

[Original Article]

研究者主導臨床試験における ローカルデータマネジメントの質と CRC 有無・ 登録症例数の関連



Association between Quality of Local Data Management and CRC Support/ the Number of Registered Cases of an Investigator-initiated Clinical Trial



村山季実子^{*1} 相原 智彦^{*2} 向井 博文^{*3} 大橋 靖雄^{*4}

ABSTRACT

Background Some problems may occur while collecting data during clinical trials. For example, case report forms (CRF) may not be submitted or data errors are present in the submitted CRF. These problems adversely affect the reliability of the results, cost of the trial, and the time required for publication. However, these problems should decrease by sufficient local data management. A few detailed studies have analyzed the process from data collection to solving problems. However, the situations in which problems tend to occur are unclear, and the local data management is insufficient.

Methods We investigated problems that occurred in a past investigator-initiated clinical trial. Percentage of CRF deficiency, incidence rate of data errors, and the number of days required for participating physicians to answer queries during the trial were calculated.

We focused on the involvement of a clinical research coordinator (CRC) and the number of patients registered per physician as factors influencing these problems; we explored the association between frequency of problems and these factors.

Results Although the trend was not monotonic, physicians who did not submit CRFs, who had many data errors, and who took a long time to answer queries tended to be those who had not received the CRC support or those who had a small number of patients registered.

Conclusion CRC support and the number of registered cases were associated with the response time required and the percentage loss of CRFs near the maximum value. Data errors tended to decrease if CRC support was provided or the number of registered cases was large. (Jpn Pharmacol Ther 2013 ; 41 suppl 2 : S103-10)

KEY WORDS clinical data management, local data management, data quality, investigator-initiated clinical trials, clinical research coordinator

^{*1}東京大学大学院医学系研究科生物統計学分野 (現 イービーエス株式会社臨床情報本部) ^{*2}相原病院乳腺科 ^{*3}国立がん研究センター東病院乳腺科・血液化学療法科 ^{*4}東京大学大学院医学系研究科生物統計学分野

Kimiko Murayama and Yasuo Ohashi : Department of Biostatistics, Graduate School of Medicine, The University of Tokyo, Japan ; Tomohiko Aihara : Breast Center, Aihara Hospital, Osaka, Japan ; Hirofumi Mukai : Division of Oncology/Hematology, National Cancer Center Hospital East, Chiba, Japan

はじめに

臨床研究のデータ収集時には、統計解析を実施する前になんらかの検討や対処が必要となる問題が発生することが多い。

問題の具体的な例として、症例報告書 (case report forms : CRF) の欠損があげられる。たとえば、長期的な追跡を行っている臨床試験において、予定された来院時期 (visit) で患者が来院せず、画像検査や臨床検査などのデータが得られなかったために、施設側が CRF の提出を見送ることがある。ほかに、患者は予定どおりに来院していても、担当医師が多忙であるなどの理由で CRF の提出が滞ることもある。

上記のような CRF の欠損の問題に加えて、回収された CRF も、そのままの内容を入力してデータセットとすることには多くの場合、問題がある。診断や測定、データの転記、データベースへの入力など、データ収集の各段階で、エラーが混入することは避けられない。こうしたエラーの例として、記入すべき項目の欠損は頻繁に発生している。定義外の文字が書かれていたり、文字が判読困難であったりする場合も、有効なデータとして利用できない。また、同一の CRF 内や異なる visit の CRF 間でデータの不整合や矛盾が生じているときも、どのデータが正しいのか判断しかねる。このように、調査票に記入されたデータにも、解釈を行う際になんらかの判断や再調査、修正を必要とするもの (データエラー) がほぼ常に含まれる。これには確認の結果、誤りではないことが判明するものも含まれるため、正確には「エラー候補」であるが、本論文では簡便のため「エラー」とよぶ。

欠損した CRF を無視し、データエラーを検討しないまま解析を行った場合、結果にバイアスが入る、あるいはバラツキが大きくなり群間差を検出することが困難となる可能性がある^{1,2)}。このため試験実施時には、未提出の CRF に対する督促や、データエラーに対する読み替え・施設問い合わせなどを行い、可能な限り多くの正しいデータを得るよう努める。臨床研究の分野では、バイアスおよびバラツキの少ない結論を導くためにデータエラーの頻度を低くコントロールする一連の活動をデータマネジメン

