

# データマネジメントについて知ろう

医師主導治験におけるデータマネジメントは何をしている？

2024年9月10日

国立がん研究センター東病院  
臨床研究支援部門 臨床研究推進部 データ管理室  
室長 池田 裕弥

# 経歴概要

- 1998年4月～1999年8月 パソコン販売の会社(かなり前に倒産)  
PCのメンテナンス(パソコンの知識を身につける)
- 2000年1月～2018年12月 CRO DM (19年)  
治験、PMS:がん、リウマチ、アルツハイマー、他  
外部活動など:CJUG、J3C、CRO協会統計DMWG
- 2019年1月～ 現在 ARO DM (5年)  
(国立がん研究センター東病院 臨床研究支援部門  
臨床研究推進部 データ管理室)  
医師主導治験、臨床研究  
外部活動など:CJUG  
研究:CDISC関連、DCT関連、臨中ネット関連

注:本資料の内容は個人のこれまでの経験より作成しております。  
その為、個人の見解となります事、ご留意頂ければと思います。

# 概要

- よくいだかれる疑問(何しているの?)
- データマネジメントがしている事
- データマネジメント業務を支える資料、プログラムなど
- データマネジメントは何を目的に実施しているのか
- なぜ、〇〇ドキュメントを作成するのか
- ドキュメントを基にした作業の紹介(一部)
- 掛かる期間のお話
- 求めたい知識やスキルのお話
- まとめ



# 雑談

下記と耳にしております。

- 昔はDMという役割はなかった。
- 役割分担などによりDM業務が確立。

プロトコール作成、各施設とのやり取り、CRA、DM、総括報告書の作成、承認申請までを手広く担っていた方もおられると耳にします。(解析部分を除けばほとんど一人で…)

# よくいだかれる疑問(何しているの?)

2024年4月のAI: Copilotに聞いてみた

## データマネジメントの言葉で想像される業務って何?

企業内に存在する全てのデータを管理する業務全般を表す言葉

具体的な業務例	説明
データの管理と活用	データを適切に管理しながら、企業の成長・発展に活用する組織的な行動すべてを指す
データの入力・分析	臨床試験により収集されたデータの入力・分析を行い、データベース化する
システム開発・運用	情報工学の知識を活かしてデータ入力やデータ管理のシステム開発・運用を行う
データの一貫性の確保	データ要件に応じた一貫性のあるアクセシビリティ、配信、ガバナンス、セキュリティを確保する活動
データマネジメントツールの利用	MDM (マスターデータマネジメント)、データ仮想化、データカタログ、セルフサービス型のデータプリパレーションやラングリングなどのツールが使われる



## 治験におけるデータマネジメントは何をしている?

治験薬の効果を統計的に検証できるように、被験者に関する症例報告書のデータ化、不足しているデータや記入ミスの洗い出し・指摘・訂正、データの不足・ミスが見つかった場合に、製薬企業やCRAに連絡等を行いながら、治験で集めたデータの確認・補足等を行う職業

