつては肺胞性肺炎を呈していたことになり、その

時点では発熱等の一般状態の悪化が外部からも判別可能な状態となっていたはずであって、そのような病歴のある被験動物を使用するなどということはG L Pに違反した非臨床試験であったという他ない。

そして、同被験動物のケースカード等の資料を見ても、どこにも肺胞性肺炎の原因菌等が検出されたという記載はない。

したがって、イレッサのドラッグデザインから肺毒性が生じることが予見されていた以上、どれだけ譲っても、この肺障害の症例は、イレッサとの関連がある肺毒性が示された症例として評価されなければならなかったというべきである。

また、被告らは、同被験動物の一般状態の悪化は、小腸病変に関連していたとするが（西丙C 4＝東丙D 5，西丙E 5 4＝東丙G 7 4 工藤意見書，西乙E 2 3＝東工藤証人主尋問調書），申請資料概要（西丙C 1＝東丙D 1）には、同被験動物の剖検結果としては、腎乳頭壊死しか記載されておらず、小腸病変についての言及は全くない。上記のとおり、同被験動物は一般状態の悪化のために早期屠殺され、この屠殺を原因として高用量群の用量減量がなされているのであって、同被験動物の一般状態悪化を招いた病変は、特に注意されて観察されたはずである。にもかかわらず、申請資料概要には、小腸病変には全く言及されていないのであって、非臨床試験当時において、この小腸病変が一般状態悪化の原因として検討されていたとは考え難い。被告会社から提出された概要書（西丙C 4）は、本件訴訟が提訴された後に作成されたものであって（2006年1月作成），明らかに跡づきのものという他ない。

工藤証人は、間質性病変であれば肺全体にびまん性に発症するはずであるとも述べるが、イレッサによってサーファクタント産生が阻害されて肺虚脱（無気肺と同義）になるとすれば、そうした肺虚脱が肺の一部に生ずること当然にあり得ることである（西甲H 5 7＝東甲G 9 3 「メルクマニュアル」

日本語版)。

さらに工藤証人は、肺胞中隔化生について、イレッサによる間質性肺炎であれば、びまん性に肺全体に障害が生じるはずであるとするが、専門家会議報告書によれば、「同一症例で部位によって、時間の経過したDADとより新鮮なDADが混在した。画像所見も含め、初期には傷害が局在的であることを示している。」(西丙L2=東丙D4 p20)とされているのであって、局所的な化生について、それが局所的であることをもってイレッサによる障害でないと結論づけることなどできないというべきである。

6 ラット6ヶ月試験の肺胞浮腫等

(1) ラット6ヶ月試験においては、高用量群オス1頭が24週目に切迫屠殺されており、肺組織所見として、中等度の肺胞浮腫と肺胞内細胞浸潤の多発、気管支には異物性肉芽腫(症)および膿瘍形成の所見が見られており(西甲B9=東甲C8)、特にケースカードでは、肺組織所見の肺胞浮腫と気管支の所見は死因に繋がるものとしてアスタリスク(*)が付されている。

これらの所見もイレッサの毒性による呼吸器症状として把握される必要があったというべきである(西甲E25=東G31 濱意見書 p19以下、西濱証人主尋問調書=東甲L102 p21以下)。

(2) これに対し、被告らは、これらの所見は誤投与によるものであるとする。

しかしながら、ケースカード等の資料のどこにも、これらの所見が誤投与であるとする記載はなく、上記のとおり、本件提訴後に提出された概要書(西丙C4=東丙D5)に至って、はじめて誤投与なる記載がされたに過ぎない。

7 まとめ

以上の通り、イレッサの非臨床試験を検討するにあたっては、イレッサのEGFR阻害というドラッグデザインから予測される毒性、特に致命的となりうる肺毒性について、十分慎重な吟味が加えられる必要があったところ、現実には、肺胞マクロファージの有意な増加、イヌ6ヶ月試験で肺毒性所見、ラット6ヶ月試験での呼吸器毒性などの所見が得られており、また、多くの屠殺処分をせざるを得ず、そして、6ヶ月試験では高用量群の用量を減量せざるを得ない状況になるなどイレッサの強い毒性が観察されていた。

被告国も、原告らが指摘するマクロファージ等の所見について、いずれも肺の炎症性変化を示唆する所見であること自体を認めている（西国第3準備書面p19＝東国第3準備書面p16）。

他方、これらの非臨床試験においては、多くの屠殺処分をしながら、その毒性所見を十分に把握できていないという欠陥があったのである。こうしたことからすれば、本来であれば、さらによくデザインした毒性試験を再試行すべきであり、そうすれば、さらにイレッサのより詳細な毒性プロファイルを得ることができたはずであるが、そうした再試験をすることもなかった。

したがって、どれだけ譲っても、臨床試験の段階においては、イレッサの毒性、とりわけ致命的となりうる肺毒性について、極めて慎重な吟味がなされる必要があったことは明白である。

第4 東京女子医大永井教授らの実験について

- 1 イレッサ承認以前において、東京女子医大永井教授らによって、イレッサがマウスの肺線維症を増強させるという実験結果が得られており、被告会社もこの情報を入手していたことは、これまでも訴状、準備書面等において詳述して

きた。

永井教授らの実験は、ブレオマイシンにより肺線維症を発症させたマウスにイレッサを投与し、溶媒単体投与群と比較してその経過を観察したものであり、その結果、イレッサ投与群は、溶媒単体投与群と比較して、「より激しい繊維化を示」したというものであった（西甲E 8＝東甲G 6 訳文 p 6）。

前々項で検討したように、肺の異常な修復（リモデリング）は、Ⅱ型肺胞細胞と繊維芽細胞との陣取りにⅡ型肺胞細胞が破れて繊維芽細胞が勝った場合に生じるところ（西乙H 3 6 の 1＝東乙F 1 4 の 1）、Ⅱ型肺胞細胞の再生、増殖にはEGFRが関与していることから（西甲E 3＝東甲F 3、西甲E 5 4＝東甲F 6 5、西甲E 5 5＝東甲F 6 6）、イレッサがⅡ型肺胞細胞の再生、増殖を抑制する結果、繊維化がより進展するという永井実験の結果は、論理的に見ても一貫した結論であった。したがって、この実験結果は、イレッサのEGFR阻害作用によって、傷ついた肺の修復過程で間質性肺炎へと進展してしまう可能性のあることを実証する結果となった極めて重要な実験であった。

しかるに、被告会社は、この実験結果を2001年10月には入手していながら、永井教授らの学会における発表を阻止し、イレッサの承認まで、この実験結果を明らかにさせなかったのである（被告会社答弁書等）。

- 2 被告会社は、永井教授らの実験結果に対して、答弁書、準備書面などで縷々反論を試みているが、これに対する再反論は、これまでの原告準備書面のとおりである。特に、永井実験では実験動物が少ないとする点については、少ない動物数であっても統計学的に明らかな有意差が認められていたことからすれば、逆に、永井実験の結論の強固性が証明されたというべきである。また、永井実験の用量が高用量であったとする点については、高用量で反応を見ることこそ動物実験においてしかできないものであり、高用量であるから信頼に値しないなどという立論は、まさに「危険性は鋭敏に」の原則を踏みにじるものに他な

らない。

そして、被告会社は、イレッサ承認後になって、石井教授らの実験（西乙H 3 4の8＝東乙G 4 4の3，4）を持ち出して、永井教授らの実験結果を否定しようと躍起になっているが、実験デザインとして見ても、永井教授らの実験は21日間観察しているのに対して、何故か石井教授らは13日間しか見ておらず、永井教授らの実験の再現実験とは言い難い。前述のとおり、肺の繊維化過程では、Ⅱ型肺胞細胞と繊維芽細胞が「陣取り」を行うのであり、観察期間の差は結論に大きな影響を及ぼすとも考えられるのである。

そして、イレッサ市販後の状況を見れば、工藤証人が関与し、科学的に画期的な調査であつとされる専門家会議調査、プロスペクティブ調査、コホート内ケースコントロールスタディなどにおいて（西工藤証人主尋問調書＝東乙L 1 6 p 1 0 5），以下のとおり、既存の間質性肺炎等が予後悪化因子として指摘されている。

「I P F等の既存がゲフィチニブ投与におけるI L D発症の危険因子の可能性が否定できない」（西丙L 2＝東丙D 4 専門家会議報告書 p 7・9 項）

「特発性肺線維症等の既存あり」（3～5行目）が有意差をもってI L Dの予後を悪化させる可能性のある因子として示唆された（西丙L 2＝東丙D 4 専門家会議報告書 p 6・2 項）

「本剤投与時に間質性肺疾患を合併している症例」（2行目）で、I L Dの発現率が高くなることが示唆された（丙C 2＝東D 2 プロスペクティブ調査 p 3・3）

「既存の間質性肺炎」がI L D発症の危険因子とされた（甲C 4＝東甲D 7 コホート内ケースコントロールスタディ報告書 p 3の主要評価項目の8行目以下）

このように、いずれにおいても永井実験が示すのと同様に、イレッサはⅡ型肺胞細胞の増殖、分化を阻害する結果、繊維芽細胞の増殖を促進してしまうという結論に親和性のある結果が出ている。こうした市販後の少なくとも我が国における調査結果は、石井教授の結論を支持するものではあり得ず、永井教授の実験結果を支持するものとなっているのである。

また、被告らは、青柴医師が再現実験に失敗したなどともするが（西丙E 32＝東甲L 107）、同実験は、従前の永井教授の実験を前提として、イレッサ投与群にさらにステロイドを投与した時に線維症が改善するか否かを見たものである。結果として、各群の間に明確な有意差（p値0.05未満）が観察され得なかったというに過ぎず、対照群に比較してイレッサ投与群は線維症スコアが上であり、依然としてイレッサが線維症を増悪する傾向にある結果となっているのである（最終ページ）。

そもそも石井教授らの実験や青柴医師らの再現実験なるものは、イレッサ市販後になって報告されたものに過ぎず、本件では、イレッサ承認当時もしくは直後における被告らの責任を問題としているのである。その当時の知見としては、永井教授らの知見があったのみであり、且つ、その結論は上記の通り、結果的にも市販後の状況と一致している。

したがって、「危険性は鋭敏に」の原則からすれば、イレッサ承認当時に出されていた永井教授らの実験結果を極めて重視し、少なくともイレッサは重篤な肺障害を発症させる可能性が高いものとして、非臨床試験や臨床試験の結果を厳密に検討しなければならなかったことは余りに明らかである。

第5 臨床試験、副作用報告に見るイレッサの安全性の欠如

1 はじめに

原告らは、イレッサの安全性に関する主張のうち、イレッサの臨床試験及び副作用報告の評価に関する主張については、西原告第2準備書面、第3＝東原告準備書面（2）、第4（臨床試験等から明らかな安全性の欠如）（以下、「臨床試験等に関する準備書面」という。）、及び西原告第5準備書面、第2＝東原告準備書面（9）、第2（国内外臨床試験及び臨床試験外使用における急性肺障害・間質性肺炎の副作用症例）（以下、「副作用報告等に関する準備書面」という）で行ってきたところである。

その後の原告らの立証並びに証人尋問の結果によって、既に行ってきた原告らの上記主張が裏付けられるとともに、それ以外にも上記準備書面では主張していなかった事実等も明らかとなり、これらはイレッサの安全性欠如とこれに対する被告会社及び被告国の予見可能性を十分に裏付けるものであった。

以下では、まず、イレッサの臨床試験及び副作用報告を評価する前提として特に注意すべき事項として、①臨床試験における有害事象の意味と重要性、②副作用報告におけるEAPの重要性について述べた上で、③イレッサの臨床試験、副作用報告に基づく安全性評価、すなわちイレッサの臨床試験及び副作用報告から致命的な急性肺障害・間質性肺炎の副作用が発生することは十分に予見可能であり、イレッサの安全性が欠如していたことについて、あらためて整理して主張する。

2 臨床試験における有害事象の意味と重要性

（1）有害事象の意味

イレッサの臨床試験の結果を評価するにあたっては、臨床試験における有害事象の意味とその重要性を正しく理解する必要がある。

有害事象とは、「医薬品が投与された患者または被験者に生じたあらゆる

好ましくない医療上のできごと」とされ、このうち当該医薬品との因果関係が否定できないものを副作用と言う（西丙D3＝東丙H3 p 1932～1933）。したがって、副作用と区別される場合の有害事象の意味としては、医薬品との因果関係が否定できるものという意味になる。

しかし、有害事象の本来の意味がそうであったとしても、臨床試験等で報告された有害事象が全て医薬品との因果関係が否定できるかと言えばそうではない。なぜなら、治験担当医師が有害事象として報告したものであっても、その中には本来副作用とされるべきものが含まれている可能性があるからである。

原告らは、既に臨床試験等に関する準備書面や濱六郎医師の意見書（西甲E25＝東G31）等において、イレッサの臨床試験の結果、有害事象として報告された症例の中に、本来副作用とされるべき症例が数多く含まれていることを指摘した。

これに対して、被告らは、臨床試験における有害事象報告は、GCPに基づいて治験担当医師が治験薬との因果関係がないとして報告したものであり、その報告は信頼できる等と反論している。

しかしながら、上記のとおり、治験担当医師からの有害事象報告は、それだけで医薬品との因果関係が否定できることを結論付ける意味を持つものではなく、治験担当医師の判断に重きを置く被告らの主張は失当である。

すなわち、以下に述べるとおり、個々の治験担当医師の判断には限界があり、有害事象か副作用かの判断は、個々の治験担当医師が判断できるものではなく、治験全体の結果やその他の情報も総合して最終的な判断がなされるべきものであり、GCPや医薬品承認制度自体もそのことを予定しているからである。

（２）治験担当医師の判断には限界があること

この点まず、個々の治験担当医師は、各医師が扱う症例数が少なく、数少ない症例だけを見て因果関係の有無を判断するのは困難であるという限界がある。

特に頻度の低い有害事象の場合、個々の治験担当医師がその有害事象に遭遇する確率は極めて低く、ほとんどの治験担当医師がそのような頻度の低い有害事象には全く遭遇しないか遭遇したとしても1例程度に過ぎないことになる。そうすると、治験担当医師がいくら経験豊富であったとしても、特に未承認の新薬のように未知の副作用が起り得る場合に、そのような頻度の低い未知の有害事象について治験薬との因果関係を判断することはほとんど不可能に近いと言ってよい。

この点については、帝京大の内藤教授が「臨床試験のクオリティ」という表題の講演の中で、「投薬中に出現したいわゆる adverse event が真の副作用、すなわち adverse reaction であるかどうかの判断は、特に治験のように各医師が少人数についての adverse event しか観察しない時には、多くの場合非常に困難であると思われる…」と述べており（西甲F35＝東甲F58 p51）、また濱証人も、個々の治験担当医は扱う症例も少なく、頻度の少ない有害事象に遭遇する確率も低いので、1例1例をだけを見て関連性の有無は判断できない旨証言している（西濱証人主尋問調書＝東甲L102 p54～55）。更に、被告側証人である福岡証人も、治験担当医師や治験責任医師がいくら経験豊富であっても、担当している患者だけを見て有害事象と治験薬との関連について判断するには限界があることを認めている（西福岡証人反対尋問調書＝東丙G58 p64）。

（3）有害事象か副作用かの最終的な判断

このように、治験担当医師の判断に限界がある以上、治験担当医師が治験薬との因果関係が否定できる有害事象として報告されたものであっても、そ

れだけで治験薬との因果関係を否定してはならず、最終的に臨床試験の全ての結果やその他に得られた副作用情報等も総合して因果関係の有無が判断されなければならない。

この点について、福島証人は、有害事象について因果関係を簡単に判断してはならず、全部カウントしてそれが本当に薬によるものかどうかをあと解析する必要がある、その時点で即断して関係あるなしを判断してはならない旨証言しており（西福島証人主尋問調書＝東甲L95 p8）、また、濱証人も、有害事象は、個々の医者が治験薬によるものかどうかの可能性を完全に否定することはできないのであり、臨床試験全体が終わって全体として見てもう一度検討した上で関連性を検討する必要がある旨を証言している（西濱証人主尋問調書＝東甲L102 p27）。

(4) 個々の治験担当医師の判断が最終判断でないことはGCPや医薬品承認制度自体が予定している

このように、個々の治験担当医師の判断が最終判断ではなく、最終的にあらゆる情報を総合して因果関係の判断がなされるべきことは、以下に述べるとおり、GCPや医薬品承認制度自体が予定していることでもある。

すなわち、厚生省（当時）は、GCPにおいて作成が義務づけられている治験総括報告書について、その構成と内容に関するガイドライン（西乙D5＝東乙H5）を定めているが、その「12. 安全性の評価」という項目中には、「…最後に、重篤な有害事象及び他の重要な有害事象を明確にすること。これは、通常、薬剤との関連が明確であるかどうかにかかわらず、有害事象のために試験完了前に脱落又は死亡した患者を十分に調べることにより検討される。」と記載されている（西乙D＝東乙H5 p15）。これは、治験担当医師が薬剤との関連を否定している有害事象であっても、有害事象のために試験から脱落又は死亡した事例は十分に調べた上で最終的に薬剤との因果

関係が判断されるべきことを意味している。

また、同ガイドラインの「12.2.2 有害事象の表示」の項目には、「…この表では、有害事象を薬剤の使用と少なくとも関連があるかもしれないと考えられる事象と、関連なしと考えられる事象に分類してもよいし、他の適当な因果関係分類（例えば、関連なし、関連があるかもしれない、おそらく関連あり、明らかに関連あり）を用いてもよい。」「このような因果関係の評価を用いた場合でも、関連性の有無の評価に関係なく、併発症と考えられる事象も含む全ての有害事象を表に含めること。当該治験又は安全性に関するデータベース全体をさらに分析することは、有害事象が薬剤に起因するか否かを明らかにすることの助けになることもある。」と記載されている（同 p 16）。これは正に、個々の治験担当医師の判断が最終判断ではなく、治験全体または安全性に関する全てのデータを分析する中で、最終的に有害事象と薬剤との関連性の有無が判断されなければならないということが示されているのである。