トと呼称している^{3,4)}。

データマネジメントは施設側で実施されるローカルデータマネジメントと、データ集積後に中央で行われるセントラルデータマネジメントに分類することができる。適切な時期の評価実施、正確な症例報告書の記載とタイムリーな提出、問い合わせへの正確な回答などのローカルデータマネジメントが十分に実施されていれば、データエラーの少ない CRF が滞りなく提出され、データ固定と解析を速やかに行うことができる。セントラルデータマネジメントによっても同等の質のデータセットを作成することも不可能ではないが、CRF 欠損やデータエラーが多いほど督促や修正の作業は増加し、施設側からの回答を待ってデータを固定するため、解析の実施が遅れてしまう。このように、CRF 欠損やデータエラーなどの問題は、解析結果の信頼性だけでなく、エラー修正に伴う人件費など、試験実施のコストおよび臨床試験の結果の公表の遅延にもつながる問題であると考えられる。

特に実施時の予算・スタッフ数が限られている研究者主導臨床試験において、信頼性の高い結果が迅速に公表されるためには、ローカルデータマネジメントの質を向上させることが重要であると考えられる。ローカルデータマネジメントの質に影響を与える要因としては、医師個人の性格や注意力、臨床試験参加の経験や意欲、試験実施時に医師をアシストする環境、CRF の項目数やレイアウトなどが考えられる。がんの第 I 相・第 II 相臨床試験においては、試験参加医師の累積の CRF 記入項目数と、CRF 記入上の問題の発生率とのあいだに単調な比例関係はみられず、これらの問題発生率に影響を与える要因のさらなる検討が求められている⁵⁾。また、国外の臨床研究においては、カルテなどの医療記録から CRF への転記の正確度は、業務補助をする者の存在の有無によらないという結果も報告されている⁶⁾ものの、日本国内で実施された研究者主導臨床研究においても同様の結果であることは確認されていない。

そこで本研究では、過去に実施された研究者主導臨床試験においてデータ収集時に発生した問題について調査し、ローカルデータマネジメントの質が試験実施時の状況とどのように関連していたかを探索した。

対象と方法

すでにデータ固定と統計解析が完了しており、かつ研究開始からデータ固定までのデータマネジメントの記録がデータセンターで管理されている臨床研究である National Surgical Adjuvant Study of Breast Cancer (N-SAS BC) 03 を本調査の対象とした。

N-SAS BC03 は日本人女性を対象とした、原発性乳がんの術後補助療法に関する多施設共同の研究者主導ランダム化比較試験である。被験者の登録は2002年11月13日から2005年12月28日まで行われ、全706名が試験に組み入れられた⁷⁾。参加施設は全71施設、担当医師は全141名であり、このうち46名はCRC (clinical research coordinator) が臨床試験業務を補助していた。経過観察は6カ月ごとに行われ、紙のCRFの提出が求められた。各施設・担当医師の登録症例数と中間解析までの合計visit数の概要を表1に示す。

N-SAS BC03 を実施した乳がん術後補助療法研究委員会は、1993年の発足以降、乳がん患者の生活の質 (Quality of Life : QOL) 向上の実現を目指し、さまざまな臨床試験を実施・計画している。現在、N-SAS BC07 まで実施されている一連の試験でデータセンターを務めるNPO 法人日本臨床研究支援ユニットは、研究コーディネーション、データマネジメント、統計解析を中心に、多様な臨床試験支援業務を行っている組織である⁸⁾。

本研究では当該試験のCRFおよび施設間問い合わせ記録を遡って調査し、研究開始から中間解析時のデータ固定までに発生したCRF欠損およびデータエラーの発生件数を調査した。加えて、データエラーの修正・確認のために行われた施設間問い合わせについて、データセンターが問い合わせ (クエリ) を送付し、回答を受領するまでの日数を調査した。

また、ローカルデータマネジメントの質との関連を探索する要素として、(1) CRC による業務補助の有無と、(2) 担当医師の登録症例数に注目した。(1) は、先行研究で報告された傾向が国内においても同様であるかを確認するため、(2) は臨床研究参加の経験などを反映すると考えられ、客観的に観察できる要因として、それぞれ選択した。

表1 登録症例数、総 visit 数の分布

	中央値	(四分位範囲)	最小値	最大値
1 施設あたり				
登録症例数	5	(2~11)	1	70
総 visit 数	59	(26~130)	2	756
医師 1 人あたり				
登録症例数	4	(2~8)	1	68
総 visit 数	22	(11~51)	1	756