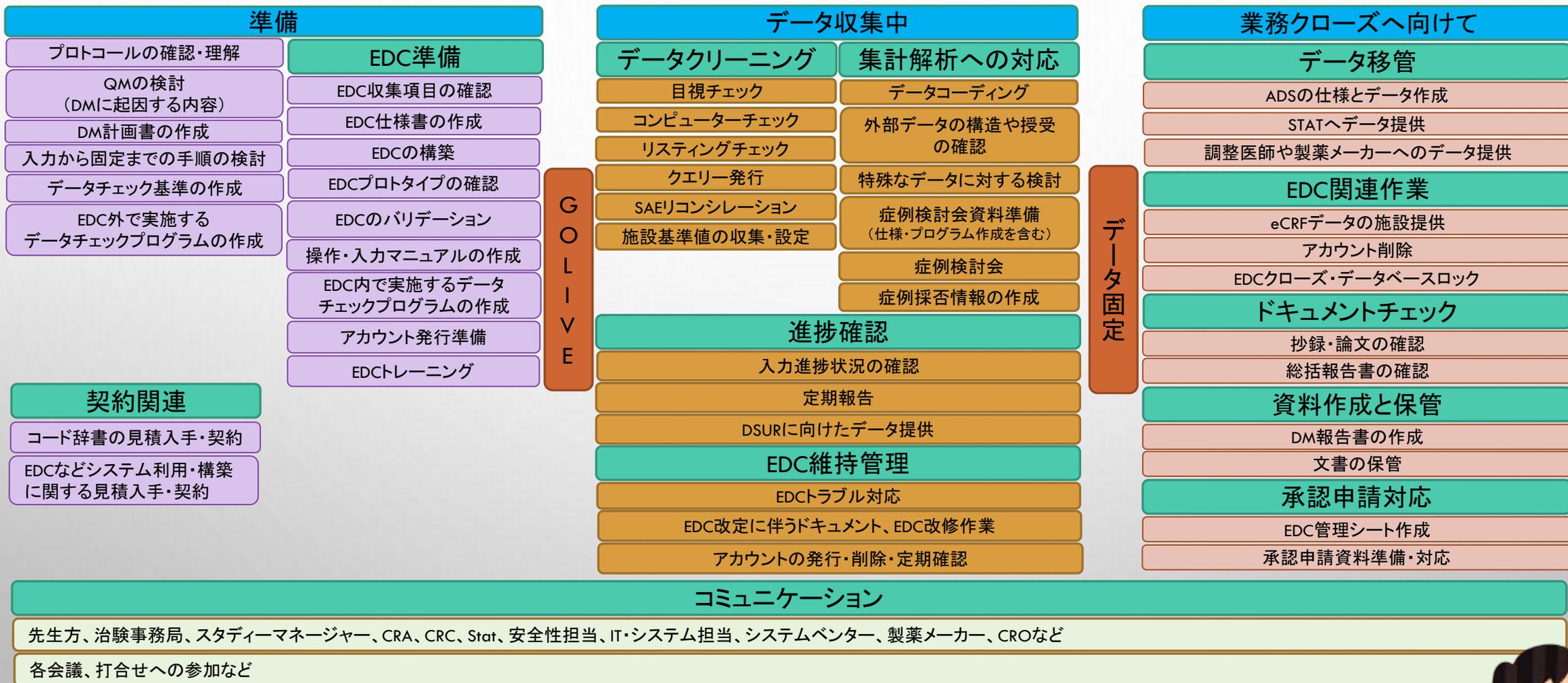
具体的な業務例	説明
症例報告書(CRF)のデータ入力	データ入力作業を派遣スタッフが行う場合には、データマネジメントは派遣スタッフの作業スケジュールの調整・管理等を行う
データのコーディング	検索や集計がしやすくなるように、症例報告書データに含まれる有害事象名や併用薬名・既往歴といったデータなどに特定のコードを割り当て、コーディングする
データのロジカルチェック	コンピュータープログラムを使用して、データのロジカルチェックを行う
データのマニュアルチェック	治験が治験実施計画書等を順守して行われたかどうか、目視で一つ一つ確認する
データの修正(クエリー対応)	マニュアルチェックおよびロジカルチェックで見つかったエラー等をクエリー(問い合わせ事項)としてまとめ、CRA(臨床開発モニター)に確認を依頼します。その後、臨床開発モニターの回答を参考にしながら、データの追記・修正を行う
データの固定	症例報告書データのエラー等が全て解消された状態を、「データの固定」と呼ぶ。データの固定が行われたら、データマネジメントの業務は完了

2024年4月のAI: Copilotへ、実際は、それだけではないんですよ

# データマネジメントがしている事(例)

注: 事例になります。各AROにおいて業務内容・範囲は異なります。

注: 一部はデータ収集中にも発生することがあります



略称の補足

EDC:Electronic Data Capture    QM:Quality Management    eCRF:Electronic Case Report Form    DSUR:Development Safety Update Report  
 DM:Data Management    SAE:Serious Adverse Event    ADS:Analysis Data Set

こんなにあるの! ?