更に、厚生省（当時）が発出した「治験中に得られる安全情報の取扱いについて」（平成7年3月20日厚生省薬務局審査課長通知）（西丙D3＝東丙H3）のQ&Aには、「治験依頼者が単独で因果関係の評価ができると考えてもよろしいのか、また、因果関係の評価に際しての治験担当医師と治験総括医師、治験依頼者の関わりについて説明願いたい。」との質問に対して、「発現した事象と治験薬との因果関係は、基本的には実際に治験を実施している治験担当医師によって評価がなされるべきである。しかし、治験担当医師により因果関係が否定された事象でも、治験依頼者が先行する治験や実施中の治験の他施設での情報等を考慮した際に因果関係が疑われる等の状況にある場合には、当該治験担当医師や治験総括医師等とも相談の上で因果関係の再評価を行っていただきたい。」との回答がなされており（西丙D3＝東丙H3 p 11）、ここでも因果関係の判断は、個々の治験担当医師の判断が

最終判断ではなく、治験全体や他の情報を持つ治験依頼者において適切に再評価されるべきことが予定されていることが示されている。

そして、イレッサの承認審査を担当した平山証人も、安全性審査における審査の手順として、まず、発生頻度の高い有害事象そのあと重篤なものを見ていき、それが終わった後に副作用を見ていくということ、有害事象については薬剤との関連を問わず見ていくということを述べており（西平山証人主尋問調書＝東甲L197p21, 22）、この点にも、有害事象と医薬品との因果関係については、治験担当医師や医薬品メーカーの判断が最終判断ではなく、承認審査段階でも更に因果関係の検討が行われるべきとの認識が示されている。

このように、有害事象と医薬品との因果関係の判断は、個々の治験担当医師の判断から、治験全体を統括しより多くの情報を持つ治験依頼者による判断、更に審査当局による判断へと段階的に行われることが予定されており、決して個々の治験担当医師の判断を鵜呑みにしてはならないのである。

（5）有害事象の重要性

以上より、たとえ治験担当医師により有害事象（治験薬との関連なし）として報告されたものであっても、それだけで治験薬との因果関係を否定することはできず、治験依頼者である医薬品メーカ及び審査当局は、あらためて全ての有害事象及び副作用情報等を総合して当該有害事象と薬剤との関連性の有無を慎重に判断しなければならず、その結果、薬剤との因果関係を完全に否定することができないと判断される場合には、全て副作用として取り扱わなければならないのである。

その意味で、有害事象は「副作用のシグナル」として十分に検討されなければならない、これを過小評価することは許されないのであり（西福島証人主尋問調書＝東甲L95p11～12）、特に有害事象死亡例は副作用死亡例

と同じように重視されなければならないのである（西濱証人主尋問調書＝東甲L102p27）。

3 副作用報告におけるEAPの重要性

(1) EAPによる副作用報告

イレッサは、承認前、臨床試験以外にEAP（Expanded Access Program 拡大アクセスプログラム）において使用され、それらEAPにおける副作用症例も数多く報告されている。

EAPは、米国において、重篤又は致死性疾患の患者で、臨床試験に不適格かつ他に治療の選択肢を有しない者に対して未承認薬の使用を認める制度であり、米国食品医薬品局（FDA）と医療機関内の倫理審査委員会（IRB）による承認、監視の下で実施される（西甲J7＝東甲I6）。

イレッサにおけるEAPは、英国アストラゼネカ社が、通常のイレッサの治験に参加できない患者を対象にイレッサ単剤の安全性評価を目的として実施したものである（乙B13の3の1等参照）。

原告らは、副作用報告等に関する準備書面や濱証人の意見書（西甲E25＝東G31）等において、これらEAPを含むイレッサの副作用報告の中に、数多くの急性肺障害・間質性肺炎の発症例が含まれており、その中には副作用死亡例も相当数存在していたことを指摘し、これら副作用報告を見ればイレッサによる致命的な急性肺障害・間質性肺炎の副作用が発生することは容易に予見可能であったことを主張した。

これに対して被告会社は、EAPはGCPに準拠して行われておらず、その副作用報告は信用性に乏しい等の反論を行っている。

しかしながら、以下に述べるとおり、イレッサにおけるEAPの副作用報告は、EAPがGCPに準拠して行われていないとの理由で軽視することは

許されず、イレッサの安全性を評価する上で極めて重要な資料（情報）として重視されなければならないものである。

（２）審査資料としての意味とその重要性

EAPにおける副作用報告を含むあらゆる副作用情報（重篤で予測できない副作用）は、全て治験関係者及び規制当局への緊急報告の対象とされている（薬事法80条の2第6項，同施行規則第66条の7，GCP省令20条2項，西乙D14＝東乙H14，西丙D7＝東丙H7 p 20，西丙D3＝東丙H3 p 1934～1935参照）。

このように、臨床試験における副作用だけでなくあらゆる副作用情報について報告義務が課せられている理由は、安全性情報という危機管理的な側面のほかに、治験実施機関、治験依頼者、審査当局の各段階において、より広い情報源に基づいて治験薬の安全性評価を行うためである。

したがって、EAPにおける副作用報告についても、イレッサによる副作用の有無の判断やイレッサの安全性評価を行う上で極めて重要な資料として位置付けられなければならない。

この点について、国立医薬品食品衛生研究所医薬品医療機器審査センターの福山圭一氏は、平成11年1月29日第1回新薬審査部門定期説明会における講演で、「有害事象は審査資料としても貴重な情報です。安全性情報という危機管理的な側面以外に、もう一つの側面として審査資料としての側面が当然あると思います。有害事象の報告というのは、治験薬の使用例で、しかも有害事象という問題のあった例がリアルタイムに私どもに報告が出てきます。審査センターで行っている審査というのは書面審査ですが、いわゆる審査資料は十分な時間をかけて企業の方で検討されて、きちんと整理されたものとしてでてくるわけです。そういう資料にもとづいて審査するということを基本にしているわけですが、そういう審査資料を補完する観点からもリ

アルタイムに有害事象情報が入ってくるというのは非常に重要な資料になりうるのではないかということで、審査資料的な価値も大きいのではないかと考えられるということです。」と述べ、治験外の副作用情報の審査資料としての価値の重要性について述べている（西乙F 2＝東乙F 1 p 182～183）。

また、イレッサの承認審査を担当した平山証人も、EAPなど治験外の症例データを確認する意味について、「より広くいろんな症例に当たって、どういう副作用が起こったのかというのを拾い上げるということも必要になってきます。その意味で、申請資料以外の部分で、見るべき副作用があれば、それを適宜取り出すということは非常に重要ですし、それから今回の治験の申請資料の中で見られたものが、ほかのところでも発生しているということになりますと、より確実性が高くなるというふうに判断できるということですので。」と証言し（西平山証人主尋問調書＝東甲L 197 p 26）、前記福山氏同様、EAPを含む治験外の副作用情報の審査資料としての重要性を認めている。更に、被告側証人である光富証人や工藤証人も、EAPの副作用情報の重要性を認めている（西光富証人反対尋問調書＝東乙L 24 p 29、西工藤証人主尋問調書＝東乙L 16 p 53～54、西工藤証人反対尋問調書＝東乙L 17 p 80）。

（3）EAPのデータは実地臨床で使用される場合に近い情報であること

更に、EAPによる副作用報告は、EAPが厳格な適格基準を定めた臨床試験とは異なり、実地臨床に近い条件で使用されることから、むしろ臨床試験における副作用報告以上に貴重な情報となり得るものである。

すなわち、医薬品は一旦承認されれば、多くの場合臨床試験における適格基準を満たさない患者にも広く使用される。臨床試験では、医薬品の有用性を判断するため、比較的状态の良い患者が選定されるが、市販後はむしろ状

態の悪い患者等にも広く使用されるため、臨床試験では見られなかったような副作用が発生する危険性がある。したがって、むしろ市販後の条件に近い形で使用されるEAPにおける副作用情報は、そのような市販後の副作用を予測する上では臨床試験における副作用情報以上に貴重な情報となり得るのである。

この点につき、福島証人は、ある意味では理想的な条件設定の下で行われるアイデアルワールドである臨床試験と、それ以外の患者にも投与されるリアルワールドである実地臨床とでは非常に違っており、その意味で、EAPで得られたデータは、むしろ実地臨床で使うときに非常に役立つデータであり、臨床試験と同等あるいはそれ以上に重んじなければならぬと証言しており（西福島証人主尋問調書＝東甲L95 p17）、別府証人も同様の趣旨を述べている（西甲E39＝東別府主尋問調書p46）。

また、FDAの担当官であるパズドゥア氏も、2004年のASCOにおいて、「EAPは、患者に対し未承認薬を提供し、かつ、より大きくより非均質な群におけるさらなる安全性の情報を獲得する効果的なメカニズムである。」と述べている（西甲J7＝東I6）。

（４）GCPに準拠していないことが副作用情報としての信頼性を低下させるものではないこと

前記のとおり、被告会社は、EAPがGCPに準拠して実施されていないことから信頼性に乏しい等と主張しているが、被告会社の主張は、そもそも何故GCPに準拠していないことが信頼性を低下させるのか、その関連性や具体的根拠が全く明らかでない。したがって、このような被告の主張は、それだけで既に全く理由がないと言うべきである。

GCPは、医薬品の承認申請等の資料とするための臨床試験の実施の基準について定めたものであるが、これは被験者保護の趣旨に加え、医薬品の有

効性は、科学的・統計学的手法によって判定される必要があり、そのための資料は、厳格な基準に基づき適切に計画された臨床試験によって収集される必要があるとの趣旨に基づくものである。他方、医薬品の安全性については、厳格な参加制限があり比較的症例数も少なく限られた条件の下で行われる臨床試験のみから収集できる安全性情報には限界があることから、GCP自身においても、臨床試験に限らず全ての副作用情報が治験関係者に通知されることが義務づけられているとおり、あらゆる副作用情報が情報源を問わず収集されることが予定されているのである。この点については、被告側証人である工藤証人も、同様の趣旨を認めている（西乙E24＝東工藤証人反対尋問調書p52～53）。

このようなGCP本来の趣旨からすれば、EAP等の個別症例報告について、その症例がGCPに準拠して実施されているか否かは、専ら被験者保護や有効性評価において必要とされているものであり、安全性情報としての「質（信頼性）」に影響するものではないと言うべきである。

むしろ、EAPは、そもそもイレッサの安全性評価を目的として実施されているのであるから、これに基づく副作用情報は当然信頼すべき情報として取り扱われなければならない。

更に、EAPは、米国FDAや各医療機関内の倫理審査委員会（IRB）の監視の下、登録制をとり、一定の適格基準や除外基準が設けられ、一定水準以上の医療機関・医師の下において使用されるなど安全性評価に資する内容の情報が提供されるべき条件の下で実施されており、そうした点に鑑みると仮に被告らの立論を前提としても、EAPに基づく副作用情報は十分に信頼できるものと言うべきである。

実際、被告ら自身も、EAPの副作用報告について、その報告書の記載内容に基づいてイレッサによる間質性肺炎の副作用発症例であると認め、これを審査報告書や添付文書に反映させているのであり、これはEAPの副作用

報告の内容が信頼に値するものとして評価していたからにはほかならない。また、後述するとおり、本裁判において、被告らが間質性肺炎発症例と認めた以外の症例についても、副作用・感染症症例報告書の記載内容に基づいて、明らかに間質性肺炎の副作用発症例であることが複数の専門家証人により認められた症例も存在するのであり、この点からもイレッサのEAPにおける副作用報告は、副作用情報として十分に信頼に値することが実証されているとすべきである。

以上より、EAPの副作用報告については、GCP本来の趣旨やその具体的な副作用報告の内容を無視して、その症例がGCPに準拠していたか否かという形式的な点のみを問題にすることには全く意味がないとすべきである。

(5) EAPによる副作用報告の重要性

以上述べたとおり、EAPにおける副作用情報は、イレッサの安全性評価における審査資料として極めて重要な価値を有しており、市販後の副作用を予測する上では、臨床試験における副作用情報と同等ないしそれ以上に重視しなければならない情報である。そして、GCP本来の趣旨やEAPがイレッサの安全性評価を目的として実施されたプログラムであること等から、EAPにおける副作用情報が十分に信頼に値するものであることも明らかである。

したがって、EAPがGCPに準拠していないことをもってEAPにおける副作用情報を軽視してよいかのような被告らの主張は明らかに失当とすべきである。

4 イレッサの臨床試験、副作用報告に基づく安全性評価（致死的な急性肺障害・間質性肺炎の副作用発生の予見可能性）

(1) はじめに

ここまで、①臨床試験における有害事象の意味と重要性、②副作用報告におけるEAPの重要性について述べてきたが、これを前提に、イレッサの臨床試験及び副作用報告に基づくイレッサの安全性評価について、以下で論ずることとする。

この点に関しては、既に、臨床試験等に関する準備書面、及び副作用報告等に関する準備書面において、個々の副作用症例の検討を含め既に詳しく主張してきたところであるが、その後の証拠調べの結果明らかとなった事実等もあり、本項では、あらためて整理してこの点に間する主張を行うこととする。

以下では、便宜上、①臨床試験に基づくイレッサの安全性評価と②副作用報告に基づくイレッサの安全性評価とに分けて論ずることとし、前者においては、主にイレッサの臨床試験における有害事象死亡例の評価及び国内臨床試験における3例の間質性肺炎発症例の評価について論じ、後者においては、国内3症例以外のEAPを含む主に海外からの副作用報告における間質性肺炎発症例の評価について論じることとする。

(2) 臨床試験に基づくイレッサの安全性評価

ア 臨床試験における有害事象死亡例の評価

(ア) 従前の主張とその後の立証

イレッサの臨床試験における有害事象例の評価に関しては、従前、主に臨床試験等に関する準備書面において主張してきた。

ここでは主に、①イレッサの臨床試験において「病勢進行による」中止例、死亡例の割合が不自然に高く、この中には実際にはイレッサの影

響による中止例，死亡例が含まれている可能性があり，被告らはこれら中止例，死亡例を子細に検討すべきであったこと，②イレッサの臨床試験の有害事象例の中には，呼吸器系の有害事象が数多く出現しているが，これらの有害事象はイレッサの毒性によるものと考えても何ら矛盾はなく，特に有害事象による中止例のほとんどがイレッサとの「関連なし」とされているのは不自然であり，これら呼吸器系の有害事象例の中にはイレッサによる急性肺障害・間質性肺炎の発症例が含まれていた可能性が極めて高いこと，③イレッサの臨床試験における副作用死亡例は2例とされているが，実際には呼吸器系の有害事象死亡例のうち相当数がイレッサによる急性肺障害・間質性肺炎による死亡例である可能性が高いこと，等を主張していた。

この時点では，原告らは，イレッサの臨床試験における個別の有害事象例について詳しい検討まで行っていなかったが，その後，被告会社から承認申請資料概要の別冊（以下，「別冊」という）が提出され，個々の有害事象例についての臨床経過の一部が明らかとなった。

そこで，原告らは，濱証人に，これらイレッサの臨床試験における有害事象例（特に有害事象死亡例）について，別冊で明らかとなった臨床経過等を踏まえて，イレッサとの関連についての考察を依頼し，その結果を意見書（西甲E 2 5＝東G 3 1）として提出するとともに，濱証人に対する証人尋問で更にこの点を明らかにした（西濱証人主尋問調書＝東甲L 1 0 2 p 2 5～6 3）。

この結果，以下に述べるとおり，イレッサの臨床試験における有害事象死亡例のほとんどが，イレッサによる急性肺障害・間質性肺炎による副作用死亡例として分類すべきものであることが明らかとなった。

（イ）濱意見書（西甲E 2 5＝東G 3 1）の内容

a 臨床試験における有害事象死亡例

濱意見書では、主にイレッサの臨床試験における有害事象死亡例について、その臨床経過等に基づいてイレッサによる副作用と分類すべきか否かについて考察されている。その内容の詳細は、同意見書 p 26～p 49 に記載されているので詳細は省略するが、各臨床試験における有害事象死亡例（34例）の概要は下記のとおりである（同意見書 p 49～51）。

第 I 相 (1839IL/0005)

- ① 300 mg群：急性呼吸窮迫症候群で死亡（消化器症状などあり 2 サイクル 42 日で中止。他に発熱，呼吸困難の後，67 日死亡）
- ② 400 mg群：肺炎で死亡（呼吸器症状・感染症が悪化。70 日で中止。90 日死亡）

第 I / II 相 (1839IL/0011)

- ① 150 mg群：鼻出血などで死亡（24 日鼻出血，30 日中止，35 日口から出血死）
- ② 150 mg群：重度食欲低下で死亡（23 日食欲低下，錯乱，幻覚，28 日中止，43 日死亡）
- ③ 225 mg群：無呼吸により死亡（22 日嘔吐，24 日中止，26 日無呼吸で死亡）
- ④ 225 mg群：肺炎・呼吸不全で死亡（皮膚・呼吸器症状，関節，神経の一連の症状あり，143 日で中止，肺炎発症 10 日後の 172 日目死亡）
- ⑤ 225 mg群：肺血栓塞栓症・呼吸困難・心不全で死亡（2 サイクルで終了 13 日目）
- ⑥ 225 mg群：肺炎と無呼吸で死亡（27 日剥奪性皮膚炎で 29 日中止，16 日後肺炎・無呼吸，2 日後死亡）
- ⑦ 300 mg群：喉の出血で死亡（皮膚・消化器症状後 56 日で終了，68 日喉出

血死)

⑧ 400 mg群：急性呼吸窮迫症候群で死亡（7日喀血・呼吸困難で中止2日後死亡）

⑨ 400 mg群：無呼吸・呼吸不全で死亡（皮膚，呼吸器，口内など一連の症状の後，37日無呼吸，44日死亡）

第Ⅰ/Ⅱ相(1839IL/0012)

① 800 mg群：急性腹症で死亡（皮膚，眼，口，声，消化器症状，41日腹痛・心停止で急死）

② 300 mg群：急性呼吸窮迫症候群で死亡（皮膚，知覚，尿路，消化器症状の後32日呼吸窮迫症候群となり中止，22日目死亡）

③ 225 mg群：肺血栓性塞栓症で死亡（2日下痢，19日肺血栓塞栓症で死亡）

第Ⅱ相(1839IL/0016日本人を含む)

① 500 mg群：肺炎で死亡（副作用に分類された例：消化器，性器，神経，皮膚症状などの後，59日肺炎で中止，即日死亡）

② 250 mg群：肺炎・無呼吸で死亡（関節，血糖，皮膚，呼吸器症状の後64日肺炎4日後死亡）

③ 250 mg群：重篤な無力症で死亡（6日無力症，16日より中止，中止日に死亡）

④ 250 mg群：喀血で死亡（咽頭炎，感染症などに続いて68日目喀血，即日死亡）

⑤ 250 mg群：肺炎で死亡（下痢，発疹などの後，14日肺炎，19日中止，30日死亡）

第Ⅱ相(1839IL/0039)

