

CRCが配置されるようになった。このことは日本の臨床試験の基盤整備の一步としては大きなものであったが、CRCが支援する臨床試験は治験のみと限定している医療機関は多く、治験以外の臨床試験へのCRCの関与割合は15%程度にとどまるという報告もある⁴⁸⁾。

本来、治験もそのほかの臨床試験も、求められる要件、科学性、倫理性、信頼性を確保しなければいけないという点では違いがないはずである。これら3要件を臨床現場で確保するためには、CRCによる関与が効果的であり、治験に限らず臨床試験全般へのCRCの関与が早急に望まれている。

(5) 臨床試験と臨床研究の違い

臨床研究は、臨床研究に関する倫理指針において下記のように定義されている。

「医療における疾病の予防方法、診断方法および治療方法の改善、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って、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」⁴⁹⁾

人に対して介入を行う臨床試験のほか、介入することなく観察のみを行う研究（観察研究）やヒト由来の材料（組織や細胞などの検体）、データのみを用いた研究も臨床研究であり、調査研究などの看護研究も臨床研究である。しかしながら、これらがすべて臨床研究に関する倫理指針の適用となるわけではなく、研究の内容により「疫学研究に関する倫理指針」や「ヒトゲノム・遺伝子解析研究に関する倫理指針」となる場合があることに注意が必要である。

医学系研究—臨床研究—臨床試験—治験の関係を図2-40に示す。

(6) 薬が世の中に出るまで—治療法開発のステップ—

ここでは、治験薬と治験以外の臨床試験で用いる試験薬を総称して、便宜的に薬とよぶ。

① 薬の候補物質の探索～非臨床試験

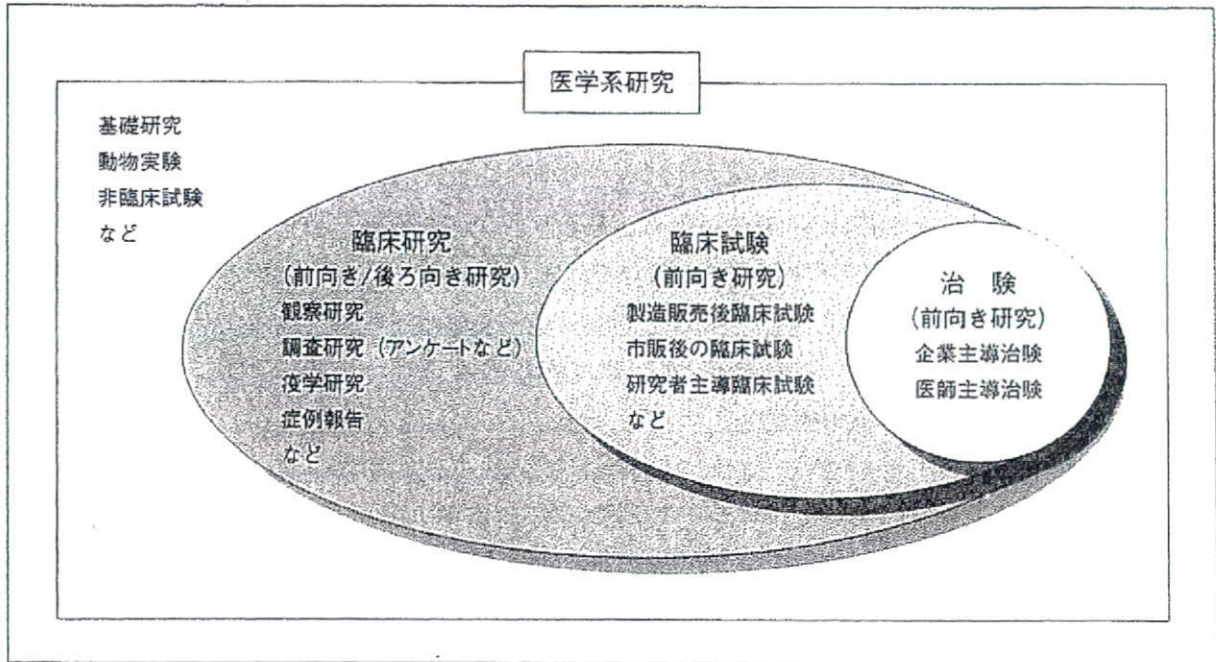
治療法の開発は製薬企業などが薬の候補物質を探すところから始まる。発見した候補物質はすぐに人で試されるべきではなく、まずは実験室において試験管の中や動物を用いて効果や安全性を試す様々な試験が実施される。これらを非臨床試験とよぶ。

② 臨床試験

i) 第I相試験

非臨床試験で得られた薬の候補物質の効果と安全性のデータをもとに、臨床試験を開始する。臨床試験は段階を追って行うが、その第1段階が第I相試験であり、主として安全性を確認する。細胞毒性のある抗がん薬の

図2-40●医学系研究—臨床研究—臨床試験—治験の関係



場合は患者を対象に行うが、そのほかの薬（以下、一般薬）では健康なボランティアを対象に行う。いずれの場合も、少量から始めて少しずつ増量して安全性を確認する試験や、体内でどのような経路で吸収され排泄されるかといった薬物動態を確認する試験（PK試験）、体内での薬の用量と薬理学的作用や臨床的効果などの関連を調べる（薬力学的な検討を行う）試験（PD試験）などを行う。これらを臨床薬理試験という。

抗がん薬の第Ⅰ相試験では、投与量の上限を規定する毒性（用量制限毒性DLT）を定め、その毒性が出現して投与が許容できる最大の投与量（最大耐量MTD）と第Ⅱ相試験で用いる用量として推奨される用量（推奨用量RD）を決定する。対象とする患者のがんの種類は特定しないで行う場合が多い。

また、すでに販売された抗がん薬を複数組み合わせる併用療法の臨床試験の場合、適応症の範囲内で用いる場合は治験という位置づけにはならない。しかしながら、組み合わせるることによる新たな毒性の出現や効果の増強、減弱の可能性があるため、第Ⅰ相試験および次に述べる第Ⅱ相試験はやはり実施しなければならない。

ii) 第Ⅱ相試験

第Ⅱ相試験では、第Ⅰ相試験で得られた情報をもとに、薬の有効性が期待できると予測される患者集団の少数の患者を対象に、治療効果を探査的に調べる。一般薬では第Ⅲ相試験で用いる用法と用量を決定するため、用量反応試験を行う。

抗がん薬の第Ⅱ相試験は、第Ⅰ相試験で決定した用量を用いて特定の

んの種類に対する有効性と安全性を評価する。有効性の評価には指標（エンドポイント）として腫瘍縮小効果を用いることが多い。

なお、患者数の少ないがんは患者数を多く必要とする第Ⅲ相試験の実施が困難なため第Ⅱ相試験のデータをもって新薬承認申請を行うことができるが、患者数の多い乳がん、非小細胞肺癌、胃がん、大腸がんなどを対象として承認申請する際には第Ⅲ相試験の結果が必要となる。しかしながら、相応の理由がある場合には第Ⅲ相試験の結果を待たずとも承認申請を行うことができ、また、これらのがんであっても、承認申請する効能効果の対象となる患者が限定されて数が少ない場合にはこの限りではない。

iii) 第Ⅲ相試験

第Ⅲ相試験では、第Ⅱ相で決定した用法と用量を用いた場合に、臨床的有用性があるか否かを第Ⅱ相試験に比べ、多くの患者を対象に検証的に調べる。標準的治療法がある場合は標準的治療法との比較により、標準的治療法がない場合は、プラシーボという有効成分が含まれず見た目が治験薬（試験薬）と見分けがつかない偽薬を用いた場合との比較により、試験を行う。参加者がどちらの治療法を受けるかについてはランダムに決定する。

