平成２９年１月１２日

臨床研究に関する倫理指針違反についての調査報告書

聖マリアンナ医科大学
臨床試験に係る調査委員会
委員長　熊　井　俊　夫

目次

第１．調査委員会発足までの経緯 1
第２．調査委員会 3
１．調査委員会の構成メンバー 3
２．調査委員会の開催日時 3
第３．調査の結果 4
１．当該臨床試験被験者の、臨床試験同意撤回と臨床試験データ削除要請等に関わる大学の対応に関する検証 4
（１）時系列の確認 4
（２）試験責任医師並びに大学の対応の検証 5
（３）カルテの改ざんについて 6
（４）虚偽の報告（シュレッダー廃棄） 7
（５）組織としての責任 10
２．当該臨床試験の調査結果について 13
（１）試験の構造の問題 13
（２）同意に関する問題 13
（３）エントリー基準に関する問題 16
（４）実施計画書を逸脱してブロナンセリンを投与した原因の究明 19
（５）利益相反について 19
（６）被験者の不利益について 21
３．当該臨床試験を基とした論文の検証結果について 21
（１）論文１の検証 21
（２）論文２の検証 26
第４．精神保健指定医資格を取り消された神経精神科医師が関わった他の臨床試験の検証 35
１．神経精神科関連臨床試験に関する調査 35
（１）調査対象 35
（２）比較試験 35
（３）調査結果 35
２．調査結果に基づく今後の対策 36
３．研究の適正な実施のための組織体制と管理者の責任 37
第５．再発防止策の検討 37
１．臨床試験における再発防止策 38
（１）臨床試験に関する学内教育の徹底 39
（２）臨床試験を適正に遂行するための学内での体制整備 39

（３）利益相反の明確化、職員倫理の徹底及び公益通報窓口の再確認..... 40
（４）大学院生指導教員への心理的圧力に関する調査 40
２．被験者（患者）に対する対応について 41
（１）病院と大学との連携強化..... 41
第６．結び..... 41

第１．調査委員会発足までの経緯

臨床試験「初発エピソード統合失調症患者の認知機能障害に対する第 2 世代抗精神病薬 blonanserin の効果：aripiprazole とのオープン比較試験」（以下、「当該臨床試験という。」）は、本学神経精神科学Ｂ医師により、統合失調症の認知機能障害に対し blonanserin（ブロナンセリン）と aripiprazole（アリピプラゾール）をランダム化して投与する比較試験として申請され、平成 21 年 3 月 6 日に承認されて同月より開始されたものである。

平成 27 年 7 月初旬、同臨床試験に参加していた一人の被験者Ａ氏から、本学神経精神科学の精神保健指定医資格取消処分を知り、「そのような不正を働く医師が実施していた試験に自分のデータが利用されるのは困るので、臨床試験への参加取消しとデータを削除してほしい」との要請があった。

当初、患者さんからの連絡ということで、大学病院医療安全管理室（以下、「医療安全管理室」という。）が窓口となり対応していたが、話の内容が徐々に臨床試験そのものになってきたため、臨床試験の担当部署である教学部大学院・研究推進課（以下、「大学院・研究推進課」という。）に窓口を移管して対応することとした。その後、Ａ氏は診療録も含めたデータの原本開示を求めてきたことから、データがすべて揃っているかＢ医師に確認したところ、すべてあるかどうか不明であるとのことであったため、これまでのやりとりから推測すると、すべての資料がそろっていないと納得されないと判断し、また当初Ａ氏本人はデータの削除を希望されていたことから、医療安全管理室の担当者からデータはすべて廃棄したと伝えている。

しかし、この経緯に納得がいかなかったＡ氏は、当該臨床試験の内容と本研究を基に発表された論文を独自に調査し、Ｂ医師に当該臨床試験に関する情報開示を求めている。

そもそも、この臨床試験は、実施計画書では２つの薬剤を被験者にランダムに投与し比較するというものであったが、実際はランダム化せずに一方の薬剤投与から開始し、そのまま単剤試験となっており、実施計画書から逸脱していたことから、平成 27 年 11 月下旬にＢ医師本人から臨床試験部会に報告が有り、生命倫理委員会で審議した結果、実施計画書違反とのことで、臨床試験を中止するよう言い渡されている。その後本学では、生命倫理委員会の他、「研究活動における不正行為の防止等に関する規程」に則り、研究倫理委員会を開催し審議がなされた。そしてさらに臨床試験・論文等の調査を行うため、学長指名による 5 名からなる内部調査委員会を発足し、これまでに 10 回ほど開催し調査が行われた。

また、平成 27 年 12 月下旬に本件について「よみドクター」というネット記事にコラムが掲載され、さらには、Ａ氏本人から厚生労働省に告発が有り、参議院厚生労働委員会において国会議員から厚生労働大臣に質問がなされている。

その後、厚生労働省から本学に対し、調査委員会には外部委員を過半数入れ、その中には弁護士、患者代表となるような一般の方を参加させるよう行政指導があった。また、調査

委員会の事務局にはこれまでの大学院・研究推進課はいわば当事者となるので他の部署に担当させること、さらには、精神保健指定医資格を取り消された医師が関わった他の臨床試験も調査するよう指導があった。

以上のような経緯から、弁護士を含む外部委員 5 名を加えた合計 8 名からなる調査委員会を発足し、平成 28 年 4 月 7 日に第 1 回目の委員会を開催した。

第２．調査委員会

１．調査委員会の構成メンバー

調査委員会の委員等は以下のとおりである。

委員長　　熊　井　俊　夫　　聖マリアンナ医科大学大学院教授

内部委員　聖マリアンナ医科大学教授

同　　聖マリアンナ医科大学教授

外部委員　弁護士

同　　精神科医師

同　　精神科医師

同　　一般代表

同　　一般代表

事務局　　総務部部長
同　　総務部参事兼総務課長
同　　総務課係員
同　　学長室室長補佐

２．調査委員会の開催日時

第１回調査委員会　　平成 28 年 4 月　7 日（木）　午後 4 時～6 時 10 分
第２回調査委員会　　平成 28 年 4 月 25 日（月）　午後 4 時～5 時 30 分
第３回調査委員会　　平成 28 年 5 月 16 日（月）　午後 4 時～5 時 30 分
第４回調査委員会　　平成 28 年 6 月　7 日（火）　午後 4 時～5 時 20 分
第５回調査委員会　　平成 28 年 6 月 20 日（月）　午後 4 時～6 時 30 分
第６回調査委員会　　平成 28 年 7 月　5 日（火）　午後 4 時～5 時 00 分
第７回調査委員会　　平成 28 年 7 月 28 日（木）　午後 4 時～5 時 30 分
第８回調査委員会　　平成 28 年 9 月 12 日（月）　午後 4 時～5 時 10 分
第９回調査委員会　　平成 28 年 12 月 26 日（月）　午前 10 時 30 分～11 時 20 分

第３．調査の結果

１．当該臨床試験被験者の、臨床試験同意撤回と臨床試験データ削除要請等に関わる大学の対応に関する検証

当該臨床試験では、同試験の被験者Ａ氏から臨床試験の同意撤回と臨床試験データの削除要請が行われたことに端を発して、臨床試験の適格性に疑義があることが明らかとなり、本委員会で検証を行った。一方、当該臨床試験ではＡ氏からの同意撤回とデータ削除要請に対する試験責任医師並びに大学の対応の適格性が問われる事態も生じており、委員会では臨床試験参加者の人権擁護、臨床試験遂行における大学の体制の観点から検証を行う必要があると考え、Ａ氏と大学関係者の対応の時系列を確認するとともに、関係者へのヒアリングを行って検証を行った。なお検証は、本事案発生当時の臨床研究に関する倫理指針に基づいて行い、今後の再発防止策については現行の倫理指針に基づいて検討を行った。本事案発生当時に大学院生であったG医師に対してもヒアリングを実施すべく連絡したが、体調不良により出向くことができないため、書面により質問事項をお送りいただきたいとの要望があったことから、書面にて質問し、書面にて回答を得ている。

（１）時系列の確認

平成 27 年 7 月 1 日当該臨床試験被験者Ａ氏から入電があり、医療安全管理室が対応した。その内容は、精神保健指定医取消し問題を知り、神経精神科のすべてに対して不信を抱いた。平成 22 年に臨床試験に参加したが、この様な医師の研究に協力はできない。自分のデータを消去してほしい。また、今まで受診してきた内容の全てを資料として自分に提出していただきたいとの内容であった。対応した医療安全管理室の担当者Ｃ氏は、データの削除は診療部長（神経精神科学教授）に依頼するが、診療内容の提出はカルテ開示となるので診療部長の判断を要するため、改めて確認して回答する旨伝え、以上の相談内容と初期対応について、所定の報告書により 7 月 6 日にＢ医師と神経精神科学教授に文書通知した。

これ以降Ａ氏からの電話には医療安全管理室が対応し、Ｂ医師に取り次ぐ形で対応が進んだ。7 月 13 日と 7 月 15 日の 2 回にわたりＢ医師はカルテ開示を想定して電子カルテの記載に修正を加え、8 月 7 日にカルテ開示(閲覧)を行った。その後、臨床試験に関する細部の質問が多々あり、大学院・研究推進課しか答えられない内容が多いと判断し、医療安全管理室を管轄する副院長と相談の上、窓口を医療安全管理室から大学院･研究推進課に移行することを決定。8 月 21 日大学院・研究推進課、医療安全管理室が神経精神科医局内でＢ医師と今後の対応について協議。その結果、Ａ氏の意向に合わせるべく「資料は全て廃棄した」と伝えることとし、8 月 24 日医療安全管理室から、臨床試験のデータの原本については、全てシュレッダーにて廃棄した旨、虚偽の報告が行われた。

（２）試験責任医師並びに大学の対応の検証

１）初期対応窓口の妥当性

当時適応されるべき臨床研究に関する倫理指針（平成 20 年全部改正、以下、「旧倫理指針という。」）では、「組織の代表者等は、苦情・問い合わせ等に適切かつ迅速に対応するため、苦情・問い合わせ等を受け付けるための窓口の設置や苦情・問い合わせ等の対応の手順を定めるなど被験者等からの苦情・問い合わせ等に対応するために必要な体制の整備に努めなければならない。」と定められている。

臨床試験の同意説明文書には、担当医師名、日中、夜間の連絡先、PHS 番号等が示されており、常時被験者からの質問や情報開示の求めに対応できる体制が整えられている。患者相談窓口に来た問い合わせについても、本件で初期対応されたごとく医療安全管理室が患者相談案件とその初期対応の状況を関係者に文書で通知し、適宜対応を求めるなどしており、この点についても院内体制として誤りはないものと思われた。本調査委員会としても、病院受診中の患者、家族の様々な申し出に対して、できる限り窓口を一本化して初期対応し、相談内容に応じて切り分けて対応することは望ましいことと考え、当該臨床試験被験者の初期対応窓口が医療安全管理室であったことについて問題はなかったと考えられる。

データの削除要請に対し、データの削除は診療部長（神経精神科学教授）に依頼するが、診療内容の提出はカルテ開示となるので診療部長の判断を要するため、改めて確認して回答する旨伝えた点は適切と考えられる。

２）情報開示への一元的対応と個人情報開示義務

旧倫理指針では、「研究代表者は、当該臨床研究に係る情報の開示等の求めに対しては、あらかじめ一元的に対応できるような手続等を定めるなど被験者及び代理人の負担をできるだけ軽減するような措置を講ずるよう努めなければならない。」と定められている。また、「研究代表者は、被験者又は代理人から、当該被験者が識別される保有する個人情報の開示を求められたときは、原則として被験者に対し、遅滞なく、書面の交付又は開示の求めを行った者が同意した方法により当該保有する個人情報を開示しなければならない。」と定めている。

当該臨床試験の研究責任者であるＢ医師および診療部長は、文書をもってＡ氏から入電があった旨医療安全管理室から報告を受けており、この時点で同氏からの要求を認識したと思われ、同指針が定める通り臨床試験で得たＡ氏にかかわる資料を遅滞なく開示する必要がある。通常、臨床試験責任医師が臨床試験のために保管したデータの内、被験者分を抜き出して開示する作業であることから、この開示は特段困難な作業とは思われない。Ｂ医師は 7 月 13 日と 7 月 15 日の 2 回にわたりカルテ開示を想定して電子カルテの記載に修正を加え、7 月 23 日資料に×印を入れ、今後データは使用しない旨の文