① 500 mg群：呼吸困難で死亡（消化器，皮膚，尿路症状の後100日呼吸困難，101日中止，106日死亡）

- ② 500 mg群：急性呼吸窮迫症候群で死亡（呼吸器，消化器，無力症，皮膚，発熱の後 12 日急性呼吸窮迫症候群，15 日まで使用，当日死亡）
- ③ 250 mg群：電撃的急性肺傷害で死亡（2 日目発熱，無呼吸，2 日目で中止，5 日目死亡）
- ④ 250 mg群：肺炎・無呼吸・低血圧で死亡（消化器，全身症状に引き続き 66 日肺炎，その後無呼吸，71 日で中止，90 日死亡）
- ⑤ 250 mg群：肺炎と敗血症で死亡（消化器，呼吸器，皮膚，全身症状の後 30 日肺炎・敗血症。32 日死亡）
- ⑥ 250 mg群：心筋梗塞で死亡（下痢増強，皮膚症状，呼吸器症状などあり，111 日中止，その後うつ発症，136 日心筋梗塞で即日死亡。併用薬関係ありか）
- ⑦ 250 mg群：心筋梗塞・DICで死亡（皮膚，消化器，呼吸器症状などあり，血痰の後，63 日心筋梗塞発症し中止，低酸素血症，不整脈，DICなどで 70 日死亡）
- ⑧ 500 mg群：急性呼吸窮迫症候群・うっ血性心不全（肺水腫？）で死亡（初日より消化器症状，呼吸困難が増強，13 日死亡）
- ⑨ 500 mg群：肺出血で死亡（副作用に分類された例：3 日目喀血，消化器症状あり，11 日死亡，死亡まで使用）
- ⑩ 250 mg群：電撃的呼吸不全で 4 日目に死亡（開始翌日低酸素血症，4 日目死亡）
- ⑪ 250 mg群：呼吸困難，脳血管障害で死亡（皮膚，中枢・末梢神経，眼，筋肉症状，頻脈，嘔声，グレード3呼吸困難など一連の症状の後，79 日脳血管障害 86 日まで使用，90 日死亡）

有害事象死亡例の集計に含まれなかった 500 mg群の重篤な有害事象死亡（詳細不明）

- ⑫ 500 mg群：肺炎で死亡
- ⑬ 500 mg群：急性呼吸窮迫症候群で死亡
- ⑭ 500 mg群：呼吸困難で死亡
- ⑮ 500 mg群：呼吸困難で死亡

以上の有害事象死亡例について、濱意見書は、そのほとんどがイレッサとの関連を否定できない有害事象死すなわち副作用死亡例と分類すべきであったとしている（同意見書 p 26～49）。

b 有害事象死亡例についての考察

これら有害事象死亡例について、濱意見書では、臨床試験の結果についての考察（同意見書 p 51以下）の中で、次のとおり述べている。

まず、これらの有害事象死亡例とされた多くの症例を観察した結果、次のような有害事象死の発症パターンが認められる。すなわち、皮膚、消化器、口・目などの粘膜、呼吸器、肝臓、代謝臓器、尿路生殖器粘膜、心、血管内皮の障害など、種々の臓器の傷害に伴う症状が多彩に出現し、重篤例は多くの場合、急性呼吸傷害が重篤化して死亡することが多い（同意見書 p 51）。

また、イレッサの作用機序、とくに6か月の反復毒性試験の結果、相前後して皮膚、粘膜、血管、臓器の多彩な症状が出現していることから考えて、死亡につながる有害事象としての急性肺傷害や出血など呼吸器系の有害事象死とイレッサとの関連は濃厚な例が多いと見る必要がある（同意見書 p 51～52）。

更に、一般的に考えてこれほど多くの有害事象死（各臨床試験において数%～十数%）をイレッサと無関係と考えることはできず、有害事象死として考えられる他の原因（併用薬剤の影響による死亡、癌以外の合併症として本人がもともと有していた心臓病や腎臓病、脳卒中などの合併症による死亡、本人がもともともっていなかった新たな病

気、例えばインフルエンザからの肺炎等に罹患しての死亡)が、数%～十数%の死亡例をもたらすことは原因となることは想定し難いことから、各臨床試験に現れた呼吸器系の有害事象死亡例を含む、多くの有害事象死亡例については、イレッサとの因果関係が否定できない(というよりむしろ相当な関係がありうる)有害反応(副作用)と評価すべきであった(同意見書 p 52～53)。

c 新規医薬品承認にあたっての安全性評価の姿勢

そして、濱意見書は、イレッサのような新規物質の安全性評価を行うにあたって、臨床試験をとりまとめる医薬品メーカー及び承認審査を行う審査当局の心構えとして、次のとおり述べている。

すなわち、臨床試験を取りまとめる医師(医薬品メーカー)、承認申請概要審査に当たる医師(審査当局)は、単に承認申請概要の記載をみるだけでなく、承認申請概要の記載からイレッサとの関連が否定された有害事象死亡例を1例1例点検する必要がある。「有害事象」は本来、試験物質との関連が完全には否定しきれない例としてとらえるべきである。臨床試験では、未知の害が現れうるのであるから、一見関連がなさそうに見えても、全く無関係とは言えず、だからこそ有害事象として扱い、最終的には厳密にはプラシーボ対象を設けた臨床試験で、出現頻度の差を見て関連性の有無を検討するのである。治験担当医師による個々の有害事象例を、関連が否定できない(副作用)か、完全に否定できる有害事象に分類する方式そのものが、未知の安全性(危険性)の評価をするための臨床試験、特にその初期のⅠ相やⅠ/Ⅱ相試験では不適切である。「有害事象」と試験物との関連を頭から否定するという考え方そのものが根本的に誤っている。イレッサの臨床試験においても、まず、これらの有害事象例をイレッサとの関連が否定できない有害反応(副作用)としてとらえ、次いで動物

実験における肺病変と十分に対比して類似していることを問題とし、EGFR阻害作用から推察しうる病変として十分にありうる病変であると考察しなければならなかった（同意見書 p 5 2）。

d なお、イレッサの各臨床試験における死亡例の割合と、そのうち有害事象死と病勢進行死との割合を示したものが以下のとおりである。

	死亡例（有害事象死／病勢進行死）		
第 I 相 (1839IL/0005)	12 (2/10)	／64例	18.8 (3.1/15.6) %
第 I 相 (V1511)	0 (0/0)	／31例	0.0 (0.0/0.0) %
第 I / II 相 (1839IL/0011)	11 (9/2)	／69例	15.9 (13.0/2.9) %
第 I / II 相 (1839IL/0012)	16 (3/13)	／88例	18.1 (3.4/14.8) %
第 II 相 (1839IL/0016)	35 (5/30)	／209例	16.7 (2.4/14.4) %
第 II 相 (1839IL/0039)	49 (15/34)	／216例	22.7 (5.1/15.7) %
計	123 (34/89)	／677例	18.2 (5.0/13.1) %

この点について、濱意見書では、イレッサの臨床試験においては、有害事象死亡例以外にも病勢進行死とされた症例が多数あり、これらについてもイレッサとの関連について検証が必要であったとして、次のとおり述べている（同意見書 p 5 3）。

すなわち、これら死亡例の多くは病勢進行による死亡とされているが、臨床試験においては、各臨床試験毎に、その評価に必要な観察期間中の生存が見込まれる患者を選定したはずであり（I相あるいはI／II相試験では12週間：84日は生存が見込めることとされた）、にもかかわらずイレッサ使用中および中止30日以内の死亡例が20%近くあり、これらの死亡例の中に、病勢進行ではなく、イレッサが関与した副作用死亡例がなかったかどうか、検証が必要である。

特に、がん患者を対象としたはじめてのⅠ相試験（0005試験）やⅠ／Ⅱ相試験（0012試験）など早期の試験で病勢進行死の割合が高い。0005試験では64人中12人の死亡例中10人（死亡例の83%）が病勢進行死とされた。0012試験でも、88人中16人の死亡例中13人（死亡例の81%）が病勢進行死とされた。Ⅰ／Ⅱ相試験でも、臨床試験のなかで、0011試験以外、すべて病勢進行死の割合が高いが、その割合の低い0011試験（18%）がむしろ本当の姿であると考えられる。日本の臨床試験では、有害事象死が133人中1人とされているが、それ自体疑問であり、副作用死亡例を除外していないかどうか検証が必要である。

（ウ）まとめ

既に述べたとおり、イレッサの臨床試験を検討するにあたっては、イレッサのEGFR阻害というドラッグデザインから予測される肺毒性、及び、イレッサの非臨床試験で得られた毒性所見を前提に、慎重かつ厳密に吟味されなければならない。

また、前記のとおり、臨床試験における有害事象については、治験担当医師の判断は最終判断ではなく、たとえ治験担当医師が治験薬との因果関係を否定したものであっても、治験依頼者たる医薬品メーカー及び審査当局は、治験全体の結果やその他の情報を総合して有害事象と治験薬との関連を判断しなければならない。したがって、治験担当医師の判断を鵜呑みにして有害事象として報告された症例について、最初から治験薬との因果関係を否定するという考え方は誤りである。

濱意見書の上記cに述べられていることも、表現の濃淡はあるが、これと同じ趣旨である。

そうすると、上記濱意見書で述べられているとおり、治験全体を観察して、全ての有害事象死亡例を1例1例検討し、更にイレッサの作用機

序や非臨床試験の結果を合わせて判断すれば、イレッサの臨床試験における有害事象死亡例のほとんどについてはイレッサとの関連が否定できない副作用死亡例と分類すべきであったのであり、医薬品メーカーたる被告会社及び審査当局である被告国は、そのように判断すべきであった。

この点については、福島証人も、その意見書及び証言の中で、臨床試験における肺に関する有害事象死のデータは、イレッサによる急性肺障害・間質性肺炎の副作用の発生を予測させる十分に注意すべきデータであったにもかかわらず、被告らはこのような「副作用のシグナル」を過小評価したと述べている（西甲E23＝東L29 p4，西福島証人主尋問調書＝東甲L95 p11～13）。

以上より、イレッサの臨床試験において現れた有害事象死亡例は、そのほとんどが副作用に分類されるべきであったものであり、少なくともイレッサによる致死的な急性肺障害・間質性肺炎の副作用の発生を予測させるに十分なデータであったと言うべきである。

イ イレッサの国内臨床試験における間質性肺炎発症例の評価

（ア）国内臨床試験における間質性肺炎発症例と発症頻度

イレッサの臨床試験において、国内臨床試験で3例の間質性肺炎の副作用発症例が現れた（乙B12の3～5）。

既に述べたとおり、抗がん剤における間質性肺炎の副作用は、一旦発症するとステロイド剤投与による治療が効を奏さなければ多くの場合死に至る極めて重大な副作用である。

国内臨床試験における間質性肺炎副作用の発症率は、少なくとも2.3%（133例中3例）と明確かつ高頻度であり、この頻度は十分に注意すべき数字であった（西甲E23＝東L29 p4，西福島証人主尋問調書＝東甲L95 p13）。

(イ) 国内 3 症例は極めて重篤な症例であったこと

上記国内 3 症例の概要は以下のとおりである。

① 乙 B 1 2 の 3 の症例

被験者等略名	T. M. 男 64 歳
医療機関所在地	神奈川県
副作用・感染症名	間質性肺炎，呼吸困難（生命を脅かす，入院期間の延長を要する，医学的措置を要する事象）
イレッサ投与期間	開始 2000/12/06 終了 2000/12/22
主な治療経過等	2000/12/06 治験薬投与開始。 2000/12/22 呼吸困難出現。 カルベニン（1g/日），プレドニン（40 mg/日），ネオフィリン（500 mg/日）の投与を行った。 胸部 CT にて PD と共に両肺下葉の間質影を認めた。 PD の判定にて試験中止。 2000/12/23 呼吸困難が急速に増悪。胸部 X 線写真にて間質性肺炎の所見あり。 プレドニンに代えてソルメドロール 1g/日（25 日まで続行）。 カルベニンに加えてミノマイシン 200 mg の投与を行った。 2000/12/24 午前 3 時より挿管・人工呼吸管理開始。 2001/01/29 死亡

担当医のコメント 臨床経過と気管支肺胞洗浄の結果より、薬剤性の急性間質性肺炎が疑われる。DLST 検査による確診は得られなかったが、薬剤投与時期と有害事象発生との関係より、治験薬が原因薬剤である可能性が高い。

呼吸困難については臨床的に改善を認めたものの、薬剤性として矛盾のない間質性肺炎が組織学的には死亡時も残存していたと考えられる。

② 乙B12の4の症例

被験者等略名	M. I. 男 年齢不明
医療機関所在地	神奈川県
副作用・感染症名	間質性肺炎，低酸素症（入院を要する事象）
イレッサ投与期間	開始 2000/12/12 終了 2001/03/05
主な治療経過等	2000/12/12 治験薬投与開始。
	2001/03/06 食欲不振，全身倦怠感が強くなったとのことで治験薬2週間休薬を希望される。
	2001/03/08 胸部 CT 上，右肺下葉に間質性肺炎を確認。
	2001/03/09 間質性肺炎治療のため入院。 間質性肺炎に対し，ソルメドロール 1g/回，3日間を開始。 治験中止。
	2001/04/13 死亡

担当医のコメント 休薬直後に確認できた間質性肺炎であるが、治験

薬との関連はあると思われる。

低酸素血症については，治験薬との関連性は多分あり。

③ 乙B12の5の症例

被験者等略名	Y. M. 女 62歳
医療機関所在地	徳島県
副作用・感染症名	間質性肺炎（生命を脅かす事象）
イレッサ投与期間	開始 2000/10/16 終了 2001/10/25
主な治療経過等	2000/10/16 ZD1839 投薬開始。 2001/10/25 呼吸困難感出現。 間質性肺炎のため入院。 間質性肺炎は治験薬を中止しても改善せず。 2001/10/26 メチルプレドニゾロンパルス療法 1g/日（～10/28）。 2001/11/07 治験より脱落。 2001/01/29 死亡
担当医のコメント	偶発症の可能性も考えられるが，患者は既に死亡されており，本剤との関連性は否定できない。

上記国内3症例は，被告らにおいていずれも死亡との関連は否定されているが，それ自体極めて重篤な症例であったことに加え，後述するとおり，少なくとも国内1症例目については，イレッサ投与が死亡に与えた影響を完全に否定することができないと評価すべき症例であった。

すなわち，国内臨床試験における間質性肺炎発症例3例のうち2例

(乙B12の3, 乙B12の5)が「生命を脅かす」副作用として報告されており, これは当時の厚生省薬務局審査課長通知「治験中に得られる安全性情報の取扱いについて」(西丙D3=東丙H3 p1933)において「死亡」を除くと最も重篤な段階にあたるという分類がなされている。そして, この「生命を脅かす」という意味は, 「その事象が起こった際に患者が死の危険にさらされていたという意味であり, その事象がもっと重症なものであったなら死に至っていたかもしれないという仮定的な意味ではない」(西丙D3=東丙H3 p1933 欄外)とされており, これはイレッサによる間質性肺炎によって, 現に「患者が死の危険にさらされていた」ことを意味するのである。

また, 福島証人は, 乙B12の4の症例についても, イレッサにより間質性肺炎が発症し, ほぼ1ヶ月後に死亡している等の経過やその不明点を踏まえて, イレッサと死亡との関連性を完全に否定すべきでない旨を証言しており(西甲E41=東福島証人主尋問調書p10), 間質性肺炎と死亡との関連性については十分な検討が必要な症例であった。

更に, 上記3例は, その臨床経過から分かるとおり, いずれも間質性肺炎に対するステロイドパルス療法が行われるほど重篤な症例であった

(西丙H46=東丙G72 p53, 西甲P103=東丙G51の1 p16, 西甲E41=東福島証人主尋問調書p5~7, 西工藤証人反対尋問調書=東乙L17 p73~74)。

特に, 1例目(乙B12の3)については, ステロイドパルス療法が奏功せず, 人工呼吸器管理まで行われたCTCグレード4に該当する極めて重篤な症例であり(西甲D7=東甲L200 p24, 西甲E41=東福島証人主尋問調書p7~8), 剖検の結果, 間質性肺炎が組織学的には死亡時まで残存していた転帰「未回復」の症例であり, 間質性肺炎の中でも極めて予後が悪いAIP(DAD)型であった可能性が高いこ

とが判明していた（西工藤証人反対尋問調書＝東乙L 1 7 p 7 7～7 8，西甲H 4 1＝東G 7 9 p 1 4）。そして，このような臨床経過や剖検の結果を踏まえると，イレッサ投与が死亡に与えた影響を完全に否定することができない症例であった（西甲E 4 1＝東福島主尋問調書 p 8～9，西甲E 4 0＝東別府反対尋問調書 p 6 8～6 9）。

（ウ）まとめ

以上より，国内臨床試験における間質性肺炎発症例についての検討結果から，まず，その副作用発症率（少なくとも2. 3%）は明確でありかつ高頻度であり，いずれも極めて重篤な症例で，特に1例目（乙B 1 2の3）については，CTCグレード4，AIP（DAD）型，転帰「未回復」で，死亡との関連を完全に否定することはできないと評価すべき症例であった。

したがって，これら国内臨床試験における間質性肺炎発症例は，それだけでイレッサによる間質性肺炎が極めて重篤で致死的な副作用であることを予見させるに十分であったと言うべきであるが，少なくとも，前述した臨床試験における多数の呼吸器系の有害事象死亡例や後述する海外からの間質性肺炎による副作用死亡例の存在を合わせて考えれば，イレッサによる間質性肺炎の副作用が極めて重篤かつ致死的なものであることは容易に予見可能であったと言うべきである。

（3）副作用報告に基づくイレッサの安全性評価

ア 国内3症例以外に被告国が間質性肺炎発症例として認めた7症例

イレッサの承認前に，上記国内臨床試験における3例の間質性肺炎発症例以外に，被告会社から，第Ⅲ相臨床試験（INTACT）やEAPにおける間質性肺炎の副作用発症例が数多く被告国に対して報告されている。

このうち、被告国が間質性肺炎発症例として認めたものは7例であり、被告国が審査報告書（1）に「2002年4月時点で海外の4症例においても、間質性肺炎が報告されている」と記載した4例（乙B13の1～4）と、被告国が「審査報告書（1）の作成から承認までに報告された間質性肺炎として評価することが適当と判断される3例」と本件訴訟における準備書面で認めた3例（乙B14の1～3）である。