抗がん薬の場合、第Ⅲ相試験では延命効果をエンドポイントとして臨床的有用性を確認する。これを真のエンドポイントといい、第Ⅱ相試験で用いる腫瘍縮小効果は代替エンドポイントあるいはサロゲートエンドポイントという。なお、プラシーボは、がんなどの進行性で治療の遅延が命にかかわる疾患での使用は倫理的配慮から通常は許されないが、標準的な治療法への上乗せ効果を調べる場合に使われることがある。プラシーボが用いられることが許容される場合については、先述したヘルシンキ宣言に記述されているので参考にされたい。

③ 新薬承認申請～製造販売許可～薬価収載～販売開始

製薬企業は、すべての治験が終了し、薬事法で定められた新薬承認申請に必要な臨床試験データがそろったら、これらを申請データとしてとりまとめ、薬の候補物質を新薬として製造して販売してよいかどうかを国に対して申請する。申請されたデータが信頼できるもので、薬の候補が社会的に価値のあるものと判断できる場合に、薬の使い方（用法用量）と対象となる病名や状態（効能または効果、適応症）を特定して製造販売許可の承認が与えられる。その後、公定価格として薬価がつけられた後、申請を行った企業は初めて新薬を販売することができ、医療機関は保険診療として新薬を使用することができるようになる。

④ 販売後の安全管理対策、標準的治療法の開発

薬はほかの商品とは異なり、販売するために国の許可が必要だけでなく、販売後にも国民の健康と安全を確保するために各種の安全管理対策

が国策として講じられている（安全管理対策についてはコラム「医薬品の販売後の安全管理対策」参照）。

なお、製薬企業は自社製品の有用性の確認のために製造販売後も臨床試験を実施するが、医師をはじめとした研究者はその領域でのよりよい治療法、標準的治療法を開発するべく、臨床試験を自ら企画して行う。これを研究者主導臨床試験という。実際のところ、標準的治療法は研究者主導臨床試験の結果から生まれることが多いにもかかわらず、研究者主導臨床試験は企業ほどの潤沢な予算なしに運営していることがほとんどである。そのため、CRCなどの雇用費を捻出することは難しく、先述したようにCRCの関与が少ないのが現状である。

b) 臨床試験コーディネーター（CRC）の役割

(1) CRCの役割と必要な知識

CRCは1997（平成9）年に施行されたGCPにより「治験協力者」として定義づけられ、医療機関への配置が進んできた。GCPの治験協力者の定義は、

「実施医療機関において、治験責任医師又は治験分担医師の指導の下にこれらの者の治験に係る業務に協力する薬剤師、看護師その他の医療関係者をいう」⁵⁰⁾

コラム

● 医薬品の販売後の安全管理対策

医薬品は服用あるいは使用した場合にからだに直接影響を及ぼすものであり、正しく使われない場合には健康被害を起こす可能性がある。そこで医薬品は、食品や電化製品などの一般の商品とは異なり、販売するために国の許可が必要だけでなく、公定価格として「薬価」がつけられ、国民は皆保険制度の下、いずれの医療機関でも同じ治療であれば同じ価格で医療を受けることができる仕組みとなっている。また、医薬品の販売後にも国民の健康と安全を確保するために、各種の安全管理対策が国策として講じられている。

製薬企業は、新薬が承認された後一定期間（4～10年）、新薬の有効性と安全性を再確認するために、新薬を使用した場合の成績などに関する調査「製造販売後調査」を行わなければならない。これを「再審査制度」という。製造販売後調査については「医薬品の製造販売後の調査及び試験の実施の基準（GPSP）」が法令の位置づけで定められており、使用成績調査の実施や「製造販売後臨床試験」の実施について規定している。そのほか、5年ごとに有効性と安全性を見なおす「再評価制度」がある。

また、医療機関の医療関係者には、健康被害など（有害事象（副作用）、医療機器の不具合など）の情報を厚生労働大臣へ直接報告する義務が薬事法により定められている。これを「医薬品・医療機器等安全性情報報告制度」という。これは、医薬品が販売され広く使われるようになった後、思わぬ副作用が起こっていないかどうかや、予測以上に副作用の発現割合が高くなっていないかどうかを一元管理して把握し、必要時速やかに国として対応が行えるようにと講じられた仕組みである。

医薬品の副作用による健康被害については「医薬品副作用被害救済制度」が適用となる場合があるが、抗がん薬のほとんどは、副作用の発現がまれではないため、同制度の対象外となっている。

というものであるが、コーディネーターというからには医師の指導を受けて業務を行うだけでは役割を果たすことができないことは明らかである。

CRCの役割は、医療機関での臨床試験の実施にあたり、臨床試験参加者の福利や人権、プライバシーを保護するとともに、臨床試験開始前から実施中、終了後に至るまで、医師、検査部や医事課、看護部や薬剤部などの関連部署、そして、治験の場合には治験を依頼する製薬企業との間の調整を行って臨床試験が円滑に進行できるようにし、その結果として倫理性、科学性の確保された信頼性の高いデータを収集することである。

参考までにいうと、米国のCRCの役割は、まさにプロジェクトマネージャーである。米国では研究者自らが試験を企画する場合、同じチームのCRCが研究に必要な予算の獲得手続きから予算管理までをも行っているという。日本のCRCは研究環境の違いもあり、まだそこまでの業務を求められることはないが、今後、日本の研究環境が米国型に近くなっていけば、プロジェクトマネージャーのスキルはCRCに必須となってくるかもしれない。

(2) CRC業務

CRC業務を表2-15に示す。詳細についてはCRC業務について書かれた成書を参考にされたい(参考資料参照)。

c) 臨床試験参加患者を受け持った場合に看護師として注意するポイント

以下に臨床試験への参加を検討している患者、参加中の患者を受け持った場合に看護師として注意すべきポイントを述べる。

(1) 患者が抱える不安の見極め

臨床試験への参加を検討する機会にある患者は、多くの場合、新たな治療法の選択肢の一つとして臨床試験を紹介される。がん患者でいえば、がんの診断、再発や転移、治療が効かなくなった、といった悪い知らせを医師から伝えられた後に、臨床試験の説明を受けることが多い。臨床試験への参加を検討している患者に不安を抱えている人が多いのは当然といえる。よって看護師は、臨床試験の説明を受けた患者に不安があるときには、その不安が何に基づくものかを見極める必要がある。臨床試験に対して不安や疑問がありそうな場合には、担当医やCRCからさらに詳しく説明を受けるように勧めるべきである。

また、いったん臨床試験に参加した後も、時に患者は再び不安になり、正しく薬を飲まなくなったり、参加の取りやめを希望することがある。そのような場合、臨床試験そのものへの不安が理由の場合もあれば、臨床試験による何らかの影響でからだの状態が悪くなっていることが理由となっている場合もある。いずれの場合も看護師は担当医、CRCに速やかに連絡しなければいけない。もし、臨床試験による悪影響でない場合で、かつ、

