医療統計学専門家の治験審査への参加による 審査の質の改善に関する検討

野村守弘^{*1,2}, 桑野寛行², 柳原喜恵², 森山健三²,
秦多恵子³, 毛利公則⁴

近畿大学医学部附属病院臨床試験管理センター^{*1}

近畿大学医学部附属病院薬剤部²

近畿大学薬学部³, 明治薬科大学薬学部⁴

Study on Improvement of Screening Quality by Having a Medical Statistics Expert on IRB

Morihiro Nomura^{*1,2}, Hiroyuki Kuwano², Yoshie Yanagihara², Kenzo Moriyama²,
Taeko Hata³ and Kiminori Mohri⁴

Clinical Trial Management Center, Kinki University Hospital^{*1}

*Department of Pharmacy, Kinki University Hospital*²

*Faculty of Pharmaceutical Science, Kinki University*³

*Faculty of Pharmaceutical Science, Meiji Pharmaceutical University*⁴

[Received November 21, 2007]
[Accepted February 22, 2008]

At Kinki University Hospital, the Institutional Review Board (IRB) was established in 1989 at the time that Good Clinical Practice (GCP) was introduced and since then, it has screened the clinical trials and clinical studies conducted by physicians and the number of persons in the screening committee has been increased. In April, 2005, a medical statistician was assigned to the board and preliminary screening was started for clinical studies, which has reinforced the screening process.

In the present study, we determined how the content of IRB screening has changed in accordance with responses made to recommendations made by the IRB, finding that the number of recommendations by the IRB had increased 1.88-fold for clinical trials and 4.62-fold for clinical studies. The fact that they covered everything from ethical to scientific aspects indicated the highly significant role they had played.

Upon classifying IRB recommendations into three categories: ethical recommendations, scientific recommendations, reliability-related and other recommendations, we found that there had been a sharp increase in scientific recommendations and that there was room for improvement in statistical screening for clinical studies conducted by physicians.

The above findings demonstrate the effectiveness of the participation of a medical statistician in the IRB. We would also like to point out the importance of the person responsible for the preparation of IRB minutes (usually a pharmacist) having a knowledge of the scientific aspects of protocols, statistical aspects in particular.

Key words medical statistician, Institutional Review Board (IRB), preliminary screening, number of recommendations from IRB, clinical studies

緒 言

平成 18 年 4 月 1 日付で厚生労働省令第 72 号¹⁾およびその関連通知^{2,3)}が公示, 発出された。これらは治験審査委員会(IRB: Institutional Review Board)の質およびその機能の向上に関するものであり, 平成 9 年 3 月 27 日に

発せられた「医薬品の臨床試験の実施の基準に関する省令(GCP)」以来, 初めて IRB の質に関する改正が行われた。従来は GCP の指示に従って IRB の審議を進めることのみに主眼が置かれ, 審査すべき項目⁴⁾が満たされていることが IRB 審査の質の評価となっていた。この評価方法が治験実施医療機関の IRB 審査において幅広く浸透してくると, 次第に審査すべき内容や中央 IRB 化

* 大阪府大阪狭山市大野東 377-2; 377-2, Ohnohigashi, Osakasayama-shi, Osaka, 589-8511 Japan

などの IRB 自体の機能**に注目が集まるようになってきた。

近畿大学医学部附属病院(以下、当院と略す)は旧 GCP 施行時から IRB を 9 名の委員で開設し、治験や医師自らが行う臨床研究に対する審査を行ってきた。その間に、IRB 審査委員の数も次第に増加し、2005 年 4 月には医療統計学の専門家を加えて全 14 名の委員を擁する IRB となった。現在の構成は、医学・薬学専門に属する者が 10 名(臨床医師 5 名と基礎医学者 1 名、社会医学者 1 名、薬剤部員 2 名、看護部員 1 名)で、外部委員が 2 名、非専門家(事務員)が 2 名となっている。また、2006 年 6 月には臨床研究の予備審査(IRB を円滑に進めるため、事前に実施する予備的な審査)も実施することによって審査の強化を図ってきた。

