

床試験であったかどうかは明確にされなかった。二〇〇五年四月の名古屋高等裁判所金沢支部判決<sup>(4)</sup>でも、あくまで他事目的に関する説明義務違反という判断しかなされず、患者を被験者と認めるまでには至っていない。二〇〇六年、最高裁決定で二審判決が確定した。法の枠外にある行為に説明義務違反の判断が下ったという意味では画期的判決だが、研究倫理の観点からは不十分な判決と言わざるを得ない。本件については後ほどさらに考察を加えることにする。

このようにICを取得せずに行われた臨床試験は後を絶たない。おそらく表面化するのは氷山の一角だろう。また、無断臨床試験以外にも、倫理的に問題を含む臨床試験は数多いと思われる。なぜ繰り返されるのか。

## 2 法規制とガイドライン

今日、日本にも臨床研究の倫理指針は存在する。それらは「ヘルシンキ宣言」にのっとって規定されているので、完全に遵守されればほとんど問題はない。しかし、倫理指針を作るだけでは守られないことは、無断臨床試験が繰り返される現実を見れば明らかである。

問題点は二つあると思う。

第一に、治験は法規制されているが、自主臨床試験は法の枠外にあることである。治験には薬事法に基づく厳格な規則があるため、今日では問題が発生する可能性は低い。一方、自主臨床試験は、厚生労働省の「臨床研究に関する倫理指針」に従うべきだとされているが、法的拘束力はなく、いわば無法状態に置かれている。金沢大学病院と神戸市立医療センターの事例はいずれも、自主臨床試験に該当する。

治験が法規制されている背景には、被験者のリスクもあるが、治験の質を確保しないと医薬品開発の国際競争に取り残されるという市場経済的な理由も大きい。しかし、被験者から見れば、実験に参加するという意味では、治験も自主臨床試験も何ら変わらない。それどころか、治験よりリスクの大きい自主臨床試験も十分あり得る。

倫理指針を空文としないためには、治験以外の研究もカバーする法規制が必要である。

研究の法規制などというと、煩瑣な手続きや他者の干渉を厭わしく思う医師・研究者の反感を招きそうだが、杜撰な研究や非倫理的な研究を淘汰し、優れた研究のみを率良く推進するためには、誰もが認めるルールに従うことが必要である。手続きが増えて一見遠回りに見えるが、規制に応じた研究基盤さえ整えば、結局はそちらの方が必要かつ有益な研究結果を、確実に、しかも低コストで手にできる。

それに、適切な法規制は、研究者の安全を守ることにもつながる。現在、診療は医師法や医療法、治験は薬事法で規定され、両方とも法的根拠を有している。ところが、治験以外の臨床研究は、これを規定した法律がなく、いわば国から公認されていない行為なのである。したがって、これが法規制されるということは、法的根拠のない不安定なグレーゾーンで日々活動している研究者の行為が、公に認められるということに他ならない。

問題点の二つ目は、倫理指針が多数あって、お互いの整合性がないことだ。

現在、治験は「医薬品の臨床試験の実施の基準に関する省令」（一九九七年、厚生労働省）、遺伝子解析研究は「ヒトゲノム・遺伝子解析研究に関する倫理指針」（二〇〇一年、文部科学省・厚生労働省・経済産業省）、疫学研究は「疫学研究に関する倫理指針」（二〇〇二年、文部科学省・厚生労働省）に従わなければならない。遺伝子治療や再生医療の研究にもそれぞれ国の指針が作られている。このような状態は、臨床研究の倫理問題が持ち上がるた

び、関係省庁が一時しのぎで指針を出し続けてきたため、整合性のない指針が乱立してしまった結果である。

そして二〇〇三年、臨床研究全体を包むものとして「臨床研究に関する倫理指針」が公表された。本来はこれが全体を包括する指針となるべきのだが、現状では他の指針に該当しない研究にしか適用されず、他の指針の上位に立つ基本原則（包括的指針）とは成り待てない。指針を乱発した後から全体を包もうとしても無理があるのである。

このように日本は倫理指針だらけになったが、相互のつながりがないため、結局どれに従えばいいのかよくわからない。どれにもまたがる研究もある一方、どれにも従う必要のない研究もありうる。穴だらけ、かつ煩雑である。煩雑なルールは守られない。このような状態は、日本の生命倫理に歴史的な文脈がないことと無縁ではなからう。

ところで、日本では治験以外は法規制されていないと述べたが、実は、すべての臨床試験を含む成文法がないわけではない<sup>(45)</sup>。「市民的及び政治的権利に関する国際規約（国際人権B規約）」の第七条<sup>(46)</sup>には、「何人も、拷問又は残虐な非人道的若しくは品位を傷つける取り扱い若しくは刑罰を受けない。特に、何人も、その自由な同意なしに医学的又は科学的実験を受けない」と規定されている。この第七条は一九五八年に国連総会で採択され、日本も一九七九年に批准している。すなわち、日本で唯一、臨床試験全般を被験者の同意なしに行ってはならないとする法的拘束力のある規制と解釈できないことはない。ただし、この規約は医学研究に特化したものではなく、「医学的又は科学的実験」が何を意味するのか説明していないため、実効性を求めるには抽象的すぎる。

やはり、米国などにならい、臨床研究に関する基本法を設けたうえで全省共通の規則を作り、臨床試験すべてに倫理審査と同意書の取得を義務づけるべきである<sup>(47)</sup>。

### 3 倫理審査の問題点

臨床試験を実施するにはICの取得が必須条件だが、これだけで十分ではない。理念的にはともかく、完全なICの取得は非現実的なものだ。ICは、被験者が研究内容を完全に理解しないと成立しないが、完全に理解することなど医療関係者ですら難しい。現実的には、最後は医師を信頼して同意するという部分が残る。このICの不完全さを補完するため、倫理審査という方法が考案されたのである。したがって、たとえ被験者がよく理解しないまま同意したとしても、被験者の人権と安全は守られることを保証するのが倫理審査委員会の最も大きな役割である。

しかしながら、今日の研究倫理審査のあり方には様々な問題が残っている。なかでも最大の問題は「審査の質をどう確保するか」ということで、ほとんど手つかずの状態が残されている。

IRBが臨床試験を承認するということは、一般社会にその適切性を保証するということである。そのためには、一般社会が認めた基準に従って審査しなければならない。ところが日本の倫理指針は、審査は求めているものの、何に基づいて判断するか、つまり審査基準については述べずに、各々のIRBに任せきりである。米国には法の下にベルモント・レポートに定めた倫理原則があり、IRBのために実用的な手引き書なども作られている<sup>(48)</sup>。しかしながら日本では、治験のIRBにすら倫理原則も判断基準もない。また、委員になるのに資格はいらないばかりか、IRB委員のための研修の機会もほとんどない。

審査のレベルを向上させるには、第一に、確固たる倫理原則と審査基準を設けること、第二に、委員のため教育の機会を設けることが必要である。

最近我々は、IRB委員用の教育プログラムを作る目的で、まずIRBの現状を知る必要があると考え、福岡県下のIRBにアンケート調査を行ったが、その結果、倫理審査の現場で奇妙な認識が形成されていることがわかった。IRBとは「治験を審査する委員会」のことだと誤解しているところが多いのである。臨床研究を審査する委員会ならずしてIRBに該当するはずだが、治験以外の臨床研究の審査だけを担っている委員会には、自分たちの委員会がIRBだという認識がないところが多かった。そういう委員会では、委員の構成についても倫理指針の要求を満たしていないところが多く、倫理性の審査のみで、科学性については審査していないと答えた委員会も少なからずあった。また、治験以外の研究はIRBの審査から漏れている可能性があることも示唆された。

治験には厳格な法規制がある一方、それ以外の臨床研究にはないことが、このような誤解や逸脱を生む原因であろう。

もつとも、いろいろ問題はあるにせよ、倫理審査委員会が「ある」というだけで、「ない」場合とは雲泥の違いであることは間違いない。IRBが存在するというだけで、少なくともピーチャーが告発したような常軌を逸した研究は発案すらされないだろう。