表2-15 ●CRC 業務

<p>臨床試験 開始前</p>	<p>1. 試験責任医師（治験の場合は治験責任医師）との調整 ①施設内業務分担 ②臨床試験実施のために必要な文書（必須文書）の作成 ③試験実施までのスケジュールの確認 2. 治験依頼者との調整（治験の場合） ①ヒアリングによる疑義事項の確認 ②保険外併用療養費（選定療養）の範囲の確認 ③治験協力費の取り扱いの確認 ④補償内容および説明資料の確認 3. 臨床試験実施計画書のレビュー ①治療および検査等が、自施設にて実施可能かどうか ②参加者に過剰な負担が課されていないかどうか ③記載内容に矛盾はないか 4. 説明同意文書のレビュー ①わかりやすく書かれているかどうか ・ 中学卒業程度の学力で理解できる内容か ・ 専門用語の使用が必要最小限で平易な言葉で書かれているか ・ 読みやすい体裁で書かれているか ②人格や心情を傷つけるような言葉が含まれていないか ③臨床試験に参加しない場合の治療法が自施設で行う治療と合致しているか 5. Case Report Form (CRF) のレビュー ①書きにくいのか ・ 原資料をもとにどのデータをどのように書くべきか明確にわかるか ・ データの数と記入欄の数が合っているか 6. 倫理審査委員会（IRB）申請手続き・審査意見への対応支援 ①審査に必要な書類の作成 ②審査意見へ対応が必要となった場合の支援 ・ 審査意見への回答書案の作成 ・ プロトコル変更が必要な場合の調整 7. 関連各部署との調整 ①診療科 ・ 臨床試験実施計画書等、臨床試験実施に必要な文書の整備 ②検査部 ・ 自施設外で行う検査への協力依頼</p>	<p>○臨床試験専用の器具や検査伝票を用いる場合 ○検査会社への検体引き渡し ③薬剤部 ・ 治験薬管理・調整の依頼（治験の場合） ・ 治験薬処方の方のオーダー設定依頼（治験の場合） ・ 服薬指導の依頼 ・ 残薬の回収および管理の依頼 ④看護部 ・ 有害事象観察・評価への協力依頼 ・ 併用禁止薬の注意喚起 ⑤医事課 ・ 保険外併用療養費（選定療養）の確認（治験の場合） ・ 治験協力費の取り扱いの確認（治験の場合） 8. 説明会の実施（対象：診療科、看護部、薬剤部ほか） ①臨床試験の概略説明 ②協力依頼 ③業務分担の確認</p>
	<p>臨床試験 実施中</p>	<p>1. 臨床試験実施の支援 ①インフォームドコンセントの支援 ・ 医師による説明の補助 ・ 患者からの質問への対応 ②患者登録手続き ③診察および検査等、来院スケジュールの管理・調整 ④問診による有害事象の観察・評価 ⑤患者からの問い合わせおよび相談への対応 ⑥重篤な有害事象発生時の報告書作成 ⑦CRF の記載 ・ 診療録からの転記 ⑧施設訪問モニタリングおよび監査への対応 ⑨臨床試験実施計画書および説明同意文書改訂時の IRB 審査手続き ⑩臨床試験の進捗管理 ⑪必須保管文書の管理 ⑫補償が必要となった場合の対応</p>
	<p>臨床試験 終了後</p>	<p>1. IRB への臨床試験終了報告書提出 2. 必須保管文書の管理 3. 規制当局による GCP (Good Clinical Practice) 実地調査への対応（治験の場合）</p>

患者の臨床試験に対する不理解から来る不安のみが理由の場合には、看護師も患者の不安感を取り除くように接するべきである。そのようなときには、たとえば「臨床試験は綿密に立てられた計画に従い第三者的立場のIRBにより審査と監視を受けながら行うので、安全性の確認は日常臨床以上に詳しく慎重に行われます」と説明するのがよいかもしれない。

なお、患者は特に理由がなくても臨床試験への参加を取りやめる権利が保証されており、取りやめたことにより不利益を被るようなことがあってはならない。しかしながら、患者がプロトコルに規定されたとおりに治療を受けてこそ、治療の評価を正しく行うことができ、より質の高い臨床試験結果が得られるのである。

(2) 臨床試験の治療が日常臨床と異なる点の把握

臨床試験に参加した場合の治療は、たとえ日常臨床でも使える薬や一般に行われる手術手技であったとしても、プロトコルに従って実施しなければいけない。薬の量を変更する場合も変更規準が設けられており、医師判断とはいえ、規準に反して薬の量を加減することはできない。検査スケジュールも同様に決められた時期と方法を守らなければいけない。しかしながら、臨床試験は日常臨床ときっちり切り分けて行えるものではなく、だからこそ、実験的要素がどこにあるかの把握が重要といえる。

よって、看護師は、臨床試験に参加する患者を担当した際には、実験的要素がどこにあるのかを把握して、プロトコルを守りながらケアを提供する必要がある。もちろん、医学的な判断によりプロトコルどおりの治療や検査を行うことが適切でない場合は、患者の安全確保と利益が優先され、この限りではないことも理解しておかなければいけない。

看護師が臨床試験を理解し臨床試験チームに加わることは、臨床試験チームの戦力アップにつながるだけでなく、患者に対し安心感を与えるという点でも意義深い。看護師のサポートにより、患者が安心して臨床試験に参加することができれば、不必要な参加の取りやめが減少し、臨床試験の質の向上にもつながる。ひいては、より早く標準的治療法を世のなかに送り出すことになるといえる。目の前の患者、将来の患者のためにも、臨床試験を正しく行うことは重要である。

〈引用・参考文献〉

- 1) Fletcher SW, et al : Report of the international workshop on screening for the breast cancer. J Natl Cancer Inst. 85 : 644-656, 1993.
- 2) Kerlikowske K, et al : Efficacy of screening mammography ; a meta-analysis. JAMA. 273 : 149-154, 1995.
- 3) マンモグラフィガイドライン委員会編 : マンモグラフィガイドライン, 第2版,

2 第Ⅲ相試験

1. 第Ⅲ相試験の意味

第Ⅲ相試験の結果が得られた後で行うべき判断は通常、①現在の標準治療を引き続き標準治療とするか、②新治療を新たな標準治療とするかのどちらかであり、これ以外の判断を得ることは目的としない。がん領域で行われる第Ⅲ相試験は①と②のどちらの判断を下すべきかを調べる目的で実施されるという意味において検証的試験 (confirmatory trial) であるといえる。試験の目的は現在の標準治療との直接比較にあるため、この第Ⅲ相試験の対照群 (control arm) には現在の標準治療を用いるのが原則であり、標準治療より劣る可能性が高いと考えられるものを対照群とすることは倫理的な観点から非常に大きな問題がある。また、第Ⅲ相試験で得られた結果を実際の臨床現場に適用する際には、専門病院のみに限らず、日常がん診療を行っている一般の市中病院にも適用されることが好ましい。そのため、一般に第Ⅲ相試験を実施するに当たっては可能な限り、これら一般の市中病院を含む多施設共同試験 (multi-center trial) として実施すべきである。

2. エンドポイント

試験の主たる結論を下す目的で測定されるエンドポイントをプライマリ・エンドポイント (primary endpoint) と呼ぶ。探索的な目的で実施される第Ⅰ相および第Ⅱ相試験では、評価すべき有効性 (真のエンドポイント: true endpoint) を間接的に反映するであろう簡便な代替エンドポイント (surrogate endpoint) をプライマリ・エンドポイントとして用いることが一般的であるのとは異なり、検証的な目的で実施される第Ⅲ相試験では、真のエンドポイントを直接的に評価できるものを

プライマリ・エンドポイントとするべきである。多くのがん領域における真のエンドポイントは延命と症状コントロールであるが、現在のところ、後者の症状コントロールを高い信頼性・再現性ととも測定する方法は確立されていない。がん領域では治療による延命効果を評価するために全生存期間 (overall survival) をプライマリ・エンドポイントとすることが一般的である。これは一般的にランダム化時点を起点として他死因を (打ち切りとして) 除外することなく全ての死因による死亡 (全死亡: any-cause death) をイベントとして定義し、このイベントが発生するまでの時間を測定したものである。

ただし、早期乳がん患者など特に予後のよい集団を対象として試験を実施する場合、全生存期間をプライマリ・エンドポイントとしてしまうと、試験期間内に十分なイベント数を観察するために現実的でないほどの登録患者数を必要としてしまう。このような特殊な領域に関しては、次善の策として全死亡に加えて再発もイベントとする無再発生存期間 (relapse-free survival: RFS)、あるいは全死亡に加えて増悪もイベントとする無増悪生存期間 (progression-free survival: PFS) などをプライマリ・エンドポイントとして用いることも当該領域におけるコンセンサスとなっている。ただし、再発や増悪などをイベントとする場合、それらが確認されるために必要な検査や来院のスケジュール・間隔が粗であるほど結果として観察される無再発生存期間や無増悪生存期間は長くなるというバイアスが生じる。したがって、これらをエンドポイントとして設定する前に、この種のバイアスの入り方自体に結果に影響を与えるほどの実質的な群間差が存在するの否か、もしも容易に存在するのであればそれを可能な限り減らすための方策を十分に検討しなければならない。