審査内容のうち倫理的問題については、委員の専門分野にかかわりなく指摘されてきたが、医学的内容については、審査対象分野の専門家もしくはそれに準ずる者が委員として参加している場合には、内容に対して十分な審査が可能であるが、そうでない場合には難しい状況にある。これに対して、試験のデザインと解析については医療統計学の専門家もしくはその取り扱いに詳しい者が審査に参加することによって、試験自体の「精度と信頼性」が確保できるようになった。当院における医療統計学の専門家としては、コホート研究など疫学・公衆衛生学分野における臨床試験・調査の立案から解析までの経験を十分に積んだ統計家(社会医学者)が審査員としての任務を遂行している。

今回はこの医学統計学の専門家を IRB 委員に、また、臨床試験の統計学的取り扱いに精通した者を予備審査の委員に加えたことで、IRB 審議内容全体がどのように変化したのかを IRB での指摘事項およびその回答から評価したので報告する。

方 法

1. 検討の手順

2003 年 4 月(2003 年度)から 2007 年 3 月(2006 年度)に当院 IRB において審査対象となった治験・市販後(製造販売後)臨床試験(以下、「治験」と略す)および医師主導自主臨床研究(以下、「自主研究」と略す)における質疑応答記録である治験意見書(図 1)およびそれに対する治験意見回答書(図 2)を調査の対象とした。

当院 IRB の審査手順は、「治験」の場合、まず、委員会開催日の 5～6 週間前に臨床試験管理センター配属の薬剤師・看護師・臨床検査技師らによるヒアリングを実施する。ここでは必要とされている審査資料の存在確認

治験 意見書	
平成 年 月 日	
治験審査委員長殿	治験審査委員 _____ 印
[薬品名] 上記の臨床研究に当たり、審査の結果、下記のとおり報告いたします。	
審査内容:	
(ア)	特に意見はない
(イ)	下記に意見を述べる
※(イ)の場合は理由・意見を必ずお書きください。	
意見:	

図 1. 治験意見書フォーム

治験意見回答書	
※太枠内のみ記入すること	
治験コード	
申請者	依頼者
Q. 質問者 _____	
A. 回答者 _____	職名 _____

* 本用紙は A4 版にて作成すること。

* 同意書・同意説明文に関する質問事項の場合、訂正版も必ず添付すること。その際、変更箇所アンダーラインを付すこと。

* 本用紙 1 枚につき 1 つの Q & A とすること。用紙不足の場合は資料を添付すること。

図 2. 治験意見回答書フォーム

と同意説明文書の記載内容に漏れや問題となる表現がないか、同意書が所定様式で準備されているかなど、当院の手順に則っていることの確認を中心に行う。委員会開

** 景山茂，高田寛治，治験審査委員会の現状と課題，第 26 回日本臨床薬理学会年会プログラム・抄録集(シンポジウム 12)

催日の4週間前を申請書などの書類と審査資料の受付期限とし、審査資料を各委員に速やかに配布する。配布後1週間から10日を目処に各委員から意見書を回収し、IRB開催前までに回答書を治験申請者・依頼者から受領してそれに基づいて審査が行われる。

「自主研究」の場合には、IRBと同様に倫理性や科学性すべての範囲を事前に審査する機能を持つ予備審査委員会を開催する。この委員は中堅(准教授・講師)の医師7名と臨床試験管理センターの薬剤師・看護師・臨床検査技師各1名の全10名で構成されている。

まず、当院IRBの審査状況を知るために年度別の審査件数を調べた。当院は総合病院ではあるが、がん治療に力点を置く医療機関であるため、がん試験が占める割合を調べてその影響も考慮した。