書を作成している。その後 8 月 7 日にカルテ開示（閲覧）を行ったが、A氏は、原本を開示せず削除したことに不信感を持ち、エクセルデータを復元させて、そのデータを開示するなどの対応を求めた。旧倫理指針に照らせば、臨床試験で得た原本の開示は行うべきであり、一部の資料について「廃棄した」「廃棄する」との説明で終始したことは適切さを欠き、A氏に無用な時間と労力を与えたことは否めない。なおカルテ開示前のカルテ改ざん行為についての検証は次項にて示す。

今日臨床試験においてはデータセンターが設置され、すべてのデータを一元管理するシステムが構築されているため、今後同様の事案が発生したときは、データセンターに保有する全ての個人情報を直ちに開示することは容易となっている。しかし、本事案発生当時データセンターは存在せず、臨床試験の資料や個人情報の管理は各研究者に任される状況にあった。そのような背景はあるにせよ、本事案においては、一部原本の不備により当時の倫理指針が定める保有する個人情報の遅滞なき開示が行われなかったと判断せざるを得ない。

（3）カルテの改ざんについて

1）経緯

平成 27 年 7 月 1 日　A氏からの入電後、7 月 6 日にB医師と神経精神科教授に文書で相談内容の通知が行われた。B医師はカルテ開示を想定し、同年 7 月 13 日と 7 月 15 日の 2 度にわたり電子カルテの記載を書き換えている。

7 月 13 日：

①平成 22 年 10 月 27 日（水）の母親に対する説明で「研究について少し話した。」との記載を削除。

②平成 22 年 10 月 29 日（金）に記載された、「ブロナンセリンの初回エピソード臨床研究について説明し、本人から文書による同意を得た。」の記載を「抗精神病薬の初回エピソード臨床研究について説明し、文書による同意を得た。」に修正。

7 月 15 日：

①平成 23 年 5 月 12 日（木）に記載された、「PANSS,QOL　DIEPSS,WELL-BEING,採血実施した。」を「PANSS,QOL DIEPSS,WELL-BEING 実施することを説明し同意を得た。採血も実施した。」に修正。

②平成 23 年 11 月 10 日（木）に記載された、「PANSS,　WELL-BEING、SQLSJ,BACS 採血実施する。」を「PANSS,　WELL-BEING、SQLSJ,BACS 採血実施することに同意をいただいた。」に修正。

8月7日、これらの変更が加えられた電子カルテをＡ氏に開示（閲覧）しているが、Ａ氏の要求により10月9日カルテの印刷物手渡しによる2回目の開示を行っている。同カルテを確認したＡ氏より、11月7日付メールにてＢ医師がカルテ改ざんを行ったと指摘を受けた。大学院・研究推進課より、Ａ氏よりカルテ改ざんの指摘を受けた旨学長に報告され、12月下旬Ｂ医師に対し学長同席のもとカルテ記載変更について確認が行われた。このときＢ医師は、「丁寧に書きなおした」ものとの認識を述べ、カルテ開示において、カルテの修正履歴も同時に印刷されるものとは想定していなかったと証言している。

２）改ざんの認識

電子カルテの誤字などを入力確定の後に気づき、記載した本人が訂正することは日常的に行われうることであるが、入力確定後のすべての修正履歴は電子的に記録されており、電子カルテの真正性が担保されている。大学病院の電子カルテも「医療情報システムの安全管理に関するガイドライン」に準拠して運用されており、カルテ開示においては変更履歴も含めて確認可能な形で印刷し、開示しているところである。この場合、特段の理由のない訂正、記載時期から遅れた訂正、当初の記載等から説明のつかない訂正等は、改ざんが強く疑われることとなる。

本件に関するヒアリングにおいてＢ医師は、カルテ記載の変更は「Ａ氏がカルテ開示を求めてくると思われたから行った」と答えており、カルテ開示対策として行ったものであることは明らかである。また「丁寧に書きなおした」ものであるとも述べているが、「ブロナンセリン」を「抗精神病薬」に書き換えたことは、２剤の割り付けをせず、当初からブロナンセリン単剤で実施したことを隠すためではないかと推測される。また、平成27年7月15日の修正は、臨床試験延長に対する同意を得られていないことについて、同意を得たことにするため書き換えたものと推測され、訂正ではなくカルテ改ざんと言わざるを得ない。

大学勤務の診療・教育に携わるベテラン医師が、電子カルテに修正履歴がつくこと、カルテの記載を安易に変更してはならないことなどのカルテ記載の常識を知らないはずはない。ヒアリングで述べた「まさか履歴がプリントアウトされるとは認識していなかった。」の言葉が真実であり、プリントアウトされないと認識して改ざんを行った可能性は否定できない。

（４）虚偽の報告（シュレッダー廃棄）

講座責任者の諭旨退職日でもある平成27年8月7日の時点をもって、本案件が臨床試験にかかわることとなってきたことから、対応窓口が医療安全管理室から大学院・研究推進課に移された。しかし、すでにＡ氏の申し出から1か月が経過しており、同人からの不満も多々メール等で送られてきている状況であった。Ａ氏からデータの開示、データの削

除を求められ、対応に苦慮した大学院・研究推進課担当者Ｄ氏と医療安全管理室担当者Ｃ氏とＥ氏は、8 月 21 日神経精神科医局でＢ医師に対し、「Ａ氏の要求は臨床試験にかかわる全てのデータの原本の提出にあるが、その全てが保管されているのか」を問うたところ、Ｂ医師は「全てあるかは不明」と回答した。この会合において、Ａ氏には「資料は全て廃棄した」と伝えることについて出席者全員が了承した。そこで、8 月 24 日医療安全管理室担当者Ｅ氏が、「臨床試験のデータの原本については、Ａ氏の意向に基づいて、全てシュレッダーにて廃棄した」とＡ氏に伝えている。

なお、「原本をシュレッダーで廃棄したことは虚偽であること」について医療安全管理室長、医療安全担当副院長、大学院・研究推進課長、大学病院長及び学長が把握した経緯については以下のとおりである。

【報告内容】（平成２７年）

日時	報告者	被報告者	報告内容	受けた指示	資料
８　月下旬	大学院・研究推進課Ｄ氏	大学院・研究推進課長	医療安全管理室よりＡ氏へ紙の原本資料をシュレッダーで廃棄したと伝えた旨と廃棄したことが虚偽であることを報告。	Ａ氏に対して一度廃棄したと伝えており、言ったことを変更するとそのことについて何回も詮索される恐れがあることから、廃棄したと伝えたことについては修正せずに進めるよう指示を受けた。	D 氏及び課長の証言に基づく
８　月下旬	大学院・研究推進課Ｄ氏	神経精神科学教授代行（学長）	医療安全管理室よりＡ氏へ紙の原本資料をシュレッダーで廃棄したと伝えた旨と廃棄したことが虚偽であることを報告。	Ａ氏に対して廃棄したと伝えたことについては修正せずに進めるよう指示を受けた。	D 氏の証言に基づく
12　月上旬	大学院・研究推進課Ｄ氏	大学院・研究推進課長	A 氏が参加した臨床試験の研究不正が発覚したことから、シュレッダーしたことが虚偽であることについてＡ氏に真実を伝えるべきと考え、その旨報告した。	学長に報告して、その指示に従って対応するようにとのことであった。	D 氏及び課長の証言に基づく

12 月上旬	大学院・研究推進課D氏	神経精神科学教授代行(学長)	A 氏が参加した臨床試験の研究不正が発覚したことから、シュレッダーしたことが虚偽であることを、A 氏に伝えるべきかどうかについて、大学院・研究推進課長から学長に報告し指示を受けるようにとのことであったため、その旨報告した。	顧問弁護士に相談するよう指示があったことから、顧問弁護士に相談し、A氏に真実を伝えた。	D 氏の証言に基づく

なお、８月初旬に対応窓口が、医療安全管理室から大学院・研究推進課に移管されたことから、医療安全管理室から大学病院長、医療安全担当副院長並びに医療安全管理室室長へは「虚偽であること」についての報告はなされていない。

１）虚偽の報告はなぜ行われたか

医療安全管理室担当者Ｃ氏自身は、Ｂ医師は原本資料をシュレッダーで廃棄していないと認識していた。しかし被験者は神経精神科の患者であり、その対応については神経精神科医の判断を尊重すべきと理解していたので、たとえ嘘であっても「資料はすべて廃棄した」と伝えることが医学的に適切な対応で、A氏に安心感を与えるとＣ氏は考えた。

大学院・研究推進課担当者D氏へのヒアリングにおいては、A氏はデータを削除するようにと申し出ているのであるから、無くなったといえば安心するのではないかと思ったと述べている。また、神経精神科の患者であるので神経精神科の医師の最終判断に従うが、D氏自身これが嘘であることは感じていたと述べている。

Ｂ医師へのヒアリングでは、医療安全管理室担当者Ｃ氏もA氏の度重なる要請に相当ストレスを感じていたのではないかと述べている。足らないものがあればすぐに追及される状況で、「シュレッダーで廃棄したといえば安心されますね」と医療安全管理室担当者Ｃ氏から発言があったので、自分も追及にストレスを感じていたため、最善の策だと感じたとＢ医師は述べている。

しかしながら、データを開示する前に削除したということに対し、A氏からはさらなる要請も続いたことから、大学院・研究推進課担当者D氏は 12 月 14 日にA氏と面談した際、原本資料をシュレッダーにかけて廃棄したことは虚偽であったことをA氏に説明している。

２）シュレッダー廃棄の虚言の責

以上の様に、関係者は全てＡ氏に虚偽の報告をすることを知っていたと理解できる。しかし、Ｂ医師の臨床試験データの整理、保管の不備、カルテ開示を想定して早々にカルテ改ざんを行った行為、不正な臨床試験の実行、当事者意識の欠如が最悪の事態を招いた主たる原因と考えられる。また診療部長も主体的な対応を怠り医療安全管理室に対応をゆだねていたことが、問題の解決を遅らせ、シュレッダーによる全資料廃棄という虚偽の報告に至る原因となったと理解される。被験者が患者であることからその対応の最終判断を医療安全管理室や大学院・研究推進課が有していたとは思われず、シュレッダー廃棄の虚言の責は、提案者が誰であるにせよＢ医師にあると言わざるを得ない。一方、医療安全管理室のＣ氏、大学院・研究推進課担当者Ｄ氏も、直接臨床試験責任医師に進言することはできなかったとしても、虚言であろうことは認識しており、各々の部門の上司に報告、連絡、相談し未然に防ぐことができた可能性があったと思われる以上、責を免れることはできない。

（５）組織としての責任

１）責任体制

大学病院に受診した患者やその関係者等からの多様な申し出に対応している医療安全管理室は大学病院の管轄下にあり、その責任者は大学病院長である。臨床試験等の審査、承認、監督は、大学の組織である生命倫理委員会臨床試験部会が行っており、大学院・研究推進課は、臨床試験等に係る事務作業を担う大学の組織である。またその最高責任者は学長である。

２）病院、大学の対応と責任

大学病院長は、毎月医療安全管理室から患者相談窓口業務活動実施報告を受けており、7 月 1 日付の入電により対応が開始された本案件については、8 月 7 日に「7 月分患者相談窓口業務活動実施報告書」の中で報告を受け把握している。同時に医療安全担当の副院長も同日付で医療安全管理室担当者Ｃ氏から相談を受けて本案件を把握し、その内容が臨床試験に関すると判断されることから、窓口を医療安全管理室から大学院・研究推進課に移管するよう指示している。

本案件の主たる対応を病院から大学に移管した 8 月 7 日は、Ｂ医師の所属する神経精神科学の教授が精神保健指定医の取消しの責により諭旨退職となった日でもあり、この日以降学長が教授代行を務めることになった。従って 8 月 7 日以降は、大学院・研究推進課から大学病院長と学長に本案件の報告がなされている。

本来７月１日の入電の時点で、データ削除に関する案件は臨床試験を計画したＢ医師とその上司である診療部長（神経精神科学教授）に報告されており、通常は診療部長ま

での間に問題は解決される。しかし、Ｂ医師が自宅謹慎処分中であったこと、同医師の臨床試験データの整理、保管の不備、カルテ改ざんを伴う不正な臨床試験の実行、当事者意識の欠如、さらには本事案に対して診療部長の主体的対応が見られなかったことが最悪の事態を招いた主因であり、その指導監督を行う大学の責は免れない。