# データマネジメント業務を支える資料、プログラムなど

## 各種SOP

データマネジメント  
システム構築・バリデーション  
アカウント発行  
コード化  
データクリーニング  
症例検討会  
他

## 他からの資料

治験実施計画書  
クオリティマネジメント手順・計画書  
解析計画書  
副作用報告書  
モニタリング報告書(必要時)  
施設基準値  
EDCトレーニング記録  
TLGs  
総括報告書  
抄録  
他

## 外部ベンダー

見積・契約書  
ベンダーからの成果物

## DMの基本となる資料

データマネジメント計画書  
データマネジメント報告書

## EDC関連資料

EDC要求仕様書  
チェック仕様書(マニュアル、コンピュータチェック)  
EDC仕様書  
バリデーション計画書  
バリデーション手順書  
バリデーションチェックリスト  
バリデーション結果一覧  
バリデーション報告書

入力マニュアル

eCRF

EDCアカウント発行トレーニング記録管理表

EDC管理シート

本番環境設定書

Golive報告書

## プログラム準備

データチェックプログラム  
入力進捗確認プログラム  
グレード判定チェックプログラム

## データ収集中に発生する資料

ロジカルチェック結果  
マニュアル・リステイングチェック結果  
クエリー

コード化結果・一覧

入力進捗確認結果

DLT確認書

症例固定報告書

データカットオフ宣言書

DCF

署名印影一覧

監査証跡

## 症例検討会に伴う資料

症例検討会資料仕様書  
症例検討会資料作成プログラム  
症例検討会議事次第  
症例検討会資料  
症例検討会議事録  
症例採否一覧

## 固定に向けた資料

QC確認記録  
QC報告書  
データ固定依頼書  
データ固定解除依頼書  
データ固定報告書  
eCRF  
EDCデータ

## データ移管に伴う資料など

外部データ仕様書  
外部データ  
解析用データセット仕様書  
解析用データセット作成プログラム  
解析用データセット  
データ授受書  
EDC取り扱い依頼書

## 成果物における記録

QC記録  
作業記録

予想より多い……



他にも資料多数あります

# データマネジメント業務を支える資料、プログラムなど 手順変更

データ収集中に治験実施計画書に修正があったら最大どの程度の資料に影響があるのか

## 各種SOP

データマネジメント  
システム構築・バリデーション  
アカウント発行  
コード化  
データクリーニング  
症例検討会  
他

## 他からの資料

治験実施計画書  
クオリティマネジメント手順・計画書  
解析計画書

副作用報告書  
モニタリング報告書(必要時)  
施設基準値  
EDCトレーニング記録

TLGs  
総括報告書  
抄録  
他

## 外部ベンダー

見積・契約書  
ベンダーからの成果物

## DMの基本となる資料

データマネジメント計画書  
データマネジメント報告書

## EDC関連資料

EDC要求仕様書  
チェック仕様書(マニュアル、コンピュータチェック)

EDC仕様書  
バリデーション計画書  
バリデーション手順書  
バリデーションチェックリスト

バリデーション結果一覧  
バリデーション報告書  
入力マニュアル  
αCRF  
EDCアカウント発行トレーニング記録管理表

EDC管理シート  
本番環境設定書  
Golive報告書

## プログラム準備

データチェックプログラム  
入力進捗確認プログラム  
グレード判定チェックプログラム

## データ収集中に発生する資料

ロジカルチェック結果  
マニュアル・リスティングチェック結果  
クエリー

コード化結果・一覧  
入力進捗確認結果

DLT確認書  
症例固定報告書  
データカットオフ宣言書  
DCF  
署名印影一覧  
監査証跡

## 症例検討会に伴う資料

症例検討会資料仕様書  
症例検討会資料作成プログラム  
症例検討会議事次第  
症例検討会資料  
症例検討会議事録  
症例採否一覧

## 固定に向けた資料

QC確認記録  
QC報告書  
データ固定依頼書  
データ固定解除依頼書  
データ固定報告書  
eCRF  
EDCデータ

## データ移管に伴う資料など

外部データ仕様書  
外部データ  
解析用データセット仕様書  
解析用データセット作成プログラム  
解析用データセット  
データ授受書  
EDC取り扱い依頼書

## 成果物における記録

QC記録  
作業記録

こんなに変わるの!?