倫理審査という方法は日本が自ら開発したわけではない。だからこそ、なおさら、この貴重な手段を世界が手にするまでの歴史をよく理解し、大切に守り発展させなければならない。

## 三 診療と研究の境

### 1 研究とは何か

#### (一) 研究の定義

診療（プラクティス）と研究（リサーチ）の関係を図示すると図2のようになる。A領域は純粋な診療行為、B領域は純粋な臨床研究を意味する。C領域は、認められた治療法どうしを比較検討する場合のように、診療と研究の要素が混在する領域である。

臨床研究全般を法規制するなら、研究（B領域＋C領域）とは何を意味するのか、はつきり定義づけておく必要がある。法の適用範囲が不明瞭のままでは、規制は不可能だからである。では、研究（臨床研究）の定義について現行の倫理指針は何と言っているだろうか。

「臨床研究に関する倫理指針」は、臨床研究の定義についてこう述べている。「医療における疾病の予防方法、診断方法及び治療方法の改善、疾病原因及び病態の理解並びに患者の生活の質の向上を目的として実施される医学系研究であつて、人を対象とするもの（個人を特定できる人由来の材料及びデータに関する研究を含む。）をいう」。

しかし、この定義では「研究」という言葉を既知のものとして用いており、それが何を指すかについては説明していない。おそらく、診療と研究の違いを明示した倫理指針は、今の日本には存在しない。これでは、指針が

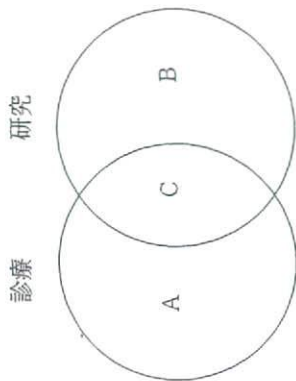


図2 診療と研究

A領域は純粋な診療（プラクティス）を、B領域は純粋な研究（リサーチ）を、C領域は診療と研究が同時に行われる場合を意味する。「診療」とは多くの場合「治療」であり、その場合、B領域は非治療的研究、C領域は治療的研究である。

特にこれが当てはまる。実験的側面が極めて強い場合でも治療といつていいのかが、そのような行為は研究に属させるべきなのか、これについては多くの議論がある<sup>(4)(5)</sup>。

このような状況の下、米国では研究を次のように定義している。

まずベルモント・レポートは、診療と研究の区別は容易ではないとしながらも、研究とは「仮説を検証し、結論を導き出すことを可能とし、それによつて、一般化可能な知識を開発したり、そのような知識に貢献したりするように考案された活動」と定義している。また、研究の内容は「目的と、目的を達成するためにデザインされた一連の方法を説明する正式の計画書」に記載されるとしている。さらに、診療と研究が同時に行われる場合の倫理審査について、「行為の中にわずかでも研究の要素が含まれるのであれば、被験者を保護するために、その行為は審査を受けるべき」だとしている。

次に、連邦行政規則第四五編第四六部（45CFR46）は、研究を「一般化可能な知識を開発したり、そのような知識に貢献したりするように考案された体系的調査」、研究対象者を「研究を実施する者が、介入したり影響を与えたりすることによりデータを得たり、識別可能な個人情報を得たりする生きた人物」と定義している。これらを総合すれば、臨床研究とは「人に介入を行つてデータを得たり、人を特定できる情報を収集したりすることにより、一般化可能な知識を生むように考案された体系的調査」となろうか。

米国のIRBの中には、臨床研究の審査を申請する研究者に向けて、45CFR46の定義をもとに、どのような行為が研究に当たるかをわかりやすく説明するフローチャートを、インターネット上に公表しているところもある<sup>(6)</sup>。

さらに、CIOMSとWHOの倫理指針では、「研究は」一般化可能な知識を獲得するため、またはそれに貢

述べているのは枝葉末節だけで、根幹は述べていないということにならないだろうか。

ただし、この指針では、「診断及び治療のみを目的とした医療行為」が対象外の行為として挙げられている。非該当部分（図2のA領域）を示すことで間接的に該当部分を定義したと言えなくもないが、それならば「診断及び治療」の正確な定義が必要になる。しかし、臨床研究を対象とする指針である以上、研究の定義から逃げてはならないだろう。

では、研究とはいかなるものか。

臨床試験については、「予防・診断・治療法の安全性や有効性を試す行為」と言われる。しかし、この定義が不十分なことは明らかである。なぜなら、認められた予防・診断・治療法といつても、安全性・有効性が100パーセントということはない。どんなに標準的とされる予防・診断・治療法でも、試してみなければわからない、すなわち実験的な側面がある。つまり、実験的側面があるから研究だとは言えないのである。標準的治療法が有効でない患者のため、未承認の（多くは安全性・有効性が確立されていない）新しい治療法を試して患者を救おうとするような場合、

献するために計画された活動」と定義されている。一般化可能な知識とは、「理論、法則もしくは関係性、あるいはそれらの基礎となる情報の蓄積からなり、科学的に受け入れられた観測方法と推測方法により確認することができる」ものである。

研究の定義は確かに容易ではない。明らかなのは、医師（研究者）の見かけ上の行為で診療と研究を区別することはできないということである。本質的な違いは行為ではなく、その目的にある。診療が目的とするのは目の前の患者個人だが、研究が目的とするのは未来の患者群だからである。

## (二) 一般の認識

では、日本では、臨床研究とはどのようなものと考えられていて、診療とどう区別されているのだろうか。

最近我々は、先に述べたIRBに対するアンケート調査の中で、治験以外の臨床研究を審査対象としている六六委員会を対象に、審査対象となる研究と一般の診療をどう区別しているか調べてみた<sup>⑧</sup>。

その結果、診療と研究の区別について議論になったことがあるIRBは少なく、「診療と研究は渾然一体となっているので分けることは不可能」とする意見や、「倫理的に問題となる事柄があるのなら治療と研究の別なく審査すべきだ」という意見もあった。

研究と見なす項目を選択肢から選んでもらったところ（複数回答可）、以下のような項目が選択された（多い順）。

未承認の薬物・医療機器を用いる場合：四一委員会

保険適応外の方法を用いる場合：三三委員会

先進医療を実施する場合：三二委員会

診断・治療という目的以外に、一般的知識の獲得という目的が存在する場合：三〇委員会

症例報告を行う場合：一七委員会

「診断・治療という目的以外に、一般的知識の獲得という目的が存在する場合」は、(一)で述べた「研究の定義」に相当する項目だが、むしろそれ以外を選択するIRBが多かった。多くのIRBが選択した「未承認の薬物・医療機器を用いる場合」「保険適応外の方法を用いる場合」「先進医療を実施する場合」は、(一)の定義からは必ずしも研究に当たらない。一方、「症例報告を行う場合」は明らかに研究に該当するが、これを選んだIRBが少ないのは、症例報告まで審査しているIRBはおそらくほとんどないからだろう。

研究と判断する基準を自由回答で求めたところ、三つのIRBから回答があった。

- ① 何らかの成果を出す目的で、綿密な実施計画に基づき行われる場合
- ② 複数の患者に同じ内容の治療を一律に行う場合
- ③ 患者に治療法の選択権が与えられていない場合

①は、行為の目的とプロトコルの存在で判断するもので、米国の連邦行政規則の定義に似ている。②は、個々の患者に最良と思われる治療を行うのではなく、複数の患者を一律の方法で治療するなら研究と判断するというもので、臨床試験の一面をとらえている。③は、無作為割付など、患者に治療法の選択権が与えられない場合は研究とみなすというもので、これも臨床試験の一面をとらえている。

今後IRBの審査基準を設けるに当たり、これらを参考に、研究とは何か、診療とどう違うか、明確に定義づけようと考えている。そこが、IRB審査の起点となるべきだと考えるからである。

本章で詳しく述べる余裕はないが、日本では病院倫理委員会（診療について検討する委員会）と研究倫理委員会（研究について審議する委員会）がしばしば混同されている。これも、診療と研究がはっきり区別できていないからとも言えるかもしれない。