当院のIRB審査は事前の治験意見書とそれに対する治験意見回答書に基づいて実施されており、指示・質問内容は文書として保管されているため、件数は明確な数値として算出することが可能である。その内容の比較にあたっては、2005年4月より医療統計学の専門家などが審議に参加したため、2005年3月以前と4月以降を境界に二分して検討した。治験意見回答書の記録に基づいて、2003年度および2004年度に実施されたものを前データとし、2005年度および2006年度に実施されたものを後データとしてまとめた。さらに回答書の内容を「倫理的指摘事項」、「科学的指摘事項」および「信頼性その他に関する指摘事項」の3つに分類し、科学的指摘事項をさらに「医学的指摘事項」と「統計学的指摘事項」に分類して検討を行った。ここで示す「倫理的指摘事項」は被験者の試験への参加や継続の意思に影響すると思われる指摘で、「科学的指摘事項」は医学薬学・統計学の知識を要する指摘であり、「信頼性その他に関する指摘事項」は誤字脱字や信頼性を担保するための指摘(登録手順や症例報告書内容など)のことを指す。

また、審査内容の活発化に伴う議事録作成時間への影響を考慮し、医療統計学の専門家が審議に参加する前後でのIRBごとおよび新規治験ごとの議事録作成時間についても比較検討を行った。

2. 指摘事項の分類基準

IRB審査での指摘事項を「倫理的指摘事項」、「科学的指摘事項」および「信頼性その他に関する指摘事項」に分類したが、その分類基準内容は以下に従った。

「倫理的指摘事項」については、基本的には同意説明文に関する指摘事項を中心とした。その内容としては「中止する場合に関する説明が簡略すぎる」との指摘が最も多く、次いで「副作用と有害事象が混在しているので、明確に区別すべき」「代諾者の範囲が不明確なので限定すること」などが挙げられた。最近ではゲノム関連

検査も増えており、Sample Bankingを行うものについては、厳密な保管手順を示すとともに新たな利用目的を確定した場合に再度IRB審査、同意取得を行うことを必須とするとの指摘もなされていた。自主研究特有のものとしては、「補償がないことを明記すべき」「個人情報はいっさいわからないように配慮しますとあるが、具体策を示すこと」「外部(データセンター)へ送る情報は識別コード化する」「適用内使用であること」などの指摘があった。以上を「倫理的指摘事項」と分類した。

次に「科学的指摘事項」については、さらに「統計学的指摘事項」と「医学的指摘事項」に二分した。統計学的な指摘では「主要評価項目は的確か」、「主要評価項目に見合う症例数設計になっているか」、「有意水準両側5%、検出力80%で非劣性であるが非劣性は片側ではないか」など専門家としての意見が多くみられた。「医学的指摘事項」では、「急性肺障害疑いに対する中止基準をもう少し明確にしてはどうか」や「長期に本治験薬(抗コリン剤)の使用には注意が必要。神経では痴呆、重症筋無力症などの副作用が懸念される。すでに痴呆のあるものは除外すべきである」など安全性にかかわるものが多くみられた。以上を「科学的指摘事項」と分類した。

最後の「信頼性その他に関する指摘事項」については、「効果安全性評価委員会の設置をすべきで、メンバーも明確にすべきである」、「研究依頼者・責任者名・連絡先を明確にすること」、「統計解析アドバイザーの表記がないので設置すること」などであった。データそのものの信頼性にかかわってくることであるため、治験ではほとんどみられず、自主研究において指摘されるものが大半を占めていた。以上を「信頼性その他に関する指摘事項」と分類した。

結 果

1. 集計結果

2003年度から2006年度までの4年間の審査件数をがん領域と非がん領域に分けて、年度ごとのがん領域の占める割合(%)を表1に示す。この算出目的は、当院が「がん治療」に力点を置いた医療機関であって、「治験」や「自主研究」がどの程度その影響を受けているかを数値で示すことにある。また、審査内容についても統計学的取り扱いが他分野に比して進んでいるがん関連の試験が多いほど、統計学的検討が増加して指摘事項への影響が大きくなると考えられるためである。「治験」と「自主研究」のプロトコル件数は163と160でほぼ同数であったのに対して、「がん試験」と「非がん試験」のプロトコル数比率をみると治験では31.3%、68.7%となっており、自主研究では72.5%、27.5%で、まったく逆の比率であった。