その当時の学長、大学病院長への報告に関する時系列は以下のとおりである。

平成 27 年

7 月 1 日：被験者Ａ氏より入電

7 月 6 日：「患者相談回答等依頼書」により、神経精神科部長に報告

8 月 7 日：「7 月分患者相談窓口業務活動実施報告書」により大学病院長へ報告

8 月 7 日：医療安全担当副院長へ相談し、対応窓口を医療安全管理室から大学院・研究推進課へ移行

8 月 24 日：Ａ氏に臨床試験のデータの原本についてはシュレッダーで全て廃棄した旨通知

8 月 26 日：大学院・研究推進課から学長に状況報告

9 月 4 日：「8 月分患者相談窓口業務活動実施報告書」により、その後の状況について大学病院長へ報告（シュレッダーで廃棄の件）

9 月 11 日：大学院・研究推進課から学長に状況報告

9 月 15 日：大学院・研究推進課から学長に状況報告

10 月 9 日：Ａ氏にカルテの印刷物手渡しによる 2 回目の開示を行う。大学院・研究推進課から学長に状況報告

11 月 5 日：大学病院長が医療安全管理室担当者Ｃ氏及び大学院・研究推進課担当者Ｄ氏を呼び現在までの状況を確認

11 月 7 日：Ａ氏からＢ医師がカルテ改ざんを行った旨メールにて指摘。

11 月 10 日：Ｂ医師から臨床試験部会に「臨床試験に関する報告書」を提出

11 月 11 日：臨床試験部会で当該臨床試験を検証

11 月 30 日：Ｂ医師から学長に「臨床試験に関する報告書」を提出

12 月 3 日：臨床試験部会で再度当該臨床試験を検証

12 月 28 日：よみドクターに本件が掲載

平成 28 年

1 月 4 日：再度大学病院長が医療安全管理室担当者Ｃ氏及び大学院・研究推進課担当者Ｄ氏を呼び現在までの状況を確認

1 月 5 日：生命倫理委員会で当該臨床試験を審議

1 月 12 日：生命倫理委員会から学長に報告
（当該臨床試験の中止・論文取り下げが妥当）

1月13日：研究倫理委員会で本件について審議
（学長指名による内部調査委員会発足・合計10回開催）
1月14日：川崎市からネット記事であるよみドクターについての事情聴取
1月26日：厚生労働省医政局研究開発振興課治験推進室より問合わせ
2月24日：文部科学省科学技術・学術政策局人材政策課研究興生推進室より問合わせ
3月10日：参議院厚生労働委員会で本件が取り上げられる
3月16日：厚生労働省へ医学部長他４名が出向き状況説明（結果を学長へ報告）
3月18日：上記結果を大学病院長へ報告
4月 7日：外部委員を含めた第1回臨床試験に係る調査委員会開催
5月12日：朝日新聞からの取材に医学部長、総務部長が対応（内容を学長、大学病院長へ報告）
6月20日：調査委員会で、Ａ氏に対してヒアリングを実施することを決定。
6月20日：（夕刻）大学からＡ氏に対して、ヒアリングを実施させていただきたい旨メールで連絡。
6月21日：（夕刻）大学からＡ氏に対して、７月５日（火）午後４時～６時の間でヒアリングを実施させていただきたい旨を連絡。
6月22日：（AM2時32分）Ａ氏からメールで連絡があり、都合が悪いので辞退したいとのこと。
6月22日：（AM6時32分）Ａ氏からメールで連絡があり、なんとか都合がついたので、お伺いいたしますとのこと。
6月22日：（AM10時50分）Ａ氏から電話連絡があり、やはり仕事の都合がつかないので、辞退したいとの連絡。それは日程変更ということではなく、これまでメール等で自分の意見・要望等は伝えてあるので、「あらためてお伝えすることはもうありません」とのことであった。
　これに対して大学からは、今回７月５日（火）は中止いたしますが、今後もし何かお話ししたいというようなことが出てきたら、改めて日程調整はさせていただきますと伝えた。
7月19日：Ａ氏からその後の進捗状況について確認のメールがあり、本学としては今月中には中間報告書をとりまとめたい旨伝えた。
　また、Ａ氏からは臨床試験の不正についてはこれまでさんざん調べてもう疲れました。ただし、嘘をつかれたのは許されないので謝罪してほしいとのことであった。
7月20日：大学からＡ氏に対して中間報告書を厚労省に送った旨メールにて連絡し、今後最終報告を待ってしかるべき対応をとりますと説明。

以上のとおり、対応窓口が移行したことにより、患者さんへの対応において責任の所在が曖昧になってしまった部分もあるのではないかと考える。そこで、本来臨床試験の責任の所在は大学であるが、今回のことを踏まえて、今後大学と病院が連携していくシステムを構築していく必要があるかと考える。

２．当該臨床試験の調査結果について

（１）試験の構造の問題

当該臨床試験の実施計画書では 2 剤の薬剤をランダム化して被験者に投与する試験計画であった。

実際にはブロナンセリンの研究が単エントリー（ランダム化されず）で先行していたことから実施計画書からの逸脱を確認した。本研究はランダム化されず、ブロナンセリンに関する臨床試験が先行して実施され、その後、アリピプラゾールに関する臨床試験が実施されたことが明らかである。

（２）同意に関する問題

１）本研究に係る実施計画書記載事項について：研究期間、対照群の設置

ア．本研究実施計画書第 1 版（初回承認実施計画書）（臨床試験部会審査日、平成 21 年 2 月 12 日、臨床試験部会承認日、平成 21 年 3 月 6 日）において、当初、臨床試験実施期間は、平成 21 年 3 月 6 日～平成 24 年 1 月 31 日、試験期間（観察期間）は、0～8 週間に設定された。

イ．実施計画書第 2 版（変更申請 1 回目）（臨床試験部会承認、平成 23 年 6 月 9 日）において、臨床試験実施期間が平成 21 年 3 月 6 日～平成 28 年 1 月 31 日までに変更され、試験期間が 0～52 週に変更された。

さらに、「認知機能検査を繰り返すと学習効果を生じる可能性があるため、慢性期の患者対照群も加えて認知機能検査と臨床症状の評価を実施し、学習効果の影響を検討したい」と追記され（1. 試験の背景の項（P2、33-35 行目）、対照群患者について、選択基準、除外基準（4. 対象患者 3）および 4）の項（P6、15 - P7、14 行目）と共に、BACS-J、PANSS および DIEPSS のみを同じスケジュールで計 4 回実施し、認知機能における学習効果の影響を検討すると明記された（6.　試験方法 2)試験のアウトラインの項　（P8、22-24 行目））。

ウ．実施計画書第 3 版（変更申請 2 回目）（臨床試験部会承認、平成 25 年 3 月 14 日）においては、実施計画書の 7. 評価項目、副次的評価項目の、機能的能力の検査に、統合失調症認知評価尺度（Schizophrenia Cognition Rating Scale-Japanese version:SCoRS）および UCSD 日常生活技能簡易評価尺度（University of California

San Diego, Performance-Based Skills Assessment-　Brief: UPSA-B）が追加された（P11、25-29 行目）。

２）本研究の実施計画書における同意取得に関する記載

本研究の実施計画書では、下記のとおり全例文書同意を得ることが記載されている。

ア．初回承認実施計画書に記載の同意取得方法について

（ア）初回承認実施計画書、5.被験者に説明し同意を得る方法の項（P5、12 - 20 行目）には、次のように記載されている。

①本試験開始前に担当医師は、試験の目的、方法さらに、被験者が本試験に同意をしない場合であっても不利益は受けないこと、同意した場合でも随時これを撤回できること、被験者の人権保護など必要な事項について被験者に十分説明し、被験者の自由意思による同意を本人より文書で得る。

②被験者が未成年の場合、被験者本人から同意が得られた場合でも、被験者本人だけでなく必ず代諾者（家族など被験者の最善の利益をはかりうる人）からも文書同意を得る。

（イ）初回承認実施計画書の 4. 対象患者 1）選択基準の項（P4　12 行目）には、⑤本人から文書同意の得られる患者と明記されている。

イ．実施計画書第 2 版における同意取得方法の追加記載

（ア）実施計画書第 2 版の 6.試験方法のデザイン 2)試験のアウトラインの項（P8、3-4 行目）には、すでに開始時と 8 週後または中止時の計 2 回の評価で試験同意を得られている症例に関しては、再同意を得ることが明記された。

（イ）実施計画書第 2 版の 6.　試験方法 2)試験のアウトラインの項　（P7、34 - P8、1 行目）に、対照患者群は、選択基準に合致し同意が取得できた患者を組み入れると、明記された。

３）ブロナンセリン患者群（0-8 週までの試験期間）における同意取得状況

初回実施計画書では、試験期間は、0～8 週と規定され、同意説明文書には、「この試験に参加された場合の予定参加期間は、計 8 週間となります」と明記されている（同意説明文書、項目 4.　この試験の予定参加期間）。

神経精神科から提出された、当該臨床試験に係る症例毎の原資料および、症例のカルテ調査を基に、患者群の同意取得状況について調査した。その結果、症例毎の原資料には、それぞれ、本人が自署した同意書のコピーが保存されていた。従って、全例（参加開始時(0 週)23 名、症例記号 BF-a～BF-w）から、初回実施計画書・同意説明文書の

記載内容で、文書同意を得たことを確認した。なお、症例記号は、症例番号等非公表のため、症例毎の差別化を図るため便宜上付けたものである。

4）ブロナンセリン対照群（0-8 週まで）における同意取得状況

神経精神科から提出された対照群 10 例（症例記号 BC-a～BC-j）の症例毎の原資料には、同意書のコピーを確認できなかった。従って、全例、文書同意を得ていなかったことが判明した。

カルテ調査を行った結果、10 例中 2 例（症例記号 BC-a および BC-b）において、外来で、研究分担医師が、本研究について説明し、研究参加の同意を得た旨の記載を確認した。従って、この 2 例からは、口頭同意を得たものと判断した。ただし、対照群の試験参加予定期間は、実施計画書第 2 版によれば、患者群と同様に、「全期間における最長期間:12 ヶ月」と明記されている(6.　試験方法 2)試験のアウトラインの項（P8、22-24 行目))。計 4 回の検査を予定していたことからも、対照群の、最長参加予定期間が 12 ヶ月であったことは、明らかである。しかし、カルテ記載からは、試験参加予定期間について、どのような説明を行ったのか、確認できなかった。

他の 8 例について、カルテ調査では、本研究について説明したことを示す記載や、本研究に参加する旨の同意を得たことを示す記載は、確認できなかった。従って、この 8 例から、本研究参加の同意を取得したという事実は、確認できなかった。

以上より、対照群において、研究参加の同意（口頭同意）が取得できたのは 2 例のみであると判断した。

5）ブロナンセリン患者群における試験期間延長時（26-52 週までの試験期間）の再同意取得状況

神経精神科から提出された症例毎の原資料、およびカルテ調査から、26 週では 11 例、52 週では 10 例が本試験を継続したことが判明した。しかし、これらの症例について、症例毎の原資料中に、再同意書のコピーは無く、26 週以降の試験延長について、全例、文書同意による再同意を得ていなかったことを確認した。加えて、カルテ調査から、26 週以降の試験延長について、全例、口頭同意による再同意を得ていなかったことを確認した。従って、52 週まで本試験を継続した症例から、再同意は得ていないことが判明した。なお、26 週時の 11 症例中 3 症例、52 週時の 10 症例中 3 症例では、同意を得て検査を実施した旨のカルテ記載があったが、試験期間延長に関する再同意を取得したと判断できる内容ではなかった。

ブロナンセリン患者群で、26 週まで試験を継続した症例 11 例、52 週まで試験を継続した症例 10 例について、変更申請 1 回目が承認された、平成 23 年 6 月 9 日の時点で、既に、6 例が 52 週までの本試験を完遂、4 例が 26 週までの本試験を完遂、残る 1 例も、既に 8 週目を超えて、本試験が進行中であったことが判明した。これら試験期間

を延長した症例は、実際には、変更申請が承認される以前に、同意を取得しないまま、8 週目を超えて、本試験を継続してしまったことが判明した。

（3）エントリー基準に関する問題

本研究の実施計画書では選択基準が以下の通り規定されている。

1）初回実施計画書、4. 対象患者の項（P4、1 行目 - P5、11 行目）に記載されている、初発エピソード患者の選択基準、除外基準は次の通りである。

DSM-IV-TR（精神疾患の分類と診断の手引き：米国精神医学会）で統合失調症、統合失調症様障害または統合失調感情障害と診断され、以下の選択基準すべてを満たし、除外基準のいずれにも該当しない患者。