なお、上記7例のうち4例（乙B13の1、乙B13の3、乙B14の1、乙B14の3）はEAPにおける症例であり、残り3例は第Ⅲ相臨床試験（INTACT）における症例であった。

これら7症例の概要は以下のとおりである。

① 乙B13の1の症例

被験者等略名	M. K. I. 女 55歳
医療機関所在地	埼玉県
副作用・感染症名	急性呼吸不全，間質性肺炎（生命を脅かす／入院を要する／機能障害に至る／医学的に重要な事象）
イレッサ投与期間	開始 2002/02/16 終了 2002/02/28
主な治療経過等	2002/02/28 急性呼吸不全及び胸部 CT にて両側びまん性間質性陰影が認められた。 その後症状は軽快
担当医のコメント	ZD1839 の投与後，腺癌の陰影は顕著に軽快したが，他の間質性浸潤影が認められた。

② 乙B13の2の症例

被験者等略名	T. H. 男 70歳
医療機関所在地	米国

副作用・感染症名	呼吸NOS, 気胸NOS, 皮下気腫 (生命を脅かす ／障害に陥る／入院を要する／医学的に重要な事象)
	心肺停止 (死に至る／障害に陥る／入院を要する／ 医学的に重要な事象)
	肺臓炎NOS (死に至る／生命を脅かす／障害に陥 る／入院を要する／医学的に重要な事象)
イレッサ投与期間	開始 2001/01/26 終了 2001/02/27
主な治療経過等	2001/01/26 ZD1839 の投与を開始した。 2001/02/21 CT スキャンにより急性両側性肺臓炎が 疑われ, 感染あるいは薬剤起因性による ものと考えられた。 2001/02/27 ZD1839 を中止し, 重傷の呼吸困難のため試験から脱落した。 2001/03/13 死亡
担当医のコメント	呼吸困難, 急性心肺停止, 両側性肺臓炎, 気胸及び 皮下気腫は ZD1839 と関連している可能性があるかと 考えている。

③ 乙B13の3の症例 (**太字は追加報告時の記載**)

被験者等略名	J. C. S. 男 60歳
医療機関所在地	米国
副作用・感染症名	間質性肺炎 (入院を要する／死に至る事象)
	報告対象なし
イレッサ投与期間	開始 2002/01/25 終了 2002/02/09
主な治療経過等	2002/01/25 ZD1839 の投与を開始した。

2002/02/09 安静時に呼吸困難が発現し，CTCgrade 3の間質性肺炎のため入院した。

ZD1839の投与は一時的に中止した。

2002/02/20 患者は間質性肺炎による呼吸不全で死亡した。

死亡診断書には，直接の死因は転移性非小細胞肺癌であると記載されていた。剖検は実施されていない。

担当医のコメント ZD1839と関連していると考えられる。

間質性肺炎はZD1839と関連しているが，病勢進展とも関連しているかもしれないと考えている。

④ 乙B13の4の症例（**太字は追加報告時の記載**）

被験者等略名 S. L. C. 女 55歳

医療機関所在地 米国

副作用・感染症名 失神，両側性肺間質浸潤，成人呼吸窮迫症候群

報告対象なし

イレッサ投与期間 開始 2000/10/02 終了 不明

主な治療経過等 2000/10/23 入院中，原因不明の両側性肺間質浸潤および成人呼吸窮迫症候群を発現。

2000/10/30 死亡

担当医のコメント 「失神」「両側性肺間質浸潤」「成人呼吸窮迫症候群」については，化学療法（カルボプラチン）および治験薬（ZD1839，パクリタキセル）との関連性あり。

本事象と化学療法（カルボプラチン，パクリタキセル）

ル) および治験薬 (ZD1839 又はプラセボ) との関連性はないと考える。

⑤ 乙B14の1の事例

被験者等略名 T. F. 男 73歳
医療機関所在地 大阪府
副作用・感染症名 肺臓炎NOS (死亡/生命を脅かす/入院を要する/機能障害に至る/医学的重要な事象)
下痢NOS, 嘔吐NOS, 発熱 (重篤でない事象)
イレッサ投与期間 開始 2002/03/29 終了 2002/05/01
主な治療経過等 2002/05/01 38℃台の発熱が続くため, 治験薬の投与中止。
2002/05/13 頃労作時の呼吸困難が出現
2002/05/16 胸部X線の結果, 両肺びまん性陰影を認め入院。
2002/05/24 肺臓炎による呼吸不全のため死亡。
担当医のコメント 本症例は, 治験薬との関連性があると考える。

⑥ 乙B14の2の症例

被験者等略名 R. W. 男 74歳
医療機関所在地 米国
副作用・感染症名 肺炎NOS, 肺臓炎NOS (入院を要する事象)
イレッサ投与期間 開始 不明 終了 不明
主な治療経過等 2002/04/17 息切れ, 胸痛, 咳嗽を訴え肺臓炎のため入院。CT スキャンにより, 閉塞性肺炎の症状を示す新たな肺浸潤が認められ

たが、肺塞栓の可能性は否定された。

2002/04/22 患者は容態が安定し退院したが、肺炎
及び肺臓炎は未回復であった。

担当医のコメント 閉塞性肺炎及び肺臓炎は ZD1839 と関連性がある。

⑦ 乙B14の3の症例

被験者等略名	W. H. 性別不明 76歳
医療機関所在地	香港
副作用・感染症名	胞隔炎NOS（生命を脅かす、障害、入院に至る事象）
イレッサ投与期間	開始 2001/12/08 終了 2002/02/17 再開 2002/03/05 終了 不明
主な治療経過等	2001/12/08 ZD1839 を投与開始。 2002/02/08 ZD1839 投与開始から9週間後、重度の呼吸困難を発症。 2002/02/11 入院。 X線及び胸部 CT スキャンを実施した結果、気質性肺炎を伴う閉塞性気管支肺炎と一致する所見を示した。 2002/02/17 本事象のため、ZD1839 投与を一時的に中断。 2002/05/05 両側びまん性胞隔炎を発症。重篤な（入院を要し、生命を脅かす）事象と考えられた。 年月日不明 本事象のため、ZD1839 投与を中止。 2002/05/14 プレドニゾロンによる治療を行い、1

0 日後症状は回復。

担当医のコメント 両側びまん性胞隔炎：ZD1839 との関連あり

上記のとおり，これらの7症例の中には，被告会社による追加報告によって副作用報告が取り下げられた症例（乙B13の3，乙B13の4）も含まれているが，被告国自身が，準備書面の中で被告会社による副作用報告の取り下げには理由がなくいずれも副作用症例として取り扱うのが適当であると認めているとおり（西被告国第4準備書面p11＝東被告国準備書面（4）p19），全てイレッサとの関連が否定できない副作用症例である（西平山証人反対尋問調書＝東甲L198p58）。

そして，この7例のうち少なくとも2例（乙B13の2，乙B14の1）は，担当医によって副作用死亡例として報告され，被告側証人も認めているとおり，イレッサの間質性肺炎による副作用死亡例であった（西福岡証人反対尋問調書＝東丙G58p69～70，西乙E20＝東西條反対尋問調書p40～41）。

また，追加報告でイレッサとの関連が否定された症例（乙B13の4）についても，前記のとおり，被告国はイレッサと間質性肺炎との関連は否定できないとして副作用症例として認めており，患者は間質性肺炎による呼吸不全によって死亡している以上，当然イレッサと死亡との関連も否定することはできない（西甲E41＝東福島証人主尋問調書p12～14）。この点については，承認審査を担当した平山証人も副作用死亡例として把握していたことを認める旨の証言をしている（西平山証人反対尋問＝東L198p61）。したがって，この症例もイレッサによる副作用死亡例であることは明らかである。

更に，追加報告で死亡との関連が否定された症例（乙B13の3）についても，当初担当医が，「患者は間質性肺炎による呼吸不全で死亡した」

と明確に副作用死亡例として報告していたにもかかわらず、その後の追加報告には「死亡診断書には、直接の死因は転移性非小細胞肺癌であると記載されていた。」とあるのみであり、剖検も実施されておらず、追加報告による死亡原因の変更には何らの理由・根拠がないと言うべきであるから、この症例についても、依然として死亡との関連が否定できない副作用死亡例として取り扱うべきである（西甲E41＝東福島証人主尋問調書p14～15）。

以上の結果、上記7例のうち、副作用死亡例は4例となる。

なお、上記7例のうち2例（乙B13の1、乙B14の1）は、EAPの国内症例（それぞれ埼玉県と大阪府）であり、うち1例（乙B14の1）が死亡例であった。

イ 上記ア以外の間質性肺炎発症例

（ア）被告国が把握しなかった間質性肺炎発症例

被告国がイレッサの承認前に把握した間質性肺炎の副作用発症例は、前記国内3症例と上記アの7症例の合計10症例のみであった。

しかし、イレッサの承認前に被告会社から被告国に報告された副作用報告の中には、上記10症例以外にも明らかに間質性肺炎の副作用発症例であると認められる症例が多数存在していた。

これらの症例については、濱証人の協力を得て副作用報告等に関する準備書面に添付した「急性肺障害・間質性肺炎を発症したと考えられる副作用症例」の一覧表にまとめており（この一覧表の39例には被告国が把握した間質性肺炎の発症例も9例含まれているため、被告国が把握しなかったものとしては30例ということになる）、このうち間質性肺炎を発症した可能性が極めて高いと考えられる症例（丙B3の54、63、67、79、115、131、132、140、141、150、

152, 164, 172, 丙B4の6, 10, 丙B5の21, 29, 38, 52)については、同準備書面において、個別に臨床経過やその評価等について詳しい検討を行っている。なお、この中には、後に被告会社によって副作用が取り下げられた症例も含まれているが、濱証人は、取り下げの経緯が不明であったり、その理由が不可解であり、取り下げには理由がなく、いずれも副作用症例として取り扱うべきであったとしている（西甲E25＝東G31 p62, 西濱証人主尋問調書＝東甲L102 p63～66）。

また、上記30例の中でも典型的にイレッサによる急性肺障害・間質性肺炎発症例であると考えられる10症例（丙B3の54, 63, 67, 79, 115, 132, 140, 152, 164, 172）については、濱証人が、その意見書においても詳しい考察を行っており、これらはいずれも明らかにイレッサによる急性肺障害・間質性肺炎の発症例であるとしている（西甲E25＝東G31 p53～62）。以下、濱意見書における上記10症例についての考察部分を抜粋する。

① 丙B3の54の症例

性別：女 年齢：51歳 情報源：カナダ

疾患名：結腸直腸癌，肺転移がある例。軽度の労作時呼吸困難があったが、酸素飽和度は室内空気で98%と全く正常であった。

副作用名：呼吸不全

重篤性・転帰：死亡・死亡

治験等：NCICが実施した第I相試験，1日量750mg

臨床経過：もともとゲフィチニブ投与開始11日目（もしくは12日目）に労作時の呼吸困難が悪化し、12日目に下痢がはじまり、17日後にgrade 4の呼吸困難のため入院。ゲフィチニブ投与を中止したが、

呼吸困難が悪化し、入院3日後に呼吸停止、肺転移等により死亡したとされている。

死亡日：2001年5月28日，6月4日入手

担当医評価：可能性あり（関与を否定できない）

企業の意見：否定はできないが、エントリー前から肺に多数の転移巣があり、病勢進行による可能性が高いと考えられる。

考察結果：下痢はしていなかったが、ゲフィチニブ投与後に出現。酸素飽和度は室内空気で98%と全く正常であったのが、わずか11日程度で呼吸困難が増悪し、20日で死亡することは病勢進行では説明がつかない。病勢進行が事実なら、そのことが、ゲフィチニブの副作用と考えるべきであるが、そう考えるより、ゲフィチニブによる急性肺傷害と考えるべきである。

また、医師がゲフィチニブに起因している可能性があると考えて報告しているのに、メーカーが「病勢進行による可能性が高いと考える」とのコメントをすることは、問題である。

② 丙B3の63の症例

性別：男 年齢：55歳 情報源：アルゼンチン

疾患名：非小細胞肺癌

副作用名：呼吸不全

重篤性・転帰：死亡のおそれ・未回復

治験等：第Ⅲ相試験（INTACT）

臨床経過：ゲフィチニブ投与開始約5か月半後に呼吸困難（grade IV）が発現し、その後増悪したため入院。入院4日目の胸部CTスキャンにより、**両肺葉の間質に浸潤**が認められた。入院後ゲフィチニブの投与は中止され、抗生物質及びメチルプレドニゾロン投与が行われた。

患者は呼吸不全と診断された。入院 8 日目患者の容態は退院できる程度に改善し、その 3 日後に行われた CT スキャンでは浸潤はほとんど完全に消失した。患者は、この事象のため試験から脱落した。

担当医評価：呼吸不全と化学療法との関係は否定でき、ゲフィチニブと関係があると判断している。

企業の意見：間質性肺炎に起因した肺浸潤による呼吸不全の可能性も考えられる。ゲフィチニブ開始後に発現し、中止後、加療により改善していることから、本剤との関連は否定できないが原疾患、併用薬との関与も無視できないと考える。

考察結果：38℃の発熱とその後グレード4の呼吸困難があり、CTで**両肺葉の間質に浸潤**が認められたので、間質性肺炎である。

カルボプラチンとパクリタキセルは1回使用しただけでその約50日後に発症しているから、担当医の評価のとおり、その関係は否定できる。また、ゲフィチニブ中止後に改善し、中止後12日にはほとんど完全に消失したので、原疾患の関与も否定できる。

企業が、根拠も示さず「原疾患、併用薬との関与も無視できないと考える」と述べるのは問題である。

③ 丙B3の67の症例

性別：女 年齢：38歳 情報源：米国

疾患名：非小細胞肺癌，骨，脳，肝転移あり

副作用名：肺浸潤NOS

重篤性・転帰：死亡のおそれ・不明

治験等：拡大治験プログラム

臨床経過：ゲフィチニブ投与開始約1か月半後に転移巣の病勢進展による腰仙部痛が悪化したため入院した，入院中，呼吸困難状態及び低酸素状

態に陥った。患者は静注のステロイド剤で治療された。入院翌日の胸部X線写真では右肺全体の微細な結節が認められ、左上葉上尖に楔形の混濁斑、心臓後方の混濁斑が認められた。入院3日目の胸部X線写真により、肺間質及び肺胞のびまん性病変が認められ、ゲフィチニブ投与は中止された。その8日後患者は死亡した。

担当医評価：肺間質及び肺胞の両側性びまん性浸潤はゲフィチニブと関連性があると考えている。

企業の意見：本剤投与後に発現した事象であるため、本剤との因果関係を完全に否定することはできない。登録時および投与開始時には肺への転移は認められなかったものの、入院翌日の胸部X線写真で結節が認められており、これは転移によると考えられている。したがって、ゲフィチニブの関与よりも原疾患の進展または転移に起因するところが多いと考えられる。

考察結果：入院翌日の胸部X線写真では右肺全体の微細な結節が認められ、そのとき、その部分は転移巣と考えられたが、その2日後（入院3日後）には、両側性に肺間質及び肺胞のびまん性浸潤が認められ、その病変が、担当医によりゲフィチニブ関連性があると考えられた。

入院翌日の胸部X線写真で認められた結節は微細なものであり、左上葉上尖には、転移巣とは考えられない楔形の混濁斑もすでに認められ、その2日後には左右とも肺間質及び肺胞のびまん性浸潤となった。まさしく間質性肺炎が左右の肺全体に広がったことを示している。

右肺全体の結節が肺癌の転移であれば、微細な結節のまま、それぞれが増大するはずで、2日後に右肺にも肺間質及び肺胞のびまん性浸潤という病変とはならない。

したがって、これら病変は、肺癌の肺転移の増大ではないし、死

亡の原因も肺癌ではない。ゲフィチニブによる左右の肺全体の間質性肺炎により死亡した。

④ 丙B3の79の症例

性別：女 年齢：68歳 情報源：米国

疾患名：非小細胞肺癌

副作用名：呼吸不全

重篤性・転帰：死亡のおそれ・回復

治験等：拡大治験プログラム

臨床経過：ゲフィチニブ投与開始約2か月半後、肺炎及び発熱のため入院。X線撮影により、右中葉に新たな浸潤、左葉の中央の浸潤の拡大が認められた。入院5日目に急性呼吸窮迫が発現し人工呼吸器を装着され抗生物質が投与された。呼吸窮迫のため、治験から脱落した。左下葉の気管支肺洗浄液からは悪性所見はなく有機体も検出されなかった。左下葉生検の結果、わずかな良性の肺組織及び反応性上皮組織を伴う気管支の組織片が認められ、間質組織には軽度の慢性活動性炎症性浸潤が認められた。入院から約3ヶ月後、患者は回復し退院した。

担当医評価：発熱および肺炎と肺炎とゲフィチニブとの関連性は否定できる。

肺炎に合併した急性呼吸窮迫とゲフィチニブとは関連の可能性があると考えている。病勢進展の所見はみられず、肺炎は十分に治療されていたと考えられるので、呼吸窮迫は、肺炎または病勢進展のみが原因で発現したとは考えていない。

企業の意見：投与後の発現であるので関連を完全に否定することはできないが、2ヶ月後であり、肺炎に伴って発症しているため、本剤の関与は否定的であり、他の因子が大きく影響していると考えられる。

考察結果：発熱は回復したと記載された翌日には急性呼吸窮迫が出現している。

また肺炎の所見が回復したとは記載されていないので、右中葉の新たな浸潤あるいは左肺中央の浸潤の拡大から急性呼吸窮迫へは連続していると推察する。発熱が一見回復したように見えたのは、発熱反応を起こすための細胞の活性もゲフィチニブの EGFR 阻害により抑制されたためかもしれないので、発熱および肺炎そのものが急性肺傷害・急性呼吸窮迫症候群の始まりであった可能性が高い。

担当医が述べた「急性呼吸窮迫とゲフィチニブとは関連の可能性はある」ということは少なくとも言える。

企業の「2ヵ月後であり、肺炎に伴って発症している」ということは否定の根拠にはならない。また、企業が、「大きく影響していると考え」「他の因子」について、企業はそれが何であるとは具体的には何も示しておらず、またそれがどうして原因といえるかの根拠も示していない。そうした具体的原因も根拠も示さずに、他の因子を原因として持ち出すのは問題である。

