

3.4. CRF 記入の手引き

3.4.1. CRF 記入の手引き

CRF 記入の手引きとは CRF の書き方を示した記載マニュアルである。記入者が記入方法を悩んだり誤ったりしそうな箇所について予め記入方針を示すことで、記載内容に統一性をもたらし、CRF の品質を高めることができる。また CRF と記入の手引きを同時に作成することで、CRF の問題点を見つけることができる。

記入の手引きは記載ルールを示したマニュアルである一方、データマネジメント側の記入方針をまとめた記録にもなり、他者とのやり取りの助けにもなりうる。

3.4.2. CRF 記入の手引きのデザイン

CRF 記入の手引きは、誰でも記入方法がわかる項目については記載せず、記載者が記載方法に迷うことが予想されるポイントにしぼるようにする。

記載方法については長い文章で説明するよりも、実例を示したりオブジェクトを利用し記載の注意を促したりするほうがわかりやすい。またカラーの方が重要なポイントを強調することができ視覚的にもわかりやすい。

CRF 記入マニュアル例

検査伝票貼り付け

臨床検査

検査日	2008年 9月 19日	
白血球数	7000	/mm ³
好中球数	3500	/mm ³
血色素数	13	g/dL
血小板数	20	×10 ⁴ /mm ³
総ビリルビン	0.7	mg/dL
クレアチニン	0.7	mg/dL

検査日が他の項目と異なる場合は、余白に記入してください。
 2008/09/20

- ・治療開始前直近の検査結果についてデータを記入して下さい。
- ・検査伝票を貼り付ける場合は、所定の位置に貼り付け、割り印またはサインをして下さい。

参考文献

- 1) Eleanor McFadden, Management of Data in Clinical Trials, second edition (2007)
- 2) 辻井敦, 臨床試験データマネジメント
- 3) 技術情報協会, 臨床試験の質向上を目指した clinical Data Management
- 4) 友次直輝, CDM 教育プログラム「データマネジメント業務とその役割」



第4章 データベースシステム

臨床研究で収集されたデータは、データセンターにて格納・保管される。そのデータは、バックアップが取られている、不正な改ざんがない等、適切な方法で保管されているという保証が必要である。また、適切なタイミングで必要な情報を効率よく抽出できる必要がある。データセンターで臨床研究を支援する際、そういった要件を踏まえた上で、データベースシステムの選択、及び構築を行う。この章では、臨床研究に必要なデータベースシステムの要件、及び構築について述べていく。

4.1. データベースシステム

4.1.1. 使用するシステム

- 研究の質、予算、データマネジメントにかかわる人員などを考慮した上で、実現可能なシステムを利用する
- 全ての機能が揃っているシステムを利用することが望ましいが、要件を満たせない場合は、不足している部分を補う為の手順を事前に取り決め、データマネジメント計画書に記載することで品質を保証するという選択もありうる
- 使用システムとしては、専用のデータマネジメント・システム（Clintrial、ClinicalWorks4/DM、Oracle Clinical など）や、一般の市販されているソフトウェア（Microsoft Excel、Access、File Maker など）が挙げられる
- 但し、一般の市販されているソフトウェア（Microsoft Excel、Access、File Maker など）を利用する場合、監査証跡が取れないものが殆どなので、品質の保証を確保する為に手順を充実させる必要がある
- EDC(Electronic Date Capture)の普及から、データセンターがサーバーを管理するのではなく、アプリケーションサービスを提供する会社と提携し Web を利用したデータ収集を外部で行う例、ASP (Application Service Provider) を利用する例も見られる
- 次のセクションで、システム選定時に確認が必要な要件を列挙する

4.1.2. システムに必要な要件

- システム選定時に確認が必要なデータマネジメント・システムの機能要件を列挙する
- 性能
 - ・ 試験のデータ量に耐えることができる
以下の数値を元にシステムの規模を見積もることが可能
 - 同時に利用可能なユーザー数
 - 主要な作業に関する最大反応時間
 - 使用できるデータベースのサイズおよび数
 - 設定可能な検索の数および複雑さ