検討項目

CRF の記載は基本的に医師によって行われることを考慮し、以下の指標を設定した。

1 試験参加医師における CRF 欠損の発生状況

被験者死亡と同意撤回による追跡中止を除いて、以下のCRFをCRF欠損と定義した。

- ・評価予定日から6カ月を経過し、次の評価時期に達しても未提出であるCRF
- ・施設間問い合わせの回答待ちで解析データセットに入力されていないCRF

提出予定CRF数に対してCRF欠損が多いほど、ローカルデータマネジメントの質が悪く考えられる。そこで以下の式を、各医師 (医師個人を添え字_iで表す) での平均的なCRF欠損割合とみなした。

CRF 欠損割合_i (%) = $100 \times (\text{医師}_i \text{の CRF 欠損の総数}) / (\text{医師}_i \text{の提出予定 CRF の総数})$

さらに、CRF 欠損割合_iとCRC関与の有無、登録症例数との関連を検討した。

2 試験参加医師におけるデータエラーの発生状況

CRF の内容の変更履歴、および内容確認のために行った施設間問い合わせの履歴をもとに、記入内容の変更もしくは内容確認が行われていたものをデータエラーと定義した。

CRF 1 冊あたりのデータエラーが多いほど、ローカルデータマネジメントの質が悪く考えられる。そこで以下の式を、医師_iの平均的なデータエラー発生率とみなした。

データエラー発生率_i (/CRF) = $(\text{医師}_i \text{のデータエラーの総数}) / (\text{医師}_i \text{の提出した CRF の総数})$

さらにデータエラー発生率_iとCRC関与の有無、

登録症例数との関連を検討した。

3 試験参加医師における施設問い合わせの回答所要時間

施設問い合わせの履歴を参照し、問い合わせ書類の送付日と、回答書類受領日のあいだの日数を回答所要時間と定義した。

クエリに対して回答所要時間が長いほど、ローカルデータマネジメントの質が悪く考えられる。そこで、各医師に対して行われた問い合わせの回答所要時間の中央値を医師_iの平均回答所要時間_iとみなした。

さらに、平均回答所要時間_iと CRC 関与の有無、登録症例数との関連を検討した。

結 果

1 試験参加医師における CRF 欠損の発生状況

全体で、提出予定の CRF の 13.1% (1016/7752) が欠損していた。CRF 欠損の発生件数の詳細を表 2 に示す。CRC 関与の有無および登録症例数の多寡によって、CRF 欠損割合_iの中央値に大きな差はみられなかった。

CRF 欠損割合_iと登録症例数の分布を図 1 に示す。ひとつの点が CRF 欠損割合_iを表しており、シンボルの違いは CRC の有無を表す。散布図の上部に該当するほど、問題発生が多かったことを意味しているが、この部分に該当する医師には CRC の関与がないことが多く、CRF 欠損割合_iが 30%を超えた 24

名のうち、16 名で CRC 関与がなかった。また、CRF 欠損割合が高い医師は登録症例数が少ない傾向もみられたが、この部分の欠損割合は個人差が大きく、登録症例数と欠損割合のあいだに単調な関係はみられなかった。

2 試験参加医師におけるデータエラーの発生状況

提出された 6736 件の CRF において 8349 件のデータエラーが発生しており、平均すると CRF 1 件あたり 1.24 件 (8349 件/6736CRF) であった。データエラーの発生件数の詳細を表 3 に示す。CRC 関与のない場合でのデータエラー発生率_iは 1.87 件/CRF であり、CRC 関与がある場合の 0.85 件/CRF に比べて 2 倍以上となっていた。また、登録症例数別にみた場合、データエラー発生率_iの中央値は 1.52 件/CRF, 1.41 件/CRF, 1.19 件/CRF, 0.57 件/CRF と、登録症例数の増加に伴って減少する傾向がみられた。

データエラー発生率_iと登録症例数の分布を図 2 に示す。CRF 欠損の場合と同様に、散布図の上部に該当する、データエラーの発生が多かった医師は、CRC 関与がないことが多かった。また、データエラーが多く発生していた医師は登録症例数が少ない傾向もみられたが、この部分のデータエラー発生率は個人差が大きく、登録症例数とデータエラー発生率のあいだに単調な関係はみられなかった。