表 1. 年度別 IRB 審査件数

分 類	年 度	がん領域	非がん領域	合計件数	比率
治験・製造販売後臨床試験	2003	13	26	39	33.30%
	2004	11	31	42	26.19%
	2005	13	34	47	27.66%
	2006	15	21	36	41.67%
		51	112	163	31.29%
医師主導臨床研究(自主研究)	2003	27	8	35	77.14%
	2004	41	15	56	73.21%
	2005	27	12	39	69.23%
	2006	23	7	30	76.67%
		118	42	160	72.50%

比率：がん領域審査件数の割合

これらの IRB 審査の過程で得られた指摘事項を「倫理的指摘事項」、「科学的指摘事項」および「信頼性その他のに関する指摘事項」に分類し、治験についてのプロトコルあたりの指示・質問件数を図 3 に、自主研究についてのプロトコルあたりの指示・質問件数を図 4 に示す。2003 2004 年度と 2005 2006 年度で、「治験」では「倫理的指摘事項」と「信頼性その他の指摘事項」には差が

みられなかったが、「科学的指摘事項」において 2005 2006 年度は 2003 2004 年度の約 3 倍の指摘事項件数があった。「自主研究」では全項目において指摘事項が増加しており、特に「科学的指摘事項」においては約 10 倍の増加がみられた。全体を通してみると、ともに 2005 2006 年度のほうが IRB による指摘数が増加しており、「治験」では 1.88 倍(図 3「全体」の 2.37/1.26 より算出)、「自主研究」では 4.62 倍(図 4「全体」の 8.68/1.88 より算出)であった。

また、科学的指摘事項についてさらに「医学的指摘事項」と「統計学的指摘事項」に分類し「治験」についての二者の内訳を図 5 に、「自主研究」についての二者の内訳を図 6 に示す。ともに 2003 2004 年度においては「医学的指摘事項」のほうが多かったが、2005 2006 年度は

倫理：倫理的指摘事項

科学：科学的指摘事項

信頼：信頼性およびその他のに関する指摘事項

図 3. 治験におけるプロトコルあたりの指摘・質問件数

倫理：倫理的指摘事項

科学：科学的指摘事項

信頼：信頼性およびその他のに関する指摘事項

図 4. 自主研究におけるプロトコルあたりの指摘・質問件数

図 5. 治験におけるプロトコルあたりの科学的指摘事項件数の内訳

図 6. 自主研究におけるプロトコルあたりの科学的指摘事項件数の内訳

「統計学的指摘事項」のほうが明らかに多くなっていた。

次に、議事録作成時間の年度ごとの変化および新規プロトコルあたりの変化を表2に示す。当院の議事録は「治験」と「自主研究」を区別なく同一のIRBで審査を実施しているため両者をあわせたものとなっており、まず、IRBあたりの議事録作成時間を調べ、そのなかから新規治験についての議事録作成時間も調査した。年度を通じて議事録の書式は変えていないが、前半の2年間のIRBあたりの作成時間に大きな変化はみられず、9.5時間および9.2時間であったのに対して、後半の2年間のIRBあたりの作成時間は18.4時間および20.4時間と大きく増加していた。しかも、後半の2年間には予備審査委員会での議事録作成にも時間を要しており、10.7時間および9.4時間が上乗せされていたことが差を大きくする原因になっている。新規治験プロトコルあたりの議事録作成時間についても類似の傾向がみられ、前半が1.2時間と1.3時間、後半が3.2時間と3.8時間であった。