- 設計 (デザイン)
 - ・ 標準化可能で設計しやすい
- データ入力 (操作インターフェイス)
 - ・ 入力しやすい画面である
 - ・ ダブルエントリー機能がある
- コンダクト (クリーニング、クエリー)
 - ・ ロジカルチェック機能がある
 - ・ CRF の回収状況管理機能がある
 - ・ DCF (Data Clarification Form) の発行機能がある
- データ出力 (検索、帳票)
 - ・ 帳票の作成機能がある
 - 各種帳票 (症例一覧表や読み合わせ帳票等) の作成機能がある
 - 【帳票作成時の設定方法について確認をする】
 - ◇ 方法は簡単か?
 - ◇ レイアウトを調整できるか?
 - ◇ どのようなファイル形式で出力できるか? など
- ・ データの検索・抽出
 - 【検索項目や検索条件の設定が容易であるか確認をする】
 - ◇ データ入力時に外部コードを参照して検索することができるか?
 - ◇ 検索条件を設定・保管しておき、データ抽出時や帳票出力時等の抽出条件として使用することができるか? など
- ・ データの抽出
 - 【抽出可能なデータ形式について確認をする】
 - ◇ XML 形式での出力が可能か?
 - ◇ SAS データセット形式での出力が可能か?
 - ◇ CSV 形式での出力が可能か? など
- 管理 (バックアップ、セキュリティ)
 - ・ データベースの修正機能がある
 - 【再編成機能の詳細について確認をする】
 - ◇ データマネジメント・システムにプロトコルの内容を設計後、測定項目の追加等により再編成が必要になった場合、既に入力済のデータに影響を及ぼすことなく再編成を行うことが可能であるか?
 - ◇ 再編成した項目のログを取得可能であるか? など

- ・ アクセス権限設定、ユーザー管理などのセキュリティ管理
 - アクセス権限の設定を臨床研究毎に細かく行う事が可能である
 - ユーザー登録の設定、グループ化したユーザー管理が可能である
 - 処理を行った履歴が取得できる
- ・ バックアップ・リカバリ管理
 - プロトコル毎に格納されている電子的なデータのバックアップ・リカバリ管理機能がある
 - 【具体的な管理方法について確認をする】
 - ◇ バックアップ・リカバリ管理が特定の権限を有する者だけが使用できる環境であるか？
 - ◇ 災害発生時のバックアップ・リカバリ方法はあるのか？ など
- ・ 他アプリケーションとのデータ交換
 - 【具体的なデータ交換方法について確認をする】
 - ◇ データ入力業務を外注してデータマネジメント・システムに一括処理で取り込むことが可能であるか？
 - ◇ どのようなファイル形式で取り込むことが可能で、その形式にあわせるためにはプログラミングが必要となるか否か？
 - ◇ データの抽出形式はどのような形式に対応しているのか？ など
- ・ 外部コードの更新・管理
 - 外部コードを取り込む機能がある
 - また、必要な時期に更新を行う機能がある
有害事象名コードや薬剤名コード等の外部コードは定期的に更新されるので、必要な時期に更新を行うことが大切になってくる
 - 臨床研究データが固定した時点で、使用した外部コードを該当プロトコルのデータとともに保管できる機能がある
 - 【具体的な更新・管理方法について確認をする】
 - ◇ 外部コード形式を変換する必要があるか？
 - ◇ 更新時期、更新者等を管理することができるか？ など
- ・ 管理情報の出力機能
 - ユーザー管理情報、データ管理情報等を特定のファイル形式で出力する機能がある
 - 【具体的な出力機能について確認】
 - ◇ 出力されたファイル形式は他の使用しているソフトウェアで読み込むことが可能であるか？ など
- ・ 監査
 - ・ 監査証跡記録機能がある

データの入力、変更、削除などの履歴（操作した日時、操作者、操作した内容などを時系列的に記録）を残すことが出来る

- ・ データ固定機能がある
 - 症例単位での固定／プロトコル単位での固定が可能である
 - 固定者、固定日、固定対象者の情報を取得可能である
 - データ固定を行うためのアクセス権限の設定が可能である
- 【具体的な固定機能について確認】
 - ◇ データ固定後の修正方法についての手順 など
- ・ 監査証跡の取得
 - データ変更履歴等の監査証跡を自動的に取得することが可能である
- 【具体的な監査証跡の機能について確認】
 - ◇ 監査証跡内容を如何なるアクセス権限を所有していても変更することができないことを確認する など

4.2. データベース構築の手順

4.2.1. 手順

- プロトコルを理解し、収集するデータ全体の把握
- 統計解析に必要な項目を把握し、それらの項目をグループ化（参:4.2.2.）
- 整理した項目を元にデータベース構造定義書を作成（参:4.2.3.）
- データ保持方法の確認（データのコード化）（参:4.2.4.）
- データベースの具体的な設計・構築：アクセス権の設定、ロジカルチェック（参:5）
- データベースの検証（参:4.4.）

4.2.2. データのグループ化

- 臨床研究では多岐に渡るデータを取り扱うことになるが、その中でも同じ領域のデータの塊をグループ化したものをドメインと呼ぶ
- ドメインとしては、患者背景（被験者特性情報）、有害事象、バイタルサイン、臨床検査、併用薬などが挙げられる
- データをドメインでグループ化することで、煩雑なデータを系統立てて収集することができる

例えば、今回の計画書を見てみると【観察・検査項目】として下記の項目が挙げられている

観察・検査項目	項目
患者背景	人口統計学的データ、病歴
有効性評価項目	トラフ血圧、24時間血圧、随時血圧、脈拍数、服薬状況
安全性評価項目	有害事象、臨床検査、心電図、身体所見、随時血圧、脈拍数
その他	前治療・併用治療薬、試験終了状況