3. ランダム化

第Ⅲ相試験はランダム化比較試験 (randomized controlled trial: RCT) として実施される。このランダム化比較試験では、何らかの確率を用いて登録された各対象者を標準治療と新治療に割り付けるランダム化 (randomization, ランダム割付: random allocation) と呼ばれる統計的手順が用いられる。

標準治療を行っている限りはほとんど必ず死亡してしまう患者集団に対し、新治療を実施すればほとんど必ず治癒するということが明らかに期待されるような状況であれば、わざわざランダム化比較試験を実施する必要はない。しかしながら、医療が進歩した現在において実際にそれほど極端に有望な新治療が存在することはまれであり、臨床的に非常に有用といえる新治療であってもその効果の大きさ自体はほどほどであることが多い。このような実際の状況においてランダム化を伴わずに群間比較を行うと、たとえ設定したプライマリ・エンドポイントにおいて臨床的には十分な群間差が結果として観察されるに至ったとしても、それが新治療と標準治療の間に治療効果の差が存在することによって生じたものであるのか、あるいは治療以外の患者属性など予後因子に関して群間で系統的な偏りが存在していたことによって生じたものであるのか、を判断することはできない。

ランダム化を行わずとも、たとえば対照群と新治療群とで既知の重要な予後因子に関してそれぞれ同数となるようにデータを集め、さらに適切な統計手法を用いて比較を行えば十分であると考えられるかもしれない。これは非介入研究を実施する際によく用いられるマッチング (matching) と呼ばれる方法である。しかしながら、マッチングを行うことが可能な予後因子は試験で測定された既知の限られたものだけで、たとえそれらの因子についてマッチングを行ったとしても、試験で測定できなかった未知の予後因子に関して群間で系統的な偏りが無いことまでは誰も保証できない。これより、マッチングでは十分な群間比較が行えないことが多くの領域でコンセンサスとなっている。

限られた対象者数で行われる臨床試験では、ランダム化を用いたとしても臨床的な観点から完全に互いに等しい群を作ることは不可能である。しかしながら、ランダム化を用いると、偶然でも起こりうる既知あるいは未知の予後因子における群間での偏りにより、どの程度の大きさの群間差がどの程度の頻度で観察されるのかを統計学的に導くことが可能となる。これに基づくことによって統計解析を行ううえで必須の確率計算が可能となり、これにより初めて実際に算出された群間差に関する p 値や信頼区間を余計な条件を与えずとも正当化できる (言い換えれば、ランダム化していない治療群間の検定で求まる p 値には実は特別な仮定の下で特別な統計的な配慮をしない限り正当性はないのだが詳細は割愛する)。また、ランダム化比較試験において観察される群間差は、全対象者に標準治療を行った場合の治療効果と全対象者に新治療を行った場合の治療効果との差として解釈可能であり、この意味において群間の比較可能性 (comparability) が保証される。

4. ランダム化における既知の予後因子の考慮

がん領域では重要な予後因子が既に分かっていることが多く、また病期など重要な予後因子が与える予後への影響のほうが、第Ⅲ相試験において評価しようとしている有望な新治療に見込まれる効果よりも段違いに大きいことがあらかじめ期待されてしまう状況も珍しくない。これより、得られた結果の信憑性を極端に劣らせないことおよび精度上昇を目的として、これら明らかな予後因子に関して群間バランスがとれやすくなる工夫 (割付調整) を付加したランダム化を一般的に用いる。これを実現する方法として、層別ランダム化法 (stratified randomization) と動的割付法 (dynamic allocation) がある。

前者の層別ランダム化法に分類されるものとして、最も一般的なものはブロックランダム化法 (block randomization) である。この方法では、割付調整したい予後因子 (割付調整因子) ごとに層を

作り、その層ごとに別々にランダム化を行う。割付調整因子が複数存在する場合には、その組み合わせごとに層を作る。さらに、各層においても各群への割付数が等しい大きさ4から10のブロックを順に並べることによって、群間バランスを保ちやすくすることが一般的である。

このブロックランダム化法は事前に全ての割付リストが決まっており、変更されることはないという意味において静的割付法に分類される。一方で、その時点までに既にどのような予後因子を持つ対象がどちらの群により多く割付されたかを表す割付履歴に依存して、新たな割付を行う割付法を動的割付法と呼ぶ。この動的割付法の中で最も一般的なものは最小化法 (minimization method) である。この方法は新たな登録者の持つ割付調整因子に関して、その時点における偏りの程度を数値化したものに基づき、現在生じている偏りを減じる方向に新たな割付がなされやすくすることで逐次的に群間のアンバランスを改善するものである。つまり、ある割付調整因子を持つ対象者が既に一方の群に偏って多く割り付けられていた場合、同じ割付調整因子を持つ新たな登録者に関しては他方の群に割り付けられやすくなる。偏りの程度の尺度として、単純には新たな登録者に等しい予後因子を持つ既登録者の予後因子ごとの群別登録数を求め、それを総和したもの(その時点までに偏って多く登録された群の総和がより大きくなる)の群間差を用いればよい。この際、偏りが減る方向に100%の確率で割付を行う方法(強制割付: compulsory allocation, deterministic allocation)は、将来の割付に対する予測可能性および統計学的な観点から好ましくなく、通常は偏りが減る方向に50%より大きく100%よりは小さい確率(例: 66%, 80%など)で割付する方法(バイアスコイン法: biased coin method)を用いる。この最小化法は、前述したブロックランダム化法と比べて、割付調整因子が複数存在する場合に、その組み合わせごとの群間バランスを保つとは限らない点において劣る一方で、その逐次的な仕組みから登録者数が少ない状況においても群間バランスが相対的に保たれやすい点において優れている。がん領域では、割付調整因子として一般的によく

用いられる施設ごとの登録患者数が少ないことや、既登録者数が比較的少ない状況で実施されやすい中間解析を試験デザインに含むことが一般的であるため、最小化法のほうが特に好まれて用いられる傾向がある。最小化法を用いた場合に検定・推定の非正当性が主張されることがあるものの、必要であれば並び替え検定など正確な統計手法を用いることも可能であり、実際には大きな問題ではない。

マッチングと同様に既知の予後因子に関して割付調整できたとしても、未知の予後因子に関しては必ずしも群間バランスを保つことができない。したがって、過度の割付調整因子を用いることはかえって危険であり、過度に割付調整するほど割付におけるランダム性が失われることで未知の予後因子に関しては逆にアンバランスが生じやすくなってしまいうことに注意が必要である。

5. 比較のタイプ—優越性と非劣性

がん治療においては毒性が避けられないため、「標準治療に比べてよい治療」は大きく2種類存在する。一つは新治療が「毒性も強いが延命効果も高い」というものであり、もう一つは新治療が「延命効果は同程度であるが毒性が軽い」というものである。前者を評価する場合には優越性試験 (superiority trial)、後者を評価する場合には非劣性試験 (non-inferiority trial) がそれぞれ用いられる。

優越性試験では、標準治療よりも毒性が強い新治療はその毒性に見合う分以上にプライマリ・エンドポイントで優れなければならない。優越性試験の目的は、標準治療と新治療がプライマリ・エンドポイントにおいて等しいことを否定できるか否かを検証することであり、これを統計的に有意に否定できた場合に限って新治療の優越性を証明することになる。

一方、非劣性試験では、毒性の軽い新治療は標準治療にプライマリ・エンドポイントで必ずしも優る必要はない。しかしながら、毒性の軽減に見合う分を超えてプライマリ・エンドポイントで劣

ることは許容されないため、「これ以上劣ることが否定できればよい治療」であるといえる差 Δ （非劣性許容下限：non-inferiority margin）を非劣性試験で否定したい差として事前に設定することになる。つまり、非劣性試験の目的は、新治療が標準治療に比べてプライマリ・エンドポイントにおいて非劣性許容下限 Δ を超えて劣ることが否定できるか否かを検証することであり、これを統計的有意に否定できた場合に限り新治療が Δ 以上は劣らないことを証明することになる。