2. 統計学的事項における問題事例

「治験」においては、2005・2006年度のなかで審査された3つの事例(事例A, 事例B, 事例C)は統計学的には承認困難なものであったが、それぞれ次に記す経過により承認されたものである。

事例Aは、局所感染症に対する外用剤の治験であった。主要評価項目が「臨床効果判定」で副次評価項目に「細菌学的評価判定」と「臨床症状判定」が挙げられていたが、IRBから「客観的評価が可能な細菌学的評価判定こそ主要評価に加えるべきであり、臨床症状判定が主要評価にも副次評価にもあることが理解し難い」という指摘がなされた。これに対して、「当該治験は、学会での評価方法に準拠したものである。主要評価における臨床効果判定は副次評価におけるスコアと検出菌の消失状況を総合的に判定したものである」との回答があった。また、「期待有効率80%、閾値有効率60%と見積もる根拠が乏しい(特に閾値有効率60%が臨床的に評価し得る率かどうか)」との指摘もなされたが、「Orphan drugでもあり臨床的においに期待され得る」という回答によってIRBで承認された。

事例Bは、術後の鎮静・鎮痛を目的とした注射剤の治験であった。必要症例数設計においてPart (pre-study)

とPart (post-study)の2部に分け、暫定的に50例の設定を先に行っていた。Partで社内文献をもとに試験的に10例を見積もり、PartでPartの結果をもとにして症例数を設定し直すプロトコルとなっていた。IRBではこのような方法を取ると都合のよい症例数設定ができる余地が生じ試験全体の精度が落ちるため、Partはpilot studyとし、Partは別プロトコルとして再計画することを推奨したが「医薬品医療機器総合機構(PMDA)からの指示による」とのことで特例的にIRBでの承認となった。

事例Cは、術後血栓予防の注射剤の治験であり、通常治療法を対照としたものである。有効性および安全性を無作為化オープン試験で検証することを目的としながら、対照との統計学的な仮説設定と検証を行わないことに対する理由を求めた。依頼者は「仮説の設定は困難」としながらも、海外ではプラセボ比較試験が実施されている事実に対して、「日本と海外の状況の違い」で処理しようとした。最終的にはデザインの甘さを残しながらも「関連各学会からの承認要望が厚生労働省宛てに出されている事実がある。医薬品総合機構の指導により作成されたプロトコルである」という理由により当院IRBでも承認されることになった。

「自主研究」における指摘内容については、表1の自主研究・非がん領域・2005, 2006にあたる19(=12+7)プロトコルのうち18プロトコルに「目標症例数の設定根拠の明記」の指摘がなされていた。根拠の記載があっても、実施可能な症例数という理由のみの設定であったり、計算しやすい副次評価項目を代理指標として数値設定しているような場合は必ず修正が要求された。前者の場合は参考にできる試験結果がないことが多いが、臨床的見地からの「期待される差」や「期待有効率、閾値有効率」を検討したうえでの症例数の設定が薦められる。後者の場合は真に改善したい指標(FEV₁など)ではあるが、バラツキがあまりにも大きく必要症例数が莫大になるため、安定して得られる代理数値(PEFなど)にて症例設計を行うケースもあった。これについても改善が求められた。また、長期にわたって行われる探索的研究においては中間解析を実施することが薦められ、多重性の配慮もあわせて指摘されていた。逐次デザインでは最大症例数を設定しておき、その数に達する前に数回の中間解析を行って効果が証明されれば、最終段階に至る前に試

表2. 年度別 IRB あたりおよび新規プロトコルあたりの議事録作成所要時間

年度	IRBあたり		予備審査あたり	小計	合計
2003年度	9.5	(1.2)	0	9.5	18.7
2004年度	9.2	(1.3)	0	9.2	(2.5)
2005年度	18.4	(3.2)	10.7	29.1	58.9
2006年度	20.4	(3.8)	9.4	29.8	(7.0)