表 2 CRF 欠損発生件数の詳細

	該当医師数	提出予定 CRF 数	CRF 欠損総数	CRF 欠損割合 _i		
				中央値	最小値	最大値
全症例	141 (100%)	7752 (100%)	1016 (100%)	9.8	0.0	92.3
CRC 関与の有無						
なし	95 (67%)	4721 (61%)	707 (70%)	10.4	0.0	92.3
あり	46 (33%)	3031 (39%)	309 (30%)	9.3	0.0	50.0
担当医師の登録症例数						
5 人以下	88 (62%)	1580 (20%)	210 (21%)	9.2	0.0	92.3
6 人以上 10 人以下	31 (22%)	1487 (19%)	209 (21%)	10.4	0.0	57.7
11 人以上 20 人以下	13 (9%)	1671 (22%)	183 (18%)	9.8	1.9	31.8
21 人以上	9 (6%)	3014 (39%)	414 (41%)	11.7	1.6	54.3

3 試験参加医師における施設問い合わせの回答所要時間

データエラーの修正・確認のために行われた施設問い合わせは 4393 件であり、施設から回答が返ってきたものは 4279 件（うち、回答の受領日が不明であった問い合わせが 90 件）であった。未回答の施設問い合わせは 197 件あり、問い合わせの送付日から本調査開始時点までに 749～2891 日が経過していた。これらは回答所要時間の定義に適さないため、回答の受領日が不明であった問い合わせとともに回答所要時間の計算からは除いた。計算が可能

であった 4189 件の問い合わせにおいて、回答所要時間の中央値は 8 日であった。回答所要時間の詳細を表 4 に示す。CRC 関与の有無によって、回答所要時間の中央値に差はみられず、登録症例数と回答所要時間の中央値のあいだに単調な関係はみられなかった。

回答所要時間と登録症例数の分布を図 3 に示す。散布図の上部に該当する、問い合わせ回答に時間を要した医師は CRC 関与がないことが多く、回答所要時間が 30 日を超えた 26 名のうち、24 名で CRC 関与がなかった。また、回答が遅れた医師には登録