# 概要

- よくいだかれる疑問
- データマネジメントがしている事
- データマネジメント業務を支える資料、プログラムなど
- データマネジメントは何を目的に実施しているのか
- なぜ、〇〇ドキュメントを作成するのか
- ドキュメントを基にした作業の紹介(一部)
- 掛かる期間のお話
- 求めたい知識やスキルのお話
- まとめ

赤字のところまで  
終わりました



# データマネジメントは何を目的に実施しているのか

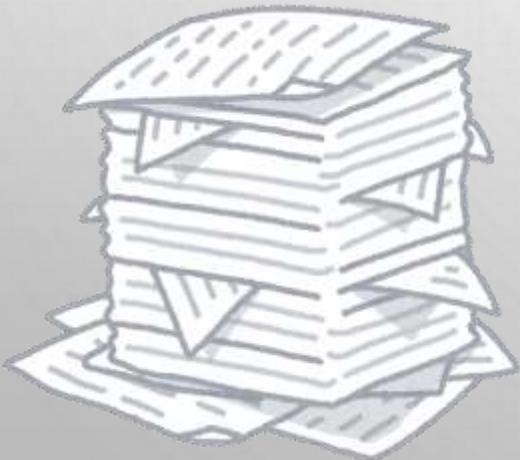
- 品質管理された信頼性、再現性があるデータの収集
- 解析に利用できるデータの収集
- 被験者保護がされたデータの収集

そのような目的が！！



# データマネジメント業務を支えるドキュメント

ところで沢山のドキュメントがありましたけど  
そんなに要らないじゃ……。  
もし、書類がなかったら、データマネジメント  
業務はどうなるのでしょうか。



ドキュメントがなかったらどうなるのか  
いくつかの例をお伝えできればと思います。



DM計画書  
EDC構築ドキュメント  
バリデーションドキュメント  
データチェック基準  
入力マニュアル  
DM報告書

# なぜ、DM計画書を作成するのか

## DM計画書の概要

業務範囲やその実施時期、作成される成果物、役割などを記載します。他にも使用するEDCや使用する辞書のバージョン、ユーザー管理方法などについても記載されています。治験実施中に、データマネジメント計画書に記載されていない業務を実施することとなった場合は、必要に応じてデータマネジメント計画書を更新します。

注:各試験によって、DM業務で実施する内容が異なるため、DM計画書に何を実施するのか定めています。

## 作成しなかった場合何が起こるか

### 例1 業務範囲の話

外部データは、ベンダーからCSVデータとして受領することで、話がまとまっていたが、いざ受領した際には、記録がなく、やむなくPDFを受領し、データ化を実施した。

### 例2 実施時期の話

Aさんは、急遽試験を引き継ぎ、情報がなかったため、他の試験のデータクリーニングタイミングを合わせ、6か月に1回実施した。実は、立ち上げ時に、試験スピードが速いため、本試験では月1回データクリーニングを実施することになっていた。