⑤ 丙B3の115の症例

性別：女 年齢：68歳 情報源：米国

疾患名：非小細胞肺癌

副作用名：呼吸困難NOS

重篤性・転帰：死亡・死亡

治験等：拡大治験プログラム

臨床経過：ゲフィチニブ投与開始 13 日後に間質性肺炎の増悪による呼吸困難の増悪のため入院した。静注の抗生物質及びコハク酸メチルプレドニゾロンナトリウムの投与を行った。しばらくの間、軽快していたが、ステロイド剤の経口投与に変更すると、重症の呼吸困難が再発

した。ゲフィチニブは約 1 か月用いられ、入院から約 3 週間後（中止して 4 日後）に死亡した。

死亡日：2001 年 12 月 13 日 情報入手日：2002 年 1 月 7 日

担当医評価：不明（未入手）

企業の意見：間質性肺炎を併発し、その増悪に起因した（死亡）と考えられるが患者の背景が不明で、現在までの情報から、本剤との関連を完全に否定することはできない。

考察結果：ゲフィチニブ投与開始わずか 13 日後の間質性肺炎と、その増悪による呼吸困難である。ゲフィチニブの投与が続けられたことから、担当医は当初、ゲフィチニブによる間質性肺炎の可能性を考えなかったのであろう。また、メチルプレドニゾロン投与で一時軽快したので、さらにゲフィチニブの投与が続けられ、重症の呼吸困難が再発したもので、その時点でゲフィチニブを中止したのでであろう。したがって、間質性肺炎がゲフィチニブによることを担当医は認めたくない事情がある。

企業の意見の趣旨は、よく分からない面があるが、間質性肺炎を併発し、その増悪に起因した（死亡）とゲフィチニブとの関連を否定することはできないとの趣旨ととれる。

⑥ 丙 B 3 の 132 の症例

性別：男 年齢：54 歳 情報源：米国

疾患名：非小細胞肺癌

副作用名：呼吸困難 N O S, 肺出血

重篤性・転帰：死亡のおそれ・未回復

治験等：拡大治験プログラム

臨床経過：ゲフィチニブ投与開始約 4 か月後に重度の呼吸困難および肺出血の

ため、入院。右葉に浸潤が確認された。気管支鏡検査で肺出血が認められ、挿管された。入院時、治験治療は中止。

最終確認日：2002年2月7日以降，2月22日入手

担当医評価：呼吸困難および肺出血は持続しており，ゲフィチニブによると考えられる。

企業の意見：否定はできないが，患者は進行度の高い肺癌患者であり，本剤による影響よりも原疾患の悪化による可能性が高いと考えられる。

考察結果：担当医が「ゲフィチニブによる考えられる。」とほとんど断定しているのは，おそらく，もともとの肺癌は縮小していたのであろう。企業は，その点を調査し肺癌そのものの進行を形態学的に確認しない限り，「病勢進行」を主張してはならない。

担当医が「ゲフィチニブによる考えられる」と明瞭に述べているのに，「原疾患の悪化による可能性が高いと考える」と，何の根拠もなく主張することは，いかに企業が危険性の情報を低く見ようとしたか，如実に示している。

⑦ 丙B3の140の症例

性別：男 年齢：63歳 情報源：米国

疾患名：非小細胞肺癌

有害事象名：肺塞栓症

副作用名：肺浸潤NOS，呼吸不全

重篤性・転帰：死亡・死亡

治験等：拡大治験プログラム

臨床経過：ゲフィチニブ投与開始19日後，息切れ増強を訴えた。CTスキャンにて，右上葉の肺塞栓の所見が示唆された。肺塞栓（grade 4）治療のため入院。入院5日後，肺塞栓は回復し退院。退院翌日，息

切れが増強し始め、再び入院。両肺浸潤（grade 4）あり。治験薬は中断。CTスキャンでは肺塞栓の所見認められず。両肺浸潤の増悪が認められ、肺胞浸潤は右上葉で著明に悪化しており、肺底部で示唆された硬化は左の方が悪化していた。再入院後 8 日目、胸部 X 線所見では、引き続き肺浸潤が示唆された。再入院後 19 日目、状態が悪化し、呼吸不全のため死亡。

最終確認日：2001 年 1 月 27 日， 3 月 5 日入手

担当医評価：肺塞栓症は無関係，検査結果では腫瘍や感染が原因とは考えられないから，薬物毒性で死亡に至ったと考えている（放射線科医は，新生物などが原因である可能性も考えている）。

企業の意見：完全に否定することはできない。しかし，他の可能性がいろいろ考えられる（詳細は省略）。

考察結果：担当医が「薬物毒性で死亡に至ったと考えている」との判断は尊重すべき。放射線科医の意見は，新生物などが原因である可能性も考えられるということであるが，ゲフィチニブの可能性を否定しているわけではない。この点，極めて誤解されやすい記載である。担当医が「薬物毒性で死亡に至ったと考えている」と明瞭に述べているのに，それをさも否定的であるかのように記載するのは，いかに企業が危険性の情報を低く見ようとしたかを，如実に示している。

なお，肺血栓塞栓についても，ゲフィチニブが血管内皮の再生を阻害することから，血栓を生じやすくなると考えられ，しかも，肺毒性（浸潤）の出現とほとんど同じ時期に出現していることから，大いに関連がありうると考える。この点は担当医の考察不足であろうし，メーカーからの情報不足に基づくものであるかも知れない。

⑧ 丙 B 3 の 152 の症例

性別：女 年齢：39歳 情報源：米国

疾患名：非小細胞肺癌

副作用名：肺浸潤NOS，アレルギー性胞隔炎

重篤性・転帰：死亡のおそれ・未回復

治験等：拡大治験プログラム

臨床経過：ゲフィチニブ投与開始 15 日後，左肺に肺浸潤が発現。プレドニゾンを使用すると軽快し中止すると再発し，回復と再発を繰り返した。経過中，首と左上肢の腫脹が増悪し，投与開始から約 2 か月後に入院。検査により肺塞栓は否定され上大静脈の狭窄を認め，上大静脈炎症候群と診断された。翌日 3 回目の肺浸潤（両側性）が発現した。その翌日，呼吸性代償不全，息切れの増悪を伴う低酸素症を発現。肺浸潤の再発との関連性が明らかに疑われるため，ゲフィチニブの投与は中止された。

メチルプレドニゾンが使用され，その 2 日後，胸部レントゲンにて肺浸潤の軽快が認められた。ステロイド剤を投与している間は肺浸潤は軽快していた。その 5 日後，右腕頭静脈閉塞のためステント術が実施された。

その翌日にゲフィチニブが再開されたところ，翌日からステント部位の血栓と上大静脈炎症候群の悪化を認め，右腕頭静脈の血管形成術，上大静脈にステント，t-PA（血栓溶解剤）静注などを実施。その 2 週間後，4 回目の肺浸潤のため再入院し，ゲフィチニブは中止された。2 日後の気管支鏡検査にて肺炎／肺臓炎と診断された。翌日の胸部レントゲンにて両肺葉に斑状の肺臓炎及び肺胞性浸潤が認められた。その 3 日後（入院 1 週間目），患者は死亡した。

担当医評価：検査で細菌はすべて陰性，好酸球がほとんど認められなかったの
でゲフィチニブと関連があるかもしてない。肺浸潤の再発と過敏

性肺臓炎はゲフィチニブと関連している。

企業の意見：肺浸潤については、投与期間中の再発であり完全に否定することはできないが、進行肺癌で肺炎を合併しているため、本剤より患者背景の影響が強いと考える。大静脈症候群は肺癌で起しやすく、担癌患者では凝固系の亢進が認められるし、血栓症もあったため関連は否定できる。過敏性肺炎は投与後の発現であり本剤との関連は否定できないが、プールで化学物質に接しており、その影響も無視できない。

考察結果：進行肺癌だが室内プールを利用するほどであり、元気であったと考えられる。プールの塩素による灼熱感や咳は他のヒトも発現していたと記載されているが、ゲフィチニブが使用されていればその影響は強く出る。

肺浸潤はゲフィチニブ使用 15 日後に発症し、ステロイド剤を使用し、ゲフィチニブを中止している間は軽快していたが、再開後 2 週間後には重篤な呼吸不全のため入院。100 %酸素吸入でもコントロールできず入院翌日に挿管された。

このように、死亡にいたる原因となった肺浸潤は中止で改善し、再投与で再現している。

さらに、治療中に病勢が進展していたとの明白な証拠はなく、むしろ治療が奏功していた、と記載されているので、進行肺癌であったこと、上大静脈症候群についても、癌そのものによる影響より、ゲフィチニブの血管内皮傷害性による血栓症と肺炎/肺臓炎に伴う影響が強かったと考えられる。

いずれにしても、ゲフィチニブによる急性肺傷害・間質性肺炎が確実な例であった。

企業は、このようにプールに行くほど元気であったと考えられ、

肺癌が縮小して癌の影響が軽減した例で、使用中止で改善し、投与再開で再発重症化したという、関連が明瞭な例も、「進行肺癌」という患者背景の影響が強いと主張している。この解釈は極めて問題である。

⑨ 丙B3の164の症例

性別：女 年齢：62歳 情報源：米国

疾患名：非小細胞肺癌

副作用名：呼吸不全，乳酸アシドーシス

重篤性・転帰：死亡・死亡

治験等：拡大治験プログラム

臨床経過：ゲフィチニブ投与開始2日後，患者は高度に進展した乳酸アシドーシスをきたし，呼吸不全のためICUに搬入された。胸部X線により，両肺の浸潤がゲフィチニブ投与開始時点より拡大しているのが認められた。患者は，呼吸不全のため治験から脱落した。その4日後，患者は呼吸不全のため死亡した。

最終確認日：2002年3月13日，3月28日入手

担当医評価：急性呼吸不全は病勢進行ではなくゲフィチニブと関連していると考えている。

企業の意見：完全に否定することはできないが，フルラゼパムとアルプラゾラムを併用しており，これらの薬剤が呼吸不全を起こした可能性も考えられる。両肺に転移のある肺癌であり，ゲフィチニブ投与2日後であるので，本字以外の原因に起因する可能性が大きい。腎機能が低下していたことが発生を助長したと考えられ，患者状態に起因するものとする。

考察結果：担当医が「急性呼吸不全は病勢進行ではなくゲフィチニブと関連し

ていると考えている。」としているのを否定するには、それなりのしっかりした根拠が必要であるが、この例についても企業は、併用薬、病勢進行、2日目であること、腎障害など、考えうる他の原因を持ち出して否定しようと試みている。しかし、併用薬剤を持ち出すなら、いつから服用を始めたのかの情報が必要だが、併用薬が開始された日も不明である。室内空気酸素飽和度が99%あった人の肺癌が1日や2日で人工換気装置を必要とするほど進行はしえない。

第I/II相臨床試験(1839IL/0011)の225mg群の1例は、1日目に2回服用後に無呼吸で中止し、26日目に吸引性肺炎で死亡した例であるが、この例と似た経過の可能性はある。

日本の市販後の調査でも、1週間以内に発症した8例中7例と大部分が死亡しているので、こうした急激な発症も、ゲフィチニブによる可能性がありうる。

企業は、4つもの可能性をあげるが、それぞれ濃厚な可能性があるとの根拠が示されていない。それらを示して、いかにもゲフィチニブが否定的であるかのように記載するのは、企業が危険性の情報を低く見ようとしたことの現われである。

⑩ 丙B3の172の症例

性別：女 年齢：73歳 情報源：ブラジル

疾患名：非小細胞肺癌

副作用名：肺浸潤NOS

重篤性・転帰：死亡・死亡

治験等：拡大治験プログラム

臨床経過：ゲフィチニブ投与開始から53日後、胸部X線にて、びまん性間質

性肺浸潤の所見を認め、入院。患者は2週間前より呼吸困難の悪化、及び乾咳（軽度～重度）を訴えていた。低酸素血症も発症し、入院中増悪。入院1週間前に行った肺生検の結果、薬剤との関連性が疑われる反応性間質性肺炎と報告された。翌日、高用量ステロイド療法を開始。突然、急性呼吸停止を発症し、集中治療室へ搬送された。その10日後、乏尿、低酸素血症を発症し、死亡。死因は間質性肺浸潤。

最終確認日：2002年4月12日，4月15日入手

担当医評価：関連性あり。（肺生検の結果，ゲフィチニブとの関連が疑われる反応性間質性肺炎と報告された）

企業の意見：間質性肺炎が疑われる症例。他に肺毒性が報告されている放射線療法，ゲムシタビンを行っているので，これらの関与も疑われるが。本剤の関与も否定できない。

考察結果：担当医が「肺生検の結果，ゲフィチニブとの関連が疑われる反応性間質性肺炎」と報告している。放射線療法，ゲムシタビンの関与は，その時期を特定したうえで主張すべきであるが，放射線療法の時期は不明である。また，放射線療法をした場合に，間質性肺傷害が生じるなら，むしろ，ゲムシタビンを使用したときに，放射線を照射した部位に生じるはず（recall現象として）だが，そのときは生じていない。この例についても企業は，他の2つの療法を持ち出して，ゲフィチニブの関与を薄めようと試みている。この例も，企業が危険性の情報を低く見ようとしたことを示している。

以上の10症例はイレッサによる間質性肺炎・急性肺障害を発症した典型的な症例であるが，これらの症例の中には，副作用名自体は必ずしも「間質性肺炎」として報告されてはいないものの，その臨床経過等の

中に「間質性肺炎」ないしこれと同義の疾患名（「間質性肺浸潤」「肺臓炎」等）が記載されており，その記載だけでも容易に間質性肺炎発症例であると判別できるものも複数存在しており（丙B3の67，115，152，172等），これらの症例については，被告側証人もイレッサによる間質性肺炎の副作用発症例であることを認めている（西福岡証人反対尋問調書＝東丙G58 p69～70，西工藤証人反対尋問調書＝東乙L17 p84～92，西平山証人反対尋問調書＝東甲L198 p67～70）。

(イ) 被告国の杜撰な審査によってこれら多くの間質性肺炎発症例が見落とされたこと

これら被告国が把握しなかった間質性肺炎発症例について，イレッサの承認審査を担当した平山証人は，間質性肺炎発症例であるか否かは報告された副作用名のみで判断したことを認めている（西平山証人主尋問調書＝東甲L197 p27）。

この点につき，平山証人は，副作用名に「間質性肺炎」の記載のないものは，別の可能性も考えられるので，そのようなものを含めると適切な判断ができなくなるとの趣旨を述べている（西平山証人主尋問調書＝東甲L197 p27）。また，その方が審査が効率的であるとも述べている（西平山証人反対尋問調書＝東甲L198 p65）

しかしながら，このような平山証人の証言は，医薬品の安全性評価に関する原則に真っ向から反するものであり，被告国がイレッサの承認審査に関していかに杜撰な審査しかしていなかったかを如実に示すものである。

すなわち，既に何度も述べてきた医薬品評価における「有効性は確実に，安全性は鋭敏に」という原則からすれば，医薬品の安全性審査に関しては，確実な副作用に関しては勿論のこと，むしろ確実でないもの

(可能性のあるもの) も含めて慎重な審査が行われなければならない。

特に、抗がん剤おける間質性肺炎という極めて重大な副作用については、慎重な審査が行われるべきであることは言うまでもなく、イレッサにおいても間質性肺炎発症例があったことが判明していた以上、全ての副作用について、間質性肺炎発症の可能性及びその重篤度等について、細心の注意を払って慎重かつ詳細な調査、検討が行われるべきであった。

前記平山証人の証言は、審査の効率性を重視して、間質性肺炎とは別の可能性が考えられる症例については、間質性肺炎の副作用とは認めないとして切り捨てるものであり、明らかに上記のような医薬品の安全性審査に関する考え方に反するものである(なお、平山証人自身もそのような考え方は個人的な見解であり、審査センター全体の考え方でないことを認めている。西平山証人反対尋問調書＝東甲L198p63)。

したがって、本来であれば、間質性肺炎発症例と疑われる症例については、更に被告会社に臨床経過や検査結果等の資料を提出させて詳しい検討がなされるべきであるが、上記のとおり、本件で間質性肺炎発症例として判明している症例は、そのような詳しい調査までしなくとも、被告会社から提出された副作用・感染症症例報告書の記載のみから、間質性肺炎発症例であると判断できるものであった。

したがって、前記平山証人の証言は、間質性肺炎か否かを副作用名のみをもって判断したというよりも、報告された副作用症例について、その内容(臨床経過等)を全く見ていなかったことを自白するものにほかならないのであり、このようなイレッサの承認審査ががいかに杜撰なものであったかは、火を見るより明らかである。

ウ 副作用報告で見られたイレッサによる間質性肺炎の評価

(ア) 上記ア、イで述べたイレッサによる間質性肺炎の副作用は、副作用報

告等に関する準備書面で述べたとおり，そのほとんどが「死亡」又は「死亡のおそれ」があるものとして報告された症例であり，イレッサによる間質性肺炎の副作用が，いかに重篤かつ致死的なものであったかは明らかである。

副作用報告等に関する準備書面添付一覧表「間質性肺炎・急性肺障害を発症したと考えられる副作用症例」の39症例のうち，国内臨床試験の2例（No.1及び2）を除く37症例について，以下の各項目毎に死亡症例数，死亡例の割合を示すと以下のとおりである。

	死亡例／症例数	死亡例の割合
全37症例	24／37	64.9%
被告国が認めた7症例	4／7	57.1%
上記7症例のうちの日本人症例	1／2	50.0%
上記一覧表の◎の症例	16／26	61.5%
濱意見書で取り上げた10症例	5／10	50.0%

被告らは，臨床試験で見られた国内3症例がいずれも回復例であったとして，イレッサによる間質性肺炎がさほど重大なものではなかったかのように主張するが，既に述べたとおり，国内3症例についても一つ間違えば十分に死に至る可能性が認められる重篤な症例であり，うち1例はイレッサ投与と死亡との関連を完全には否定できない症例であったことに加え，上記副作用報告において相当数の間質性肺炎発症例が見られ，そのうち相当高い割合（50～60%）で死亡例が存在するなどイレッサによる間質性肺炎が極めて重篤かつ致死的な副作用であることは明らかであった。

このように，被告らは，イレッサによる間質性肺炎を評価するにあた

っては、国内臨床試験の結果だけでなく、上記副作用報告に見られた間質性肺炎発症例についても合わせて判断する必要があったのであり、これらイレッサによる間質性肺炎の副作用を全体として見れば、イレッサによる間質性肺炎が極めて重篤かつ致死的な副作用であったことは容易に判明していたと言ふべきである。