これらの情報をドメインに分けると、下記のように分けることができる

ドメイン	略語	項目
被験者特性情報	DM	人口統計学的データ(生年月日、性別、身長体重など)
原病歴・既往歴	MH	病歴
バイタルサイン	VS	トラフ血圧、24時間血圧、随時血圧、脈拍数(収縮期・拡張期)
曝露	EX	服薬状況(投与薬剤、投与量、期間など)
併用薬	CM	併用治療薬(投与薬剤、投与量、期間など)
有害事象	AE	有害事象
臨床検査	LB	臨床検査(検査日、測定項目、測定結果、単位など)
心電図	EG	心電図(測定日時、方法、体位、項目(QT 間隔など)、結果など)
身体所見	PE	身体所見(実施日、結果など)

4.2.3. データベース構造定義書の作成

- グループ化されたデータ定義を元に、データベース構造定義書を作成していく
- データベース構造定義書におよそ含まれる項目名を列挙する
 - ・ ドメイン名、項目名 (Variable)、SAS Name / SAS Label、Mandatory 必須 (Y/N)、Data Type、Data Length、コードリスト名など

例えば、被験者特性情報、併用薬は下記のような定義書が考えられる

【被験者特性情報】

Variable	SAS Name	必須	Type	Length	コードリスト名など
試験 ID	STUDYID	Y	Char	8	
ドメイン	DOMAIN	Y	Char	2	DM
被験者 ID	USUBJID	Y	Char	8	
生年月日	BRTHDTC	Y	Char	19	
性別	SEX	Y	Char	1	
身長	LENGTH	N	Num	3,1	
体重	WEIGHT	N	Num	3,1	

【併用薬】

Variable	SAS Name	必須	Type	Length	コードリスト名など
試験 ID	STUDYID	Y	Char	8	
ドメイン	DOMAIN	Y	Char	2	CM
被験者 ID	USUBJID	Y	Char	8	
併用薬連番	CMSEQ	Y	Char	19	
併用薬記載名	CMTRT	Y	Char	200	
併用薬正式名	CMMODIFY	N	Char	200	
服用理由	CMINDC	Y	Char	1	
服用量	CMDOSE	Y	Num	10	
服用量単位	CMDOSU	Y	Char	10	
服用頻度	CMDOSFRQ	Y	Char	1	
服用経路	CMROUTE	Y	Char	1	
報告日	CMDTC	Y	Char	19	
服用開始日	CMSTDTC	Y	Char	19	
服用終了日	CMENDTC	N	Char	19	
服用終了状況	CMENRF	N	Char	1	

4.2.4. データの保持方法(データのコード化)

- データを保持する際、内容が同じ、或いは外部辞書が存在する場合は、それらの辞書類を使用することで、データの共通化を図ることができる

- コード化には、各種団体、企業などが公開しているコード(シソーラス辞書)を利用する方法と、各データセンター内独自のコード(コードリスト(マスタ))を設定する方法に分けることができる

- 各種団体、企業などが公開している代表的なシソーラス辞書を下記に記載する

- ・ JAPIC: 財団法人 日本医薬情報センターの医療用医薬品集
- ・ MEDIS: 財団法人 医療情報システム開発センターの標準マスター(病名、手術・処置、臨床検査、医薬品など)
- ・ MedDRA: 財団法人 日本公定書協会 JMO 事業部の
症状、徴候、疾患などに対応する医学用語集

- データセンター独自のコードリストの例を下記に記載する
文字列ではなくコードで登録する

- ・ 「いいえ・はい」の場合、「いいえ 1、はい 2」
- ・ 「男・女」の場合、「男 1、女 2」など

先ほどの【併用薬】の場合は、下記のような例になる(一部項目、省略)

Variable	SAS Name	必須	Type	Length	コードリスト名など
併用薬正式名	CMMODIFY	N	Char	200	JAPIC2009
服用理由	CMINDC	Y	Char	1	1:高血圧、2:悪心/吐き気、3:感冒
服用量単位	CMDOSU	Y	Char	10	1:mg、2:mL
服用頻度	CMDOSFRQ	Y	Char	1	1日服用回数
服用経路	CMROUTE	Y	Char	1	1.経口、2.吸入、3.皮下注、4.静脈注、5.筋注、6.坐薬、7.外用
報告日	CMDTC	Y	Char	19	ISO8601 yyyy-mm-dd
服用開始日	CMSTDTC	Y	Char	19	ISO8601 yyyy-mm-dd
服用終了日	CMENDTC	N	Char	19	ISO8601 yyyy-mm-dd
服用終了状況	CMENRF	N	Char	1	1.調査期間前、2.継続中、3.調査期間後、4.不明