### 例3 成果物の話

話し合いによりQC記録を成果物として残す予定であったが3年前に話したことなので忘れてしまい、QC未実施となった。

### 例4 役割の話

チェック仕様の署名が、室長と、部長と、チームリーダーと存在する。何故バラバラなのでしょう？

### 例5 辞書バージョン更新の話

辞書のバージョン更新を想定していたが、普段辞書の更新をしないため辞書の更新を失念した。

### 例6 ユーザー管理の話

EDCを使用する際、同じ役割なのに、違う権限が付与されている。また、トレーニング記録があるケースとないケースがある。

経験は皆違います。  
思い込みで実施せず、何を実施予定かなど、計画を記録しておくのは大事ですね。  
この試験では●●実施する？  
の健忘記録にもなります。



# なぜ、EDC構築ドキュメントを作成するのか

## EDC構築ドキュメントの概要

治験実施計画書に定められた項目をデータ収集するため、システムとして必要な要件を定めた書類です。

例：変数名、データタイプ(文字、数値、日付)、選択肢(複数、単数)、システム上に表示する言葉(性別、薬剤名などの入力してほしい内容の概要)

システム上チェックをしたい要件、入力画面の発生条件、SDVの要否、電子署名の要否、ロール権限(どのような権限で機能を利用できるか)

レポート権限(誰がどのレポートを利用できるか)

試験によっては、600以上の入力項目と例で示したような設定の作成が必要となります。

## 作成しなかった場合何が起こるか

### 例1 思い込みによる作成ミス

治験実施計画書通り作成したつもりだったが、収集項目が足りなかった。(作成、設定したつもり)

コンピューターチェックを作成したが、登録時のチェックを作成し忘れ、本来登録できない症例も登録できてしまう。(必要なチェックの作成忘れ)

### 例2 コミュニケーションエラー

治験実施計画書に則った収集項目の設定を担当の方へ依頼したが、足りなかった。(治験実施計画書の読解力による必要項目の差)

治験実施計画書に則ったコンピューターチェック作成を担当の方へ依頼したが、思っていたチェックとは異なるものが作成された。

体表面積を算出するプログラムを作成してもらったが、算出式が異なっていた。

### 例3 設定内容の失念

試験中頃に体表面積を算出するプログラムが気になったが、どの計算式でプログラムしたか、失念した。

最初に●が入力された段階で、入力欄が展開されるよう設定したが、どのような順番でプログラムが動き、展開されるよう設定したか記憶が曖昧になっており、質問者へ説明できなかった。

### 例4 どこまでを構築したかがわからなくなる

他の業務が入り、しばらく中断せざるを得なくなり、一月後に作成を再開した。どこまで何を作成したか、ウル覚えでしか覚えていなかったため、状況を把握するところから業務開始となった。

他の人が作成を引き継ぐこととなったが、何がどこまで作成されているのか、わからない

EDC構築ドキュメントは、覚えることが困難な量の設定などを、他の方とコミュニケーションするのに利用したり、チェックリストとして利用できますね。



# なぜ、バリデーションドキュメントを作成するのか

## バリデーションドキュメントの概要

人間なので、ミスはあります。構築したEDCやプログラムが問題なく動いているか、EDC構築ドキュメントの要望に沿って作成されているか、操作・動作確認やテストデータを用いて確認をする際に、その根拠資料として作成される資料になります。

書類には、バリデーション計画書、UAT計画書、UAT手順書、作業記録、UAT報告書、バリデーション報告書などがあります。

## 作成しなかった場合何が起こるか

### 例1 本当に確認したのか

データ収集中に間違った挙動が発見されました。  
テストはしたということですが、証拠がありません。該当箇所は修正しましたが、他の部分は問題なく動くのだろうか・・・。  
手当たり次第にEDCの確認をしました。〇〇の動作確認はしましたか、と質問されたが、記録が多すぎて該当箇所の動作確認記録を見つけ出せませんでした。

### 例2 確認漏れは？

A: EDCの動作は問題なかったですか？  
B: はい。大丈夫です。  
A: 〇〇部分は確認が難しかったと思いますが、どんな確認をしましたか？  
B: それは・・・、してないです・・・。