## 2 診療のIC、研究のIC

「ヘルシンキ宣言」は、臨床研究の対象者に、研究の「目的、方法、資金源、起こりうる利益相反、研究者と研究組織との関係、予想される利益と危険性、および研究に伴う不快感」について説明することを定めている。

では、なぜ、このような情報を伝えなければならないのだろうか。なぜ、無断で臨床研究（ここでは特に臨床試験）を行ってはならないのだろうか。なぜ、診療のICだけでは不十分で、別に研究のICを取得する必要があるのだろうか。

「ヒポクラテスの誓い」にあるように、「自分の力の限り病人を助けるために治療に当たり」「病人にとって無益有害なことは決してしない」のが医師の大前提であるが、この大前提に立つかぎり、臨床試験は、医師にとって本来は行ってはならない行為である。なぜなら、臨床試験は、被験者本人のためではなく、将来の患者のために行われるものであり、被験者にとって不必要な処置を行う可能性、被験者にとってベストではない介入行為を行う可能性、被験者に害を及ぼす可能性があるからである。

しかし、「医学・医療は進歩すべきである」という理念を是とするかぎり、臨床試験は避けられない。実際に患者に試してみなければ、新しい治療法の安全性・有効性はわからないからである。医学の進歩が要求するこのような新しい状況に、古典的な「ヒポクラテスの誓い」だけでは対応できない。

そこで、この「誓い」の限界を克服するために考案されたのが、ICの取得という条件であった。患者にとっては不必要で有害かもしれない行為を正当化するには、患者が十分理解したうえで自発的に参加しているという保証が求められたのである。また、他者あるいは社会の利益のために行われる研究という行為は、患者の利益のためだけに行われる診療行為とは性質の異なるものなので、診療のICだけでは足りず、別個ICの取得が必要なのである。

以上より、研究のICを診療のICとは別に取得する必要があることは明白である。しかし、診療のICと違い、研究のICは取得が難しい。無断で研究を実施しようとする医師が後を絶たない一つの原因はこの難しさであろう。

臨床試験では「医師と患者の関係」と「研究者と被験者の関係」がしばしば併存している。医師と患者の目的は、どちらも患者のQOL向上であることが普通で、一致しやすい。一方、研究者と被験者の目的は別の方向を向いていることが多い。研究者の目的は研究結果を得ることだが、被験者はより優れた治療に出会えることを求めているからだ。ICを取得しようとする医師は、研究者としては研究に参加してもらいたいと思っても、医師の立場に立つと必ずしも患者のためになるとは限らない研究に参加させることに躊躇する。ここに一種の利益相反状態（利益と義務の衝突）が発生する。つまり、研究は、研究者の興味と医師の義務との衝突によって生ずる矛盾を必ず内包しているのである。そのような意味で、倫理的な問題が全くない研究はあり得ない。研究であれば全て倫理審査の対象となるのはこのためである。

ICの取得においては、この矛盾を含めて研究とは何かを被験者に伝えなければならない。研究のIC取得の難しさはここにある。この困難を克服するには、診療のIC以上に、医師（研究者）と患者（被験者）の信頼関

係が要求される。

### 3 金沢大学事件判決が示すもの

金沢大学の事例は、臨床試験のICについて司法が下した判決と受け取られがち、患者遺族の主張が通ったかのように見られがちだが、厳密に言うところではない。

原告遺族の主張は、「患者は知らないうちに比較臨床試験<sup>(8)</sup>の対象にされ、患者人格権が侵害され精神的被害を被った」ということである。これに対して被告病院は、患者に行った医療行為は「比較臨床試験」に該当しないと主張し、遺族側と対立した。我々が判決文を読む限り、この行為が図2のC領域に含まれる臨床試験であったことは明らかだ。ところが裁判所は、本件の行為を「比較臨床試験」と呼ぶことに躊躇している。その結果、このIC取得義務違反を説明するのに、治療以外の目的（他事目的）についての説明義務違反という、回りくどい論理を述べることになった。

もし、臨床試験の定義に該当すると明確に認めていれば、そのような回りくどい論考は必要なかった。臨床試験であるにもかかわらずICを取得しなかったということは、たとえ法規制はないとしても「ヘルシンキ宣言」に違反しているからである。

前節に例示した裁判のうち、愛知県がんセンターの事例は治験（明らかに研究）なので、「それが研究かどうか」は争点にはならなかったが、金沢大学の事例は治験ではなく法の枠外の臨床試験であったからか、本質的には愛知県がんセンターの事例同様の臨床試験であったにもかかわらず、臨床試験であるという明らかな見方をしていない。

法の枠外の臨床試験について司法がどう判断するか注目され、確かにIC取得義務違反は認められたわけだが、その論拠は事件の本質を突いていない。被告の行為を臨床試験だとはつきり認めることを避けたからだ。その結果、当該患者を被験者として扱った判決とはならず、あくまで治療法選択における患者不在の問題としてしか扱われなかった。

しかし、本事例における医師の行為は、ベルモント・レポート、連邦行政規則令CFR6、CIOMS倫理指針のいずれの定義に照らしても、さらに、我々の調査で提示された基準に照らしても、研究（臨床試験）であったことは明白である<sup>(9)</sup>。ところが、裁判所はそういう見方をしなかった。

この判決により、「日本の研究には法的根拠がない」ことがあらためて示されたことになる。これでは、再び同様の事件が発生した時の判断はその時の裁判所次第ということになり、将来に不安を残したままなのである。

### おわりに

現在、様々な団体が競って臨床試験に関する教育制度や認定制度を作ろうとしており、将来の混乱の源となりそうなので危惧している。

本章で述べたように、日本には臨床試験を規制する基本的な枠組みができていない。それがなくところに闇雲に教育・認定制度を作っても、混乱するばかりである。ここは、「研究者と被験者双方に最も有益なシステムは何か」という基本に立ち返って考えるべきだ。

まず、治験だけ特別扱いするのではなく全ての臨床研究を（それが無理なら臨床試験だけでも）法規制する。

法規制すると研究がやりにくくなると思われがちだ。しかし、治験以外の臨床試験が法の枠外に置かれている現在では、それらの研究が法的に認められているのかいないのか（公認されているのかいないのか）あやふやである。研究者にとっても被験者にとっても、安心して臨床試験を行える基盤がないのである。是非とも、基本法で全てを括り、その下に共通の倫理指針を設けなければならない。

そして次が教育である。臨床試験を実施するに当たって最も基本的な条件は、科学的合理性・IC・倫理審査の三つである。これらの条件を質とともに保証するため、研究者はもちろん、研究協力者、IRB委員など、臨床試験に関わる人々の教育と資格認定を行う。特に、ほとんどの医師が何らかの形で臨床研究に関わる現実を考えると、倫理も含めた臨床研究教育を医学部や医学系大学院で十分行う必要がある。

本文でも述べたが、臨床試験には、医師と患者の良好なコミュニケーションが一般診療以上に要求される。医師にとっては専門教育以前に人格の形成が必要だ。また、被験者となる一般の人々に対して、臨床試験についてわかりやすく説明し、正しい理解を得ることも重要だ。

以上のような諸条件が整わないかぎり、「人が人を研究材料にする」という極限的行為が正当化されることはないのである。

## 注

(1) 厳密に言う、企業主導の臨床試験には治験でないものもあり、また今日では治験を医師主導で行うこともあるが、大部分の臨床試験はこの二つに分類できる。

(2) ヒポクラテス（原典、常石敬一訳、解説『ヒポクラテスの西洋医学序説』（小学館、一九九六年）二二六―二二七頁。

(3) ジェンナーの実験では近所の少年が、華岡青洲による全身麻酔の実験では母と妻が被験者になっている。

(4) Grodin MA. Historical Origins of the Nuremberg Code. Annas GJ & Grodin MA, ed. *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*, Oxford University Press, 1992, pp. 121-144.