A．選択基準（初発エピソード患者）

① 初発の精神症状が少なくとも 1 か月以上 5 年以内持続している事。

② 抗精神病薬の服薬が合計 16 週未満かつ同意取得の 12 週間前から投与直前まで抗精神病薬を服薬していないこと。

③ 同意取得時 16～40 歳の男女。

④ 外来通院中あるいは入院中の患者。

⑤ 本人から文書同意の得られる患者。

B．除外基準（初発エピソード患者）

① 同意取得前に blonanserin か aripiprazole の投与を受けたことがある患者。

② 神経変性疾患を有する患者。

③ アルコールまたは薬物乱用歴のある患者。

④ 精神遅滞合併患者。

⑤ 症状性あるいは器質性精神障害（認知症を含む）合併患者。

⑥ 三環系抗うつ薬を投与中の患者。

⑦ 過去 3 か月以内に電気けいれん療法を受けた患者。

⑧ 過去 3 か月以内に抗精神病薬の持効製剤が投与されていた患者。

⑨ 措置入院あるいは医療保護入院中の患者。

⑩ 妊婦、授乳中の女性患者、妊娠している可能性のある女性患者、同意取得後から 8 週間以内に妊娠を希望している女性患者。

⑪ 自傷行為または自殺企図の可能性の高い患者。

⑫ Blonanserin および aripiprazole の添付文書において禁忌にあたる患者。

⑬ その他、試験担当医師が不適当と判断した患者。

2）実施計画書第 2 版、4. 対象患者の項（P6、15 行目 - P7、14 行目）に記載れている、

対照患者群の選択基準、除外基準は次の通りである。

C．選択基準（対照群患者）

① Blonanserin または aripiprazole を単剤で 3 か月以上投与され、用量の変更がなく、精神症状が安定している患者。

② 同意取得時 16～40 歳の男女。

③ 外来通院中あるいは入院中の患者。

④ 本人から文書同意の得られる患者。

D．除外基準（対照群患者）

① 神経変性疾患を有する患者。

② アルコールまたは薬物乱用歴のある患者。

③ 精神遅滞合併患者。

④ 症状性あるいは器質性精神障害（認知症を含む）合併患者。

⑤ 三環系抗うつ薬を投与中の患者。

⑥ 過去 3 か月以内に電気けいれん療法を受けた患者。

⑦ 過去 3 か月以内に抗精神病薬の持効製剤が投与されていた患者。

⑧ 措置入院あるいは医療保護入院中の患者。

⑨ 妊婦、授乳中の女性患者、妊娠している可能性のある女性患者、同意取得後から 8 週間以内に妊娠を希望している女性患者。

⑩ 自傷行為または自殺企図の可能性の高い患者。

⑪ Blonanserin および aripiprazole の添付文書において禁忌にあたる患者。

⑫ その他、試験担当医師が不適当と判断した患者。

3）エントリー基準違反症例および、その他の疑義症例について

ブロナンセリン患者群（初発エピソード患者群）23 名は、症例記号 BF-a～BF-w で表示し、ブロナンセリン対照群（対照患者群）10 名は、症例記号 BC-a~BC-j で表示した。

ア．ブロナンセリン患者群

BF-c：前医でリスペリドンあるいはジプレキサを、通算で 18.7 週相当の期間服薬しており、A－②を満たさない。

BF-m：前医で処方されたジプレキサを試験参加後もしばらく服薬していたため、ブロナンセリンと併用期間が 2 週間存在しており、A－②およびブロナンセリン単剤治療という試験の目的に合致しない。カルテには、ブロナンセリンを上乗せして使用する旨の記載があり、担当した医師は、明らかに併用投与であることを理解していた

ものと判断した。

イ．ブロナンセリン対照群

BC-a：ブロナンセリン単剤治療ではなくレボメプロマジン（抗精神病薬）とカルバマゼピン（抗てんかん薬）が併用されており、Ｃ－①を満たさない。レボメプロマジンとカルバマゼピンは共に、併用禁止薬に指定されている（実施計画書第 2 版、併用禁止薬の項 6）-①-1)（P9、25 行目、30 行目）。

BC-d：ジプレキサが頓服用（服薬する時間が決められておらず、必要なときに使用する薬）として処方され、実際に服用したとのカルテ記載があることからＣ－①を満たさない。

BC-e：入院中、不穏状態でレボメプロマジン、リスペリドン（ともに抗精神病薬）を適宜使用したことがカルテ記載から確定でき、単剤治療の基準を満たさない。また、ブロナンセリン投与量が 8 mg → 4 mg → 24 mg → 20 mg → 12 mg →16 mg → 12 mg → 4 mg と変更されており、Ｃ－①を満たさない。

BC-g：ブロナンセリン服用が 3 か月未満でありゾテピン（抗精神病薬）も併用投与されており、Ｃ－①を満たさない。

BC-h：ブロナンセリンの投与量が、4 mg → 8 mg → 16mg と変更されており、Ｃ－①を満たさない。

ウ．ブロナンセリン対照群における、エントリー基準違反以外の疑義症例

ブロナンセリン対照群では、エントリー基準を満たさない 5 症例以外に、カルテ調査の結果、3 症例で、適格性に疑義のあることを確認した。

BC-f：神経精神科から提出された、症例毎の原資料に加え、カルテ調査を行っても、いかなる症例で、いつ検査を実施したのか、特定できなかった。従って、この症例は、神経精神科から申告された症例と同一症例であることを、確認できなかった。

BC-i：カルテ調査の結果、この症例は、別の独立した臨床試験、「統合失調症に長期投与された抗不安薬の減量中止が認知機能に及ぼす影響に関する研究」（承認番号 1668、実施責任者、Ｂ医師：承認日、平成 22 年 3 月 10 日；実施期間、平成 22 年 3 月 10 日～平成 24 年 12 月 31 日；終了）に、同意し、参加した症例であることが、判明した。本試験において、対照群を設定した目的から判断して、対照群には不適格な症例であると判断した。

BC-j：カルテ調査の結果、この症例は、症例 BC-i と同様、別の独立した臨床試験、「統合失調症に長期投与された抗不安薬の減量中止が認知機能に及ぼす影響に関する研究」（承認番号 1668、実施責任者、Ｂ医師：承認日、平成 22 年 3 月 10 日；実施

期間、平成 22 年 3 月 10 日～平成 24 年 12 月 31 日；終了）に、同意し、参加した症例であることが、判明した。症例 BC-i と同様、本試験の対照群には不適格な症例であると判断した。

以上、ブロナンセリン対照群 10 例中、同意取得ができ、かつ対照群として適格であることが確認できたのは、症例記号 BC-b、1 例のみであった。

（4）実施計画書を逸脱してブロナンセリンを投与した原因の究明

カルテ調査およびＢ医師に対する聞き取り調査より、当該臨床試験 1 例目から実施計画書を逸脱して乱数表を使わず、意図的にブロナンセリン投与群に振り分けられていた。この理由としてＢ医師は 2 つ挙げている。①ブロナンセリン 1 例目が劇的に効果をあげたため、科学的興味がわいてブロナンセリン投与群に組み入れた。②大学院生がおり、学位論文を仕上げなければならなかったのでブロナンセリン投与群に偏ってしまった。このうち①についてはカルテ調査の結果、改善はみられるものの“劇的”とは言いがたく、この理由は否定される。②については当時大学院生のＧ医師が在籍しており時期的にも研究開始から論文受理までの期間に組み入れた被験者の数が論文作成に必要な症例数と考えられる。よって②が実施計画書を逸脱してまでブロナンセリン投与群に組み入れた理由として合理的である。

（5）利益相反について

本学における利益相反の管理状況については、平成 19 年 5 月 1 日付けで利益相反管理規程が制定され、利益相反の報告について、各研究者が自主申告を行うこととしたが、各研究者に対する周知については、学内ホームページに掲載しただけで、周知徹底がなされていたとは言いがたい。また、利益相反委員会自体も開催されておらず、各研究者から自主申告のあった報告書を年度ごとに所管部署である人事課が件数等を取りまとめ、原議書により利益相反管理委員会委員長の決裁を経て理事長へ報告がされているのみであった。なお、この方法では、研究者の自発的申告を待つという受け身のシステムであり、研究者が自ら申告しない限り全く把握することができず、また、委員会自体も特に問題となる案件がない場合開催されておらず対応が不適切であった。今後は少なくとも半年に１回程度の定期的な委員会の開催と、研究者全員からの報告を義務づけるような体制を構築する必要がある。

なお、平成２２年度からの申告状況は次のとおりである。

利益相反開示報告書　提出者数および件数

	第１次報告書のみ		第１次および第２次報告書		合　計	
	人　数	件　数	人　数	件　数	人　数	件　数
Ｈ２２年度	３３	３３	１	２	３４	３５
Ｈ２３年度	６５	１００	０	０	６５	１００
Ｈ２４年度	２５	２６	０	０	２５	２６
Ｈ２５年度	３０	４１	０	０	３０	４１
Ｈ２６年度	３３	５３	５	８	３８	６１
Ｈ２７年度	１２７	１４５	１１	１９	１３８	１６４
Ｈ２８年度	５３	６１	４	４	５７	６５

また、本件における利益相反の状況についても調査した。Ｂ医師はブロナンセリン・アリピプラゾールの製薬会社から講師謝金及びコンサルティング業務委託費等の名目で支払いを受けているが、旧倫理指針において、「研究者等は、臨床研究を実施する場合には、被験者に対し、当該臨床研究の目的、方法及び資金源、起こりうる利害の衝突（利益相反）、研究者等の関連組織との関わり、当該臨床研究に参加することにより期待される利益及び起こりうる危険、必然的に伴う不快な状態、当該臨床研究終了後の対応、臨床研究に伴う保証の有無その他必要な事項について十分な説明を行わなければならない。」とあるが、今回においては、製薬会社から講師謝金等の資金を得ていることについて被験者へは説明がなされていない。また、学内規程に則り利益相反の窓口である人事課へは申告がなされていなかった。このことについてＢ医師からヒアリングを行った結果、これらの謝金は当臨床試験との直接的な関連はなく、また「利益相反開示報告書に「企業と連携している場合は」との記載があり、その内容には抵触していないことから申告書は提出しなかった」旨の報告を受けている。しかしながら、利益相反の開示に関する基準額には、年間 100 万円以上の講演料・謝礼金等を受け取った場合には報告が義務づけられているにもかかわらず、申告がなされていなかった。

以上のような結果について、①当該臨床試験が実施計画書から逸脱し、両剤の比較試験が単剤試験になっていたこと、②単剤試験に変更するについて、実施計画書の変更申請が行われず、真相が隠蔽されていたこと、③エントリー基準が遵守されていなかったこと、④カルテが勝手に改ざんされていたこと、⑤利益相反について、開示報告書が提出されなかったことについては、Ｂ医師が講師謝金や業務委託料を受け取っている製薬会社の利益になるようにそれらのことを行なったのではないかという疑念が完全にぬぐ

いきれず、当調査委員会としてはこれ以上の調査は困難であるが、そのように疑われてもやむを得ない事実経過であると考える。

なお、製薬会社に確認したところ、ブロナンセリンにおいては、パンフレットにデータの一部が利用され、MRによって希望する医師等に対して配布していたとのことであり、報告書の論文の検証結果で述べるとおり、多数の不適格症例を包含したまま、まとめられたデータをそのような形で利用した事については、研究者としてのモラルを逸しており、はなはだ不適切であったと考える。また、「当該論文を引用したガイドラインは、広告宣伝とした制作物も無く、プロモーションには用いておりません。」との回答を製薬会社から得ている。

以上のとおり被験者への説明及び学内規程に則った申告がなされていなかったことについては、はなはだ不適切であったと考えるが、冒頭述べたように本学において利益相反については教職員への周知が十分になされているとは言いがたく、特に研究者に対してはこれまで定期的な研修会を開催するなどの周知徹底がなされておらず、今後は研究者全員に対して研修会の出席を義務づけるなど、徹底した周知が必要と考える。

なお、神経精神科医師が関わった他の臨床試験についても調査した結果、１件について製薬会社から研究費を受け入れていたが、他の臨床試験については資金提供は一切受けていなかった。