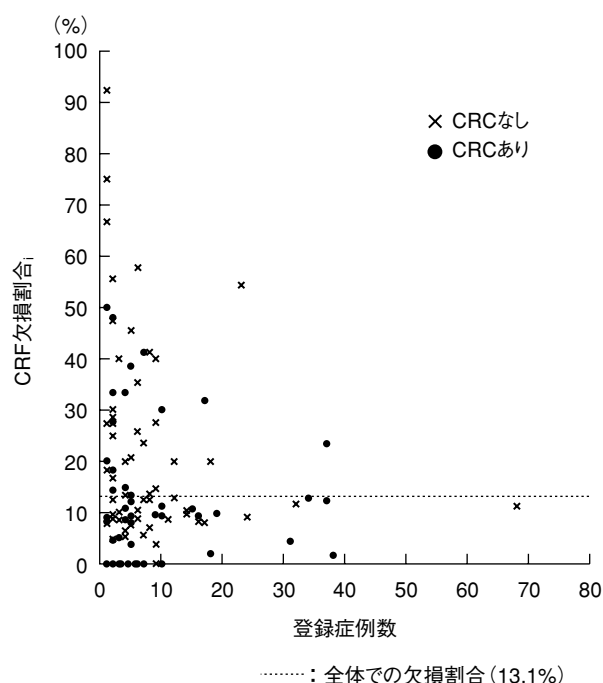


図 1 CRF 欠損割合_iと登録症例数の分布

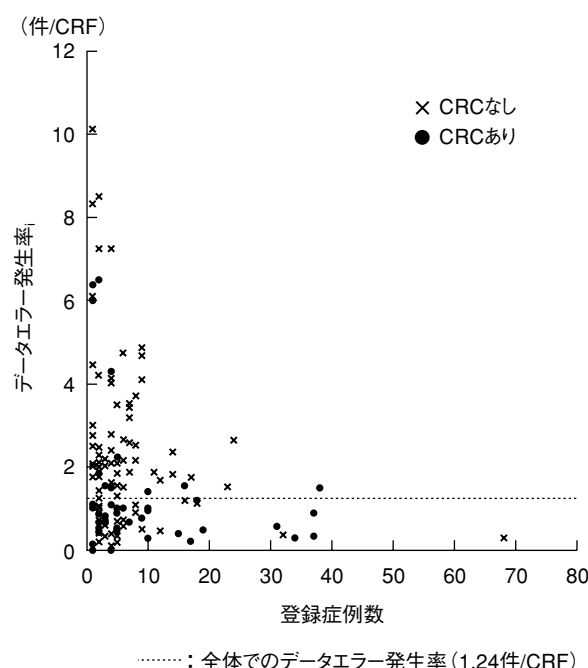


図 2 データエラー発生率_iと登録症例数の分布

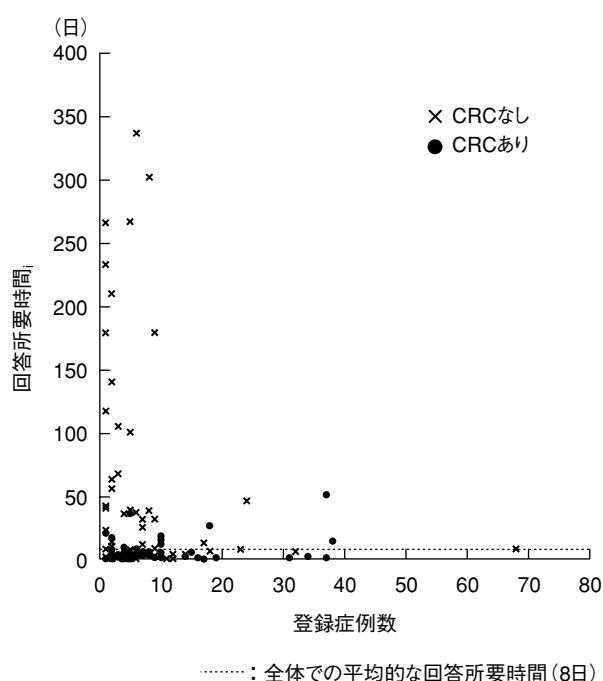
表 3 データエラー発生件数の詳細

	該当医師数	提出 CRF 総数	データエラー 総数	データエラー発生率 _i		
				中央値	最小値	最大値
全症例	141 (100%)	6736 (100%)	8349 (100%)	1.41	0.00	10.13
CRC 関与の有無						
なし	95 (67%)	4014 (60%)	6011 (72%)	1.87	0.00	10.13
あり	46 (33%)	2722 (40%)	2338 (28%)	0.85	0.00	6.50
担当医師の登録症例数						
5 人以下	88 (62%)	1371 (20%)	2474 (30%)	1.52	0.00	10.13
6 人以上 10 人以下	31 (22%)	1276 (19%)	2131 (26%)	1.41	0.28	4.86
11 人以上 20 人以下	13 (9%)	1488 (22%)	1842 (22%)	1.19	0.21	2.36
21 人以上	9 (6%)	2601 (39%)	1902 (23%)	0.57	0.29	2.64

表 4 回答所要時間の詳細

	該当医師数	問い合わせ* 総数	平均回答所要時間 _i		
			中央値	最小値	最大値
全症例	141 (100%)	4189 (100%)	8	0	1349
CRC 関与の有無					
なし	95 (67%)	3019 (72%)	8	0	1349
あり	46 (33%)	1170 (28%)	8	0	960
担当医師の登録症例数					
5 人以下	88 (62%)	1266 (30%)	6	0	1349
6 人以上 10 人以下	31 (22%)	1153 (28%)	15	0	1030
11 人以上 20 人以下	13 (9%)	793 (19%)	3	0	332
21 人以上	9 (6%)	977 (23%)	15	0	660

* 未回答, 受領日不明を除く

図 3 回答所要時間_iと登録症例数の分布

症例数が少ない傾向もみられたが, この部分の回答所要時間は個人差が大きく, 登録症例数と回答所要時間のあいだに単調な関係はみられなかった。

考 察

CRC 関与の有無および登録症例数の分布ごとに見たとき, CRF 欠損割合_i, 回答所要時間_iでは中央値に差がみられなかった一方で, 値の分布は異なっており, 高い CRF 欠損割合_iを示した医師, および

回答所要時間_iが長かった医師は CRC 関与がない傾向がみられた。CRC が行う試験実施スケジュール管理, データセンターからの問い合わせ対応などの業務⁹⁾によって, CRF の多くが欠損してしまうことや問い合わせ回答が長期化することが避けられている可能性が示唆された。