(5) 同。

(6) クロト・ベルナルド、三浦俊栄訳『実験医学序説』（岩波書店、一九七〇年）一六七―一六八頁。

(7) 同、一六八―一六九頁。

(8) 同、一六九―一七〇頁。

(9) Grodin MA. Historical Origins of the Nuremberg Code. Annas GJ & Grodin MA, ed. *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*, Oxford University Press, 1992, pp. 121-144.

(10) 「ドイツ人体実験指針」生命倫理百科事典翻訳刊行委員会編『生命倫理百科事典』（丸善、二〇〇七年）三〇七―三〇八頁。

(11) 次世代の人間を「改良」するため、「優れた」人間を選別し「劣った」人間を排除することを目的に、公衆衛生学と優生学が結びついて生まれた似非学問。ナチス・ドイツの人種政策に利用された。

(12) アイリントン・ウェルカム、渡辺正訳『アルトニウムファイル（上・下）』（翔泳社、二〇〇〇年）。

(13) 河井智康『原爆開発における人体実験の実相』（新日本出版社、二〇〇三年）。

(14) ドストエフスキ『罪と罰』の主人公が金貸し老婦の殺害を正当化するのに用いた論理。

(15) Grodin MA. Historical Origins of the Nuremberg Code. Annas GJ & Grodin MA, ed. *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*, Oxford University Press, 1992, pp. 121-144.

(16) 「ニュルンベルク綱領」生命倫理百科事典翻訳刊行委員会編『生命倫理百科事典』（丸善、二〇〇七年）三〇七―三〇九頁。

(17) 治療者の倫理規範の原型「ヒポクラテスの誓い」に相当する、研究者の倫理規範の原型と言える。

(18) 世界医師会「ヘルシンキ宣言」、生命倫理百科事典翻訳刊行委員会編『生命倫理百科事典』（丸善、二〇〇七年）三〇八―三〇九頁。

(19) 一九八二年に原案が作られ、一九九三年、二〇〇二年に改訂版が公表された。



- (20) 国際医科学機関評議会・世界保健機構(協力)「人を対象とする生物医学研究についての国際的倫理指針」生命倫理百科事典翻訳刊行委員会編『生命倫理百科事典』(丸善、二〇〇七年)三二二―三二四頁。
- (21) Rothman DJ. *Strangers at the Bedside*. Basic Books, 1991.
- (22) Beecher HK. Ethics and Clinical Research. *New England Journal of Medicine* 274, 1966, pp. 1354-1360.
- (23) 審査委員会を設置した施設はそれ以前にも存在したが、連邦指針として公布されたのはこれが初めてだった。国際指針として初めて倫理審査を取り入れたのは「ヘルシンキ宣言」の東京改訂版(一九七五年)である。
- (24) Rothman DJ. *Strangers at the Bedside*. Basic Books, 1991.
- (25) Jones JH. *Bad Blood: The Tuskegee Syphilis Experiment (New and Expanded Edition)*. Free Press, 1993.
- (26) 一見したところ非介入研究に見えるが、「治療法を与えない」という人為的介入を行った「人体実験」に他ならない。
- (27) 被験者保護のための国家委員会「ベルモント・レポート: 研究被験者保護のための倫理原則および指針」生命倫理百科事典翻訳刊行委員会編『生命倫理百科事典』(丸善、二〇〇七年)三〇八―三〇九頁。
- (28) 「DHHS 被験者保護規則」生命倫理百科事典翻訳刊行委員会編『生命倫理百科事典』(丸善、二〇〇七年)三〇九―三一〇頁。
- (29) 一九九七年、ICの概念は医療者の努力義務として医療法に盛り込まれた。
- (30) 常石敏一『医学者たちの組織犯罪: 関東軍第七三部隊』(朝日新聞社、一九九四年)。
- (31) 例えば、「ヘルシンキ宣言」のすべてを読んで理解している医師はまれで、これを「ヒポクラテスの誓い」の現代版(医師の職業倫理を述べたもの)であるかのように誤解している者が多い。
- (32) 日本商事が開発した骨状癌治療薬ソリアジンが抗がん剤との相互作用により発光直後に多数の死者を出した事件。
- (33) 薬事法に基づく厚生省令「医薬品の臨床試験の実施の基準」(新GCP)が公布された。
- (34) 『朝日新聞』(朝日新聞社、二〇〇七年七月二十九日)。
- (35) 『判例タイムズ』六五七号(判例タイムズ社)一六〇頁。
- (36) 薬の臨床試験は三つの相から成り、第二相は薬の効果を患者で初めて調べる段階を指す。
- (37) 『判例時報』一七三三三号(判例時報社)七〇頁。
- (38) 仲正昌樹・仁木恒夫・打出善義「人体実験」と患者の人格権——金沢大学付属病院無断臨床試験訴訟をめぐる——(御茶の水書房、二〇〇三年)。

- (39) 仲正昌樹・安西明子・打出善義・仁木恒夫「人体実験」と法——金沢大学付属病院無断臨床試験訴訟をめぐる——(御茶の水書房、二〇〇六年)。
- (40) 『判例タイムズ』二二〇九号(判例タイムズ社)二五三頁。
- (41) 「TKC法律情報データベース:LEX/DBインターネット」文獻番号二八〇一九三三。
- (42) 光石忠敏「臨床試験はどうあるべきか: 被験者・患者のおかれる立場」青藤孝雄監修・神山有史編集『生命倫理学講義: 医学・医療に何が問われているか』(日本評論社、一九九八年)一五三―一八六頁。
- (43) 国際連合総会「市民的及び政治的権利に関する国際規約」生命倫理百科事典翻訳刊行委員会編『生命倫理百科事典』(丸善、二〇〇七年)三〇八頁。
- (44) 笹栗俊之「臨床試験: 系統立てすべてに法規制を」『朝日新聞』(朝日新聞社、二〇〇七年八月二十九日)。
- (45) Greenwald RA, Ryan MK, Mulvihill JE, ed. *Human Subjects Research: A Handbook for Institutional Review Boards*. Plenum Press, 1982/阿岸鉄三・今里嘉夫・土肥修司・羽賀通信訳『被験者保護ハンドブック: アメリカIRBの活動』(地人書館、一九八七年)。
- (46) Barkert EA, Andur RJ, ed. *Institutional Review Board Member Handbook, 2nd ed.* Jones & Bartlett Publishers, 2006.
- (47) 笹栗俊之「IRBメンバー教育研修プログラムの開発を目的とするIRB実態調査」『厚生労働科学研究費補助金臨床研究基盤整備推進研究事業「臨床研究基盤整備の均てん化を目指した多目的教育プログラムと普及システムの開発」(主任研究者: 山本禮一郎)平成十八年度総括・分担研究報告書』(二〇〇七年)一四―一五頁。
- (48) 予防・診断・治療の総称として用いた。
- (49) Miller FG & Rosenstien DL. The Therapeutic Orientation to Clinical Trials. *New England Journal of Medicine* 348, 2003, pp. 1383-1386.
- (50) Grunberg SM & Cefalu WT. The Integral Role of Clinical Research in Clinical Care. *New England Journal of Medicine* 348, 2003, pp. 1386-1388.
- (51) The University of South Dakota Office of Human Subjects Protection (<http://www.usd.edu/oorsch/compliance/SOPs/GA102-AIsithumansubjectresearchflowchart.pdf>)
- (52) 柴田智美「無断臨床試験事例判決の問題点に関する考察: 『研究』と『治療』の境界の検討」九州大学大学院医学系学府医学専攻修士論文(二〇〇七年)。