（６）被験者の不利益について

本臨床試験においては、被験者に対してブロナンセリンとアリピプラゾールをランダムに投与して比較する試験であり、その旨被験者に対して説明し同意を得ているが、実際にはブロナンセリンのみを割り付けており、そのことについて被験者に対して説明がなされていない。ブロナンセリンはすでに承認され平成 20 年 4 月に市販された日常診療で用いられている第 2 世代抗精神病薬で、その適応内で試験が行われたことから、被験者にとっての直接的な影響はなかったと思われるが、実施計画書から逸脱して一方の薬剤を割り付けられたことについては、不信感や不安を与えたことで、被験者にとっては不利益を被ったのではないかと考える。また、このことについてＢ医師の説明では、被験者の方々には口頭で謝罪と説明をさせていただいたとの事であったが、被験者Ａ氏のみならず全ての被験者に対して書面により謝罪と説明をする必要がある。

３．当該臨床試験を基とした論文の検証結果について

本臨床試験を基に、作成された論文 2 編について、その科学的適切性について検証を行った。なお、論文の検証は、聖マリアンナ医科大学データセンターの協力を得て実施した。

（１）論文１の検証

１）データセンターにおける論文１に関する検証作業の内容

データセンターでは、神経精神科学教室から提出された、基礎データ（エクセル等の電子媒体基礎データと、評価表等の紙ベース基礎データ）から、論文 1 の Table 1 ~ 6 が作成された過程についての検証作業を行った。

作業手順は、次の通りである。

手順 1 : BACS-J、SQLS-J、PANSS、CGI、DIEPSS の各評価について、紙ベース基礎データを基に、全ての項目について、データセンターで確認・再計算を行った。

手順 2 : データセンター再計算値と紙ベース基礎データ、および電子媒体基礎データとを比較した。

手順 3 : データセンター再計算値に基づき、論文記載の解析方法に従って統計解析を行い、その結果を、論文記載値と比較した。

2）論文 1 の Table 1 ~ 6 の検証

①Table1 について論文値及びデータセンター再計算値の比較

論文に記載されている数値及びデータセンターで紙ベース基礎データを元に再計算した PANSS の値は、ほぼ一致したが、僅かながら差異も認めた。

PANSS は、30 項目から構成され、1 回の検査あたり、4 因子を評価する。0 週の症例数が 23 例、8 週の症例数が 20 例であったので、総評価因子数は、172 であった。この中で、データセンター再計算値との間に差異を認めたのは 17 箇所であり、総評価因子数の 9.9%であった。これらの差異は、不規則に起こった、素点の得点の見間違えや素点の合計の計算ミスなど、ヒューマンエラーによるものと推測した。

Education level 及び Duration of untreated psychosis、Benzodiazepine dose については、提出された基礎データだけでは、検証できなかった。

②Table2 について、提出された基礎データだけでは、8 週時の併用薬名・使用量を把握することができず、検証できなかった。

③Table3 について論文値及びデータセンター再計算値の比較

Table3 の Baseline 及び Endpoint における論文値及びデータセンター再計算値の比較したところ、Baseline における Raw score 及び z-score について、値に差異を認めるところはあるが、全体的に大きな値の差は認めなかった。

BACS-J では、1 回の検査あたり、7 項目の Raw score を求め、さらに、標準値を用い、z-score を計算する。0 週の症例数が 23 例、8 週の症例数が 20 例であったので、総評価項目数は 301 であった。この中で、データセンター再計算値との間に差異を認めたのは 5 箇所であり、総評価項目数の 1.7%であった。従って、Raw score の差異は、不規則に生じた、計算ミスなどヒューマンエラーによるものと推測した。

z-score について、総計算項目数は 301 であった。この中で、データセンター再計算値との間に差異を認めたのは 56 箇所であり、総計算項目数の 18.6%であった。

データセンターにおける検証作業過程で、z-score の計算に用いた標準値が 2 種類存在したことが判明した。すなわち、8 週まで本研究に参加した 20 例のうち、16 例(80%)では確かに神経精神科から提出された標準値を用いていたが、残りの 4 例(20%)については精神医学 50 巻：913-917(2008)記載の標準値を用いたことが判明した。

精神医学 50 巻:913-917(2008)の論文は、論文 1 で引用されている論文である。20～49 歳に限ると、健常者データ数は、20~29 歳が n=31、30~39 歳が n=15、40~49 歳が n=17 であった。一方、神経精神科から提出された標準値の健常者データ数は、20~29 歳が n=110、30~39 歳が n=81、40~49 歳が n=72 であり、データ数が多いことが判明した。従って、本来、健常者データ数が多い神経精神科から提出された標準値を用いるはずであったが、本研究の分担研究者間で、標準値の取扱いについて、十分な総意が得られていなかった、あるいは、十分な注意を払わなかったため、4 例だけ精神医学 50 巻:913-917 (2008) に記載の標準値を用いたのであり、意図的に 2 種類の標準値を用いたものではないと推測した。この事実を踏まえると、z-score について、データセンター再計算値との間に、18.6%もの差異を認めたのは、z-score 計算時のヒューマンエラーもさることながら、z-score 計算に用いるべき標準値を 20 人中 4 人だけ別の値を使用した事が、大きな要因である。

以上のことから推測されるように、変化量（Change）に関して、Raw score、z-score、効果量（Effect size）、*t* 検定によって算出した *p* 値の、いずれにおいても、論文記載値とデータセンター計算値とに差異を認めた。

Results 内の記述" z-scores of the letter fluency score and the Tower of London score were significantly increased after treatment with blonanserin in the first-episode group (all $p<0.05$)"について、データセンター再計算値と *p* 値が揃っており、論文の記述通りであることを確認した。

同段落内の記述" No BACS-J subscale scores showed a significant change in the control group during the 8-week interval (all $p<0.05$)に関して、*p* 値に差異は認めるところはあるが、全ての項目において $p<0.05$ となり、論文の記述通りであることを確認した。

④Table4 について

Table 1、3、5、6 の検証より、論文記載値とデータセンター計算値の差異が、僅かであることが判明した。そこで、解析ソフト SPSS へ数値入力を行うことで得られる Table4 について、データセンターにおいて、同様の作業を改めて行う必要性はないものと考えた。

⑤Table5 について論文値及びデータセンター再計算値の比較

Table5 の論文値及びデータセンター再計算値の比較したところ論文値及びデータ

センター計算値に、僅かの差異を認めるが、全体的に大きな値の差を認めなかった。

SQLS-J は、1 回の検査あたり、合計 30 項目を評価する。0 週の症例数が 23 例、8 週の症例数が 20 例であり、総評価項目数は、1290 であった。この中で、データセンター再計算値との間に差異を認めたのは 11 箇所であり、総評価項目数の 0.9%であった。従って、これらの僅かの差異は、素点の得点の見間違えや素点の合計の計算ミスなど、ヒューマンエラーによるものと推測した。

CGI-S は、1 回の検査あたり 1 評価である。0 週の症例数が 23 例、8 週の症例数が 20 例であり、総評価数は 43 であった。この中で、データセンター再計算値との間に差異を認めたのは 6 箇所であり、総評価数の 13.9%であった。

DIEPSS は、1 回の検査あたり 1 評価である。0 週の症例数が 23 例、8 週の症例数が 20 例であったので、総評価数は 43 であった。この中で、データセンター再計算値との間に差異を認めたのは 5 箇所であり、総評価数の 11.6%であった。

CGI-S および DIEPSS では、データセンター再計算値との間に 10%を上回る差異を認めたが、SQLS-J と同様に、素点の得点の見間違えや素点の合計の計算ミスなど、ヒューマンエラーに起因するものであると推測した。

Results 内の記述" Significant improvements were found in all items on the PANSS, SQLS-J, and CGI-S (all $p < 0.01$)"に関して、SQLS-J の 2 項目について、論文記載値では $p = 0.006$ および $p = 0.008$ が、データセンター計算値では、$p = 0.028$ および $p = 0.010$ となり、論文の記述とは異なる値であったが、本論文では p 値 0.05 を規準としていることから、データセンター計算値においても有意差ありと判定され、論文記述に誤りは無いことを確認した。

またそれに続く記述、" There was a significant difference in the DIEPSS total score between baseline and endpoint in the first-episode group ($p < 0.05$)"について、データセンター計算値では $p = 0.014$ となったが、論文記述通り $p < 0.05$ であることを確認した。

⑥Table6 について論文値及びデータセンター再計算値の比較

Table6 の論文値及びデータセンター再計算値の比較したところ First-episode group 同様、値に僅かの差異を認めるところはあるが、全体的に大きな値の差を認めなかった。 PANSS は、30 項目から構成され、1 回の検査あたり 4 因子を評価する。0 週の症例数が 10 例、8 週の症例数が 10 例であったので、総評価因子数は、80 であった。この中で、データセンター再計算値との間に差異を認めたのは、3 箇所であり、総評価項目数の 3.8%であった。

また Results 内の記述" There were no significant differences in each PANSS score between baseline and endpoint in the control group"に関して、論文記載値と同様 $p > 0.05$ となることから、論文の記述通りであることを確認した。

3）承認実施計画書記載内容、実際に実施した臨床試験内容と、論文記載内容の整合性に関する検証

① 論文では、This prospective, single-blind, open-label study (page 91, right column, lines 2) と記載されているが、当該臨床試験の承認実施計画書の内容と一致しておらず、適切な記載ではない。

② 論文では、All participants gave informed consent after the study procedures had been fully explained と記載されているが、上記 II.-4 に述べたように、カルテ調査から、ブロナンセリン対照群 8 例では、同意を得た事実を確認できなかったので、適切な記載ではない。

③ 論文では、The clinician and psychologists who provided the clinical ratings were blinded to the assigned procedure of study patients と記載されているが、ランダム化されていなかったので、盲検下で検査を行ったとは言えず、適切な記載ではない。

4）その他の不適切記載

① 論文では、Part of this work was supported in part by Health and Labour Sciences Research Grants and grants from the Japan Society for Promotion of Sciences と記載されているが、本研究が、厚生労働科学研究費および日本学術振興会科学研究費の助成を得て実施された事実は確認できなかった。従って、適切な記載とは言えない。

なお、このことについて、研究責任者であるＢ医師からヒアリングをした結果、学外者のＦ先生からの依頼で記載したとのことであったことから、調査委員長がこの学外者のＦ先生に確認したところ、平成 28 年 6 月 27 日、Ｆ先生から「当該論文を作成する際に添削や推敲を行った。その際に私が受けている両科研費で国際学会などに参加して得た新知識をこの論文に使い、反映させたので記載してもらった。」との回答をいただいている。また、両論文がＦ先生の両科研費の業績に使われたかどうかは調べてみないとわからないとの回答も得ている。

5）論文 1 検証結果の総括

① 論文 1 の基になった当該臨床試験の実施に際して、承認された実施計画書からの逸脱が明らかになった。論文 1 では、文書同意取得がなされていない症例、エントリー基準違反の症例など、多数の不適格症例を包含したまま、臨床データを取りまとめ、統計学的解析を実施した点において、既に、倫理的適格性・科学的適格性が損なわれている。そこで、平成 28 年 1 月 5 日に開催された生命倫理委員会から学長に対して「１．同意書を取得していない被験者がいることから、当該臨床試験は中止すること。２．承認されていない計画に基づいて作成された論文を撤回することが望まれる。３．今後詳細な調査を行うにあたり、本案件全容を調査する調査委員会を立ち上げるこ

と。」との報告がなされており、本調査委員会としても生命倫理委員会の審議結果は妥当であり、当該論文は撤回するべきと考える。また、当論文は日本神経精神薬理学会が昨年公表した統合失調症薬物治療ガイドラインに引用されていることから、早急な対応が望まれる。

② 論文 1 に記載されている数値とデータセンターで素点を元に再計算した数値に、僅かな差異を認めた。これらの僅かな差異は、意図的な操作によって生じたものではなく、不規則に起こった、素点の得点の見間違えや素点の合計の計算ミスなど、ヒューマンエラーに起因するものと推測した。しかしながら、これらの誤りは、注意深く検査データを取り扱うこと、あるいはデータ入力後に多重チェック（ダブルチェック、トリプルチェック）を実施していれば、データ解析のいずれかの段階で、容易に気づかれて然るべき問題である。従って、基礎データの取扱い方に不適切な点があったことは、否めない。