(イ) また、被告らは、EAPの副作用報告を考慮するとなれば、EAPの登録患者数（被告会社によると平成14年7月までで1万5243例、西丙K1の4＝東丙E1の4 p3参照）から考えられる間質性肺炎の発症頻度は低い数字となり、むしろイレッサの間質性肺炎について過小評価することになりかねない旨主張するようである。

しかしながら、EAPの副作用報告数は、その登録患者数に比較して明らかに報告率が少ない（臨床試験における副作用報告率と比較すると7分の1程度に過ぎない）のであり、EAPにおいては、そもそも本来副作用として報告されるべきであるのに報告されていない症例（暗数）が相当多数あることが推定される（西乙E24＝東工藤証人反対尋問調書p90～92）。したがって、そのようなEAPにおける副作用報告については、形式的に報告された副作用症例数をもって発症頻度が低いとの議論をすることは全く適切でない。

そして、日本人のEAP症例に限ってみれば、使用患者数296例（西甲O8＝東甲K53，西甲O58＝東甲K55）に対して、2例の間質性肺炎発症例（乙B13の1，乙B14の1）が報告されており、うち1例が死亡例（乙B4の1）である。そうすると、それ自体の頻度としても、0.67%の間質性肺炎発症率と0.33%の死亡率ということになり、決して無視できない頻度となる。これに上記のとおり相当の暗数が存在することを考慮すれば、EAPにおける間質性肺炎発症例は、その発症頻度においても十分に注意すべきものであったと言ふべき

である。

(ウ) 以上のとおり、本件では、イレッサ承認前に、多数の間質性肺炎の副作用症例が報告されていた。その中には数多くの副作用死亡例も含まれており、イレッサによる間質性肺炎の副作用が極めて重篤かつ致死なものであることは明らかであった。

にもかかわらず、被告国が把握した間質性肺炎発症例は、前記アの7症例のみであり、その他の副作用症例は、すべて「症例の集積を待って検討」として、イレッサの安全性評価にあたって考慮しなかったのである。

このような被告国の対応について、福島証人は、「言語同断」であり、「こんなことは、今までの薬害の歴史の中では本当にあったのか」というほど「信じ難い」ことであり、「医師として許し難い」ことであると非常に厳しい表現で明確に証言しており（西福島証人主尋問調書＝東甲L95p15～16）、イレッサの安全性審査がいかに杜撰なものであったかはもはや明白である。

(4) イレッサによる間質性肺炎の評価のまとめ

ア 極めて重篤かつ致死的である

イレッサの臨床試験で認められた間質性肺炎の副作用発症例である国内3症例は、それ自体極めて重篤かつ致死的な症例であったことに加え、それ以外にも承認前に多数の間質性肺炎の副作用発症例（その中で被告国が把握した症例は7例のみ）が報告されており、その中には間質性肺炎による副作用死亡例も多数かつ相当高い割合（50～60%の割合）で存在していた。

このように、臨床試験及び副作用報告で見られたイレッサによる間質性肺炎は、極めて重篤かつ致死的なものであることは明らかであった。

イ 発症頻度も高頻度（十分に注意すべき頻度）である

イレッサによる間質性肺炎の日本人における発症頻度は、少なくとも国内臨床試験において2.3%（3/133）と明確かつ高頻度であることが判明していたのであり、日本人のEAP患者でも0.67%（2/296）、うち死亡率0.33%（1/296）とのデータが出ており（しかも、前述のとおり、EAPの副作用報告率は臨床試験の副作用報告率の7分の1程度に過ぎず、相当の暗数が存在することが推定されるから、EAPにおける発症頻度の数値は実際よりも相当低いものと見なければならぬ）、このような間質性肺炎の発症頻度は十分に注意すべきものであった。

なお、上記間質性肺炎の発症頻度に関するデータが十分に注意すべき数値であったことは、以下に述べるとおり、統計学的な考察からも裏付けられる。

すなわち、ある副作用の発症頻度を判定する場合において、母数となるべき症例数が少ないときには、そこで得られた副作用の発生率は、（神のみぞ知る）真実の発生率と乖離する幅が大きくなる。すなわち、同じ1%の発生率でも、100例中1例の場合と10万例中1000例の場合とでは、自ずと1%という数値の信頼性が異なるのである。例えば、症例数が500例の場合では、500例中5例（1%）の発生率における95%信頼区間（95%の確率で信頼できる発生率の幅）は、0.325～2.32%と相当な幅があることに注意する必要がある（西甲F53＝東甲G101 p113～114）。

このように、ある副作用について、比較的少ない症例数における発症頻度を評価する場合には、その真実の発症頻度には相当の幅があり、実際には、その幅の上限値が真実の発症頻度である可能性もあることを念頭において、安全性審査を行うことが求められるのである（西平山証人反対尋問調書＝東甲L198 p81）。

そこで、イレッサにおける間質性肺炎の副作用について、上記データに基づいて95%信頼区間の上限を計算すると、国内臨床試験における間質性肺炎発症率(3/133)では上限値が6.45%、日本のEAP患者における間質性肺炎による副作用死亡率(1/296)では上限値が1.86%となり、この数値はそれぞれ、市販後のプロスペクティブ調査における間質性肺炎発症率5.8%、副作用死亡率死亡率2.5%という数値と近似する。したがって、上記承認前の発症頻度に関するデータは、そのような市販後の発症頻度をも十分に予見しうるデータであったと言いうるのである(西甲P105=東甲L196, 西平山証人反対尋問調書=東甲L198p84~85)。

5 まとめ

以上述べてきたように、イレッサの臨床試験を検討するにあたっては、まず、イレッサのEGFR阻害というドラッグデザインから予測される肺毒性、イレッサの非臨床試験で得られた毒性所見を前提に慎重かつ厳密に吟味されなければならない。

また、有害事象と治験薬との因果関係については、治験担当医師の判断を鵜呑みにせず、治験全体の結果や他の副作用情報を総合して判断する必要があり、副作用については、臨床試験における副作用報告のみならずEAPなど他の副作用情報も重視して安全性評価を行わなければならない。

これらを前提として、イレッサの安全性を評価すれば、第1に、イレッサの臨床試験において現れた有害事象死亡例は、そのほとんどが副作用に分類されるべきものであり、少なくともこれら有害事象死亡例のデータはイレッサによる致死的な急性肺障害・間質性肺炎の副作用の発生を予測させるに十分なデータであった。

そして、イレッサの国内臨床試験やEAPを含む副作用報告で認められた間

質性肺炎の副作用発症例は、いずれも極めて重篤かつ致死的なものであり、その発症率及び死亡率も高頻度（十分に注意すべき頻度）で、これらのデータは市販後のプロスペクティブ調査等に基づく発症率、死亡率をも十分に予見するものであった。

以上より、イレッサは、極めて重篤かつ致死的な間質性肺炎の副作用を高頻度で発症させるものとして、その安全性が欠如していたことはイレッサ承認前の段階で既に明らかになっていたと言ふべきである。

第3節 イレッサ承認後の有効性評価

第1 第Ⅲ相試験に見るイレッサの有効性の欠如

1 はじめに

これまで検討してきたとおり、イレッサは市販前においても、本来の抗ガン剤の有効性の指標である延命効果は全く確認されずに、腫瘍縮小効果のみによって有効性が判断されたが、その腫瘍縮小効果においてさえ、決してそれまでの抗ガン剤を越えるものではなく、むしろ、劣っているものでしかなかった。他方、致死的な急性肺傷害・間質性肺炎という毒性を有するものであったことが、ドラッグデザイン、非臨床試験の結果からも予見しえ、また、臨床試験段階における副作用情報からは確定的に認識しえたのであって、旧ガイドラインにおける腫瘍縮小効果を前提とした承認制度においても、イレッサは、その有効性、安全性のバランスを欠き、有用性を認め得ないものであったことは明らかである。

そして、イレッサの市販後に至っては、承認条件を充たすために行われたV

1532試験をはじめ、各種第Ⅲ相試験において延命効果を示すことができず、イレッサに有効性がないことがさらに明確になった。

したがって、もはや、イレッサの薬剤としての命脈は尽きており、市場に置かれることは許されないものである。

2 V1532試験について

(1) 試験の概要

V1532試験は、2003年9月から2006年1月に登録され、1レジメンまたは2レジメンの化学療法治療歴を有する進行／転移性（ⅢB期／Ⅳ期）又は術後再発の非小細胞肺癌患者を対象に、ゲフィチニブ投与群とドセタキセル投与群を無作為割り付けして、それぞれの生存期間を比較した、市販後第Ⅲ相比較臨床試験である。

この試験は、「手術不能又は再発非小細胞肺癌に対する本薬の有効性及び安全性の更なる明確化を目的とした十分なサンプルサイズを持つ無作為化比較試験を国内で実施すること」というイレッサの承認条件を充たすために行われた試験であった（西甲A1＝東甲A2 p4）。

この試験の目的は、イレッサがドセタキセルに比べて生存期間の点で劣っていないこと（非劣性）を証明することにあつたが、イレッサの非劣性を証明することができなかった。

既に述べたとおり、抗ガン剤の第Ⅲ相試験においては延命効果の確認が最も重要であり、本来、独立した2つの第Ⅲ相試験で延命効果が確認されなければ承認取消となる（西甲H13＝東甲F51 p211，西甲H14＝東甲F52 p87～88，西乙E20＝東西條証人反対尋問調書 p113）。

イレッサについては、本来2つの第Ⅲ相試験が要求されるにもかかわらず、V1532試験1つしか行っていない点をおくとしても、そのただ1つの試験においてすら延命効果を示し得なかったのであるから、当然に承認が取り

消されなければならない。

しかも、V 1 5 3 2 試験はイレッサの承認条件とされた試験であり、その試験で延命効果を示すことができなかつたのである。先述したとおり、抗がん剤における真のエンドポイントが延命効果であることのみならず、添付文書において承認条件であると記載されているという重みからしても、承認条件をみたまず、延命効果を示すことができないイレッサについては、当然に承認取消となるべきものである。

(2) 被告側証人である福岡証人は、V 1 5 3 2 試験の結果について、「この試験において・・・非劣性を証明することはできませんでした。しかしながら、ドセタキセルとの優越性の検定におきまして、差はないということでございました。」と述べて、あたかもイレッサとドセタキセルの延命効果に差がなかったかのように印象づけようとしている（西福岡主尋問調書＝東丙 G 5 7 p 3 4）。

しかし、このような証言は、科学的根拠に基づかないことは明白である。すなわち、「臨床試験の統計解析に関するガイドライン」（西甲 D 1 9 = 東甲 H 1 4）では、「ここで、“有意差のないこと”をもって直ちに同等であるとするのは明らかに間違いであり、有意差のないことは統計的に“同等”を保証するものではない。」と記載されている。また、「Japanese Journal of Clinical Oncology 投稿に際しての統計解析結果のレポートに関するガイドライン」（西甲 F 3 3 = 東甲 F 5 4 p 8）においては、「医学論文でもっともよく見られる統計学的基本概念についての誤解は“統計学的有意性（有意差）”についてであるが、それは「統計学的有意差がない」ことを「差がない」つまり同等であると勘違いしていることである。」「『差がないこと』を積極的に示したい場合は『同等性の検定』と呼ばれる手法を用いる必要がある」と記載されている。これは、福岡証人も認めざるを得ないところであ

った（西福岡反対尋問調書＝東丙G 5 8 p 4 3）。

すなわち、福岡証人は、「有意差がない＝同等」と考えることが誤りであることを熟知しながら、あえて上記のような証言を行い、裁判所の判断を誤らせようとしていたのである。これは、福岡証人が、イレッサを擁護するためであれば科学的妥当性すら無視するという姿勢で証言していたことを示すものである。

(3) さらに、被告側証人である光富証人は、「非劣性が証明されないということは、劣っていないことが証明されなかったということですがけれども、逆に劣っていることが証明されたというわけでもありません。」「あの試験の結果からいえることは何もありません。」と証言する（光富主尋問調書＝東乙L 2 3 p 7 6）。

しかし、かかる証言も、科学的妥当性を欠くものであることは明白である。

たしかに、V 1 5 3 2 試験によって、イレッサがドセタキセルに劣ることが証明されたわけではない。しかし、すでに繰り返し述べているように、医薬品はその有効性が科学的に証明されない限り、無効として扱われなければならない（西甲F 1 7＝東甲F 3 3）。つまり、V 1 5 3 2 試験において延命効果の証明に失敗したことによって、イレッサは無効として扱われなければならないことが明確に示されたのであって、「あの試験の結果からいえることは何もありません」というのは全くの詭弁である。

なお、同証人は、「仮に、この試験において、ドセタキセルに対する非劣性が証明されたとしたら、イレッサは医薬品としての有効性も証明されたというふうにおっしゃるんじゃないですか。」との質問に対して、「それはもちろん、その場合は証明されたと思います。」と証言し（光富反対尋問調書＝東乙L 2 4 p 4 3）、他方、「劣性が証明されない限り、この試験からは何もいえないと、こういうことになるわけですか。」との質問に対して「そ

うですね。だから研究としては、結論が出ないという研究になります。」と証言した（光富反対尋問調書＝東乙L 2 4 p 4 2）。

これは、V 1 5 3 2 試験で非劣性の証明に成功すれば有効性の根拠となるが、失敗しても有効性は否定されないという、被告会社にとってはきわめて都合のよい考え方である。しかし、このような考え方に立てば、市販後第Ⅲ相試験でどのような結果が出ようともイレッサの有効性は否定されないことになってしまうのであり、市販後第Ⅲ相試験で延命効果が確認されなければ承認取消となるという確立した考え方（西甲H 1 3＝東甲F 5 1 p 2 1 1，西甲H 1 4＝東甲F 5 2 p 8 7～8 8，西乙E 2 0＝東西係証人反対尋問調書p 1 1 3）と明らかに矛盾する。

以上からすれば、光富証人の証言態度も、有効性評価についての科学的原則を無視したきわめて偏ったものというほかない。

これらの証言からいえることは、光富証人は、ドセタキセルに対して非劣性が証明されればそれはイレッサの有効性が証明されたことを意味し、逆に無治療に対して勝つかどうかを見ない限り、イレッサの有効性を否定することはできないという旨の考え方を有しているということである。

しかし、このような考え方はまさに「いいとこ取り」というべきご都合主義的な考え方である。このような光富の考え方は、「有効という証拠がない限り、それは常に無効だと思っておくべきである。」（西甲F 1 7＝東甲F 3 3）という“有効性は確実に”との原則を踏みにじるものである。

(4) また、光富証人は、主尋問において、同試験において非劣性が示されなかった理由について、投薬のクロスオーバーが多かった旨証言した。

しかし、C o x 回帰モデルによるV 1 5 3 2 試験の解析結果について、西甲C 6＝東甲E 1 0－7：5 枚目のスライドでは、0 以上はドセタキセル群が有効で、0 以下はゲフィチニブ群が有効との解析がなされているところ、

治療効果の95%信頼区間の下限が、10か月程度までは0を上回っている。

つまり、この分析結果からは、10ヶ月程度まではドセタキセル群が有効であることが示されている。

また、有意差はないが、イレッサ群の治療効果がよくなる傾向を示してくるのは、20か月以降であり、したがって、後治療（イレッサ群においてはイレッサ以外の治療）が進むにつれてイレッサ群の成績がよくなるのであった。

したがって、後治療の影響（クロスオーバー）がイレッサに不利に働いたとは言えない。このことは、光富証人も反対尋問において認めざるを得なかった（光富反対尋問調書＝東乙L24p52）。

なお、V1532試験における後治療の影響については、その後発表された追加解析結果（西甲C9＝東甲D11）において、

- ① ドセタキセル群で後治療としてイレッサが用いられたグループと、イレッサ群で後治療としてドセタキセルが用いられたグループを比較した結果、両群ほぼ同様の傾向で推移しているが、後半（22か月以降）では、イレッサからドセタキセルに切り替えられた群の方が優れる傾向にあること（スライド21）
- ② 後治療が行われなかった患者の全生存期間のサブグループ解析で、一貫してドセタキセルの方が優れる傾向が認められ、生存期間中央値でゲフィチニブ群4.1か月に対しドセタキセル群8.7か月、1年生存率でゲフィチニブ群が12.2%に対しドセタキセル群が27.0%と、いずれもドセタキセル群が大きく上回っていること（スライド22）
- ③ 後治療として他の化学療法（ドセタキセル及びイレッサ以外）が行われたサブグループの解析結果でも、一貫してドセタキセル群の方が優れる傾向が認められ、生存期間中央値は11.5か月対17か月、1年生存率は47.5%対61%で、いずれもドセタキセル群が上回っていること（ス

ライド24)

などが示され、『後治療の影響によってドセタキセルに有利な結果が出た』などとは言えないことがより明確となっている（西乙E24＝東工藤証人反対尋問調書p42）。

(6) まとめ

同試験において、延命効果が示されなかった以上、イレッサには有用性がないことは明らかである（西甲D35＝東甲H19p85）。よって、承認条件として行われた試験において有効性及び有用性が証明できなかったのであるから、この一事をもってしても、イレッサは承認取消しとなるべきものである。

東日本訴訟で被告側の証人となった西條長宏証人は、「癌治療における国際化」という論文（西甲H13＝東甲F51）において「薬剤の承認後、薬の survival benefit の確認のため、independent な phase III study を二つ要求される。これで survival benefit がなければ承認取消しとなる」と述べている。さらには、西條証人は、「臨床試験の早い段階でサロゲート・マーカーをもって有効性が示唆されても、最終的には第Ⅲ相試験で延命効果が証明されなければ臨床応用されないことは言うまでもない。」（西甲H14＝東甲F52）とさえ述べている。

その上、西條証人は、「イレッサの場合も、この第Ⅲ相試験によって延命効果を示すことが必要であって、それができなければ承認が取り消されるべきだということになりますね。」との質問に対して「そうですね。」と証言している（西乙E20＝東西條反対尋問調書p113）。

そして、被告側の証人である西條証人でさえ、イレッサの有用性は「統計学的には証明されていない」とまで証言せざるを得なかったのである（西乙E20＝東西條反対尋問調書p129～130）。

3 INTACT試験について

(1) 試験の概要

INTACT 1は、イレッサの延命効果を検証するため、化学療法歴のない進行非小細胞肺癌（NSCLC）の患者1093名を、既存化学療法剤であるシスプラチン，ゲムシタビンとイレッサ併用群と，シスプラチン，ゲムシタビンとプラセボ併用群に分けて行った第Ⅲ相の無作為二重盲検化比較試験である。

INTACT 2は，INTACT 1と同デザインの試験であり，化学療法歴のない進行非小細胞肺癌の患者1037名を，既存化学療法剤であるカルボプラチン，パクリタキセルとイレッサ併用群と，カルボプラチン，パクリタキセルとプラセボ併用群に分け比較したものである。