MedDRA (Medical Dictionary for Regulatory Activities) とは？

MedDRA は、医療に関する国際間の情報交換を迅速かつ的確に行うために、国際的に共通する用語集として医薬品規制ハーモナイゼーション国際会議 (ICH) において作成された、症状、徴候、疾患などに対応する医学用語集である。SOC (器官別大分類)、HLGT (高位グループ用語)、HLT (高位用語)、PT (基本語) 及び LLT (下層語) の5階層構造を有する用語集である。日本語版は MedDRA/J (ICH 国際医薬用語集日本語版) であり、MedDRA の英語版とそれに対応する日本語が対になったものである。MedDRA/J は半年毎に改訂されており、2008年9月に Version 11.1 にバージョンアップされている。日本では、厚生労働省の指示により、薬事法に基づく感染症定期報告、医薬品副作用・感染症症例報告、治験薬副作用・感染症症例報告及び安全性定期報告に使用する副作用名等の用語については、MedDRA/J を使用すること、承認審査資料(その他の臨床試験資料を含む)、再審査申請資料及び再評価申請資料等に使用される副作用名等の用語については、可能な限り MedDRA/J を使用することとされている。MedDRA を使用する場合には、バージョン番号を付けなければならない。年2回の更新の為、どのバージョンを用いたのかを管理することが重要である。

4.3. 標準化

- 共通している部分は標準化していくことで作業の効率向上、ひいては品質の向上を図ることができる。
- 標準化の作業は実務に沿ったものが望ましく、一度構築すれば未来永劫変更してはならないというものではなく、品質向上の為であれば、むしろ積極的に修正を行っていく必要がある。但し、安易な変更は避けなければならない。その為にも、標準化は最初によく考える必要がある。また、変更修正を行う場合の手順を事前に設定しておくべきである。

4.3.1. データ構造の標準化

- データ構造の標準化としては、データベース構造定義書の標準化が挙げられる。
- 各ドメインで使用頻度の高い項目を列挙し、標準データベース構造定義書を準備することで、作業の効率化を図ることができる。

4.3.2. コードの標準化

- 基本的には既存のシソーラス辞書 (参: 4.2.4) を利用することが望ましいが、単純なコード化であったり、施設に依存した内容の場合、一般の辞書類では定義されていない部分もある。

- ・ その場合は、データセンター固有のコードを作成し、施設内でコードを統一するのが望ましい。また、それらのコードがデータセンター内での標準コードリスト（マスタ）となる。
- ・ 具体的なマスタ類については、「4.2.4. データの保持方法（データのコード化）」で挙げた通りである

CDISC (Clinical Data Interchange Standard Consortium) とは？

臨床データの国際的業界基準の開発に取り組む公開された非営利団体である。医薬品・生物製品開発における治験（臨床研究）データ・メタデータの電子的取得、交換、申請、アーカイブのサポートを行う。CDISC は、医薬品業界におけるデータ品質の向上と製品開発促進を目的としてベンダーに中立的でプラットフォームに依存しないグローバルスタンダードの開発を目指している。活動は、CDISC の理念に賛同する企業及び個人の活動によって成り立っている。2000 年に組織され、その活動は米国を起点にヨーロッパ、日本およびインドに広がり、各地において地域グループが活動を行っている。

CDISC は現在、その役割や対象に応じて SDTM を含む複数のモデルを有しているが、将来的には、既存の全ての CDISC モデルを 1 つの CDISC 標準として集約することを目指している。また、さらにその後は、患者データ（電子カルテ）や治験・臨床研究で使用しているシステムとのデータ相互運用を可能にする統合ドメイングループ構築モデル（Biomedical Research Integrated Domain Group、「BRIDG」）を構築し、データの統合を目指す。BRIDG は、FDA や NCI（国立癌研究所）との共同プロジェクトである。

CDISC の活動や今後の方針、また、関連する文書およびインプリメンテーションガイドは、全て CDISC のホームページ (<http://www.cdisc.org>) で公開されている。（次頁へ続く）

【標準様式】（一部）

- ・ Study Data Tabulation Model (SDTM)
Standard for the Exchange of nonclinical Data (SEND)
 - 臨床試験データおよび非臨床（動物）試験データ (SEND) の Case Report Tabulations (CRT) として規制当局への申請に用いられる。
 - 各ドメインの項目が定義されている
- ・ Operational Data Model (ODM)
 - オペレーショナルデータのトランスポート、電子データのアーカイブおよび Define.xml として試験データの申請に用いられる。
- ・ Clinical Laboratory Model (LAB)
 - 臨床検査機関および試験依頼者間のデータの交換に用いられている。
 - XML, ASCII, SAS または HL7 V3 メッセージを含む様々なフォーマットで実装されている。
- ・ Analysis Dataset Model (ADaM)
 - 解析データセットを規制当局へ提出するために SDTM と一貫した General Considerations document を発行。
 - 数種の解析手法に特化したデータセットの例示が含まれている。