③ BACS-J の z-score を計算する際に、4 例では、他の 16 例と異なる標準値を用いていたことが判明した。本研究の分担研究者間で、標準値の取扱いについて総意が得られていなかったことを示す結果である。この事実を踏まえると、z-score について、データセンター再計算値との間に 18.6%もの差異を認めたのは、z-score 計算時のヒューマンエラーもさることながら、z-score 計算に用いるべき標準値を 20 人中 4 人だけ別の値を使用した事が大きな要因である。BACS-J を用いた認知機能評価は、論文 1 において、論文の主題と密接に関連する重要評価項目である。にもかかわらず、重要な基礎データの取扱いに十分な注意が払われておらず、基礎データの取扱い方が極めて杜撰であったと言わざるを得ない。

④ 上記の事実から推測されるように、t 検定によって算出したいずれの p 値においても、論文記載値とデータセンター計算値とに僅かの差異を認めたものの、データセンター再計算値に基づく有意差検定結果は、論文記載の有意差検定結果と一致し、論文記載には誤りがないことを確認した。すなわち、論文記載の結論は、確かに論文記載の統計解析方法により導き出されたものであることを、神経精神科から提出された基礎データに基づく検討により確認した。従って、論文 1 の作成過程で、データに改竄、捏造、盗用のないことを確認した。ただし、盲検化していないのに盲検化していると記載するなど、不適切な表現が見られた。

（2）論文 2 の検証

論文 2 の検証では、まず論文記載の臨床試験計測症例数の推移、欠測データの補完方法について検証し、次に Table 1~ 3 について、作成過程について検証した

1）論文 2 における臨床試験継続症例数の推移について

本研究では、全観察期間は 1 年間であった。すなわち、0 週（服薬開始前）、8 週、26

週、52 週に、臨床症状評価・臨床検査を実施した。まず、神経精神科から提出された基礎データに加え、カルテ調査を実施し、臨床試験継続症例数の推移を検証した。

本研究の進行に伴って、臨床試験継続症例数が徐々に減少したことが判明した。すなわち 0 週では 23 例であったが、8 週で 20 例、26 週で 11 例、52 週では 10 例に減少した。加えて、同一症例が、26 週と 52 週において、いずれも BACS-J を拒否したことを確認した。予定した SWNS-J、SQLS-J、PANSS、BACS-J、DIEPSS、臨床検査を、52 週まで全て完了できたのは、9 例であったことが判明した。従って、Results 内の記載、14 (60.9%) remained in the study after 1 year, and 11 (47.8%; five men and six women) completed all evaluations at the end-point は、データセンターの調査結果と一致しないことが判明した。

一方、本研究と同じ研究グループが発表した総説では、投与開始後 1 年目の治療継続率は 60.9%（14/23 例）と記載している。従って、14 (60.9%) remained in the study after 1 year という記述は、正しくは、1 年目の治療継続症例数と記載すべきところを、1 年目の臨床試験継続症例数と混同し、記載したものと推測した。

論文 2 では、11 症例(男性 5 例、女性 6 例)を対象に completer analysis を実施したとの記載があるが、上記のように、真の completer analysis の対象者は 9 例であった。

論文記載の 11 例のうち、真の対象者 9 例を除く残り 2 例について、詳細に調査した。残り 2 例のうち 1 例は、26 週と 52 週の BACS-J 以外の全検査を 52 週まで完了した症例であることが判明した。このような特異な経過から、当該症例を 52 週までの全検査完了症例と誤認したものと推測した。

他の 1 例は、実際には 8 週時の検査までは完了したものの、26 週と 52 週の検査は実施していないことが判明した。神経精神科から提出された当該症例の原資料ファイルには、26 週と 52 週の検査結果は含まれていなかった。加えて、カルテ調査でも、26 週と 52 週に、臨床検査を含め、予定された検査を実施したという記録は確認できなかった。

論文には記載が無いが、次項に述べるように、本研究では、欠測データの補完方法に、最終観察繰越法(Last Observation Carried Forward :LOCF)を適用したことを確認した。LOCF 法に基づいて基礎データを見直すと、当該症例は、PANSS では、8 週のデータで 26 週と 52 週の欠測データを補完したことを確認した。SWNS-J、SQLS-J、BACS-J、DIEPSS、臨床検査では、8 週のデータで 26 週の欠測データを補完したことは確認できたが、8 週のデータで、52 週の欠測データを補完していないことが判明した。すなわち、52 週の欠測データは、LOCF 法を適用して補完したものではないと推測できたものの、これまでに収集した情報からは、具体的な補完方法は確定できず、当該症例を、52 週までの全検査完了症例と誤認した原因は確認できなかった。カルテ調査から、当該症例は、26 週および 52 週時にも、外来通院を継続していたことが判明した。この事実が、何らかの原因で 52 週までの検査完了症例と誤認することの素地と

なった可能性がある。

2）論文 2 における欠測データの補完方法、最終観察繰越法(Last Observation Carried Forward :LOCF)について

論文 2 の基になった、当該臨床試験の初回承認実施計画書には、14. データの集計および統計解析法の項に、「解析対象集団は、試験薬を 1 錠以上服薬し、BACS-J を試験薬投与後実施し逸脱例を除いた対象とする」と、明記されている。従って、そもそも、逸脱例は本研究の解析対象とはならない。

本研究では、論文に記載が無いが、神経精神科から提出された基礎データの検証から、欠測データの補完方法として LOCF 法を適用したことを確認した。LOCF 法は、欠測値を 1 つの値で補完する、単一代入法の 1 つである。

しかし、LOCF 法が妥当であるためには、「脱落後の結果変数の推移は最後に観察された値のまま変化しない(中止後に値が変わらない)」という非常に強い仮定を必要とする(松山、計量生物学 25 巻:89-116(2004))。いわゆる NAS(National Academy of Sciences;全米科学アカデミー)レポート(The Prevention and Treatment of Missing Data in Clinical Trials(2010))でも、「前提となるべき仮定が正当化できない限り主要解析として LOCF 法を使用すべきでない」と述べている(Recommendation 10、pp. 100)。

本研究のような抗精神病薬の臨床試験では、全ての症例に対して治療薬の効果があるかどうかはわからない。すなわち、治療薬の効果が症例によって大きく異なり、改善する症例も、悪化する症例も、全く影響を受けない症例もいることが想定され、LOCF 法が妥当であるとは到底言い難い。

加えて、本研究では、研究対象薬物の安全性評価も重要な研究目標である。LOCF 法の前提条件から考えて、LOCF 法を適用して欠測データを補完すると、安全性評価が正確に行えない。更に、本研究では 0 週(服薬開始前)の検査を受けたものの、8 週観察時まで本臨床試験を継続できなかった 3 例について、0 週の検査データを用い、LOCF 法を適用し、8 週、26 週、52 週の欠測データを補完し、薬効評価および安全性評価に使用した。0 週の検査は、服薬前すなわち研究対象薬物を服用していない時期に実施されたものである。従って、一度も研究対象薬物を服用していない時期の 3 症例のデータを、LOCF 法を適用し、薬効評価・安全性評価に使用することには、科学的妥当性がなく、明らかに、初回承認実施計画書記載事項（P12、4 - 5 行目）に反する。

本研究では、研究開始時の研究対象は、23 例であった。その後の経過で、臨床試験継続症例は、8 週では 20 例、26 週では 11 例、52 週では 10 例であった。試験開始時には、データは、のべ 92 症例分が取得できると想定されたが、実際には、のべ 64 症例分であった(64/92, 69.6%)。従って、本研究では、全取得データの約 30%が、実際には欠測データであったことが判明した。本研究では、欠測データは LOCF 法により補完され、統計解析が行われた。約 30%という欠測データ率は、LOCF 法を適用するに

は、高すぎると言わざるを得ない。

本論文では、Statistical analysis の項に、The analysis was performed on the 'intention-to-treat'(ITT) basis (n = 23)と記載している。ITT 解析の趣旨は、薬剤 A と B による治療法を、2 群にランダム化割り付けして比較する試験において、例えば薬剤 A 服用群に割り付けられた場合、その後たとえ薬剤 A の服用を中止したり、あるいは何らかの理由で薬剤 B を服用しても、あくまで A 群として解析することを意味する。

しかし、上述のように、本論文の基になった臨床試験実施計画書には、「解析対象集団は、・・・逸脱例を除いた対象とする」と明記されているので、本研究の立案段階で、ITT 解析（治療意図の原理による解析）ではなく、いわゆる per protocol 解析（実施計画書に適合した対象集団についての解析）を採用したことが明らかである。従って、ITT 解析の趣旨に基づき統計解析を行うことは、本臨床試験実施計画書記載の解析方法を遵守しなかったことに、他ならない。本論文の記載からは、LOCF 法を用い、欠測データを補完し、統計解析を行うことを、ITT 解析と、誤って理解していた可能性が推測できる。

以上より、本研究では、LOCF 法に関する理解不足が要因となり、LOCF 法を適用する際の条件に、必要十分な配慮がなされなかったことが否めない。加えて、LOCF 法を用い、逸脱症例も解析集団に包含させたことは、明らかに実施計画書から逸脱している。

3）論文 2 における Table1 ~ 3 の検証

①論文 2 Table 1 について論文記載値及びデータセンター再計算値の比較

次に、神経精神科から提出された基礎データから論文の Table1 が作成された過程の検証作業を行った。

論文記載値及びデータセンター再計算値の比較をしたところ、本論文については、論文記載値に関し、次のように 2 段階の比較検証を行った。

第 1 段階では、神経精神科から提出された、論文作成に使用した Excel データに基づき、データセンターにて再計算した値と、論文記載値を比較した。

第 2 段階では、SWNS-J、SQLS-J、PANSS、BACS-J、DIEPSS、臨床検査の全ての検査を 52 週まで完了した症例であることをカルテ調査結果も加えて確認した、真の completer analysis 対象者 9 人の素点だけに基づきデータセンターにて再計算した値と、論文記載値を比較した。

まず上記に示す第 1 段階の比較検証を行った。LOCF 法を適用し、欠測データを補完した割合が約 30%と高率だったので、実際に取得した検査データとデータセンター再計算値を比較した。

SWNS-J は、1 回の検査あたり、6 項目の評価をする。0 週の症例数が 22 例(SWNS-J のデータが無い症例が 1 例)、8 週の症例数が 20 例、26 週の症例数が 11 例、52 週

の症例数が 10 例であったので、総評価項目数は 378 であった。この中で、データセンター再計算値との間に差異を認めたのは、44 箇所であり、総評価項目数の 11.6%であった。

SQLS-J は、1 回の検査あたり 3 項目の評価をする。0 週の症例数が 23 例、8 週の症例数が 20 例、26 週の症例数が 11 例、52 週の症例数が 10 例であったので、総評価項目数は 192 であった。この中で、データセンター再計算値との間に差異を認めたのは 24 箇所であり、総評価項目数の 12.5%であった。

PANSS については、神経精神科から素点の Excel データが提出されていないため、データセンターにて PANSS 評価表に基づき確認した素点を用い、各項目の合計点を再計算した。

BACS-J については、神経精神科から提出された Excel データは、z-score のみを含むことが判明した。そこで、データセンターにて、BACS-J 評価表に基づき確認した素点を用い、各項目の合計点を再計算した。提出されたデータの z-score についてのみ、データセンター再計算値との比較を行った。BACS-J は、1 回の検査あたり、composite score を除き、7 項目の評価をする。0 週の症例数が 23 例、8 週の症例数が 20 例、26 週の症例数が 10 例(BACS-J 拒否が 1 症例)、52 週の症例数が 9 例(BACS-J 拒否が 1 症例)であったので、総評価項目数は 434 であった。この中で、データセンター再計算値との間に差異を認めたのは、60 箇所であり、総評価項目数の 13.8%であった。

DIEPSS では、0 週の症例数が 22 例(1 例、0 週の検査結果が無い)、8 週の症例数が 20 例、26 週の症例数が 11 例、52 週の症例数が 10 例だったので、総評価項目数は 63 であった。この中で、データセンター再計算値との間に差異を認めたのは、4 箇所であり、総評価項目数の 6.3%であった。