一方, データエラー発生率_iは, CRC 関与のある場合および登録症例数が多い場合は低くなる傾向がみられた。CRC は CRF の作成補助においても支援を行っており⁹⁾, これらの実施支援によってデータエラーの発生が抑えられている可能性が示唆された。

また, これらの問題が多くみられた医師は, 登録症例数が少ない傾向もみられた。臨床試験参加経験が豊富である医師や参加意欲の高い医師ではデータエラーの発生は少ないと考えられ, 登録症例数がこれらの経験や意欲を反映している可能性が考えられる。しかし登録症例数が少なければ提出予定 CRF 数や提出 CRF 数も少なくなるため, 数件の CRF 欠損やデータエラーが割合や率として大きく反映された可能性も考えられる。

CRF 欠損割合, エラー発生率, 回答所要時間に共通してみられる傾向として, 問題が多く生じると予想される状況 (CRC が関与しない, 登録症例が少ない) においても CRF 欠損などの問題の発生が低く抑えられている医師が多数存在しており, 特に CRF 欠損割合_i, 回答所要時間_iでは中央値に差がみられなかった。CRC 関与の有無, 登録症例数だけでは CRF 欠損やデータエラーの発生頻度, 問い合わせ回答の遅れを説明しきれないと考えられる。

そのため、これらの問題の減少を目的になんらかの介入を行う場合でも、一律に登録症例数や CRC 関与の有無などで標的集団を規定するのではなく、試験初期や過去の臨床試験で発生したデータエラーの情報を利用するほうが適切である可能性がある。

データエラーの特性は、疾患分野、対象とするイベント、プロトコルで規定された治療の複雑さなどに依存すると考えられ、今回対象とした臨床試験がデータ収集時に発生する諸問題を網羅し、代表性を持っていることは保証されない。そのため、今後、今回対象としたものとは性質の異なる臨床試験においてもデータエラーの評価が必要である。具体例として、減量や中止基準が複雑である化学療法の臨床試験や、イベント判定にデータエラーが多く含まれる循環器疾患や進行癌での臨床試験などがあげられる。また、ローカルデータマネジメントに影響を与える要因もさらに細分化していき、試験参加医師の臨床試験の経験数や、CRC の帰属や経験年数の詳細とデータエラーの関連も検討していくことが求められる。

本研究の解釈時には、データエラーの定義に注意する必要がある。本研究で対象とした試験は原則的に施設訪問を行っておらず、データセンターで受領した CRF のみからデータエラーを検出した。原資料閲覧を行ったうえで検出されたデータエラーではないため、本研究でのデータエラーの総数が、原資料と CRF での内容の不一致の件数と等しいことは保証されない。データエラーに関する研究では、こうした原資料と CRF の不一致をデータエラーと定義する^{10~14)}、あるいはダブルエントリー時の入力情報の不一致をデータエラーと定義する^{10,15,16)}場合もあるが、今回定義したデータエラーとは発生箇所や性質が異なるため、発生頻度を単純に比較することはできない。

また、本研究の限界として、データエラーと各要因との因果関係は明らかではない点があげられる。いくつかの要因ではデータエラーの発生とのあいだに一定の傾向がみられたが、介入を行ってこれらの要因を変化させた場合にデータエラーが減少することは確認されていない。

結 論

CRC 関与の有無と登録症例数は、CRF の欠損割合および回答所要時間の最大値付近の分布と関連がみられた。また、CRC 関与のある場合、登録症例数が多い場合において、CRF 記載時のデータエラー発生率は減少する傾向がみられた。CRC 関与の有無・登録症例数の増加と問題発生頻度のあいだの因果関係については、今回の調査のみでは明らかではない。

抄 録

背景 ほとんどの臨床試験では、データを収集する過程でなんらかの問題が発生する。たとえば症例報告書 (CRF) が提出されなかったり、回収された CRF に誤りが疑われるデータが存在していることなどがあげられる。

データ収集時の問題は、解析結果の信頼性、試験実施のコスト、結果の公表までの時間などに悪影響を及ぼすと考えられる。十分なローカルデータマネジメントの実施によって、これらの問題発生が減少することが期待される。現在、データが収集されてから固定されるまでの過程を対象とした研究は何件かみられるが、どのような状況で問題発生が多いのか、つまりローカルデータマネジメントの実施が不十分であるのかは明らかではない。