6. ITT (intention-to-treat) 解析

検証を目的とした第Ⅲ相試験で群間比較を行う際には、ITT (intention-to-treat) 解析と呼ばれる考え方にしたがって試験の主たる解析を行うことが一般的である。実際どのような治療をしたかに関わらず（たとえ治療をしなかったとしても）割り付けた治療群として解析するというのが ITT の原則である。統計的には、ランダム化した全患者 (all randomized patients) を解析に含めるのがさらに好ましく、最も厳密には、たとえある対象が手違いで割付とは異なる他方の治療を受けようとも、解析から除外したり他方の群に再割り付けしたりせず、ランダム化の際に割り付けられた群通りに群間比較を行うことになる。

がん領域では、ランダム化した全患者から不適格患者を除いた全適格患者 (all eligible patients) に対して ITT 解析を行ったものを主たる解析とすることも一般的である。適格規準は定義として時に曖昧であるがゆえに、ランダム化後に不適格者を除くことは結果にバイアスを生む可能性がある。一方で、不適格者を含む全ランダム化解析では結果の信憑性が劣ってしまう危険性もがん領域では高いため、全適格患者に対する ITT 解析を主たる解析に用いることも十分に正当化される。もちろん、不適格者が非常に多い場合は、結果の妥当性が問題となるため、その状況に応じた適切な解析対象集団を定義する必要がある。またそのような場合、異なる定義を用いた場合にどの程度結論が変わりうるのかを評価するために適切な感度分

析 (sensitivity analysis) を行うことが勧められる。

7. 仮説検定, α エラー, β エラー, サンプルサイズ設計

仮説検定は帰無仮説を否定することで対立仮説を採択するという一連の統計的手順である。たとえば優越性試験では、一般的に群間差がない（新治療がよくない治療であることに対応）という帰無仮説を否定することにより新治療の標準治療に対する優越性、同様に非劣性試験では、一般的に群間差が Δ 以上であるという帰無仮説を否定することによって新治療が標準治療と比べて Δ 以上は劣らないこと、がそれぞれ証明される。仮説検定では p 値が求められる。 p 値は帰無仮説が正しい確率を表すのではなく、帰無仮説が正しい場合に観察されたデータより等しいか極端なものが得られる確率を表す。事前に設定した有意水準（ここで α とする）を観察されたデータから求めた p 値が下回った場合には、統計的有意と判断して帰無仮説を否定できる一方で、これを p 値が上回り統計的有意と判断できなかった場合には、そのまま否定されなかった帰無仮説が正しいという結論が得られないことに注意が必要である。また、 p 値の大きさは登録患者数や追跡期間に依存するという性質があるため、 p 値が非常に小さいからといって群間差が非常に大きいという解釈はすべきでない。

真には帰無仮説が正しいのにこれを否定して対立仮説を採択してしまう誤りを α エラーと呼ぶ。この α エラーを犯す確率は事前に設定した有意水準に等しくなる。 α エラーの場合に害を被るのは真にはよくない治療を受けることになる患者ひいては国民であるため、新薬治験の場合は規制当局、それ以外の場合は当該研究領域の研究者コミュニティ全体の監視の下で厳しく守られるべきである。 α エラーは消費者リスクと捉えることもできる。一般に第Ⅲ相試験で設定される α エラーの大きさは5%ないしは2.5%である。

一方、真には対立仮説が正しいのに、帰無仮説を否定できないことでこの対立仮説を採択できな

い誤りを β エラーと呼ぶ。また、真には対立仮説が正しい状態で、これを正しく採択できる確率は検出力 (power) と呼ばれる。 β エラーを犯す確率を β とすると検出力は $1-\beta$ となる。一般的に登録患者数が少ない試験で偶然を超える統計的有意な結果を得るためには当然、極端に大きな群間差を観察する必要がある。つまり、臨床的には十分に意味のある差が実際には存在しようとも、登録患者数が少ない試験では統計的有意な結果を得る確率は小さくなってしまう。生存期間をプライマリ・エンドポイントとする第Ⅲ相試験では、厳密にはイベント数によって検出力の大きさが定まる。つまり、登録患者数を多くするだけでなく、各患者に予定される追跡期間を長くした場合にも検出力は一般に大きくなる。検出したい群間差に対して 80% の検出力を保つために必要となるイベント数を表 1 に示す。表にはハザード比ごとに 2 群の生存曲線を比較する際に必要となるイベント数を示したが、サンプルサイズ設計の際によく用いられる指数分布の仮定の下で、このハザード比は両群の生存期間中央値の比に対応する。試験計画段階で適切なサンプルサイズ設計法を用いて、当該試験に必要な十分な登録患者数、登録期間、追跡期間などを決めることは非常に重要である。登録患者数が不足し検出力を適切な範囲に保てない場合には、何の結論も出せない結果が出る可能性があることから倫理的な問題も生じうるため、試験実施の可否も含めて試験計画を再検討すべきである。 β エラーの場合には、本当はよい治療を受けられないことでやはり患者ひいては国民が害を被ることは考慮すべきであるが、標準治療が既にあ

る場合には α エラーの場合に比してその害は大きくない。むしろせっかくよい治療を開発した製薬企業、よい治療法を考案した研究者が損をすることになるため、「生産者リスク」とも捉えることができ、研究実施主体が決めることになる。状況にも依存するが、一般に第Ⅲ相試験で設定される β エラーは α よりも多少ゆるくして 10~20%、検出力に換算して 80~90% である。登録患者数、追跡期間、解析の実施時期などはプロトコールに事前に明記しなければならない。

8. 生存時間解析

イベントが発生するまでの時間を測定する場合、試験期間内に全ての患者さんのイベントを観察できるとは限らず、一部の患者さんの正確な結果が不明となってしまふことは避けられない。これは打ち切りと呼ばれ、生存時間を評価するに当たってはこの打ち切りの発生を適切に考慮できる特殊な統計解析手法が必要とされる。観察された生存時間データを要約する 5 年生存割合などの年次割合や生存期間中央値 (median survival time : MST) を求める際には Kaplan-Meier 法と呼ばれる推定法、群間比較を行う際にはログランク検定 (log-rank test) と呼ばれる検定法がそれぞれ一般的に用いられる。ただし、ログランク検定はある特定の時点での両群の生存割合や生存期間中央値を比較するものではなく、正しくは両群の生存曲線全体を比較するものである。

表 1 2 群の生存曲線を比較する際に必要となるイベント数

ハザード比	必要イベント数 (両群計)	
	片側 $\alpha=0.05$	片側 $\alpha=0.025$
1.25	497	631
1.33	300	380
1.50	151	191
1.67	95	121
2.00	52	66

対応するハザード比 (生存期間中央値の比に相当) を検出力 80% で検出するために必要なイベント数

9. 中間解析

循環器疾患などと比べると、中規模であるがん第Ⅲ相試験であっても登録患者数は数百人以上となるため、もしも試験途中で群間差が偶然を超える明らかなものとなった場合には、試験を中止して速やかに結果を公表しなければならない。そのため、予定された追跡期間よりも早期の試験中止の検討を意図した群間比較を試験途中で実施することが一般的となっている。このような群間比較を中間解析 (interim analysis) と呼ぶ。

何度も検定を行うことによって偶然有意になってしまう確率が増えてしまうことは多重性の問題と呼ぶ。中間解析を行うことは試験期間中に複数回繰り返して群間比較の検定を実施することに対応し、これによって生じる多重性の問題により、真には群間差が存在しない状況でも偶然に誤って有意な結果が得られてしまう危険性がより高まる。そのため、中間解析を実施するに当たってはこの多重性に対する配慮が必須となる。中間解析での多重性への配慮として国際標準となっていることは、①事前に定めた少数回の中間解析を実施すること、②経時的な多重性の調整を行うことである。①に関しては状況にも大きく依存するが、おおむね、登録中に1回、登録終了後に1回の中間解析が行われることが多い。②の中間解析に特化した多重性調整法としてはいくつか提案された方法が存在するが、代表的な方法の一つとして α 消費関数 (alpha-spending function) を用いる Lan-DeMets 法がある。中間解析以外の場面で有名な、多重性調整のための簡便法であるボンフェローニ法では、全体の α を検定の回数で分割することで個々の検定に対して用いるべき α を求めるのに対し、この α 消費関数法では群間比較を行うたびに全体の α