## 執筆者一覧

- 山崎喜代子 (やまざき きよこ) はしがき・第2章  
西南学院大学人間科学部教授。生物学 (動物発生)。
- 北垣 徹 (きたがき とおる) 第1章  
西南学院大学文学部准教授。社会学 (知識社会学・社会思想史)。
- K. J. シヤフナー (Karen J. Schaffner) 第3・4章  
西南学院大学国際文化学部教授。  
社会倫理学 (神学・アメリカ文化史)。
- 河島 幸夫 (かわしま さちお) 第5章  
西南学院大学法学部教授。政治学 (ドイツ政治と宗教)。
- 山本 裕子 (やまもと ゆうこ) 第6章  
西南学院大学人間科学部准教授。  
社会福祉学 (医療ソーシャルワーク)。
- 中馬 充子 (ちゅうまん みつこ) 第7章  
西南学院大学人間科学部教授。保健学 (保健教育史)。
- 堀口 良一 (ほりぐち りょういち) 第8章  
近畿大学法学部准教授。歴史学 (安全運動史)。
- 新福 尚隆 (しんぷく なおたか) 第9章  
西南学院大学人間科学部教授。精神医学 (社会精神医学)。
- 片桐 英彦 (かたぎり ひでひこ) 第10章  
医療法人片桐クリニック理事。産婦人科学 (内分泌・不妊)。
- 笹栗 俊之 (ささぐり としゆき) 第11章  
九州大学大学院医学研究院教授。  
臨床薬理学 (分子薬理・臨床薬理・研究倫理)。
- 柴田 智美 (しばた ともみ) 第11章  
九州大学大学院医学系学府医科学専攻修士課程。  
臨床薬理学 (研究倫理)。

- (53) 笹栗俊之「研究倫理教育の徹底を：薬理学教育に求められるもの」『日本薬理学雑誌』一二四巻(二〇〇四年)一八二～一八六頁。
- (54) 伊藤友希「倫理委員会における症例コンサルテーション機能の調査及び検討」九州大学大学院医学系学府医科学専攻修士論文(二〇〇六年)。
- (55) ヒポクラテス(原典)、菅石敬一訳・解説『ヒポクラテスの西洋医学序説』(小学館、一九九六年)二二六～二二七頁。
- (56) Levine R.J. *Ethics and Regulation of Clinical Research (Second Edition)*, Yale University Press, 1986.
- (57) ほとんどの臨床試験がそうだが、比較対照群を置いた臨床試験のことを指す。
- (58) 柴田智美「無断臨床試験事例判決の問題点に関する考察：『研究』と『治療』の境界の検討」九州大学大学院医学系学府医科学専攻修士論文(二〇〇七年)。

## 7

## 臨床試験審査委員会 (IRB)

### 7-1. IRBの歴史と概念

臨床試験審査委員会 (IRB) とは、当該臨床試験 (治験) に直接関与する者から独立した第三者によって試験の是非を審議する会議である。最も重要な任務は被験者の権利と安全を守ることだが、非科学的な研究を行うことは被験者や社会に負担をかけるため非倫理的であるという観点から、研究の科学的根拠も厳格に審査すべきである。したがって、研究の倫理性と科学性の両方について審議する能力を持たなければならない。

臨床研究を他人が審査するという手続きは、今日では当然のこととされている。しかし、歴史的にみれば、多くの人々の犠牲を伴った長い困難の後に人間社会がようやく獲得した、かけがえのない方法であることを忘れてはならない<sup>1)</sup>。

臨床研究を審査するというコンセプトを初めて取り入れたのは、米国国立衛生研究所 (NIH) である。1950年代より、NIHの臨床研究センターは、NIHで行われる全研究を委員会で倫理審査する手続きをとっていたが、1964年、NIH長官ジェームズ・シャノン<sup>2)</sup>は、公的資金で行われる全研究を審査する方針を打ち立てた。1966年、ヘンリー・ビーチャーによる非倫理的臨床研究の告発<sup>2)</sup>をきっかけとして、研究倫理に関する米国世論が高まり、臨床研究はもはや専門家集団だけのものではなく、一般社会が認めるものでなければならなくなった。これに対し、米国食品医薬品局 (FDA) はインフォームドコンセントに重点を置いた政策を

発表した。NIHはインフォームドコンセントには限界があるとし、研究施設の同僚による審査という手続きを指示した。1974年、「タスキギー梅毒研究」の暴露<sup>1)</sup>などに後押しされて成立した国家研究法により、倫理審査システムは法的に確立した。倫理審査という方法は、1975年の改訂でヘルシンキ宣言に初めて導入され、世界に広まることとなった。以上が、研究倫理審査の始まりである。

米国では倫理審査委員会は研究施設単位に設けられたため、施設内審査委員会 (Institutional Review Board, IRB) と呼ばれた。1960年代の米国の臨床研究は、少数の被験者を対象に一施設内で行うものが大半であったため、施設毎に審査することになったのである。後に日本もこれに倣ったため、日本の委員会も「IRB」と通称される。

一方、英国やヨーロッパ諸国では施設単位ではなく、施設とは独立した委員会を地域毎などに設置することが多く、研究倫理委員会 (Research Ethics Committee, REC) などと呼ばれている。どちらの設置形態にも長所と短所がある。施設内委員会では、独立性を保つのが容易でないこと、多施設共同研究の審査が困難なことなどの問題があり、一方、独立委員会では委員会ショッピング (研究者にとって都合のよい委員会を選べる) などの問題がある。今日ではこれらの折衷様式などさまざまな方法が試みられている。

また、ヒトゲノム解析を含む試験など、特定の専門的事項に関する審査が必要な研究に備

え、親委員会の下に専門部会を設置する場合が増えていく。

## 7-2. IRBの現状

### (1) IRBの設置

治験および製造販売後臨床試験（以下、「治験等」）を実施する医療機関の長は、治験等に関する調査審議をIRBに行わせなければならないことが、GCPにより定められている。IRBの設置は今日かなり自由化され、試験実施医療機関の長以外に、民法により設立された法人、特定非営利活動法人、医療関係の学術団体、学校法人、独立行政法人、国立大学法人、地方行政法人がIRBを設置できることになっている。自機関にIRBを設置できない場合や、研究の内容からみてそうしたほうがよいと判断される場合には、他機関のIRBに審査を依頼したり、意見を聞いたりすることができる。

一般の臨床試験の場合は、厚生労働省の「臨床研究に関する倫理指針」に従ってIRBを設置する。設置や依頼の条件は、治験等とほぼ同様である。

### (2) IRBの構成

簡単にいえば、IRBは、①医学・薬学の専門家、②医学・薬学の専門家でない者（非専門家）、③外部委員、から構成される。GCPと「臨床研究に関する倫理指針」では若干異なるため（表III-3）、治験等も一般臨床試験も審査するIRBは両方とも満たす必要がある。

非専門家として、現状では事務員や弁護士などを加えている場合が多いが、医学・薬学の専門外であれば誰でもよいわけではない。見識の高い一般人として、主に倫理的観点から被験者の人権保護に関する審議を期待される。「臨床研究に関する倫理指針」では「人文科学の有識者」と明記されている。

外部委員は委員会の第三者性を確保するために加えられる者であり、GCPでは「実施医療機関（および）治験審査委員会の設置者と利害関係を有しない者」とされる。現状では外部委員にも医学・薬学の専門家でない者が当てられる場合が多い。しかし、研究の科学性については、実質的に審議するためには、外部の専門家も委員に加えられるべきであろう。

表 III-3. IRB の構成

「医薬品の臨床試験の実施の基準に関する省令」（第28、29条より抜粋・要約）

- ・ 治験について倫理的及び科学的観点から十分に審議を行うことができること。
- ・ 5名以上の委員からなること。
- ・ 医学、歯学、薬学その他の医療又は臨床試験に関する専門的知識を有する者以外の者が加えられていること。
- ・ 実施医療機関と利害関係を有しない者が加えられていること。
- ・ 治験審査委員会の設置者と利害関係を有しない者が加えられていること。
- ・ 次に掲げる委員は、審査の対象となる治験に係る審議及び採決に参加することができない。
  - 1) 治験依頼者の役員又は職員その他の治験依頼者と密接な関係を有する者
  - 2) 自ら治験を実施する者又は自ら治験を実施する者と密接な関係を有する者
  - 3) 実施医療機関の長、治験責任医師等又は治験協力者