4.4. データベースの検証

- データベースの設計が完了し、それに伴う資料、データベース構造定義書、Annotated CRFなどが出来ればデータベースの準備は完了する
- データベース自体の検証方法について
 - ・ 出来上がったデータベースに対して本当に正しいデータが格納できるのか、データマネジメント・システムの入力画面との整合性など検証を行う必要がある
- 画面上のチェックのプログラムの検証方法について
 - ・ また、データマネジメント・システム上で入力規制、ロジカルチェックをかけている場合は、それらのプログラムが正しく動作するのかが確認する必要がある
 - ・ 勿論、不正データが入力された場合に正しく処理されるか（データが入力されない、エラーが出力される等）の確認が重要になる
 - ・ セキュリティ（アカウント発行、権限設定）についても、正しく設定されているのかが確認する必要がある
 - ・ また、最終的に出力されるデータ抽出（Data Setの作成）が正しく行われるのかも確認する必要がある
 - 確認方法の例
 1. データ抽出仕様の設定
 2. データマネジメント・システムで入力
 3. データ抽出の実行
 4. 抽出されたデータの内容を確認項目が正しいか？入力された内容が正しいか？など

参考文献

- 1) 技術情報協会, 臨床試験の質向上を目指した Clinical Data Management
- 2) 技術情報協会, 臨床試験におけるデータマネジメント実務集
- 3) CDISC Study Data Tabulation Model(SDTM) Implementation Guide Version3.1.1
- 4) R.K.Rondel・S.A.Varley・C.F.Webb, クリニカル・データマネジメント
- 5) 大橋靖雄 荒川義弘, 臨床試験の進め方
- 6) 社団法人 日本薬学会, 薬学用語解説

第5章 データクリーニングの方法と要点

データクリーニングとは、CRF 作成時や入力時に発生する様々なエラーを検出し、必要に応じて施設への問い合わせを行い、正しいデータへ修正するプロセスのことである。データクリーニングは、データの品質を保証するための重要なステップであり、クリーニングの早期実施は、エラーの再発防止にもつながる。本章では、データクリーニングの方法や、注意点等について述べる。

5.1. データクリーニングの種類

5.1.1. ロジカルチェックとマニュアルチェック

- データクリーニングの手法には主に、「ロジカルチェック」「マニュアルチェック」「併用チェック」がある。

- ロジカルチェック

データマネジメント・システムに入力されたデータに対して行う、コンピュータによるチェックである。また、あらかじめデータマネジメント・システムに制限をかけておくことで、入力時にチェックをかけることも可能である。一定の法則性があるチェックに対して有効である。

- マニュアルチェック

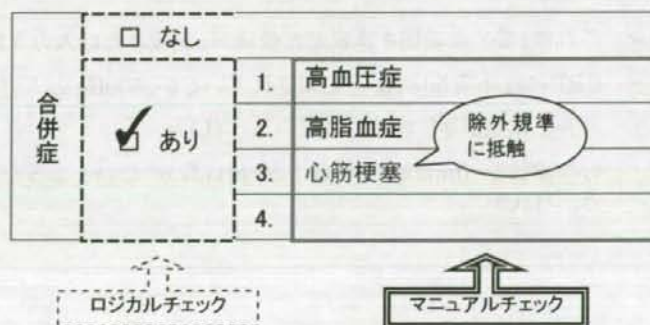
人（目視）によるチェックである。コメントデータや医学的知識を要するチェックに適しており、ロジカルチェックプログラムを組むのが難しいチェック内容もマニュアルチェックで行う場合がある。

- ロジカルチェックとマニュアルチェックの併用

ロジカルチェックにより抽出されたデータや、一覧表、グラフ等に対してマニュアルチェックを行う方法である。

併用チェックの例)

まず、ロジカルチェックで合併症“あり”が選択されている症例を抽出する。そして抽出された症例の合併症名についてマニュアルチェックを実施し、除外規準に該当する事象や、不適切な表現が記載されていないか等を確認する。



- どちらの手法にもメリット・デメリットがあるため、試験の規模や phase、データの複雑さ等を考慮し、適切な手法を選択する必要がある。

表 1. 各チェックのメリット、デメリット

	メリット	デメリット
ロジカルチェック	<ul style="list-style-type: none"> ・ 大量のデータ処理が可能 ・ 一貫したチェックが可能 	<ul style="list-style-type: none"> ・ プログラミングの知識が必要 ・ プログラムの作成、検証に時間を要する ・ 医学的な内容をチェックするには不向き
マニュアルチェック	<ul style="list-style-type: none"> ・ 医学的な内容や、コメントデータ等の妥当性がチェックが可能 	<ul style="list-style-type: none"> ・ レビュー結果に個人差がある ・ 見落としといったヒューマンエラーが生じる ・ チェックに時間を要する
併用チェック	<ul style="list-style-type: none"> ・ 適したチェック項目であれば、エラー候補の絞り込みが可能となり、効率のよいチェックが可能となる 	<ul style="list-style-type: none"> ・ ひとつのチェックでは、作業が完結しない ・ プログラムの作成、検証に時間を要する場合がある