から少しずつ分割を行って経時的に α を消費していく。どのように分割して消費していくかは用いる関数 (α 消費関数) によって定まる。最もよく用いられる関数は O'Brien-Fleming タイプの α 消費関数である。この関数を用いると試験初期に行われる検定ほど α を消費しないため、試験初期ほど極端に大きい差が観察されない限り試験を中止しない傾向がある。 α 消費関数法の利点として、事前にプロトコルで規定されていた中間解析の時期と実際の実施時期が多少前後した場合、あるいは途中で何らかの必要性により中間解析の実施回数が変更された場合にも、適切に多重性を調整できることがあげられる。

当該試験を実施する研究者が未成熟な中間解析の結果を知ってしまうことでその後の試験運営に甚大な影響が与えられてしまう可能性があり、これによって時には試験の意義自体が失われてしまうことになりかねない。したがって、中間解析を実施するに当たっては、研究者とは独立な組織 (独立データモニタリング委員会: independent data monitoring committee: IDMC) を構成し、この組織外には試験の途中結果を開示しないまま、試験の継続可否も含めて審査されることが一般的となっている。

■文献

- 1) 福田治彦, 新美三由紀, 石塚直樹 (訳): 米国 SWOG に学ぶがん臨床試験の実践—臨床医と統計家の協調を目指して, 医学書院, 東京, 2004 [Green S, Benedetti J, Crowley J: Clinical Trials in Oncology, 2nd Ed, CRC Press, 2003]
- 2) 椿 広計, 藤田利治, 佐藤俊哉 (編): これからの臨床試験—医薬品の科学的評価—原理と方法, 朝倉書店, 東京, 1999
- 3) Piantadosi S: Clinical Trials: A Methodological Perspective, 2nd Ed, Wiley, 2005

第十一章

診療と研究の境

— 臨床試験の倫理 —

笹栗 俊之・柴田 智美

「外科医も生理学者もネロ皇帝も、生物を切り刻むことに専心している点はまったく同一である。彼らを区別するものは何であろうか。もしもそれが思想でなかったならば。」

— クロッド・ベルナール (三浦尚栄訳) —

はじめに

医師が人を対象として行う実験的な医学研究を臨床試験という。二十年近くも前のことだが、被験者にとって危険性のみ大きい「人体実験」的な臨床試験が当然のように行われていることを学会などで知って驚いた覚えがある。その後、インフォームド・コンセントや倫理審査が普及したため、今日ではそのような研究は減ってい

る。ただ、一部の研究を除けば法規制がないため、倫理性に欠ける研究が密かに行われたいという保証はない。

臨床試験には、医薬品などの承認申請のため主に企業主導で行われる「治験」と、主に承認済みの医薬品などを使った医師主導の「自主臨床試験」がある⁽¹⁾。先に「一部の研究」と言ったのは治験のことで、治験は薬事法で厳格に規制されているため、倫理的に大きな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、この指針には法的拘束力がないため、これに該当する研究では今日でも問題が後を絶たない。

一九九八年、金沢大学医学部附属病院は卵巣がん患者に無断で化学療法の臨床試験を行った。次の年、患者の人格権が侵害されたとして遺族が提訴し、遺族側の主張を基本的に認めた二審判決が確定している。二〇〇七年には、神戸市立医療センター中央市民病院で、同意書を得ずに乳がん患者四人に抗がん剤の臨床試験が行われていたことが明らかになった。同意書の取得は試験の計画書にも記されていたが、忙しさのため得ていなかったという。

前例があるにもかかわらず、同様の事態が繰り返されている。表面化する事例はおそらく氷山の一角だろう。ここで考えたいのは、その理由と対策である。

本章では、はじめに、医学研究倫理の歴史的背景を振り返ることにより、日本では研究倫理がどのように扱われてきたかを考える(第一節)。次に、日本の現状と問題点を整理し、臨床研究に対する包括的な規制法の必要性を述べる(第二節)。最後に、適切な規制を設けるには、一般診療行為と臨床試験の違いを明確にする必要があるため、診療(プラクティス)と研究(リサーチ)の境界について考察する(第三節)。

一 歴史的背景

1 戦争と医学

医の倫理の起源という、紀元前四世紀の「ヒポクラテスの誓い」⁽²⁾がすぐ思い浮かぶが、これはあくまで治療者としての医師の倫理の起源であって、研究者としての医師の倫理が論じられるようになるのは、はるか後世、医学が科学の一つとして発展を始めた十九世紀になってからである。

ジェンナーによる種痘の人体実験やバスツールによる狂犬病ワクチンの人体実験など、十八世紀から十九世紀にかけて、臨床医学研究に実験的方法が次第に用いられるようになった。しかし「ヒポクラテスの誓い」のような治療者の倫理規範だけでこれに対応することはできなかった。研究は、必ずしも目の前の患者の健康を目的とするものではないからである。しかし、当時の研究では、家族や隣人など研究者の周囲にいる人々が被験者となることが多く、被験者になることの同意は自然な形で得られていたのだと思われる⁽³⁾。

研究に言及した倫理規範のうち最も古いものとして、新しい治療の試みにおける同僚への相談の必要性を述べた英国の医師パーシバルの綱領(一八〇三年)⁽⁴⁾や、非治療的な研究における自発的同意の必要性を述べた米国の医師パーモントの綱領(一八三三年)⁽⁵⁾などが知られている。これらをそれぞれ、倫理審査およびインフォームド・コンセント(IC)の起源と見なせないこともないが、大きく普及することはなかった。

フランスの著名な生理学者ベルナールは、一八六五年に著した『実験医学序説』に、「たとえその結果が科学

にとつてきわめて有益、すなわち人々の健康に役立つことであっても、被験者にとつて害にしかならない実験は決して行つてはならない」と書いており、ヒポクラテスの「無加害の原則」が、研究者としての医師の倫理へと拡大したことを示している。ただし、ベルナールは、「(死刑囚に対する実験は)何らの苦痛を与えず、何らの不都合をも引き起こさない限り、十分許されてよい」、「いかに動物にとつて苦痛であり、また危険であろうと、人間にとつて有益である限り、(動物実験は)あくまで道徳にかなつてゐる」とも述べており、今日では認められない倫理観も持っていたようだ。

そのようななか、一国の倫理指針として作られたものに、一九〇〇年のプロシア帝国宗教・文部・医学省令「すべての大学病院、集合診療所、病院の施設長への指示」がある。これは一八九二年のナイセル事件を機に発令された。この事件は、ブレスラウ大学教授ナイセルが梅毒ワクチンの研究で健康な子供や若春婦を梅毒に感染させたもので、その方法の是非をめぐつて論争が起きていた。この政令では、同意を得ずに診療目的以外の医学的介入を行うことを禁じている。ドイツではその後、ワイマール共和制時代の一九三一年、やはり人体実験スキャンダルを機に先進的な「人体実験に関する指針」が作られたが、残念なことに、ナチズムの台頭により反故と化してしまつた。

二十世紀に入ると臨床研究は急激に巨大化し、しばしば大量の被験者を要求するようになった。そうなるに、研究者と被験者の個人的なつながりは薄れ、多くの被験者は研究者にとつて名も知らぬ患者群となる。また、何よりも大きな変化は、被験者個人の直接的な利益にならない研究が増加したことである。以前は人体実験といえども多くは治療的研究であり、実験の成功は被験者個人の幸福に直結していたが、医学の発展に伴い「最大多数の最大幸福」が求められるようになると、目の前にいる患者の利益より、その背後に控える患者予備群の利益が