方法 本研究では、過去の研究者主導臨床試験においてデータ収集時に発生した問題を調査した。試験参加医師ごとに CRF 欠損割合、データエラー発生率、平均回答所要時間を求めた。試験実施時の要因は、今回は clinical research coordinator (CRC) の関与の有無と、その医師の登録症例数に注目し、問題発生頻度との関連を検討した。

結果 関係は単調ではないものの、CRF 欠損が多い、データエラーが多い、施設問い合わせに対する回答が遅い医師では、CRC が関与していなかったり、登録症例数が少ない傾向がみられた。

結論 CRC 関与の有無と登録症例数は、CRF の欠損割合および回答所要時間の最大値付近の分布と関連がみられた。また、CRC 関与のある場合、登録症例数が多い場合において、CRF 記載上のデータエラー発生率は減少する傾向がみられた。

【謝辞】 試験実施時の記録の提供にご協力くださるとともに、データマネジメントの実務に関してご指導くださいました野村由美子様、小倉千春様をはじめ、CSPOR データセンターの皆様にお礼を申し上げます。

文 献

- 1) Green S, Benedetti J, Crowley J, et al. 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めざして. 医学書院 2004 ; xviii : 234.
- 2) 新美三由紀, 福田治彦, 山本精一郎ほか. データセンターにおけるデータマネジメント. 癌と化学療法 1998 ; 25 (2) : 253-9.
- 3) 福田治彦. 研究者主導のがん多施設共同臨床試験におけるデータマネジメント. 日小児臨薬理会誌 2003 ; 16 (1) : 76-82.
- 4) 新美三由紀, 加幡晴美, 福田治彦. 公費臨床試験におけるデータマネジメントの実際. 癌臨床研究・生物統計研究会誌 2001 ; 21 (1) : 74-7.
- 5) 新美三由紀, 福田治彦, 奥山抄織ほか. がん臨床試験における CRF 記入上の問題の分析. 癌臨床研究・生物統計研究会誌 1998 ; 18 (1) : 62-7.
- 6) Vantongelen K, Rotmensz N, van der Schueren E. Quality control of validity of data collected in clinical trials. EORTC Study Group on Data Management (SGDM). Eur J Cancer Clin Oncol 1989 ; 25 (8) : 1241-7.
- 7) Aihara T, Takatsuka Y, Ohsumi S, et al. Phase III randomized adjuvant study of tamoxifen alone versus sequential tamoxifen and anastrozole in Japanese postmenopausal women with hormone-responsive breast cancer : N-SAS BC03 study. Breast Cancer Res Treat 2010 ; 121 (2) : 379-87.
- 8) NSAS-BC-CSPOR が支援している研究 (homepage on the Internet). NPO 法人日本臨床研究支援ユニット (cited 2013 Apr 30). <http://www.csp.or.jp/cspor/kenkyuu/nsasbc.html>
- 9) 矢田菜穂子, 石倉健司, 友常雅子ほか. CRC による臨床試験の支援とローカルデータマネジメントの質の向上. 日小児臨薬理会誌 2010 ; 23 (1) : 138-40.
- 10) Nahm ML, Pieper CF, Cunningham MM. Quantifying data quality for clinical trials using electronic data capture. PLoS One 2008 ; 3 (8) : e3049.
- 11) Strayhorn JM. Estimating the errors remaining in a data set : techniques for quality control. J Amer Stat 1990 ; 44 (1) : 14-9.
- 12) West M, Winkler RL. Data base error trapping and prediction. J Am Stat Assoc 1991 ; 86 : 987-96.
- 13) Mcintegart DJ, Jadhav SP, Broun T, et al. Checks of case record forms versus the database for efficacy variables when validation programs exist. Drug Inf J 1999 ; 33 (1) : 101-7.
- 14) Goldberg SI, Niemierko A, Turchin A. Analysis of data errors in clinical research databases. AMIA Annu Symp Proc 2008 ; 242-6.
- 15) Rostami R, Nahm ML, Pieper CF. What can we learn from a decade of database audits? The Duke Clinical Research Institute experience, 1997-2006. Clin Trials 2009 ; 6 (2) : 141-50.
- 16) Mullooly JP. The effects of data entry error : an analysis of partial verification. Comput Biomed Res 1990 ; 23 (3) : 259-67.

* * *