(2) INTACT試験は，イレッサの大規模第Ⅲ相臨床試験で，世界で初めて結果が公表された試験であり，イレッサの承認直後である2002年8月にその結果が公表された。

この試験において，イレッサは延命効果を証明できず，生存期間中央値，一年生存率，病勢進行までの中央値のいずれを取っても，イレッサ投与群は，プラセボ投与群に比して生存期間を延長する効果がなかったことが判明したのである。

のみならず，イレッサ投与群は，プラセボ投与群と比較して寿命短縮の傾向さえ見られたのである（西濱主尋問調書＝東甲L102 p67～68，西甲E25＝東甲G31 p63頁表5-1）。

西甲E25＝東甲G31 p64図5-1は，公表論文などからINTACT試験での有害事象等をグラフにしたものがあるが，このグラフから，イレッサ群はプラセボ群に対して有害事象死亡率が高く，また，有害事象によるイ

レッサ中止の割合は、有意な用量依存関係が認められることが分かる。これは、プラセボ群に比較してイレッサ群の方が寿命を縮める傾向にあることを裏付ける結果であり、イレッサの安全性には問題があることが示されたと言える。

このように、INTACT試験においては、イレッサの延命効果は証明されず、かえって危険性に問題があることが示されたと言い得る結果が判明したのである。

なお、被告側証人らはINTACT試験はイレッサの上乗せ効果を見た試験であるから、これで有効性が否定されたからといって、イレッサ単剤での有効性が否定されるものではないとの証言をしている。

しかし、イレッサそのものの延命効果がないことが、上乗せによって延命効果が示されなかった原因である可能性は否定できない。そして、既存抗がん剤との相互作用によってイレッサ群の延命効果に影響が生じたという根拠は明らかとなっていないこと、及び上記のとおりイレッサ群に寿命短縮傾向があったことを考えると、INTACT試験は、たとえ上乗せ試験であることを考慮しても、イレッサそのものの延命効果に重大な疑問を生じさせるものであったといえる。のみならず、その後に行われたイレッサ単剤での第Ⅲ相試験でイレッサが延命効果の証明に失敗を重ねていることを考え合わせれば、INTACT試験は、イレッサそのものの延命効果を否定する重要な資料の一つというべきである。

4 ISEL試験について

(1) 試験結果について

ISEL試験は、局所進行あるいは転移した非小細胞肺癌に対する2次・3次治療としてのゲフィチニブ（イレッサ）の生存に対する効力について評価することを目的として1692もの症例で行なわれた大規模な第Ⅲ相

臨床試験であった。I S E L 試験はプラセボ対照の試験であり、症例数も相当数に上ることから、イレッサの有効性について、信頼性の高い臨床試験である。

そして、I S E L 試験において、イレッサは、延命効果を示すことができなかった。I S E L 試験のように症例数がかかり多い試験においては、統計学的な検出力が高まり（西光富主尋問調書＝東乙L 2 3 p 7 8）、有意差が検出されやすい上、対照がプラセボであったにもかかわらず、なおイレッサは延命効果を示すことはできなかったのである。

このように、イレッサは、もっとも信頼性が高く、しかも、有意差が出やすい試験において、統計学的に延命効果を示すことができなかったということを確認しなければならない。

(2) サブグループ解析について

ア 被告会社は、I S E L 試験においては、東洋人のサブグループ解析において統計的に有意な生存期間の延長がみられるなどとして、その有効性が否定されるものではないと主張する。

しかし、このような主張に科学的合理性はない。I S E L 試験のサブグループ解析は、日本人におけるイレッサの有効性を証明するものではなく、その結果を前提としても、やはりイレッサは無効なものとされなければならない。

イ まず、I S E L 試験を受けた「東洋人」の中には、日本人は1人も入っていない。日本人がその対象に入っておらず、後述するようにブリッジング試験も行なわれていないイレッサにおいては、日本人の含まれない『東洋人』のサブグループ解析について統計学的に有意な生存期間の延長が見られたとしても、それは、イレッサの有効性評価について何ら積極的意味を有するものではない。

ウ そして、そもそも「サブグループ解析」なるものは、仮説としての意義しか持ち得ず、上述したとおり厳密な有効性を要求する薬剤の有効性の証拠には到底なり得ない。このことは、多くの文献で指摘されているところである。

すなわち、「学会・論文発表のための統計学」（西甲 P 1 4 = 東甲 F 9 p 1 0 0）では、「ICHの統計のガイドラインでは、サブグループ解析については層別した効果の測定、あるいは交互作用解析に重点を置き、サブグループごとの検定結果は、参考程度に考えるべきであるとされています。」と記載されている。

「『臨床試験のための統計的原則』について」（西甲 P 1 5 = 東甲 H 3 p 2 9）でも、「部分集団的解析・・・は探索的であるため、探索的であることを明確に確認しておくべきである。」「探索的解析である場合、これらの解析結果は慎重に解釈すべきである。試験治療の有効性（若しくは有効性のないこと）、又は安全性に関する結論は、どのようなものであっても、探索的な部分集団的解析のみに基づいては受け入れ難い。」とされている。

「Japanese Journal of Clinical Oncology 投稿に際しての統計解析結果のレポートに関するガイドライン」（西甲 F 3 3 = 東甲 F 5 4 p 9 以下）においても、「予め設定された仮説と解析計画に従って行われた解析のみが“確定的データ解析”とみなされるが、そのみが結論的な結果を与えることができる。」「多くの研究者はこれら2つの主要なタイプの解析（または研究）の違いを誤解または混同している。検定で同じように“有意”な（ $P < 0.05$ ）結果が得られても、それが解析（データ収集）の前に立てられた主要な仮説の検定の場合と、事後的に（ad hocに）行われたサブグループ解析における副次的な仮説の検定の場合とは全く違った意味合いを持つ。データ解析の中で得られたサブグループに対する解析の結果は、

他の別の研究で確認 (confirm) されるまでは常に注意して扱われなければならない。』と記載されている。

これらのことは、被告側証人である福岡証人も認めざるを得なかったことであり、また、西條証人も、イレッサについて東洋人に対する延命効果が証明されたわけではない旨認めている (西乙 E 20 = 東西條反対尋問調書 p 117)。

エ なお、被告会社は、東洋人というグループを含む人種によるサブグループ解析が、試験実施計画の当初から予定されていたと主張する。

しかし、たとえ当初から予定されていたとしても、上記のとおりサブグループ解析の結果によって有効性を証明することはできないし、そもそも、以下のとおり、東洋人というサブグループ解析は当初から計画されたものではなく明らかに後付の解析であった。

西 C 1 = 東甲 C 5 p 1 ~ p 6 は、「試験計画書」の統計解析部分 (Statistical Analysis Section of Protocol) であり、p 3 冒頭に「プロトコール作成日に関する資料」と記載され、同頁では、2003年4月23日に被告会社の臨床試験チームリーダーの Marianne Curdno (マリアンヌ・カードナー) が、同 p 4 及び p 5 では、4月29日に Nick Thatcher (ニコラス・サッチャー、治験責任医師である) が試験計画を確認したことが示されている。

この試験計画では、解析のファクターについて以下のように記載されている (同 p 1)。「比較治療群について、以下のファクターを考慮に入れた長期的統計解析を行った。すなわち、性別 (男 vs 女)、喫煙 (喫煙歴なし vs 喫煙者または喫煙歴あり)、前化学療法 of 失敗理由 (化学療法に抵抗性を示したかそうでないか)、化学療法の回数 (1回 vs 2回)、Performance Status (0,1vs2,3) である」。

つまり、ここでは「人種」は記載されていない。

ところが、この試験計画書はその後改定され、2004年12月9日には、統計解析計画書の補遺が作成された（「サブグループ解析の計画日に関する資料（統計解析計画書）」西甲C1＝東甲C5 p6）。このp7では統計解析書の補遺が作成されたことを受けて、統計解析用ソフト「SAS」がE-mailで送付されたことが示されており、p8は最終改定された統計解析用ソフト「SAS」の中身である。

これをみると、解析対象のサブグループ8として、「Oriental」、すなわち「東洋人」が加えられており、東洋人は、2004年12月9日にサブグループとして追加されたものと推測できる（2004年12月9日以前に「東洋人」をサブグループに加えていたと主張するのであれば、被告会社はその根拠となる試験実施計画書を提出すべきである）。

ちなみにISEL試験において、最初の患者が登録されたのは、2003年7月15日、最後の患者登録は2004年8月2日、最終データ入力は2004年10月29日であって、いずれも2004年12月以前である。

そして、ISELの延命効果なしとの初回解析結果が公表されたのは、2004年12月17日である（FDAは同日、この結果を受けた声明を公表している。西甲J3＝東甲I1）。

要するに、東洋人がサブグループとして加えられたのは、最終データ入力から約2月後、初回解析結果公表のわずか8日前なのである。

すなわち、全体としての延命効果の証明に失敗したことを知った被告会社が、少しでもよい材料を求めて様々なサブグループによる解析を行った結果、「Oriental」でたまたま有意差が検出されたこと分かったことから、急遽発表前に補遺を追加したことが明らかであり、まさに探索的な後付解析であることを示すものといえる。

オ さらに、このサブグループ解析においては、各試験群の背景因子に重要な差異があることが指摘される。

すなわち、まず、一般に、非喫煙者は喫煙者に比較して生存期間が長くなるのが通常であるところ、ISEL試験サブグループにおいて、東洋人非喫煙者と喫煙者とのプラセボ群同士で生存期間中央値を比較すると、プラセボ群の東洋人非喫煙者の生存期間中央値は4.5か月であるのに対し、同じく喫煙者のそれは6.3か月と非喫煙者の方が喫煙者よりも著しく短くなっている（西丙K4の5＝東丙E4の5 p 17上のスライド、西丙K3の9＝東丙E3の9・2枚目のTable 1、西濱証人主尋問調書＝東L102 p 69, 70）。

また、「診断からランダム化までの期間」という予後に強く関係する背景因子が、プラセボ群よりもゲフィチニブ群の方が長かったのである。診断からランダム化までの期間が長いということは、自然の経過が長いということの意味している可能性があり、長期に生存する傾向がある。したがって、ゲフィチニブ群に生存期間延長に有利となる偏りが存在したといえる（西丙K4の6＝東丙E4の6別添資料12-2「東洋人及び非喫煙者の患者背景に関する資料」、西甲E25＝東甲G31 p 66、西濱証人主尋問調書＝東甲L102 p 70, 71）。

以上のように、重要な背景因子の偏りを調整しないデータでは、東洋人に対する効力の可能性について語ることは許されるべきではない。

カ 以上のとおりであり、いかなる観点から見ても、東洋人のサブグループ解析を根拠にイレッサの有効性を論じることなど全く非科学的であることは明白である。

5 SWOG0023試験について

(1) SWOG 0023試験は、未治療の患者に対して、シスプラチン及びエトポシド並びに放射線の同時併用療法を実施し、ドセタキセルで地固め療法を実施した後、ゲフィチニブもしくは、プラセボを使用し、寿命延長効果を確認するために行なわれたプラセボ対照ランダム化第Ⅲ相臨床試験であり、2001年7月に登録が開始された試験である。

同試験においても、イレッサは、統計学的に延命効果を示すことができなかったのみならず、統計学的にもイレッサ群が有意差 ($p = 0.013$) をもって短命であったという試験結果だったのである。

すなわち、同試験においては、ゲフィチニブ群 (118人) の生存期間中央値は、23か月であるのに対し、プラセボ群 (125人) の生存期間中央値は35か月であり、ゲフィチニブ群はプラセボ群よりも12か月も短命だった (西甲E49 = 東甲G51)。

つまり、イレッサは、標準薬との比較ではなく、プラセボ群と比較して、統計学的に有意に生存期間中央値が12か月間も短くなったということであり、イレッサには、延命効果があるどころか、寿命短縮効果があるということが統計学上証明されたのである。これは、イレッサの害作用の強さを明確に示す結果ということができる。被告側の西條証人でさえ、少なくとも放射線化学療法対照症例については、イレッサの生存に対する否定的な影響を示したものである可能性が十分にある旨証言している (西乙E20 = 東西條反対尋問調書 p 120)。

(2) しかるに、被告会社は、同試験は途中で中止されており、実験結果は出ていないなどと主張する。

しかし、同試験は、上述したように、ゲフィチニブ群の生存期間中央値が、プラセボ群の生存期間中央値よりも劣るということが明確になったからこそ、試験が早期に中断されたものである。

そして、アメリカでは、このSWOG 0023試験の結果と上記ISEL試験の結果とを受けて新規患者へのゲフィチニブの投与を禁止する措置を講じたのである（西甲E 49＝東甲G 51，西濱主尋問調書＝東甲L 102 p 71以下）。

その後、2007年に発表された追加解析結果（西甲E 49＝東甲G 51）でも、上記のとおりイレッサ群118人、プラセボ群125人という相当規模での解析が行われ、イレッサ群が著しく生存期間を短縮させたというきわめて明確な結果が出ているのであって、試験は途中で中止されており実験結果は出ていない、などという主張は苦し紛れの言い訳というほかない。

(3) また、被告らは、同試験での投与方法は我が国の医療現場で広く行なわれている標準的なイレッサの投与方法ではないという反論をする。

しかし、SWOG 0023試験で示された明確な害作用が、非放射線化学療法併用例では現れるおそれがないという科学的根拠は示されていない。そうである以上、非併用例でも害作用が現れる危険性があることを考慮して安全性を評価する必要があることは当然である。

また、放射線化学療法後にイレッサを投与するような症例は決して少なくない。西被害者稲垣丈夫も、放射線化学療法後に必ずしも癌が増殖したわけでもないのに、イレッサが投与されている。

上記のとおり、アメリカにおいてはSWOG 0023試験のデザインのようなイレッサの投与方法だけを禁止したのではなく、新規患者への投与自体を禁止したのであり、このことはイレッサそれ自体の有用性を否定したからに他ならない。

被告らの主張は、1つの臨床試験で失敗した場合には、当該試験デザインによる投与が否定されるだけであり、他の投与方法の有効性は「否定されていない。」という論理であるが（前述のV 1532試験についての光富証人

の供述も同様) , こうした考え方は, 「有効性は確実に」「危険性は鋭敏に」の原則を無視して根本的に誤っていることは何度も述べてきたとおりである。

6 INTEREST試験について

(1) INTEREST試験は, 上述のV1532試験と類似のデザインの試験であり, 欧米で行なわれた試験である。

(2) 被告会社は, INTEREST試験において, ドセタキセルに対する非劣性が証明されたとしてイレッサの有効性の根拠として主張するようである。

しかし, 承認条件は, 「国内で」臨床試験を実施することを要求している(西甲A1=東甲A2)。INTEREST試験は欧米で実施されたものであり, 被験者に日本人は含まれていない。日本国内で日本人を対象に行なわれ, しかも承認条件とされたV1532試験において, 延命効果を証明できなかった以上, INTEREST試験がイレッサの有効性を根拠づける根拠となるはずもない。

被告側証人である光富証人でさえ, INTEREST試験の結果を考慮しても, なお臨床試験のエビデンスには乏しいという評価をせざるを得ないということを認めている(光富反対尋問調書=東乙L24p39)。

そして, 西條証人も, INTEREST試験の結果が出た後でも, イレッサについては, 「グレードC」(行うよう勧めるだけの根拠が明確ではない)という評価に変わりはない旨供述している(西乙E20=東西條反対尋問調書p128~129, 西甲F42=東甲L133)。

7 IPASS試験について

(1) I P A S S 試験は、臨床背景因子により選択されたアジアの進行非小細胞肺癌患者を対象に、ファーストライン治療としてのイレッサの有効性等をカルボプラチン／パクリタキセル併用化学療法と比較した試験で、主要評価項目は無増悪生存期間であった。

(2) 同試験において、イレッサはカルボプラチン／パクリタキセル併用化学療法に対して、無増悪生存期間において優越性を証明したとされる。

しかし、同試験の主要評価項目である無増悪生存期間は、既に述べたとおり、全生存期間の代替指標、サロゲートエンドポイントに過ぎず、それ自体、有効性の指標としての位置づけは相当に低いものでしかない。そして、同試験においては、対象患者が1200例を越えており（西甲C11＝東甲D12）、全生存期間についての統計学的な解析も十分可能で、これを主要評価項目とすることが可能であったと考えられるにもかかわらず、敢えて、全生存期間を副次的な評価項目とするデザインがされている。このようなデザインを敢えて採用したのは、穿ってみれば全生存期間ではイレッサの優越性を証明できないと考えたからではないかとも推測される場所である。

そして、事実、中間解析の結果では、サロゲートエンドポイントに過ぎない無増悪生存期間では非劣性を示したとされながらも、真のエンドポイントである全生存期間では、解析途中であるとはいえ、両群で「同様の傾向」、つまり、ほとんど差が見られないとされているのである（西甲C11＝東甲D12）。

したがって、こうしたサロゲートエンドポイントに過ぎない無増悪生存期間の結果をもって、イレッサに有用性が認められたなどと主張することは決して許されないことである。

(3) また、同試験の対象患者のうち、日本人は約20%・233例に過ぎ

ず、これは、V1532試験の半分以下である。日本人以外のアジア人が80%を占める試験では、承認条件（「国内で実施」）を満たすことにならないのは当然である。

そして、同試験での患者適格基準は、化学療法未治療であることに加えて、癌腫は腺癌で且つ非喫煙者に絞られているのであって、イレッサの適応である「手術不能または再発の非小細胞肺癌」からすれば、相当に患者範囲が狭くなっている。

したがって、同試験は、どれだけ譲っても日本におけるイレッサの有効性の根拠とはならないのである。

この点、工藤証人も日本人に対する有効性が証明されたかという質問に対しては「日本人に対して言えるかですか・・・これはだから、アジア人に対して言えるということになりますね」と言葉を濁した（西乙E24＝東工藤証人反対尋問調書p46）。

また、工藤証人は、IPASS試験においてはイレッサの延命効果が示されるかどうかは、遺伝子変異に依存すると述べているが（西乙E24＝東工藤証人反対尋問調書p46～47）、背景因子によって対象患者に大幅に絞りを加えている（「進行非小細胞肺癌患者で、組織型が腺がん、かつ喫煙歴のない、または軽度の喫煙歴（10 Pack-Year以下で少なくとも15年以上禁煙している）を有する患者」を対象としている）という点や、対象がファーストラインであることも考えると、IPASS試験は、日本におけるイレッサの適応、つまり、「手術不能または再発非小細胞肺癌」の全体をカバーするデザインとはなっていない（西乙E24＝東工藤証人反対尋問調書p48）。