### 例3 システムの品質を説明できない

A: システムが問題なく動いていることを教えてくださいませんか？  
B: はい。確認し、問題なく動作しました。  
A: どのような確認をしたのか、それらを書面でお願いいたします。  
B: えっ、書面・・・。記録は何も・・・。  
A: そうですか・・・。それでは、システムの品質を保証できませんね。試験データの収集や監査証跡、アカウント管理は問題なく機能したのでしょうか？  
B: 問題はないと聞いています。  
A: では、その根拠資料をお願いできますでしょうか？  
B: えっ、資料はないです。  
A: それでは、システムは正しく動いたことは証明できませんね。利用するデータが保証できないので、解析結果も保証できません。  
B: えっ！？

システムが保証できないと、データが保証できず、それに基づいた解析結果も保証できないのですね。  
システムのバリデーションドキュメントは、解析結果の保証のためにも必要ですね。



# 余談 データ収集中にEDCの修正が発生した場合

どのような経緯でどの部分のEDC修正になるかによりますが、DMドキュメントとしては概ね下記に影響があります。

## 各種SOP

データマネジメント  
システム構築・バリデーション  
アカウント発行  
コード化  
データクリーニング  
症例検討会  
他

## 他からの資料

治験実施計画書  
クオリティマネジメント手順・計画書  
解析計画書  
副作用報告書  
モニタリング報告書(必要時)  
施設基準値  
EDCトレーニング記録  
TLGs  
総括報告書  
抄録  
他

## 外部ベンダー

見積・契約書  
ベンダーからの成果物

## DMの基本となる資料

データマネジメント計画書  
データマネジメント報告書

## EDC関連資料

EDC要求仕様書  
チェック仕様書(マニュアル、コンピュータチェック)  
EDC仕様書  
バリデーション計画書  
バリデーション手順書  
バリデーションチェックリスト  
バリデーション結果一覧  
バリデーション報告書  
入力マニュアル  
αCRF  
EDCアカウント発行トレーニング記録管理表  
EDC管理シート  
本番環境設定書  
Golive報告書

## プログラム準備

データチェックプログラム  
入力進捗確認プログラム  
グレード判定チェックプログラム

## データ収集中に発生する資料

ロジカルチェック結果  
マニュアル・リステイングチェック結果  
クエリー  
コード化結果・一覧  
入力進捗確認結果  
DLT確認書  
症例固定報告書  
データカットオフ宣言書  
DCF  
署名印影一覧  
監査証跡

## 症例検討会に伴う資料

症例検討会資料仕様書  
症例検討会資料作成プログラム  
症例検討会議事次第  
症例検討会資料  
症例検討会議事録  
症例採否一覧

## 固定に向けた資料

QC確認記録  
QC報告書  
データ固定依頼書  
データ固定解除依頼書  
データ固定報告書  
eCRF  
EDCデータ

## データ移管に伴う資料など

外部データ仕様書  
外部データ  
解析用データセット仕様書  
解析用データセット作成プログラム  
解析用データセット  
データ授受書  
EDC取り扱い依頼書

## 成果物における記録

QC記録  
作業記録

こんなに影響あるの!?



# なぜ、入力マニュアルを作成するのか

## 入力マニュアルの概要

EDC画面の説明や入力の仕方が記載されています。

EDCの操作マニュアル(パスワードの再発行、監査証跡の見方など)が含まれる場合もあります。

入力マニュアルは、入力の際だけでなく、解析担当の方も、どのような形でデータが収集されているのかを確認するために利用しています。

## 作成しなかった場合何が起こるか

### 例1 入力したいのに入力できない

年月日のうち日が不明であった。他で「UN」を入力できたので、入力したいができない。なぜ？

入力欄やVISITを追加したいのに、追加方法がわからない。

### 例2 入力データにばらつきが

小数点第二位を四捨五入してEDCに入力してほしかったが、小数点は切り捨てられて入力されていた。

### 例3 入力データがそのまま解析に利用できず、加工が必要になる

選択肢の内容に該当するが、詳細と思い、その他欄に詳細を入力した。解析時に、その欄を選択肢に人が判断しての読み替えが発生した。複数の記載をその他欄に記載してもらったが、記載の区切りが、「」、「\_」、「」などまちまちで、プログラム処理できず、目で見ても振り分けることになった。