「臨床研究に関する倫理指針」（第3.倫理審査委員会（5）細則より）

1. 医学・医療の専門家等自然科学の有識者、法律学の専門家等人文・社会科学の有識者及び一般の立場を代表する者から構成され、かつ、外部委員を含まなければならない。また、男女両性で構成されなければならない。
2. 審議又は採決の際には、自然科学分野だけではなく、人文・社会科学分野又は一般の立場を代表する委員が1名以上出席していなければならない。
3. 臨床研究機関の長など審査対象となる臨床研究に携わる者は、当該臨床研究に関する審議又は採決に参加してはならない。ただし、倫理審査委員会の求めに応じて、会議に出席し、説明することはできる。
4. 臨床研究機関の長は、必要に応じ、会議に出席することはできる。ただし、当該者は倫理審査委員会の委員になること並びに審議及び採決に参加することはできない。

表 III-4. IRB の役割

## 1. 試験の計画段階での審査

新たに試験計画が提出されたとき、およびすでに承認された試験計画が変更される場合に審査する。当該計画が倫理的・科学的に妥当かどうか、また当該機関における実施が適切であるかどうかを審議し、意見を述べる。

## 2. 試験の実施中の審査

試験が適切に行われているかどうかを定期的に（ふつう1年に1度）審査し、試験の継続が妥当かどうかを審議する。

## 3. 試験の終了後の審査

試験が適切に行われていたかどうかを審査し、意見を述べる。

## 4. 試験に何らかの問題が生じたときの審査

重篤な有害事象の報告がなされた場合など、当該試験に関わる何らかの問題が生じた場合に、当該医療機関で試験が適切に行われているかどうかを調査し、試験を継続して行うことの適否を審議する。事態の緊急性に応じて速やかに意見を述べなければならない。

なお、委員会の第三者性を確保するため、施設長自身が審議に加わることはできない（本来IRBメンバーに施設長を加えるべきではない）。GCPも「臨床研究に関する倫理指針」も施設長の審議・採決への参加は禁止している（表III-3）。日本では、治験等以外の臨床研究の審査はいわゆる「倫理委員会」に任せているところが多い。それ自体は問題ないが、その委員会にIRBとしての機能があるかどうか確認が必要である。特に、病院長や学部長などが審議・採決に加わっている場合が現状ではかなりあるので、改められるべきである<sup>3)</sup>。

## (3) IRBの任務

IRBの役割は、臨床試験計画およびその実施状況が倫理的・科学的に適正かどうかを臨床試験（治験）実施計画書や被験者説明文書などの資料をもとに審査し、さらに、試験を担当する医師の資質をも含めて当該施設で実施するのが適切かどうか判断することである。

審査は一度だけで済むわけではない。治験等に関しては、表III-4の4つの場合にIRBで審議するよう、GCPで定められている。一般の臨床試験でも計画段階での審査（表III-4の1）はおおむね実施されていると思われるが、その後の審査（表III-4の2～4）については徹底されていないのが現状である。

表III-4に示すように、IRBには実施中の試

験を監視する役割があるため、新規申請課題がなくても毎月1回など定期的に開催すべきである。審議時間は内容にもよるが、新規課題なら1時間程度は費やして審査すべきである。会議の時間は限られているため、重要問題の審議に十分な時間を割り、被験者説明文書の語句修正などに多くの時間を費やすべきではない。治験等については治験事務局と治験依頼者の間で事前ヒアリングを行うことが多いので、書類の不備や表記の誤りがIRBで初めて指摘されることは少なくなっている。

IRB委員は、試験計画の内容について会議の前に十分把握しておくべきであり、内容を理解するために会議の時間を多く費やしてはならない。IRBで実質的な審議を行うためには、会議前に十分な調査を行う必要がある。特に研究の科学性についてはその場で判断することは困難なことが多く、事前調査が必須である。しかしながら、委員が事前に十分な調査を行うことは現実的には難しい。徹底した事前調査（あるいは予備審査）を行っているIRBはごく一部に限られている。申請課題ごとに担当委員を決めて事前調査を十分に行わせる方法も良いのではないかと思われるが、担当委員の負担増などの理由により、採用しているIRBは少ない<sup>3)</sup>。

もしも事前調査が十分になされていれば、依頼者や責任医師などの当事者をIRBに同席させる必要はないが、同席させていけば委員の質

問への答えが直ちに得られるので、審査の迅速化という意味ではそれも有効である。ただし、会議の独立性を保つため、同席を許可するのは説明と質問の時間に限り、審議と採決の場に研究の当事者を入れてはならない。

採決の方法は各 IRB に任されているが、大部分の IRB では全員一致を旨としており、多数決で決めるところはほとんどない（また、そうすべきではない）。採決結果には、承認、条件付承認、却下、保留（審議不十分）、取り消しなどがあり、判断理由とともに施設長に報告される。

#### (4) IRBの意見はどのように扱われるか

IRB の基本的な立場は試験実施機関の施設長の諮問委員会である。施設長は IRB の採決結果や意見を聞き、依頼者や責任医師にそれを文書により通知しなければならない。試験の実施は適切ではないという意見を IRB が述べた場合、施設長は、試験の依頼を受けたり、試験の実施を承認したりしてはならない。試験の継続が適切ではないという意見を述べた場合、施設長は、試験の契約を解除するか中止しなければならない。また、試験が適切に行われていないという意見を述べた場合、施設長は必要な措置を講じなければならない。このように、施設長は IRB の否定的な意見には逆らうことはできない。反対に、IRB が試験の実施や継続を承認した場合であっても、施設長が何らかの理由によりそれを受け入れられないと判断した場合は拒否することができる。

### 7-3. 残された問題

今日の研究倫理審査のあり方には2つの大きな問題がある。審査委員会の設置様式に関わる制度上の問題と、審査の質の問題である。前者は、施設単位か施設とは独立させるかという、先に述べた問題である。特に、今日のように多施設が共同で実施する臨床試験が増えてくると、施設単位の審査だけで対応するのは難しい。

中心に位置づけられる IRB で審査した後、各施設の審査にかかる2段階審査などの方法が試みられている。

一方後者、すなわち「審査の質をどう確保するか」という問題は、日本ではほとんど手つかずの状態に残されている。IRB が臨床試験を承認するということは一般社会にその適切性を保証することであるからには、一般社会が認めた基準に従って審査すべきである。ところが、GCP や「臨床研究に関する倫理指針」は審査の中身までは規定していないので、何をどう審査してどう判断するかは各 IRB に任されている。米国には国家研究法の下に通称「ベルモントレポート」という研究倫理原則があり、IRB のための実用的な手引書なども作られている<sup>4, 5)</sup>。しかしながら日本には、治験等を除けば規制法はなく、それどころか倫理原則もない。また、委員になるのに資格は要らないばかりか、教育・研修の機会すらほとんどない<sup>註)</sup>。

第一に、確固たる倫理原則と審査基準を設け、これに基づく実用的なマニュアルを作ること、第二に、IRB の登録制や連絡会を設け、IRB メンバーに教育・研修の機会を与えることが、審査のレベル向上のために必要である。

被験者を守る上で、倫理審査はインフォームドコンセントと並んで最も重要な手続きである。最近では倫理審査が形骸化しているという批判を耳にするが、形骸化するほど当然のことになったのだとすれば、それほど悪いことではないのかもしれない。IRB が存在するというだけで、少なくともピーチャーが告発したような常軌を逸した研究は発案すらされなくなるだろう。ただし日本では、倫理審査が法令で義務づけられているのは治験等のみである。一般の臨床試験でも審査を法的に義務づけるべきだと思われる<sup>6, 7)</sup>。

註：今回改正された「臨床研究に関する倫理指針」（2009年4月施行）には、「倫理審査委員会の設置者は、倫理審査委員会委員の教育及び研修に努めなければならない」という規定が加