以上より、IPASS試験は、イレッサの有効性の根拠とはなり得ないものである。

8 まとめ

以上のとおり、イレッサは、承認条件であるV1532試験で延命効果を証明することができず、現時点において統計学的に有用性はない（西乙E20＝東西條証人反対尋問調書p130）のであるから、その承認は取り消されるのが当然である。

このように、市販前においてもイレッサの有用性は否定されるべきであったが、市販後の大規模臨床試験の結果は、いずれもイレッサの有効性を証明できたとは言えないものであって、イレッサには有効性が認められないことが明らかとなったのである。

第4節 イレッサの承認後の安全性評価

前述したとおり、イレッサによる間質性肺炎の副作用が極めて重篤かつ致命的なものであることは明らかであり、イレッサの安全性が欠如していたことはイレッサ承認前の段階で既に明らかになっていたと言ふべきである。そして、承認後、わずか6年足らずの平成20年3月現在において、734人というこれまでに類を見ないほど多数の副作用死亡者数を出したということが、イレッサには安全性が欠如していたということを何より雄弁に物語っている。

以下では、こうしたイレッサの市販後の被害状況に加え、市販後に調査されたプロスペクティブ調査、コホート内ケースコントロールスタディ等の結果から見て、イレッサの安全性の欠如が実証されたことについて検討する。

第1 イレッサは他の抗がん剤に比して高度の危険性を有する薬剤であること

1 イレッサによる副作用死亡者

イレッサは、平成14年7月に承認後、平成20年3月末までの約5年半強の期間に734人の副作用死亡者を生み出した。特に承認直後の平成14年7月から同年末までに180人、翌平成15年中に202人、平成16年中に175人も多くの副作用死亡者を生み出したことは紛れもない事実である（西甲P141＝東甲L183）。

これに対し、被告らは、抗がん剤に常に副作用死の危険が内在していることを強調し、イレッサが他の抗がん剤に比して高度の危険性を有する薬剤でないかのような主張を繰り返してきた。

しかし、以下に述べるとおり、イレッサは、承認後、他の抗がん剤に比してはるかに多くの副作用死と疑われる症例が報告されており、その結果に鑑みれば、イレッサは他の抗がん剤に比して極めて高度の副作用死の危険性を有する薬剤であることは明らかである。

2 独立行政法人医薬品医療機器総合機構への副作用症例報告

製薬企業は、副作用によるものと疑われる症例等を知ったときは、薬事法第77条の4の2第1項の規定により厚生労働省に対して報告することが義務づけられているが、平成15年7月の薬事法改正により同法第77条の4の5第3項の規定に基づき、平成16年4月からは独立行政法人医薬品医療機器総合機構（以下、「総合機構」という。）に対して報告することが義務づけられた。

総合機構は、上記副作用症例報告につき、そのホームページにおいて、平成16年4月以降の総合機構が受理した報告の全てをラインリスト（新掲載様式）として公表している（なお、平成16年度以前の報告は、旧掲載様式として公表しているが、平成16年度以降の新掲載様式とは異なる形式で公表されており、公表された旧掲載様式が受理した報告の全てを公表したものであるか

どうかは、ホームページ上明らかでない)。

この総合機構が公表する副作用報告症例は、法律に基づき報告されているものであり、副作用の実態を把握するにあたり、信頼性の高いデータといえるものである。

3 副作用報告症例の死亡例の数の比較

そこで、上記総合機構のホームページで公表されている副作用報告症例のうち、報告が法律で義務づけられ、新掲載様式に移行した平成16年度における、イレッサを除く抗がん剤で副作用報告に係る死亡例が多い上位10品目を列挙すると以下のとおりである(西甲P141=東甲L183)(以下、一般的名称、当該死亡例の数の順に記載)。

- ① テガフル・ギメラシル・オテラシルカリウム 50
- ② パクリタキセル 43 (西甲P145=東甲L186)
- ③ ドセタキセル水和物 40 (西甲P146=東甲L187)
- ④ シスプラチン 28 (西甲P100=東甲L117)
- ⑤ メシル酸イマチニブ 26
- ⑥ リツキシマブ 26
- ⑦ 塩酸イリノテカン 25
- ⑧ 塩酸ゲムシタビン 19 (西甲P147=東甲L188)
- ⑨ リン酸フルダラビン 15
- ⑩ シクロホスファミド 13

これに対し、イレッサの平成16年度における副作用報告に係る死亡例数は「140」であった(西甲P141=東甲L183別表1, 西甲P149=東甲L193, 西甲P155=東甲L203)。すなわち、次順位の副作用報告

に係る死亡例数が多い抗がん剤と比しても、イレッサの副作用報告に係る死亡例は3倍近い数字となっており、イレッサという薬剤に内在する危険性が高度なものであることは明らかである。

4 肺癌という癌腫に限定した場合の副作用死亡例の数

そもそもイレッサは、肺癌のみを対象として投与される抗がん剤である。これに対し、上記列挙された抗がん剤は、肺癌のほか、胃癌等他の癌腫においても投与される抗がん剤である。肺癌と他の癌腫では、その疾患部位・機序等が異なる以上、抗がん剤の副作用を把握するにあたり、同列に扱って比較することに合理性はなく（本件訴訟に証人として出廷した医師工藤翔二氏も、法廷において、肺癌だけで比較するのが合理的であるとの証言をしている、西乙E23＝東工藤証人主尋問調書p29・13行以下）、抗がん剤の副作用を把握するにあたって、肺癌患者に絞った比較をする必要がある。

そこで、上記上位10品目に列挙される抗がん剤のうち、肺癌にも投与される抗がん剤である②パクリタキセル、③ドセタキセル水和物、④シスプラチン、⑧塩酸ゲムシタビンの上記死亡例につき、総合機構のホームページに掲載される症例一覧（パクリタキセルにつき西甲P145＝東甲L186、シスプラチンにつき西甲P100＝東甲L117、ドセタキセル水和物につき西甲P146＝東甲L187、塩酸ゲムシタビンにつき西甲P147＝東甲L188）から、「原疾患等」の欄に、「肺非小細胞癌」、「肺小細胞癌」、「肺腺癌」、「肺扁平上皮癌」、「肺の悪性新生物」とされているもののみを列挙すると以下のとおりとなる。

② パクリタキセル 16

③ ドセタキセル水和物 16

④ シスプラチン 4

⑧ 塩酸ゲムシタビン 9

これに対し、イレッサの平成16年度における副作用報告に係る死亡例数「140」は、全て総合機構のホームページに掲載される症例一覧の「原疾患等」の欄に「肺非小細胞癌」、「肺腺癌」、「肺扁平上皮癌」、「肺の悪性新生物」とされるものである（西甲P155＝東甲L203）。

すなわち、副作用死亡例発生頻度を合理的に比較するための分子にあたる死亡例の数において、イレッサは他の抗がん剤との間に10倍近い差があり、イレッサに内在する危険性がよりいっそう高度なものであることは明らかである。

5 抗がん剤の使用者について

以上のイレッサの圧倒的な副作用死亡者の数につき、イレッサの使用患者数が多いことを理由に挙げる主張もある（西乙E24＝東工藤証人反対尋問調書p39・6行以下、但し根拠はない）。

しかし、イレッサは現在に至っても、標準的治療薬として認められている薬剤ではなく、肺癌患者で最初に抗がん剤治療を受ける場合（ファーストライン治療）の標準的治療方法は、

- a シスプラチン＋イリノテカン
- b シスプラチン＋ドセタキセル水和物
- c シスプラチン＋ビノレルビン
- d シスプラチン＋塩酸ゲムシタビン
- e カルボプラチン＋パクリタキセル

といったプラチナ製剤と新規抗がん剤の2剤併用療法となっている（西乙E18＝東乙L10 p7）。

すなわち、非小細胞肺癌患者は、抗がん剤治療を試みる場合、まずシスプラチンやカルボプラチンといったプラチナ製剤とゲムシタビンやドセタキセルと

いった新規抗がん剤を併用して使用して、その後にイレッサの投与を受けているのである（イレッサの添付文書上でも「化学療法未治療例における有効性及び安全性は確立していない」と明記されており、上記標準的治療の先行を前提としている）。

つまり、イレッサを使用する肺癌患者は、通常、前提としてシスプラチンもしくはカルボプラチンといったプラチナ製剤の投与を受け、同時に標準的治療に列挙される新規抗がん剤を先行して使用している患者なのである。とすると、イレッサを使用する患者のほとんどは、通常プラチナ製剤及び新規抗がん剤を先に投与されているのであり、10倍程度の乖離を生み出すほどイレッサの使用患者のみ圧倒的に多いという実態は存在しえない。

そもそも、平成16年度以外においても、本項の冒頭で述べたとおり、イレッサによる急性肺障害等副作用報告に係る死亡例の数は、徐々に減少しているとはいえ、年次（1月1日から12月31日までの間）別で、平成14年は約半年間（承認後の7月16日から12月31日まで）で180人、平成15年は202人、平成16年は175人と、むしろ承認直後の平成14年、平成15年は、平成16年度よりもより多い副作用死亡者が存在していたのであり（東甲L183）、上記肺癌患者のみを比較した10倍程度の差は、もっと大きかったものであることが容易に推測できる。

もはや母数の違い、などの理由付けでは説明できない大きな格差が明らかになっているのである。

6 まとめ

以上は、承認後に発生した実態である。

イレッサにより発生した副作用死亡例が他の抗がん剤と比しても圧倒的に多いという事実は、承認当時からイレッサが極めて高度の副作用死の危険性を有していたことの証左であり、もはや抗がん剤一般に内在する副作用死では説明

できないレベルのものである。

第2 プロスペクティブ調査について

1 調査の概要

プロスペクティブ調査とは、ゲフィチニブの副作用発現頻度及び危険因子（発生危険因子、予後因子）をできるだけ速やかに明らかにする目的で、平成15年6月から平成15年12月の間に登録された3322例について副作用発現頻度及び危険因子の検討が行なわれた調査である。

2 イレッサによる急性肺障害・間質性肺炎の副作用について、大規模かつ詳細な個別症例の検討が行われたプロスペクティブ調査の結果（西丙C2＝東甲D1）の信頼性は高い。

そして、このプロスペクティブ調査の結果によれば、イレッサによる急性肺障害・間質性肺炎の発症率は5.81%（193例／3,322例）、そのうち死亡率は2.5%（83例／3,322例）、急性肺障害・間質性肺炎からの死亡転帰は38.6%（85例／220例）とされており（丙C2 p2, 14）、イレッサによる急性肺障害・間質性肺炎の発症率は極めて高い発症率を有しており、また致死率も極めて高いという特徴を有するのである。特定の副作用について、これだけの高い発症率・致死率を示す薬剤は、他に例を見ないといっても過言ではない。

なお、イレッサの危険性情報は市販後に多くの被害者を生み出しつつ蓄積されてきており、プロスペクティブ調査が行われた平成15年下半期当時には、それなりの情報が把握されていた。したがって、イレッサが市販におかれた直後では、当然、プロスペクティブ調査において把握された発症率、死亡率を上

回っていたであろうことは想像に難くない。この点については、被告側の坪井証人も同様の可能性があることを認めている（西丙E 4 9 の 1 = 東坪井反対尋問調書 p 2 6 ~ 2 8）。

なお、プロスペクティブ調査の結果が公表される以前である専門家会議報告においては、これより遙かに少ない数字が上げられていたが、これは母数の患者数が単なる推計で到底信頼に値しないものであり、また、分子の発症患者数も副作用報告数を前提にしているに過ぎず、報告されない暗数があったことを否定できない程度のものに過ぎなかった（西工藤証人反対尋問調書=東乙L 1 7 p 4 3 ~ 4 6）。

第3 コホート内ケースコントロールスタディについて

1 試験の概要

コホート内ケースコントロールスタディは、第1に、非小細胞肺ガン患者のイレッサ投与例における急性肺障害・間質性肺炎（ILD）の発症を、化学療法投与例と比較することによって、相対リスクを推定すること、第2に、治療中の非小細胞肺ガン患者における急性肺障害・間質性肺炎の発症率を推定することを目的に行なわれた試験で、1レジメン以上の化学療法歴を有し、イレッサあるいは化学療法を受ける予定の進行または再発の非小細胞肺ガン患者を対象とした。そして、事前に規定された進行または再発の非小細胞肺ガン患者のコホートにおける急性肺障害・間質性肺炎発症例、及びコホートより無作為抽出した急性肺障害・間質性肺炎非発症例（コントロール）を対象としたコホート内ケースコントロールスタディの方法で行われた。本試験は、2003年11月から2006年2月に実施され、4473件を登録して終了した。これは、本来、6000件のコホート群を集積する予定であったところ、本試験の主要

目的である急性肺障害・間質性肺炎発症の相対リスク推定に必要な急性肺障害・間質性肺炎発症件数が120件を超えたため終了したものである。

- 2 本試験結果によれば、非小細胞肺癌患者のイレッサ投与例における急性肺障害・間質性肺炎発症の相対リスクは、化学療法投与例に対し、3.23倍という結果であった（西甲C4＝東甲D7 p20）。

とりわけ、投薬開始後28日以内で比較した場合、非小細胞肺癌患者のイレッサ投与例における急性肺障害・間質性肺炎発症の相対リスクは、化学療法投与例に対し、3.80倍と極めて高くなることが判明した（95%信頼区間1.90～7.60）。

このように、イレッサ投与による急性肺障害・間質性肺炎の発症は、化学療法に比べ、統計的有意差が存在することが明らかとなった。そして、イレッサは、他の化学療法に比較して、格段に急性肺障害の危険性が高い薬剤であることが科学的に確認された。

- 3 なお、被告会社は、コホート内ケースコントロールスタディにおけるイレッサによる治療関連死亡率が1.6%であるという趣旨の発言を法廷内で繰り返し、また、被告会社が発行に関与した雑誌等でも繰り返している。

しかし、イレッサ投与による死亡率を見るのであれば、複数回コホート登録された患者数であるのべ症例数ではなく、コホート解析対象の初回登録数、つまり患者の絶対数である1482例を分母として見る必要があり、その場合は約2%（0.02024）が治療関連死亡率となる（西工藤反対尋問調書＝東乙L17 p46～49，西甲C4＝東甲D7）。

よって、コホート内ケースコントロールスタディにおいてイレッサによる治療関連死亡率は1.6%ではなく2%が正確である。

第4 まとめ

以上のとおり、イレッサによる間質性肺炎の副作用が極めて重篤かつ致命的なで、イレッサの安全性が欠如していたことは、イレッサ承認前の段階で既に明らかになっており、その安全性の欠如が、承認後、わずか6年足らず間に、734人というこれまでに類を見ないほど多数の副作用死亡者数を出したという結果として表れたのである。

そして、イレッサの安全性の欠如は、これまでのべてきたとおり、市販後において、他剤との比較やプロスペクティブ調査、ケースコントロールスタディの結果によっても、より明確に実証されたのである。

第5節 イレッサの有用性結論

イレッサの有用性が欠如していることについて、承認・市販前の時点における知見と市販後の知見について検討を加えてきた。

イレッサの市販前においては、本来の抗ガン剤の有効性の指標である延命効果は全く確認されないまま、旧ガイドラインに基づくものとして腫瘍縮小効果のみによって有効性の有無が判断された。しかしながら、イレッサのIDEAL試験等に基づく腫瘍縮小効果は、決してそれまでの抗ガン剤を越えるものではなく、むしろ、劣っているものでしかなかったと言っても過言ではない。

他方、イレッサが致命的な急性肺傷害・間質性肺炎という毒性を有するものであったことは、イレッサのドラッグデザイン、非臨床試験の結果からも予見されたものであり、臨床試験段階における副作用情報からは確定的に認識しえたものであった。しかも、致命的な急性肺傷害・間質性肺炎が市販後のような極めて高頻度で発

症することも、イレッサの承認時における情報から十分判明していたのである。

したがって、旧ガイドラインにおける腫瘍縮小効果を前提とした承認制度においても、イレッサは、IDEAL試験の結果からの有効性の見込みと危険性とを比較しても、全くそのバランスが欠如しており、有用性を認め得ないものであったことは明らかである。

そして、市販後においては、イレッサは、INTACT 1, 2, ISEL, SWOG 0023試験と立て続けに延命効果を証明することができず、そして、旧ガイドラインを前提として我が国の承認条件となったドセタキセルとの比較国内第Ⅲ相試験（V1532試験）においても、延命効果を証明できなかった。

被告側証人である光富証人も、イレッサは臨床試験のエビデンスに乏しいと述べざるを得ず（西光富証人反対尋問調書＝東乙L12 p 39）、同様に、被告側証人である西條証人に至っては、イレッサは統計学的には有用性が証明されていないことを認めざるを得なかったのである（西乙E20＝東西條証人反対尋問調書 p 130）。

これまで述べてきたとおり、抗ガン剤の有効性は延命効果によって「検証」されなければならない、その検証にあたっては、「客観性を保った一定の評価を行うためには統計学的手法以上に優れた方法は今日まだ知られていない。」（「新医薬品の臨床評価に関する一般指針について」（西乙D25＝東乙H28））のである。これは臨床試験の指針として発出された厚生労働省の通知であり、被告国の公式見解である。被告側証人である西條証人が、「イレッサは統計的には有用性が証明されていない」（西乙E20＝東西條証人反対尋問調書 p 130）と述べざるを得なかったことは、単に統計学的な問題に留まるのではなく、医薬品の有用性評価として、イレッサは有用性が検証されていない、すなわちイレッサは有用性を欠いていることを認めたことに他ならないのである。

そして、市販後、イレッサの急性肺傷害・間質性肺炎の毒性が極めて高頻度に発症し、多くの被害者を生み出してきたが、それがイレッサ自体の毒性によるもので

あることがコホート内ケースコントロールスタディの結果等によって科学的に確認されるなど、イレッサが、他の抗ガン剤と比較しても極めて毒性の強い物であることが白日の下に曝されていった。

このようなイレッサが、有効性と安全性のバランスを欠き有用性を持たない医薬品（否、医薬品とは呼べない単なる化学物質という他ない。）であることは、既に科学的にも決着の付いた議論である。

2008年（平成20）年3月時点において少なくとも734名もの尊い命を奪ったイレッサは、もはや我が国の市場から撤退されなければならないことは明らかである。

以 上