5.1.2. エラーの種類

- 一言に“エラー”と言っても、その種類は多岐に渡る。それぞれの特徴を理解し、適切な方法でチェックを行うことが重要である。
 - 以下に、エラーの種類とその例を示した。例の文末についているアルファベットは適切であると思われるチェックを示している（L:ロジカルチェック、M:マニュアルチェック、C:併用チェック）。ただし、CRFの構成やデータベースの構造によって、どちらのチェックが適切であるかが変わってくることも念頭に置いて欲しい。
- ✓ 欠測；記載必須項目が記載されていないエラー。
 - ・ 性別、検査結果、試験薬投与日等が記載されていない。(L)
 - ✓ 外れ値；妥当な範囲を逸脱した数値が、記載または入力してしまったエラー。
 - ・ CRFには体重 50.0Kg と記載されているが、500Kg と入力してしまった。(L,C)
 - ・ 身長と体重が逆に記載されている。(L)
 - ・ 好中球数を (/mm³) で記載しなければいけないところを、(%) で記載している。(L,C)

- ✓ 日付の整合；時系列の前後関係に不整合があるエラー。
 - ・ 登録日が2008/2/1なのに、初回投与日が2007/2/4と記載されている。(L)
 - ・ ありえない日付(2008/6/31など)が記載されている。(L)
 - ・ 未来の日付や100年以上前の日付が記載されている。(L)

- ✓ 項目間の整合；同一項目内に不整合があるエラー。
 - ・ 既往歴の有無で「無」が選択されているが、病名が記載されている。(L)
 - ・ 「プロトコル完了」が選択されているが、「中止理由」が選択されている。(L)

- ✓ ページ間の整合
 - ・ 有害事象の処置にて薬剤名が記載されているが、併用薬ページには該当する薬剤名が記載されていない。(M,C)
 - ・ 投与延期理由に有害事象名が記載されているが、有害事象ページには該当する事象名が記載されていない。(M,C)

- ✓ 適格性、逸脱；不適格症例や、プロトコルからの逸脱に関するエラー。
 - ・ 除外規準に「間質性肺炎のある患者」と規定されているが、合併症欄に「間質性肺炎」との記載がある。(M)
 - ・ 臨床検査を、投与2日前～投与当日に行わなければならないところを、投与5日前に実施していた。(L)
 - ・ 食欲不振Grade3が出現した場合、次コースは減量しなければならないところを、減量せずに投与していた。(M,L,C)

5.2. チェック仕様書

- チェック仕様書とは、ロジカルチェック、マニュアルチェックにてどのようなチェックを行うかを定義したものである。
- チェック仕様書を用いて確認を行うことにより、全症例のデータに対して一貫したチェックを行うことが可能となり、データ品質の均質化につながる。

5.2.1. チェックの範囲

- クリーニングの目的は、データの品質を、予め規定した品質基準以上に保つことであり、エラーを完全になくすことではない。
- 重箱の隅をつつくようなチェックは、データセンターだけではなく、施設側の負担となることを忘れてはならない。
- 例えば、誤字脱字等はよくあるミスであるが、これを全て施設へ問い合わせることに意味があるだろうか。治験のようにCRAがついていれば、SDVの際に見つけて修正をしてもらうことも可能であるが、データセンターのように中央モニタリ

ングを行っている場合はそうはいかない。

- このように、両者にとって負担となるようなオーバーキルなチェックにならないよう注意しなければならない。
- データの顧客は誰か、本当に求められている品質はどの程度なのかということを念頭におき、品質基準を設定することが重要である。

5.2.2. チェック仕様書の作成

- 規定されたチェック仕様書のフォーマットはない。組織内で使用しやすいフォーマットを考える必要がある。以下に、チェック仕様書のテンプレートの一例を示す。

表 2. チェック仕様書

Page	項目名	方法	チェック内容	エラー抽出条件
患者背景	体重	L	記載されている	PEWEIGHT = blank
臨床検査値	検査日	L	投与2日前～投与当日の日付が記載されている	EXSTDTC-LBDTC>2 or EXSTDTC-LBDTC<0
投与状況	投与延期理由	M	投与延期理由に有害事象名が記載されている場合、有害事象ページに該当する事象が記載されている	