わった。ただし、IRB 委員が教育を受けたり研修したりする機会はいまのところ極めて乏しいため、教育・研修プログラムを早急に作る必要がある。

#### ■文献

- 1) Rothman DJ : Research, Human : Historical Aspects. Post SG, ed, Encyclopedia of Bioethics, 3<sup>rd</sup> ed, pp. 2316-2326, Macmillan Reference, 2003 [笹栗俊之 (訳) : 研究 (人を対象とする) : 歴史的側面. 生命倫理百科事典翻訳刊行委員会編, 生命倫理百科事典, pp. 914-923, 丸善, 2007]
- 2) Beecher HK : Ethics and Clinical Research. N Engl J Med 274 : 1354-1360, 1966
- 3) 笹栗俊之, 柴田智美, 上口 愛, 他 : 福岡県における研究倫理審査委員会の実態調査—IRB メンバー教育システムの構築を目指して—. 臨床評価 36 (2) : 393-419, 2008
- 4) Greenwald RA, Ryan MK, Mulvihill JE, ed : Human Subjects Research, A Handbook for Institutional Review Boards, Plenum Press, 1982 [阿岸鉄三, 今里嘉夫, 土肥修司, 羽賀道信 (共訳) : 被験者保護ハンドブック—アメリカ IRB の活動, 地人書館, 1987]
- 5) Bankert EA, Amdur RJ, ed : Institutional Review Board Member Handbook, 2<sup>nd</sup> ed, Jones & Bartlett Publishers, 2006 [初版=栗原千絵子, 斉尾武郎 (訳) : IRB ハンドブック, 中山書店, 2003]
- 6) 笹栗俊之 : 「臨床研究法」の制定に向けて. 日本医事新報 4234 : 22-25, 2005
- 7) 笹栗俊之, 柴田智美 : 診療と研究の境—臨床試験の倫理. 生命の倫理 2—優生学の時代を越えて [山崎喜代子 (編)], pp. 297-328, 九州大学出版会, 2008

(笹栗 俊之)

## 2 サブグループに対する 治療開発のための臨床試験 デザイン

山本精一郎 (国立がんセンターがん予防・検診研究センター情報研究部JCOGデータセンター統計部門)

### P o i n t

- 個別化治療とは、何らかのマーカーに基づいてサブグループを定義し、そのグループに対して効果のある治療を行うことである。個別化治療であっても、それが有効であるかどうかは臨床試験による検証が必要となる。
- Phase III Trialでall comers design、  
(a) ある薬剤が対象者全体の全生存期間を延長するような標準治療となり得るかどうかだけでなく  
(b) 試験薬投与前に腫瘍組織で調べた薬剤に対する反応(マーカー)によって治療効果が予測できるかを主要な目的として研究デザインすることも可能である。
- マーカーが高いサブグループほど治療効果が大きい場合(linear仮説)や、マーカーが高いサブグループにだけ治療効果がある場合(threshold仮説)、どのサブグループに新治療を導入するかは、そのサブグループに対する治療効果がどの程度あるかということに依存する。

分子標的薬などの開発が進むにつれ、個別化治療という言葉をよく耳にするようになった。個別化治療とは、何らかのマーカーに基づいてサブグループを定義し、そのグループに対して効果のある治療を行うことである。個別化治療であっても、それが有効であるかどうかは臨床

試験による検証が必要となる。個別化治療の臨床試験デザインとして、大きく分けて図1の3つが提案されている<sup>1)</sup>。このうちのall comers designについて、本稿では深く掘り下げてみたい。

All comers designは、まずはじめにマーカーを調べるが、マーカー結果によらず、ランダム化



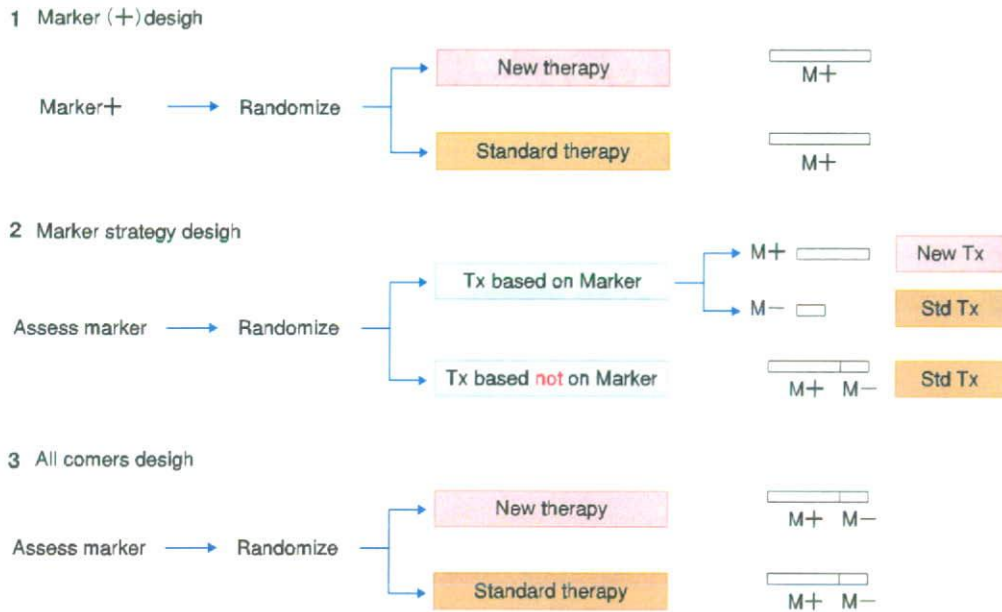


図1 ● 個別化治療のための臨床試験デザイン

して一方には新治療、もう一方には標準治療を行うものである。マーカー結果によらず両群を比較することによって、対象集団全体に対し新治療と標準治療の治療効果のプライマリな比較ができるだけでなく、マーカー (+) 群、(-) 群のそれぞれで新治療の治療効果を調べることができる。つまり、マーカー結果と治療の間に交互作用があるかどうか、言い換えるとマーカーが治療効果の予測因子かどうかを調べることができる。全体で治療効果をみた後に単にサブグループアナリシスをするのではなく、さらに進めて、検証的なサブグループアナリシスを組み込むにはどうすればよいか、感受性仮説などの proof of concept をどのように行っていけばよいか、について統計的な試みがいくつか提案されているので、本稿ではそれらの方法を議論するこ

とによって、個別化治療の臨床試験デザインについて理解を深めていただきたいと思います。

Simon<sup>2)</sup> は、事前に規定したマーカー結果によって定義されるサブグループでの検定に対して、 $\alpha$ エラー<sup>\*1</sup>のコントロール、つまり検定の多重性<sup>\*2</sup>を調整するようなデザインを組めばいいの

**\*1  $\alpha$ エラーと $\beta$ エラー**

統計的検定において、帰無仮説が正しいときに誤ってこれを棄却してしまう確率を $\alpha$ エラーとよぶ。 $\alpha$ エラーの水準は5%とすることが慣習的である。また、対立仮説を誤って採択できない確率を $\beta$ エラーとよぶ。 $1-\beta$ エラーが検出力となる。

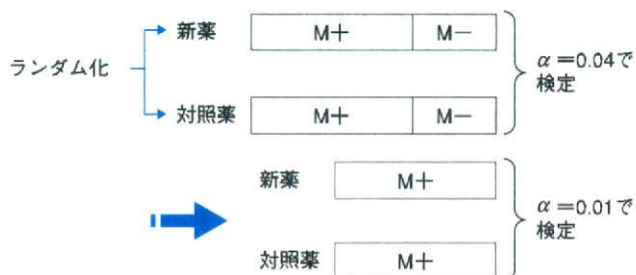
**\*2 検定の多重性**

1つの研究のなかで検定を何度も繰り返すと $\alpha$ エラーが大きくなることをいう。検定の多重性を調整して試験全体の $\alpha$ エラーを名義水準通りに抑えることを多重性の調整という。臨床試験における検定の多重性とその調整の例として、中間解析と最終解析における複数回の検定、複数のエンドポイントの検定、などが挙げられる。

図2 ● Simonの提案

(Simon R, JCO 2005. より引用)

- 事前に規定したサブグループに対して多重性を調整
- 「マーカーと治療の交互作用を調べるデザイン」で2つの仮説を検証する
  - ・ 全対象者での群間比較： $\alpha=0.04$
  - ・ マーカー(+)のサブグループでの群間比較： $\alpha=0.01$