臨床検査データに関しては、神経精神科から臨床データの原資料が提出されていないため、データセンターでは検証作業を行えなかった。

神経精神科 Excel データとデータセンター再計算値との間に、平均すると、11.1%の差異を認めた。BACS-J の z-score 算出については、論文 1 の検証でも述べたように、異なる 2 種類の基準値を用いたことにより、データセンター再計算値との差異が 13.8%と、他の検査の不一致率より高率であった。これらの差異は、不規則に起こった素点の得点の見間違えや素点の合計の計算ミスなど、ヒューマンエラーによるものと推測した。

PANSS の Total を除いて、論文の値と再計算の結果がほぼ一致した。Total の値の不一致の理由は、上記のように、不規則に起こった、素点の得点の見間違えや素点の合計の計算ミスなど、ヒューマンエラーによるものと推測した。一方で、他の値についてはほぼ一致していることから、患者群 23 例中、8 週目の検査未実施例（3 例）、26 週目の検査未実施例（13 例）、52 週目の検査未実施例（14 例）と判断した症例に

ついては、LOCF 法を適用して欠測データを補完し、Table 1 を作成したものと判明した。なお、上記 IV で示したように、1 例においては、提出されたデータだけでは 52 週の欠測データ補完方法を特定できなかった。

次に第 2 段階の比較検証として、論文記載値と 52 週目まで、予定された全ての検査を完了した 9 例の各検査結果を比較した。その結果、予想されたことではあるが、全ての平均値及び標準偏差値において、論文記載値と再計算値は一致しなかった。加えて、t 検定のボンフェローニ補正後の p 値についても、論文記載値とデータセンター再計算値は一致しなかった。

統計解析結果について、Result 内の記述 "After Bonferroni correction, significant improvements from baseline to end-point were observed for the self-control, physical functioning, and social integration subscale scores on the SWNS-J, and the psychosocial condition subscale score on the SQLS-J (all $P < 0.05$)"について、p 値は揃わなかった。52 週目の SWNS-J の self-control 、physical functioning 及び 52 週目の SQLS-J の psychosocial condition score について、n = 9 の解析では $p > 0.05$ となり、論文の記述と一致しなかった。

同段落内の記述 "Furthermore, significant improvements were found for the total score on the PANSS and for the composite score on the BACS-J (both $P < 0.05$)." について、52 週目の PANSS の total score では $p < 0.001$ で一致したが、一方、52 週目の BACS-J の composite score は論文記載値 $p = 0.034$ であるのに対し、n = 9 では $p = 0.194$ であり、$p > 0.05$ となることから、論文記述と一致しなかった。

同段落内の記述 "After Bonferroni correction, significant improvements were observed for all factor scores on the PANSS and for the letter fluency and the Tower of London scores on the BACS-J (all $P < 0.05$)." について、52 週目の PANSS については、n = 9 では、p 値の不一致はあるものの、ほぼ $p < 0.05$ であることが確認されたが、8 週目の PANSS の cognitive についてのみ $p > 0.05$ となり、論文記述と一致しなかった。52 週目の BACS-J の letter fluency と Tower of London についても同様に、$p > 0.05$ となり論文記述と一致しなかった。

本論文中の記載"Similar results of the clinical assessments were obtained through a completer analysis performed with the 11 patients について検証するため、52 週目の検査データに注目し、データセンターでの n=9 の統計解析結果と、論文記載の統計解析結果を比較した。上記のように、両解析結果に一致を認めたのは、SWNS-J では、6 評価項目中 4 項目、SQLS-J では、3 評価項目中 2 項目、PANSS では、6 評価項目中 6 項目、BACS-J では、8 評価項目中 5 項目であった。総合すると、一致したのは、全 23 評価項目中 17 項目（17/23,73.9%）であった。従って、約 2/3 の評価項目で、completer analysis と n=23（あるいは n=22）の統計解析結果が一致したことから、本論文では、" Similar results of the clinical assessments were

obtained through a completer analysis performed with the 11 patients と記載したものであると推測した。

②論文 2 Table 2 について論文記載値及びデータセンター再計算値の比較

論文記載値及びデータセンター再計算値の比較を示す。また比較検証は Table 1 と同様に、論文記載値に関し、2 段階の比較検証を行った。

上記３）－①に記載した方法で、第 1 段階の比較検証を行った。平均と標準偏差については、ほぼ値が一致した。一方で、p 値については値が一致しなかった。例えば、8 週目の DIEPSS の p 値は、データセンター再計算値 0.111、一方、論文記載値 0.037 (0.037 は、0.111 の 1/3 の値)、26 週目の Low-density lipoprotein の p 値は、データセンター再計算値 0.953、一方、論文記載値 0.853、52 週目の bodyweight の p 値は、データセンター再計算値 0.160、一方、論文記載値 0.016、などであり、p 値を計算した全 30 箇所のうち、8 箇所(26.6%)で、誤りが判明した。これらの誤りは、転記ミスや t 検定のボンフェローニ補正後の p 値の計算ミスなど、ヒューマンエラーによるものと推測した。

次に第 2 段階の比較検証として、論文記載値と 52 週目まで予定された全ての検査を完了した 9 例の各検査結果を比較した。その結果、Table 1 での比較検証同様、全ての平均値及び標準偏差値において、論文記載値と再計算値は一致しなかった。加えて、t 検定のボンフェローニ補正後の p 値についても、論文記載値とデータセンター再計算値は一致しなかった。

論文中の記載 “Results from the safety assessments demonstrated that the values of total cholesterol, body weight, and BMI were transiently increased (all $P < 0.05$)”について、p 値の値は揃わなかった。

同段落内の記述 “The DIEPSS total score was significantly increased at 8 weeks ($P < 0.05$), but no significant difference was evident between baseline and end-point.”について、8 週目における p 値は $p > 0.05$ となり論文記述と一致しなかった。

論文中の記載”In a completer analysis, the values of all items on the safety assessments remained within the normal range, and ANOVA indicates no significant changes in each value or in the DIEPSS total score during 1 year について検証するため、データセンターでの n=9 の統計解析結果と、論文記載の統計解析結果を比較した。上記のように、両解析結果に一致を認めなかったのは、8 週目の total cholesterol、DIEPSS、26 週目の BMI、52 週目の bodyweight であり、全検定項目数 30 のうち、4 箇所(4/30、13.3%)であった。加えて、本文中の ”the values of all items on the safety assessments remained within the normal range”という記載は、正しい記載とは言えないと判断した。

③論文 2 Table3 について論文記載値及びデータセンター再計算値の比較

神経精神科から提出された論文 2 の Table 3 作成用のデータセットには、症例の入力順番が他のデータセットと異なることが判明した。そこで、まず、これらのデータセット間で、入力順番を一致させる作業を行った後、当初の目的であった、Table 3 の検証を行った。

以上のことを踏まえて、PANSS と SWNS-J についてピアソンの相関係数の検定を行った。上記で述べた神経精神科が使用した SWNS-J の 9 名のデータが、順番に入力されたデータ（以下データ 1 と記述）とデータセンターで再計算したデータ（以下データ 2 と記述）について、ピアソンの相関係数の検定を行い、論文記載値との比較を行った。

まず PANSS の Negative と SWNS-J の比較検証を行った。その結果、上記のデータ 1 における検定結果が論文記載値と一致した。次に、PANSS の他の 4 項目（Excitement、Cognitive、Positive、Depression/Anxiety） に対しても同様に検定し、論文記載値との比較を行ったところ、データ 1 に関する検定結果が論文記載値と一致した。

さらに BACS-J、SQLS-J に対して PANSS とのピアソンの相関係数の検定を行い、データの検証を行った結果、上記の SWNS-J 同様、データ 1 に関する検定結果が論文記載値と一致した。

4）承認実施計画書記載内容、実際に実施した臨床試験内容と、論文記載内容の整合性に関する検証

① 論文では、All participants gave informed consent after the study procedures had been fully explained と記載されているが、本臨床試験の試験期間延長に対し、論文記載の ITT 解析をしたとする全例 23 例、その中で試験を完遂したとする 11 例ともに試験期間延長（26 週， 52 週での検査）の再同意が得られていないため、不適切な記載である。

5）その他の不適切記載

① 論文では、This work was supported in part by Health and Labour Sciences Research Grants, Comprehensive Research on Disability, Health and Welfare, and grants from the Japan Society for Promotion of Sciences と記載されているが、本研究が、厚生労働科学研究費、障害保健福祉総合研究および日本学術振興会科学研究費の助成を得て実施された事実は確認できなかった。従って、適切な記載とは言えない。

なお、このことについてのヒアリング結果は、論文 1 の 4）その他の不適切記載で述べた内容と同様である。

6）論文 2 検証結果の総括

① 論文 2 の基になった当該臨床試験の実施に際して、承認された実施計画書からの逸脱が明らかになった。論文 2 では、文書同意取得がなされていない症例、エントリー基準違反の症例、データが欠測した症例など、多数の不適格症例を包含したまま、臨床データを取りまとめ、統計学的解析を実施した点において、既に、倫理的適格性・科学的適格性が損なわれている。そこで、当論文は論文 1 と同様、生命倫理委員会の審議結果が妥当であり、当該論文は撤回すべきと考える。

② 論文 2 記載値とデータセンター再計算値に、差異を認めた。特に、Table 2 では、p 値計算箇所 30 のうち、実に 8 箇所(26.6%)で、論文記載値に誤りがあることが判明した。これらの差異は、意図的な操作によって生じたものではなく、不規則に起こった、素点の得点の見間違えや素点の合計の計算ミス、転記ミスや t 検定のボンフェローニ補正後の p 値の計算ミスなど、ヒューマンエラーによるものと推測した。しかしながら、これらの誤りは、注意深い検査データの取扱い方、あるいはデータ入力後の多重チェック（ダブルチェック、トリプルチェック）により、データ解析のいずれかの段階で、容易に気づかれて然るべき問題である。

Table 3 の検証の項で指摘したように、症例の入力順番が異なるデータセットが複数存在したことから、多くの臨床データを一元管理していたとは言えず、基礎データの取扱い方に不適切な点があったことは、否めない。データ管理面での責任の所在が不明確であったとことが窺える。

本研究における臨床試験継続症例数について、実際の数とは異なる症例数が、論文に記載されていた。症例毎の臨床試験経過を丹念に追跡すれば、データ解析の過程、あるいは研究結果を論文化する過程で、容易に気づいて然るべき問題である。従って、本研究の責任者および分担研究者が、臨床データの取扱いにおいて、極めて杜撰であったと言わざるを得ない。

③ 本研究の承認実施計画書には、「解析対象集団は、逸脱例を除いた対象とする」ことが明記されている。従って、そもそも、逸脱症例を統計学的解析に含めることは、認められない。

それを踏まえた上で、論文 2 では、欠測データの取扱い方、補完方法に関する記載がない。データセンターでの検証過程で、LOCF 法を適用し、欠測データを補完したことが判明した。本研究では、薬効解析（Table 1 および 3）のみならず、安全性評価(Table 2)も重要な課題であり、欠測データ率が約 30%と高率であったことを合わせ考えると、LOCF 法を本研究の主要解析に適用することには、慎重であって然るべきであった。特に、3 症例では、0 週(服薬前)の検査データを 8 週、26 週、52 週の補完に使用したことが判明した。一度も服薬していない時期のデータを、薬効解析・安全性評価に使用したことは、極めて不適切であったと言わざるを得ない。本研究における LOCF 法の適用限界を含め、LOCF 法に関する理解が、甚だ不十分であったと判断せ

ざるを得ない。

④　以上から、基礎データの取扱法、管理方法、データ解析方法において、いずれも不適切な点が認められたものの、論文記載の Table 1、2 および 3 は、いずれも、神経精神科から提出された基礎データに基づき、論文記載の統計解析方法により、導き出されたものであることを確認した。従って、論文 2 の作成過程で、データに改竄、捏造、盗用のないことを確認した。

第４．精神保健指定医資格を取り消された神経精神科医師が関わった他の臨床試験の検証

１．神経精神科関連臨床試験に関する調査

指定医取消しとなった医師で、現在当大学病院に勤務しているものが実施責任者をつとめている臨床試験について調査した（当該臨床試験を除く）。

（１）調査対象

終了：８試験

実施中：11 試験

登録前：2 試験

（２）比較試験

13 試験（代表施設：11 施設、分担施設：2 施設）

（３）調査結果

①すべての患者において同意書は保管されていた。

②17 試験において、ケースレポートフォームが作成されていない：データベースの欠如（４試験以外）。

③当大学病院が代表施設として行った 11 比較試験における割付票（登録票）の保存がない。

④登録前の 2 試験は現時点では問題は認められていない。

⑤実施中の 11 試験のうち

- 比較試験 6 試験では、割付が行われている形跡が確認できず、実際に行われている治療に偏りが見られる試験も認められたため、生命倫理委員会による中止勧告要と判断した。
- 比較試験以外の 5 試験においては、年次報告書が生命倫理委員会へ提出されていない、予定登録数到達前に実施計画書に予定されていない統計解析を行ったという