()は新規プロトコルあたりの議事録作成時間 単価: 時間

験を終了すべきとの指摘があった。

考 察

当院での「治験」と「自主研究」における指摘事項の増加率の違いは、科学的指摘事項、特に統計学的指摘事項の数の違いに起因することがわかった。この違いは統計解析担当者もしくはアドバイザー(以下、解析担当者と略す)が配置され、十分機能している臨床試験がどうかによって左右される面が大きい。「治験」ではほぼ機能しているのに対して、「自主研究」では十分機能しているとはいえない状況がみられる。一般的に「がん領域」の試験には解析担当者が配置されている場合が多く、「非がん領域」では逆に少ない。「自主研究」では「がん領域」の試験が多いにもかかわらず「治験」の場合より指摘事項が増加していることから、解析担当者を増やすことで機能レベルをアップする必要性が示唆された。同時に、医療機関での審査において医療統計学に造詣が深い者を参加させる必要性も示唆された。

指摘内容からみると、2003 2004 年度は特定の項目が脱落している場合や説明表現が好ましくない場合など、同意説明文書における「倫理的項目」に相当する部分に重点が偏っていた。しかし、2005 2006 年度は統計学的症例数の設定根拠を明確にするなど「科学的項目」に指摘が集中していた。医学統計学の専門家が IRB 委員として審査に参加することによって指摘数が「治験」で 1.88 倍、「自主研究」で 4.62 倍に増加しており、内容的にも倫理性から科学性までをカバーしたものへとシフトしていた。このことから、医学統計学の専門家が IRB 委員として審査に参加することの意義はかなり大きいと考えられる。最近では、この医療統計学専門家の指摘に刺激を受けて、統計学非専門家からも「と、1」の注釈もしくは説明がない」や「FAS の定義をもっと明瞭にすべき」などの意見が出されるような状況になりつつあり、今まで注目せずに素通りしてきた指摘点が次第に増加している。

「治験」において問題となった 3 事例は統計学的見地からは素直に受け入れがたいプロトコルデザインではあったが、「絶対的症例数が少ない」や「学会の要望、医薬品総合機構の指示」などの社会的背景によって実施されやすいデザインが優先される場合が存在することが浮き彫りになった。このような問題が明らかになること自体が発展的であり、従来ならば「医薬品医療機器総合機構の指示による」という理由だけで問題提起されることなく承認されていたところである。

IRB 審査によって「治験」や「自主研究」の科学性が重視されるようになって、IRB 事務局側が受けた最も大きな影響は議事録作成時間であった。医療統計学の専門家が IRB に参加する前後で約 3 倍(表 2 の合計時間 58.9/18.7 より算出) 治験新規プロトコルに限れば約 2.8 倍(表 2 より 7.0/2.5 より算出)に増加し、被審査者と審査委員との質疑応答の記録に多大な労力と時間を要することが明らかになった。記録テープを再度聴き起こす単純作業ではなく、内容を理解していないと適格な議事録の作成は不可能である。その意味で、これからの IRB 事務局担当者(当院は薬剤師)には、プロトコルの科学性についての知識と判断力が必須であると思われる。

引用文献

- 1) 医薬品の臨床試験の実施の基準に関する省令の一部を改正する省令。厚生労働省令第 72 号、平成 18 年 3 月 31 日。
- 2) 厚生労働省医薬食品局長通知、「医薬品の臨床試験の実施の基準に関する省令の一部を改正する省令」(治験審査委員会の質及び機能の向上関係)の施行について、薬食発第 0401001 号、平成 18 年 4 月 1 日。
- 3) 厚生労働省医薬食品局審査管理課長通知、「医薬品の臨床試験の実施の基準に関する省令の一部を改正する省令」(治験審査委員会の質及び機能の向上関係)の施行に関する留意事項について、薬食審査発第 0401001 号、平成 18 年 4 月 1 日。
- 4) 医薬品の臨床試験の実施の基準に関する省令、厚生省令第 28 号、平成 9 年 3 月 27 日。