### 例5 業務を引き継いだけれど、どのような基準で入力しているかわからない

なぜ、小数点を四捨五入しているのか。その他欄への入力方法がバラバラで、そのまま良いのか、ルールがわからず、なぜそのようになっているのか、解析への影響など調査する必要が発生した。

### 例6 EDC画面だけでは伝わらない

頻繁にクエリーが発生する項目がある。

例 1日投与量が欲しいところ、1回投与量が入力されるなど各試験でほしい情報が異なるところが、他試験の情報につられて入力される。

EDCの操作方法だけでなく、解析でどのようなデータが必要なかが記載されているのですね。入力されているデータが解析時に困ったり、また頻繁にクエリーが発生する箇所は、入力マニュアルの改定が必要ですね。



# なぜ、データチェック基準を作成するのか

## データチェック基準の概要

EDCに入力されたデータの品質を確認・確保するためのチェック基準です。

この場合の品質とは、解析に利用するデータが漏れなく揃っているか、入力データ内の不整合はないか、GCP違反の症例はないか、選択除外基準に抵触する症例が収集されていないか、個々の症例と全症例のデータを見渡します。

チェックはコンピュータチェック、マニュアルチェック、コンピュータチェックを補助資料とした目視チェックなどがあり、チェック数は300項目を超えていきます。

昨今は、QMの導入により、CRAで確認しているところは、DMでは確認しないケースもあります。

## 作成しなかった場合何が起こるか

### 例1 解析に必要なデータが想定外、もしくはない

試験デザインより投与時の入力方法が複雑だった。入力が正しくされているか、だれも確認できなかったデータがある。  
必要なデータが入力されないことに気が付かず、解析へデータを提供した。(欠損か未入力かわからない)  
小数点第2位四捨五入でデータが欲しいところ、整数値のみ入力されたまま、解析データに用いる事になった。(想定していた解析ができない)

### 例2 逸脱したデータを解析に利用してよいか

固定時に逸脱データを確認したが、どのような経緯で逸脱なのか、もしくは誤入力なのか、DMの記録から判断できない。(CRAで記録されている場合は、問題ない)

### 例3 EDC内の整合性を見落とし

「有害事象に使用」と記載があるのに、有害事象に該当する記載がない。  
死亡日以降の検査値が記載されている。

### 例4 担当者によってチェック方法が異なり、一様なチェックがされていない

Aさん:〇〇まで確認しました。

Bさん:そこまでは、見てないですね。必要ですか？

必要かどうかは、試験によりますが、必要なチェックなのか、過剰チェックなのか

基準がないと、それぞれの担当でバラバラなチェックが実施されてしまいますね。一定のデータ品質確保のためにも、チェック基準はあったほうが良いですね



# なぜ、DM報告書を作成するのか

## DM報告書の概要

DM計画書に沿って実施した内容を報告書にまとめ、逸脱がある場合には逸脱した内容が記載された書類になります。

個々の詳細な記録は業務時に発生した資料にあります。それらを見に行かなくとも、本資料で実施した業務の概要を知ることができます。記載の一部となりますが、入カマニュアルの作成版一覧、使用したシステム、アカウント権限、用いた辞書のバージョン、データ固定日などの日付、成果物の一覧などが纏められています。承認申請時はPMDAへの提出資料となります。

## 作成しなかった場合何が起こるか

例1 一つの資料で、どのようなデータマネジメント業務をしたかを  
確認できない

データ固定をいつしましたか。 →この書類です。  
EDCはいつ稼働開始しましたか。 →この書類です。  
UATが終えたのはいつですか。 →書類を探してきます。

例2 忘れたところに確認依頼がある

5年前の〇〇試験ですが、いつデータ固定しましたか。  
→確か〇〇ごろです。資料探して確認します。  
コード辞書はどのバージョンでしたでしょうか。  
→確か〇〇ですが、EDCが稼働してないのでもうわかりません。  
逸脱はなかったでしょうか。  
→覚えてないですが、資料の山から探してみます。