問題点はあったが、その他には問題は認められなかった（3 試験）。

⑥終了した 8 試験のうち、予定登録数到達前に実施計画書に予定されていない統計解析を行い、学会、論文発表されているものを認めたが（2 試験）、その他には問題は認められなかった。

２．調査結果に基づく今後の対策

平成 20 年度に「臨床研究に関する倫理指針」が、平成 20 年 7 月 31 日付で改正され、平成 21 年 4 月 1 日から施行される旨通知するとともに、平成 20 年度に講習会を 5 回開催し、延べ 933 人が受講している（Ｐ8：「臨床試験の講習会開催状況」参照）。翌年度以降も大学以外の附属病院でも講習会を開催するなど定期的に開催しており、受講しない場合は臨床研究に関われないこととしていたことから、ある程度の教育は行われていたと考えられるが、それが神経精神科内のすべての研究者には周知徹底されていなかった。今回問題となった臨床試験以外にも、6 つの研究で割り付けがなされていなかったことは、当時は臨床研究データセンターも発足しておらず、モニタリングもなされていなかったことから、あくまでも研究者からの年次報告のみであり、組織として臨床試験に対するチェック体制に問題があったものと考える。

そこで、具体的な対策として考えられるのは以下のとおりである。

1）今後、新たに開始される神経精神科関連臨床比較試験は、割付票（登録票）の管理を含めて臨床研究データセンターがデータ保存をチェック、もしくは請け負うよう、臨床研究データセンターに要請する。

2）今後、新たに開始される神経精神科関連臨床試験は、すべて臨床研究データセンターがデータマネージメント、モニタリングを行うよう、臨床研究データセンターに要請する。

3）実施中の比較試験 6 試験では、上記の理由で中止勧告を行うことを生命倫理委員会に要請する。

4）現在進行中で、かつ生命倫理委員会からの中止勧告がでなかったもの 5 試験については、今後は臨床研究データセンターが関与し、データマネージメント、モニタリングを行うことを条件に継続を審議することを生命倫理委員会に要請する。

5）今後、新たに生命倫理委員会へ提出される神経精神科関連臨床試験については、実施可能性、統計解析の妥当性、データマネージメント、モニタリング、監査の必要性などにつき、より詳細に審査を行うことを要請する。

6）今回は神経精神科医師が関わった臨床試験の検証を行ったが、今後は他の診療科において実施されている臨床試験についても全件において、まずは自主点検という形で全件調査を行うべきであると考える。

3．研究の適正な実施のための組織体制と管理者の責任

平成 26 年度までは、臨床試験の年次報告や終了・中止報告は行われていたが、あくまでも自主申告であり、組織としてのチェック体制に問題があったと考える。

また、モニタリングなどの研究の適正な実施に関する対応がとられていなかったことから、研究者が適正な理解の元、研究を遂行することを想定した運用であった。平成 27 年度からは、「人を対象とする医学系研究に関する倫理指針」に沿って研究データの管理とモニタリングを実施していることから、研究不正や誤った研究遂行への適正な管理が進んでいる。

なお、本件における管理者それぞれの責任等については以下のとおりである。

1）学長の責任

臨床研究機関の長の責務として、これらの問題が起こることを想定した形での臨床研究の自己点検が行われておらず、評価が不十分であったことについての責任がある。

2）病院長の責任

臨床試験を管轄している部署が病院ではなく、大学に所属していることから、臨床試験の実施については、病院長の責任を問うことはできないが、臨床試験は患者さんを対象としているものであり、病院として関与するという責任と必要性があったのではないかと考える。

3）神経精神科部長の責任

部下の臨床試験に常に目を光らせているという、いわば当然の管理監督責任を果たしていなかったことについて重大な責任がある。

第５．再発防止策の検討

まず再発防止策を検討するに当たり、今回の事案が起きた責任について、関係者毎に整理した結果は以下のとおりである。

1）病院長及び医療安全管理室室長

初期段階から対応窓口が病院の組織から大学の組織へ移管したことから、病院長並

びに医療安全管理室室長の積極的な関与が見られなかった。しかしながら、医療安全管理室からの報告が現状１ヶ月分をまとめた形で受けているのは、重大な問題を見落とすことも考えられ、事案によっては通常の報告とは別に、単独で報告する必要があると考える。

２）学長

Ｂ医師より不適切な臨床試験が行われていたことについての報告があるまでは、精神保健指定医の問題に端を発した神経精神科医への根拠のないクレームに近いとの認識であったことから、Ａ 氏からの要望事項への個々の対処方法について指示をしており、大きな問題があるとは気づいていなかった。しかしながら、臨床試験においては学長が最高責任者であり、その研究が適切に実施されているかどうかのチェック体制を整えていなかったことについては、責任を免れる事はできない。

３）神経精神科部長

本来であれば研究者の所属する部署の長として、最後まで責任を持った対応をするべきであったが、昨年発覚した精神保健指定医の問題で、８月７日付けで諭旨退職となっていた。

４）Ｂ医師

研究責任者として、実施計画書から逸脱した研究を行い、また被験者に対する対応に誠実さが欠けていたことが重大な問題であったと考える。

５）医療安全管理室担当者並びに大学院・研究推進課担当者

虚偽と感じていたことをそのまま被験者に説明したことについては、許されることではなく、Ｂ医師がどのように言ったにしても、事実を説明すべきであった。また、Ｂ医師と同様に被験者に対する対応に誠実さが欠けていたことが重大な問題であったと考える。

なお、関係者に対する処分については、大学において人事給与委員会がその任をまかされており、調査報告書に基づいて学内規程に則って審議がなされるべきであると考える。

１．臨床試験における再発防止策

今回問題となったのは、臨床試験が実施計画書どおりに実施されていなかったことであるが、臨床試験において不正な研究を防止するための再発防止策として、以下の対応が必要と思われる。

（1）臨床試験に関する学内教育の徹底

学内教育については教育機会の確保だけでは不十分であり義務化が必要である。そこで、臨床試験の講習会の受講を義務化する体制として、現在では、臨床試験の申請を行う際に受講番号を記載する方式をとっている。

また、平成 27 年度からは年 1 回以上臨床試験の講習会受講を義務付けており、臨床試験に申請する際（研究分担者も含む）には、受講済みであることを求めている。

その他、e-learning での教育方法として CITIJapan の受講を義務化しており、平成 27 年 7 月から 28 年 10 月 21 日現在で「01. 責任ある研究行為：基盤編を 3,075 名、02. 人を対象とした研究：基盤編 3,058 名」の研究者および研究補助者（事務担当者を含む）が受講しており、一定の教育が行われているが、今後さらなる学内教育の徹底が望まれる。

なお、平成 20 年度以降の講習会開催状況については以下のとおりである。

臨床試験の講習会開催状況

年度	講習会回数	受講者数（延べ人数）	備考
平成 20 年度	5 回	933 人	
平成 21 年度	5 回	447 人	附属病院での開催含む
平成 22 年度	5 回	372 人	附属病院での開催含む
平成 23 年度	6 回	346 人	附属病院での開催含む
平成 24 年度	6 回	371 人	附属病院での開催含む
平成 25 年度	6 回	678 人	附属病院での開催含む
平成 26 年度	7 回	715 人	附属病院での開催含む
平成 27 年度	30 回	931 人	附属病院での開催含む （ビデオ講習含む）

（2）臨床試験を適正に遂行するための学内での体制整備

1）平成 27 年 4 月から学内でデータセンターの運用を開始した。学内研究者からは料金を取らずに運用中である。データセンターのメンバーとしては、センター長・副センター長に臨床研究に精通している研究者の配置および生物統計に詳しい者やデータマネージャーを配置している。

2）臨床試験部会で、「侵襲介入試験」は、データセンター利用を条件としている。臨床試験部会にはデータセンター職員が、予備審査会から本会議まで、オブザーバーとして参加しており、どのような案件が進行中であるか把握している。

３）データセンター職員は、研究者が独自に設置するモニタリングの手法とは別に、臨床研究の進捗を定期的にチェックするように指示されており、問題のありそうな研究についてはデータセンター長経由で学長まで報告が上がる経路が作られている。実務的には、データセンター運営会議および臨床試験部会長に情報が上がり、その精査を部外者であるモニター担当者（現在、他大学所属の CRC・臨床薬理学講座助教）に依頼する形を作っている。（研究に問題がありそうだと思われる場合は、本学データセンター職員が協力して確認を行う）その位置づけは、米国式のモニターではなく、欧州式の「部外者による早期の監査」に近いものとしており、その内容は学長まで報告が上がる事になっている。また、そのような形で第三者が監視する形式になっていることを研究者に周知し、抑止力として利用する事としている。
（現在すでに同モニター担当者は本学の複数の試験をモニター中。）

４）モニタリングの結果、問題がありそうな試験が確認されたときは、データセンターのメンバーや臨床試験部会の部会員が一旦研究を止めたうえで、詳細な調査を行い、結果を学長まで報告する。なお、研究の中断が可能な場合は中断し、中断が困難な場合は研究不正が明らかな場合は中止とし、疑わしい場合は中断可能な段階で研究を止めることで対応する。

５）以上、上記で述べた体制については、現実の運用面において、実効あるものとするための周知徹底が望まれる。

（３）利益相反の明確化、職員倫理の徹底及び公益通報窓口の再確認

現時点において、利益相反の報告については自主申告となっており、研究者に対する周知が徹底されていない。今後は利益相反の対象とならない研究者についても、全件報告を義務付ける等必ず報告を行うような体制を確立することが必要であると考える。また、データセンターの整備、病院と大学の連携強化のみで、カルテの改ざん、被験者への虚偽の説明など不幸な事態に至った今回の様な事案の発生を防げるわけではない。病院の公共性を重んじ、法令及び諸規則を遵守するという職員倫理の徹底、電子カルテの修正や改ざんに関する職員教育、公益通報窓口の再確認を含む危機管理体制の整備を行う必要がある。

（４）大学院生指導教員への心理的圧力に関する調査

研究責任者からの大学院生の学位論文を作成するために研究不正を行ったとの証言では、大学院生を指導する教員は４年で学位を取得させるためのプレッシャーを感じながら指導を行っていることが憂慮されることから、指導教員に対してアンケート調査を

施行し、実態の把握を行うことが必要と考える。

２．被験者（患者）に対する対応について

（１）病院と大学との連携強化

臨床試験参加者からの試験参加中止希望、データ削除希望、情報開示等、臨床試験に係る被験者からの申し出に対して通常は臨床試験責任医師、その医師が所属する部門長により適切に対応されるものであろうと認識し、臨床試験責任医師もその上司も適切な対応を取らない事態は想定されていなかった。今回の事案を教訓として、患者相談窓口でとらえられた臨床試験被験者からの申し出については、臨床試験責任医師、部門長のみに連絡して対応を求めるのではなく、臨床試験を審査、監督する臨床試験部会にも同時に情報が届けられる体制に改める必要がある。

すでに大学病院では、平成 28 年度から臨床倫理委員会を設け、委員会の主務の一つに、患者相談窓口で確認された患者及びその家族、関係者が抱く倫理的課題に関する即応が加えられており、臨床試験被験者の個別具体的な臨床倫理問題について対応する体制を整備している。

第６．結び

今回の件は精神保健指定医資格不正取得に端を発した案件であるが、臨床試験が実施計画書どおりに適切に行われていなかったことが最大の問題である。臨床試験の責任医師である研究者の臨床試験に対する認識が甘く、単に診療の延長線として捉えていたのではないかとも思われ、今後は再発防止策で述べたように、臨床試験に関わる研究者すべてに対して、教育研修を徹底するとともに、臨床試験を監査する体制を充実させ、「人を対象とする医学系研究に関する倫理指針」に則って臨床研究を進めることを徹底すべきである。

また、被験者からの問い合わせに対して当初から事実通りの正確な説明とともに、資料を開示していればこれほど大きな問題に発展することはなかったと思われ、問い合わせへの対応については複数で確認し、適切な対応を行う体制を整えるよう改善すべきである。