優先され始める。

こうして医学研究が大量の被験者を要求していたところへ、第二次世界大戦が勃発した。

戦争という極限状態では、人間の尊厳を守り抜くことは極めて難しく、人権は容易に蹂躪される。医学研究という側面から見ると、第二次世界大戦は、大量かつ容易に被験者を供給する恰好のシステムとして稼働した。また、軍人の利益を追求する「戦場の医学」が、被験者を求めてもいた。ナチス・ドイツでは、優生学や人種衛生学に基づく「劣等民族」排除政策を背景に、「いずれ始末される」人間として、強制収容所の囚人たちが人道に反する人体実験の犠牲となつた。日本では「抗日分子」として捕らえられた人々や戦争捕虜が、細菌戦兵器開発などを目的とする生体実験の犠牲となつた。

一方、戦勝国であつた米国でも、原爆開発計画(マンハッタン計画)に伴い、放射性物質の人体への影響を調べるため、プルトニウムを静脈注射したり、大量の放射線を人体に照射したりする実験を、一般市民を被験者として行つていたことが今では明らかになつてゐる。広島・長崎への原爆投下を、広義の人体実験ととらえることもできよう。現代の戦争は、医学研究にも多大な影響を及ぼすのである。

2 倫理規範の形成

ナチス・ドイツの戦争犯罪を裁いたニュルンベルク国際軍事裁判の中で、戦争裁判でありながら被告のほとんどが医師(被告二三名中、二〇名が医師)という極めて特異な裁判が米国により開かれ、ナチスの医師らによる非人道的行為が裁かれた。検察団の訴追に対して弁護団は、「大きな善(多数の人々の救済)のためには、小さな悪(多少の人命の犠牲)は許される」というラスコーリニコフ的な論理をも用いつつ被告らを弁護したが、判

決は、七名を絞首刑、五名を終身刑、四名を禁固刑に処した。

この裁判の最も大きな歴史的意義は、人体実験に関する最初の国際的な倫理規範「ニュルンベルク綱領」⁽¹⁶⁾が、一九四七年の判決に伴って作られたことにある。ナチスの残虐な人体実験を裁くに当たって、人体実験そのものを罪悪とすると、医学研究全体を否定することになってしまう。なぜなら、人を対象とする研究は医学の進歩にとって必須だからである。米国の医学者アイビー（生理・薬理学）とアレクサンダー（精神神経医学）の二人が起草したこの綱領は、医学研究における人体実験の必要性を認めつつ、「否認できる人体実験とは何か」を示した世界初の倫理規範である。

全一〇項目のうち、第一項「被験者の自発的な同意が絶対に必要不可欠である」と第九項「被験者は、実験を中止させる自由を有するべきである」の二つは、今日でいうICの概念を示しており、特に注目される。

「ニュルンベルク綱領」は医学研究倫理の原型⁽¹⁷⁾であり、後に作られた数多くの法令や指針に多大な影響を与えた。一九六四年、世界医師会が採択した「ヘルシンキ宣言」⁽¹⁸⁾は、人体実験を行うに当たって守るべき具体的な手続きを示したもののだが、基本理念は「ニュルンベルク綱領」を踏襲している。ただ、「ニュルンベルク綱領」は、ナチスの医師たちが行ったような非治療的人体実験を対象としており、より頻繁に行われる治療的研究を想定していない。「ヘルシンキ宣言」は治療的研究をも対象に加え、さらに同意の絶対性など、「ニュルンベルク綱領」の問題点を改訂して作られた。初版以来五回にわたる改訂を経て、三二項目からなる最新版（二〇〇〇年エジンバラ改訂版）となり、今日も、人を対象とする研究の世界的な基本原則となっている。さらに、国際医学団体協議会（CIOMS）と世界保健機構（WHO）は、「ニュルンベルク綱領」と「ヘルシンキ宣言」をもとに、特に途上国への適用に際して不足する部分を補い、詳細な指針「人を対象とする生物医学研究の国際倫理指針」⁽¹⁹⁾を

作成している。

一方、時代を遡るが、ニュルンベルクの医師裁判を行った米国は、ナチスの非人道的行為を自国の問題としてとはとらえなかった⁽²⁰⁾。

第二次大戦中、米国では医学も軍事研究の一環としてとらえられ、兵器研究とともに国家プロジェクト的な色彩が濃くなっていたが、戦後もその方向が継続された。規制らしい規制もなく、一方、医学研究の国家予算はうなぎ登りに増加したため、臨床研究は精力的に推し進められ、研究者たちは二十年にわたる「黄金時代」を謳歌する。先に述べたように、マンハッタン計画の一環として放射性物質の人体への影響を調べるプロジェクトに多数の米国民が巻き込まれていたことが後年明らかにされたが、当時は一般の人々は知るよしもなかった。

しかし、米国の研究者にとって「古き良き」時代は、ひとりの人物により終止符を打たれることになる。一九六六年、ハーバード大学医学部教授であったピーチャーは、「ニューイングランド医学雑誌」に「倫理と臨床研究」と題する論文⁽²¹⁾を发表する。この論文でピーチャーは、一九五〇～六〇年代に米国内で行われた二二の人体実験を「非人道的」として例示し、告発した。この中には、某州立学校の知的障害児らを人為的に肝炎に罹患させて行った研究や、高齢の入院患者に無断で癌細胞を静脈注射した実験などが含まれる。

この告発に促されるように、国立衛生研究所（NIH）は一九六六年、公衆衛生局を通じ、連邦政府が出資した研究すべてを包括する倫理指針を发表した。この指針で注目に値するのは、研究計画を審査する委員会（施設内倫理審査委員会：IRB）の設置を、国の方針として初めて研究施設に求めたことである。NIHは研究のICには限界があると考えていた。被験者にとって研究内容の完全な理解は実現不可能と見たのである。倫理審査委員会の設置は「ヘルシンキ宣言」にも当時は盛り込まれておらず、画期的な新手段の提案であった⁽²²⁾。

しかし指針に法的拘束力はなく、以後も「人体実験」スキャンダルが相次いで持ち上がる。ついに医学研究の倫理は連邦議会に持ち込まれ、激しい論争の末、一九七四年、医学研究全般にわたる規制を目指す初の法律「国家研究法」が成立する⁽²⁸⁾。

この法律の成立には、モデルンおよびケネディ両上院議員が果たした役割も大きかったが、当時スキャンダルを巻き起こしたタスキギー梅毒研究⁽²⁹⁾が追い風となった。これは、公衆衛生局がアラバマ州タスキギーで四十年にわたって行ってきた研究で、梅毒に罹患した黒人住民を無治療のまま観察し、梅毒の自然経過を研究するというものであった。被験者が徴兵されると梅毒が治療されてしまうため徴兵されないように手配し、さらに、有効な治療薬ペニシリンが入手可能になっても与えなかった⁽³⁰⁾。

「国家研究法」は、臨床研究を実施する機関にIRBの設置を義務づけるとともに、「生物医学・行動学研究における被験者保護のための国家委員会」を設置した。この委員会により臨床研究の倫理基準が検討され、一九七九年、「研究対象者保護のための倫理原則および指針」通称ベルモント・レポート⁽³¹⁾が発表される。ベルモント・レポートは、研究に関する責任の所在を明らかにするため、研究と診療は明確に区別されるべきことをまず述べる。そして、倫理規範は、わずか三つの原則「人格の尊重」「恩恵」「正義」にまで凝縮され、これらはそれぞれ「ICの確保」「危険性と利益の評価」「被験者の公正な選抜」として臨床研究に適用された。ベルモント・レポートの原則主義は、多項目からなつて煩雑だったそれまでの諸原則の欠点を克服し、必要かつ十分な判断基準を提示するものとして、極めて高い評価を獲得した。

連邦政府諸機関は、ベルモント・レポートを根本原則として臨床研究の諸規則を作成するが、機関ごとに異なるルールを設けては煩雑なため、基本的に全ての機関が保健福祉省(DHHS)の連邦行政規則第四五編第