- **Page、項目名**：どの項目に対するチェック内容であるかがわかるように記載する。
- **方法**：ロジカルチェック(L)、マニュアルチェック(M)、併用チェック(C)のどの方法で行うのかを記載する。表1に示した各チェックのメリット・デメリットや試験の規模、マンパワーを考慮する。
- **チェック内容**：表現系を統一して記載しなくてはならない。表2のように true (CRFに記載もしくはDBに入力されているべきデータの内容) で記載することが望ましい。また、容易で、かつ誰が読んでも同一の解釈が得られるような表現を心がけなくてはならない。曖昧な表現(“妥当である”や“矛盾がない”等)は出来る限り避けるべきである。そして、1つのチェック内容から1つの回答が導き出されるように作成する。例えば「体重/身長が記載されている」というチェック内容では体重と身長の両方が未記載の場合エラーとするのか、それとも体重もしくは身長のどちらかが記載されていればエラーにはならないのかといった問題が発生してしまう。このように、1つのチェックから複数の回答が出るようなチェック内容では運用時に問題が発生するため避けなければならない。
- **エラー抽出条件**：チェック内容を条件式で表している。ここでも表現系を統一し、表2のように false (チェック内容とは逆の記載方法。エラーが出るときの条件を記載。) で記述する。言葉で表現したチェック内容では、プログラムを組む際に誤解や解釈の違いが生じることがある。チェック仕様書の作成者とプログラム作成者が異なることもあるため、このように条件式で日付や数値の関係性を正確に記載することが大切である。

- チェックリストの作成手順の1つの方法として、まずはすべてのデータパターンを考慮してチェックリストを作成する。次にデータクオリティーに応じたチェックを検討し、チェック項目を削除するなどして簡略化していく。また、ロジカルチェックは目視では単純ミスが発生するような点、記載有無などに絞る、目視を極力避けるためにロジカルチェックを増やすなどを検討する。このようにすべてのチェック項目をリストすることによって、後に発生するかもしれないイレギュラーデータへの対応も事前に検討ができるため有用である。



実習 チェック仕様書を作成してみよう！

5.2.3. チェック仕様書のレビュー

- チェック仕様書のドラフト版ができたなら、試験担当者間でレビューする。
- レビューの実施者は1人ではなく、複数人で行うことにより、様々な方向からの確認が可能となる。
- レビューする際は、チェックに漏れがないか、重複するチェックはないか、チェック方法（マニュアル/ロジカル）は適切であるか、CRF記載方法や入力ルールに適合しているか等に注意して行う。

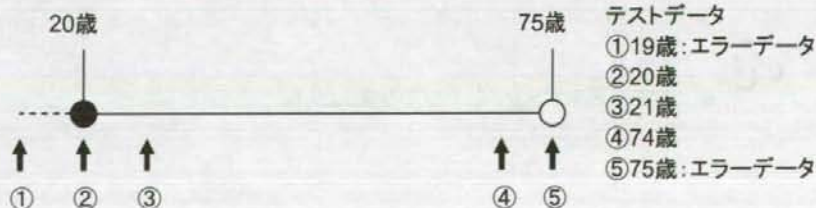
5.2.4. ロジカルチェックプログラムの検証

- ロジカルチェックプログラムの作成が終了したら、プログラムがチェック仕様書通りに正しく動作するかを確認する。
- 検証の方法として、テストデータを入力し、実際にプログラムを実行するという確認方法がある。テストデータとして、「想定したエラーが、エラーとして抽出されるようなデータ」、「整合がとれているデータが、エラーとして抽出されないようなデータ」の2種類を入力し、ロジカルチェックプログラムを実行後、実行結果を確認する。確認結果に不具合を見つけた場合は、プログラムの修正をし、再度ロジカルチェックプログラムを実行し、実行結果を確認するという作業を、不具合がなくなるまで行う。
- 検証のプロセスや結果は、きちんと保存しておくことが重要である。

テストデータの例)

チェック内容: 年齢は20歳以上75歳未満である。(20 ≤ AGE < 75)

上記のようなチェック内容について検証を行う場合、①～⑤のテストデータを入力し、実行結果を確認する。



5.3. データクリーニングの方法

記載済みのCRFを受領後、予めデータマネジメント計画書で定めたタイミングでデータクリーニングを実施する。

データクリーニングによってCRFデザインやプロトコルの問題点等、設計段階の不備が発見されることも少なくないため、早期発見・早期解決のためにも、できるだけ早い時期からデータクリーニングを開始することが望ましい。

また、データクリーニングの作業に品質のバラツキが生じないように、実施担当者は、事前に、チェック仕様書の内容、個々のチェック項目の意義と目的を十分に理解しておくことが必要である。

クリーニング作業中にチェック仕様書の不備を見つけた場合には、必要に応じて仕様書の内容を修正し、関係者間で情報を共有すること。

臨床研究では、1試験に関わる関係者が少数であることが多いため、自分だけが正しい理解を持っていると問題ないと考えがちであるが、長期間に及ぶ試験も多く、1スタッフが試験終了までの責任を持って担当できるとは限らないことに注意して、記録を残すことを心がける。

5.3.1. ロジカルチェックの方法

ロジカルチェックの方法には、データマネジメント・システム内のロジカルチェック機能を利用する方法や、SAS・EXCEL (VBA) といったソフト上でプログラムを構築する方法等がある。