ではないかという提案を行った(図2)。これはマーカーと治療の交互作用を調べるデザインで、2つの仮説を検証する。つまり、はじめに全対象者をランダム化して、2群間比較の検定を行い、さらにマーカー(+)だけで群間比較の検定を行う。ポイントは、全対象者について行う検定は $\alpha=0.04$ でやっておいて、そのあとマーカー(+)のサブグループだけで検定するときには $\alpha=0.01$ で行うというものである。この論文のなかでは、厳密な統計的な議論はしておらず、簡単な提案のみ行っている。

これをさらに進めたものとして、Freidlin & Simon<sup>3)</sup>は“Adaptive Signature Design”という方法を提案している(図3)。この方法は、マーカーを試験開始前に特定できない場合に特に有効である。この方法は次の3ステップからなる。全員をランダム化し、新治療群と対照群に分けて治療を行うことは同様である。

a) まず、割付を行った全対象者に対して $\alpha=\alpha_1$  ( $<0.05$ )で2群間比較をする。全体の対象者で差があるかをみる部分である。

b) 試験前半(1st stage)で登録された対象者のデータに対し、さまざまなマーカーを用いて治療効果の異なるサブグループを探索的に探し出す(training set)。次に後半の対象者(2nd stage)のなかから同一のサブグループを抽出する(test set)。すなわち後半の対象者のなかでこの薬剤に対して“感受性の高い”と考えられる対象者のサブグループを同定する。

c) サブグループを探索した1st stageのデータと独立である2nd stageのサブグループを用いて治療効果が異なるかどうかの検証を行う。すなわち、2nd stageのサブグループに対し、 $\alpha=\alpha_2$ で治療効果の群間比較の検定を行う(ただし、 $\alpha_1+\alpha_2=0.05$ となるように $\alpha_1, \alpha_2$ を定める)。

彼らは全対象者に対する検定を行う際の $\alpha_1$ を全体の $\alpha$ の80%、つまり $\alpha=0.05$ であれば $\alpha_1=0.04$ とし、探し出した“感受性の高い”と思われるサブグループに対する検定を行う際の $\alpha_2=0.01$ とすることを勧めている。この方法を用いると、試験全体の $\alpha$ エラーは0.05で抑えることができる。 $\alpha_1=0.04$ にすることによって、全体の対象者

- “Adaptive Signature Design”
- マーカーを試験開始前に特定できない場合
- 全被験者を対象に  $\alpha = 0.04$  で群間比較して有意差なしの時
  - ・ 試験前半で登録された被験者を対象として、予測因子を探索
  - ・ 試験後半で登録された(予測因子の探索とは独立な)被験者を対象に、予測因子に基づくサブグループにおける群間差を有意水準 0.01 で検定

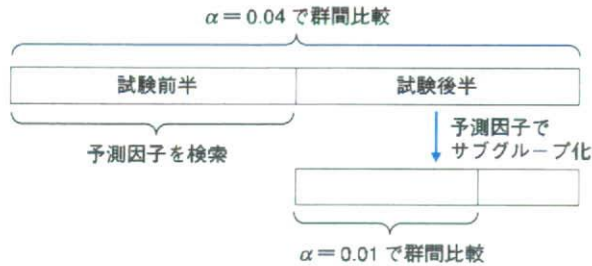


図3 ● Freidlin & Simonの提案

(Freidlin B & Simon R, Clin Cancer Res 2005. より引用)




に対する検出力は少し落ちるが、2nd stageの“感受性の高い”サブグループに対する検定を加えることによって、試験全体の検出力は増加することが期待される。“感受性が高い”(=十分な治療効果が期待できる)ことが本当であれば、そのサブグループでの群間差は全対象者での群間差よりも大きくなることを期待されるため、 $\alpha_2 = 0.01$ でもそれなりの検出力をもつであろうということである。また、 $\alpha_1 = 0.04$ で行う全対象者に対する検定の検出力を $\alpha = 0.05$ で検定する際の検出力と同様となるようサンプルサイズを多少増加させることもできる。その場合には“感受性の高い”サブグループに対する検定を加えた2つの検定を行うことによって、通常的设计よりも有効な薬剤を誤って捨てる確率は少なくなることが期待できる。新規薬剤が多くの対象者に効果がある場合には、通常的设计と同じ検出力をもち、マーカーで規定される少数のサブグループにしか効果がない場合にもそのサブグループでの感受性が高く治療効果が大きければ群間差を検出でき、試験全体としての検出力の増加が

見込めるためである。

National Cancer InstituteのこのSimonらのグループは、microarrayの解析やpharmacogenomicsの解析に積極的に取り組んでおり、そのなかでどのようにして予測因子を探索するかということに多くの経験をもっている。1st stageの対象者のデータを用いてさまざまなマーカーによって規定される感受性の高いサブグループを探索的に探す、というアイデアはそのなかから出てきたものと思われる。感受性の高いサブグループが見つかるかどうか、そのサブグループの割合がどうかといった結果に依存する部分や、1st stageと2nd stageのサンプルサイズ比をどう設定するか、など試験計画のデザイン時に最適化できない部分も多いが、試験開始前に感受性の高いサブグループが十分同定できない場合には有効な方法であると考えられる。

次に筆者が所属する日本臨床腫瘍研究グループ(Japan Clinical Oncology Group; JCOG)で現在計画されている研究デザインを紹介する。これは進行癌に対する化学療法の効果を調べる

表1 ● 薬剤感受性に関する仮説

治療と感受性の関係を3レベルで考える	
一様な効果(交互作用なし) uniform effect ・薬剤の治療効果は感受性によらない	
閾値モデル threshold model ・薬剤の効果はマーカー値がある値を超えると発現する	
直線モデル linear model ・薬剤の効果はマーカー値と直線的に関連する	
ポイント	
治療とマーカーの交互作用として感受性仮説を検討	
想定した感受性仮説が正しければサンプルサイズは減少する	
<b>感受性仮説の証明と臨床的意義は別</b>	
帰無仮説は、治療効果なしと交互作用なしの両方が定義できる	

phase III trialを計画する際に検討されたデザインである。研究の主要な目的は、

a) ある薬剤が対象者全体の全生存期間を延長するような標準治療となり得るか。

であるが、

b) 試験薬投与前に腫瘍組織で調べた薬剤に対する反応(マーカー)によって治療効果が予測できるか。

ということも大きな関心である。もちろん、“腫瘍組織で調べた薬剤に対する反応”自体は単なるマーカーに過ぎず、生存期間延長といった治療効果と必ずしも一致するわけではない。bの仮説、すなわち“マーカーの値によって、治療効果が異なる”ことを本稿では“感受性仮説”とよぶことにする。

例となった実際のJCOG試験では、主要な目的であるa)を優先させるために、通常の2群間比較を数百例の規模で行い、副次的な目的として(多重性の調整は行わず)、b)の仮説を調べることにした。ここでは思考実験として、b)の“感受性仮説”を主要な目的として検証することを考えてみる。治療への“感受性”との関連を期待するマーカーの値が連続的であり、3段階に分けて治

療効果との関係を調べることにする(表1)。3段階に分けると、さまざまな感受性仮説が検証しやすいからである。感受性仮説の例を以下に挙げる。

- ①linear仮説：マーカー値が「低」「中」「高」となるにつれ治療効果が大きくなる。
- ②threshold仮説1：マーカーが「中」以上であれば治療効果が発揮される。
- ③threshold仮説2：マーカー「高」であれば治療効果が発揮される。  
感受性がない場合にも以下の2つの仮説が考えられる。
- ④uniform仮説：マーカー値が「低」「中」「高」に対し、治療効果が一定。
- ⑤全体の帰無仮説：どのサブグループに対して治療効果がない。

通常、臨床試験で行われる検定は、“⑤全体の帰無仮説”に対して“④uniform仮説”を対立仮説として行われる。繰り返しになるが、統計的に言うと感受性仮説は治療とマーカーの生存期間に対する交互作用のことを指す。つまり、治療効果があっても感受性がないことを示すuniform仮説は、統計的な交互作用がないことを示す。