四六部(45CFR⁽³²⁾)の一部を共通する規則(通称コモン・ルール)として採用するようになり、今日に至っている。

3 日本への波及

ICは生命倫理における最大のキーワードだが、その概念は「ニュルンベルク綱領」によつて初めて世界へと発信された。今日、生命倫理学は大変広範な対象を扱う学問に発展しているが、その起源は、戦争中の非人道的人体実験への反省から生まれた研究倫理だと考えることができる。戦後になされた反省の結果、一九六〇―七〇年代、欧米では臨床研究に対する法規制や倫理指針が次第に整備されたのである。

一方、当時の「異議申し立て運動」の高まりの中で、「消費者としての患者」の権利も求められるようになり、これに応じて、医学研究のみならず一般の診療行為へと生命倫理は適用範囲を拡大していった。一九七三年の米国病院協会「患者の権利章典」、一九八一年の世界医師会「患者の権利に関するリスボン宣言」はその表れである。

つまり、世界的な文脈としては、生命倫理は人体実験の倫理(ニュルンベルク綱領)に始まり、次いで臨床研究へと拡大し(ヘルシンキ宣言)、やがて診療一般の倫理へと拡大されたのである。日常診療よりも危険度の高い医学研究の方が先に取り上げられたことは、当然の成り行きといえよう。しかしながら、欧米でも、臨床研究の法規制が実現するまでの道程は決して容易なものではなかった。米国の「国家研究法」は、医学界の猛烈な反発により当初は挫折を余儀なくされたが、長い闘いの末、やつと成立に漕ぎ着いたのである。

ところが、日本の生命倫理にはそのような歴史的な文脈がない。ICの概念は、一九八〇年代半ば頃導入された

が、医療の現場が求めて勝ち取ったわけではなく、欧米で用いられていた方法を受け身で手に入れただけであった。しかも導入されたのは医療行為に関するICに限られ^②、医学研究における被験者の自己決定権を意味するものではなかった。さらに、本来、臨床研究全体を対象とするものであるべき倫理審査委員会は、日本では治験だけのルールとして、これまた文脈なく、しかも倫理的な理由というよりも、「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により導入された。

つまり、生命倫理は、日本へはその歴史の前半部分（研究倫理史）を飛び越して、日常診療の倫理として海外から突如として脈絡なく伝わり、より危険度が高い臨床試験を対象とする研究倫理は、それ以前に解決済みでなければならなかったにもかかわらず、後回しにされ、未だに取り残されたままなのである。このことは、本来ならば国民の手でしっかり決着をつけるべきであった七三一部隊の責任がうやむやにされてしまい、何の反省もなされなかった日本の戦後史^③と無関係ではないだろう。生命倫理という言葉は誰でも知っているが、日本に本当の意味で生命倫理が定着していないのは、こんなところに原因があるのではなかろうか^④。

歴史的な文脈の欠如と一致して、日本の研究倫理は戦後四十年以上にわたってほとんど放置され、その間、数多くの人々が臨床試験による健康被害に遭ってきたと考えられる。日本に倫理指針を作る動きがやると始まったのは一九八〇年代だが、自発的に始めたと言うより、医薬品開発競争上の圧力や諸々の不祥事により取り掛からざるを得なかったと言うべきである。

一九八九年に治験のルール「医薬品の臨床試験の実施に関する基準」（旧GCP）が初めて作られたが、法的拘束力がないうえ内容にも不備が多く、一九九三年にはソリアジン薬害事件^⑤の発生を許してしまう。一九九七年に治験はようやく法規制されたが、それ以外の研究領域は法の枠外に置かれたままである。その後、無断遺伝子

解析研究が明らかになったことなどを契機に臨床研究の倫理指針が作られ始めたが、何か問題が起こって初めて関係各層がその部分だけに継ぎを当てるような作り方をしたため、日本の倫理規制は無計画で系統立っていない。

次節では、このような指針の作り方が生んだ現状と問題を見ていくことにする。

二 日本の現状と課題

1 相次ぐ無断臨床試験

医師が人を対象として行う実験的な医学研究を臨床試験という（図1）。実験的な研究とは、自然に置かれた状態に対して何らかの人為的介入を行い、その後何が起こるかを観察する研究のことである。人に対して介入を行うため、大きな倫理的問題を生じやすい。

日本も、二十年前に比べれば、臨床試験の倫理規制は強められている。しかし今でも問題はなくなり、特に、ICを取得していない臨床試験が今後も絶たない。最近では、神戸市立医療センター中央市民病院で、同意書を得ずに乳がん患者四八人に抗がん剤を使った臨床試験をしていたことが明らかになった^⑥。同意書の取得は試験の計画書にも記されていたが、忙しさのため得ていなかったという。

過去明るみに出た無断臨床試験のうち、以下に挙げる事例は、司法の判断を仰いだ事件としてよく知られている。



図1 臨床試験とは

医学研究のうち、人を対象とする研究を臨床研究といい、その中で、介入を伴う前向きな研究（実験的研究）を臨床試験という。臨床試験のうち、医薬品や医療機器の製造販売の承認を国から得るために行う研究を治験という。

験について事前に全く説明はなされず、もちろん同意も取得していなかった。遺族が、愛知県と医師本人に対して損害賠償請求訴訟を起こし、ICを取得せずに行った治験について争われた唯一の訴訟となった。二〇〇〇年三月、名古屋地方裁判所は、治験薬投与を受けた患者を被験者と認めたとうえで、「ヘルシンキ宣言」や旧GCPなどを引用し、被験者保護の観点から同意取得義務違反という判断を下した^②。本治験では、同意を取得していなかったばかりでなく、治験プロトコルに大幅に違反していたうえ、治験データの捏造・改竄が行われていた。

(三) 金沢大学附属病院無断臨床試験訴訟^③

一九九七年十二月、金沢大学附属病院で卵巣がん摘出手術を受けた患者は、翌年一月、シスプラチン等による術後化学療法を開始すると担当医から説明され、同意した。しかし、投薬開始後、薬の腎毒性により腎機能が悪化し、別の薬剤に切り替えられた。説明された程度より副作用が強いと感じた患者が他の医師に相談したところ、高用量化学療法の臨床試験の被験者とされていたことがわかった。臨床試験については担当医から何も知らされていなかったため、患者は病院に強い不信を抱き、退院して別の病院で治療を受けたが、一九九八年十二月に死亡した。翌年六月、患者の夫は、患者は承諾のないまま臨床試験の被験者とされ、人格権を侵害されたとして、国を相手取って訴訟を起こした。治験以外の臨床試験についてICの必要性が争われた初の裁判となった。しかし被告側は、承認済みの市販薬のみを用いているため臨床試験には当たらない、したがって治療のICだけでよい、と主張した。二〇〇三年二月、金沢地方裁判所は、保険診療の枠内で行える行為であっても、治療とは異なる目的がある以上、治療についてのICだけでは不十分で、治療以外の目的についての説明が必要であったとして、被告側に対し説明義務違反の判決を下した^④。遺族側の主張は一部認められたが、患者が受けた行為が臨

(一) 東北大学附属病院インスリン・ブドウ糖負荷試験訴訟

一九七〇年、バセドウ病による周期性四肢麻痺で東北大学附属病院に入院中の四十五歳男性に対し、「麻痺時の体内水分の移動」を研究テーマとしていた内科の医局員が、麻痺を誘発するインスリン・ブドウ糖負荷試験を実施したところ、低カリウム血症によると思われる急性心停止を来して患者が死亡した。ICは得ておらず、「医師の人体実験の犠牲となった」として遺族が国を訴えた。一九八七年三月の仙台高等裁判所判決^⑤は、試験後に生じた副作用に対する回復措置に注意義務違反があったとして賠償金の支払いを命じたが、負荷試験の実施は日常の診療の範囲内であり、患者から日常的な診療行為について包括的な同意を与えられていた医師に裁量の誤りはなかったとし、負荷試験についてのICを得ていないことなど倫理上の問題については判断を避けた。

(二) 愛知県がんセンター抗がん剤治療訴訟

一九八八年五月、四十五歳の卵黄嚢腫瘍患者を対象として、愛知県がんセンターでシスプラチン誘導体254Sの第二相試験^⑥が実施され、副作用（重篤な骨髄抑制）のため患者は死亡した。治