データセンターで所有するシステムやソフトに応じて、また、プログラムを構築するスタッフの能力に応じて、最も都合のよい方法を選択する。

通常、プログラム実行により、ロジカルチェック項目でエラー条件にあてはまった箇所の一覧がエラーメッセージとともに出力される。

入力画面上にロジカルチェックプログラムを構築した場合は、エラー条件にあてはまる入力が行われた場合、入力者宛てにエラーメッセージが表示される。

5.3.2. マニュアルチェックの方法

マニュアルチェックの方法は、プリントアウトしたチェック仕様書上で、マニュアルチェック項目をひとつひとつ確認し、エラーの有無やエラーの内容について記録する方法や、これを電子ファイル上で行う方法等がある。

何れの方法も、チェックをいつ誰が実施したかの履歴を残すことが重要である。



実習 マニュアルチェックをやってみましょう！

5.3.3. エラーの解決方法

ロジカルチェックおよび、マニュアルチェックで発見されたエラーの解決方法には以下の2つの方法がある。

- ① データセンター内部で解決する
- ② ①で解決できなかった疑義事項を施設へ問い合わせる

①の方法で解決できるエラーの例には以下のようなものがある。

- これ以上の問合せは不要であると判断できるエラー
 - ・ QOL アンケートの欠測や不整合（事後に確認することが無意味）
 - ・ 測定できなかった理由が既にコメントされている欠測
 - ・ 施設から逸脱である旨の連絡を受けている問題点
- CRF の修正を行わずに、データマネジメントによってデータの修正が可能と予め定義された疑義内容
 - ・ self evident data corrections (SEDCs) と呼ばれる手法

SEDCs は施設への問合せ数を減らし、関係者の負担を軽減、効率化をはかることを目的として考えられた手法である。

データの本質に関わらない修正で、正しい答えが明白である疑義についてのみ、データマネジメントによる修正を許容するという方法で、この方法によって修正される例には以下のようなものがある。

- ・ 有無が空白で、内容が記載されているとき、“有”を入力する
- ・ 前後関係から明らかに年の記載が間違っていると判断できるとき、年を修正する（主に1月）
- ・ 明らかな誤字・脱字を修正する
- ・ 略語を正式名称に読み替える
- ・ 男性で、妊娠の有無が“なし”を“該当せず”に修正する

臨床研究では、CRC 等のサポート体制が乏しい施設も多いため、担当医師の負担を考慮して施設への問合せは必要最小限に留めたい。

SEDCs は非常に有用な手段と考えられるが、データの改竄や個人の判断による不適切なデータ修正が行われないために、SEDCs の運用に際しては修正可能な範囲を事前に定義・明文化することが重要であり、事前に研究代表者より内容の承認を得ることが望ましい。

更に、GCP 準拠が求められる臨床研究の場合、答申 GCP8-1-11-4 「…治験依頼者が指名した者によって行われた CRF の変更又は修正においては、それらが文書に記録され、必要なものであり、かつ治験責任医師が承認したものであることを保証

するための手順書を作成しておかなければならない。」の遵守事項から、すべての修正履歴をとり、事後に修正内容について責任医師に承認を得ることが求められるので、これに対応できるような手順を予め用意する必要がある。

5.3.4. クエリーの発行と管理

「6.3.3 エラーの解決方法」において、①（データセンター内部で解決）で解決できなかった疑義事項は、施設の当該症例担当医師へ問合せを行う。

この際、クエリー(Query)と呼ばれる特別な様式を用いて問合せを行うのが一般的である。

問合せの前に、内容に漏れがないか、正しい箇所を誤って指摘していないか等を改めて見直すこと。

リソースが許せば、チェック実施者以外のスタッフによるダブルチェックを行うことが望ましい。

問合せの要否に医学的な判断を要する場合には、身近な専門家(医師等)に相談するのも有効な方法である。

クエリーの間合せ文面は、確認事項を分かりやすく記載すること。

不明瞭な文面は、回答者を悩ませるだけでなく、誤った回答を招き、問合せが再度必要となるなど、データセンターの負担も増大させる。

CRF のどの部分に対する問合せなのかを明確にし、問合せを行う相手側に理解しやすい文章を、読み手の立場になって考えなければならない。

また、回答を決め付けるような表現(「60”でいいでしょうか?」)ではなく、正しい回答を導く表現(「不明瞭ですが正しい値は?」)で記載することを心がける。

クエリーは通常、FAX、郵送、モニターによる手渡し、メールなどの手段により施設へ発行される。

クエリ(回答済)の原本は CRF と同様にデータセンターへ提出され、施設では複写を保管してもらう。

データセンターへ回収されたクエリ(回答済)の原本は、CRF と同様に入手庫の記録を残し、回答漏れがないよう管理する。



実 習 クエリーを作成してみましょう!

5.3.5. クリーニング作業終了後の評価

試験終了後、クリーニング作業を振り返り、エラーの発生状況等について以下のような分析を行うことにより、現行のプロセスにおける問題点を見つけることができる。