例3 担当した人がいなくなったら、わからなくなる

5年前の〇〇試験ですが、いつデータ固定しましたか。 →資料探してみます。  
〇〇業務ではこの作業してましたでしょうか。 →わかりませんが、調査します。  
逸脱はなかったでしょうか。 →フォルダの中のファイルを調査してみます。

例4 解析に使用されたデータの品質(収集からデータ固定までの経緯)を  
伝えるのが大変

EDCに入力して頂きました。チェック基準はあります。EDCをUATした記録はこれです。  
EDCの入カマニュアルはこちらで、症例検討会は実施しました。コード化はこの辞書で……。  
いろいろな資料やデータから伝えなければならない。

当時担当した方でなくても、  
DM報告書があると、どのよう  
な業務をしていたのか、逸脱  
はあったのかなど、説明する  
ことができますね。  
便利ですね



# なぜ、ドキュメントを作成するのか まとめ

- ドキュメントに記載されている内容(品質)で業務を実施している
- ドキュメントがあるから再現性(ある水準の品質の再現性)が取れる
- 解析に利用できるデータを収集するため
- DMメンバーや他セクションの方とコミュニケーションを図るため

要約しすぎてますが  
そういうことなんです



# 承認申請の際に PMDAに提出する 資料

参考までに、となりますが、  
次のスライドでお伝えします。



# 承認申請の際にPMDAに提出する資料

## 各種SOP

データマネジメント  
システム構築・バリデーション  
アカウント発行  
コード化  
データクリーニング  
症例検討会  
他

## 他からの資料

治験実施計画書  
クオリティマネジメント手順・計画書  
解析計画書  
副作用報告書  
モニタリング報告書(必要時)  
施設基準値  
EDCTレーニング記録  
TLGs  
総括報告書  
抄録  
他

## 外部ベンダー

見積・契約書  
ベンダーからの成果物

## DMの基本となる資料

データマネジメント計画書  
データマネジメント報告書

## EDC関連資料

EDC要求仕様書  
チェック仕様書(マニュアル、コンピュータチェック)  
EDC仕様書  
バリデーション計画書  
バリデーション手順書  
バリデーションチェックリスト  
バリデーション結果一覧  
バリデーション報告書  
入力マニュアル  
eCRF  
EDCアカウント発行トレーニング記録管理表  
EDC管理シート  
本番環境設定書  
Golive報告書

## プログラム準備

データチェックプログラム  
入力進捗確認プログラム  
グレード判定チェックプログラム

## データ収集に発生する資料

ロジカルチェック結果  
マニュアル・リスティングチェック結果  
クエリー  
コード化結果・一覧  
入力進捗確認結果  
DLT確認書  
症例固定報告書  
データカットオフ宣言書  
DCF  
署名陰影一覧  
監査証跡

## 症例検討会に伴う資料

症例検討会資料仕様書  
症例検討会資料作成プログラム  
症例検討会議事次第  
症例検討会資料  
症例検討会議事録  
症例採否一覧

## 固定に向けた資料

QC確認記録  
QC報告書  
データ固定依頼書  
データ固定解除依頼書  
データ固定報告書  
eCRF  
EDCデータ

## データ移管に伴う資料など

外部データ仕様書  
外部データ  
解析用データセット仕様書  
解析用データセット作成プログラム  
解析用データセット  
データ授受書  
EDC取り扱い依頼書

## 成果物における記録

QC記録  
作業記録

他にも要求がありましたら提出しています。

予想より多い……



# 概要

- よくいだかれる疑問
- データマネジメントがしている事
- データマネジメント業務を支える資料、プログラムなど
- データマネジメントは何を目的に実施しているのか
- **なぜ、〇〇ドキュメントを作成するのか**
- ドキュメントを基にした作業の紹介(一部)
- 掛かる期間のお話
- 求めたい知識やスキルのお話
- まとめ

赤字のところまで  
終わりました



# ドキュメントを基にした作業の紹介(一部)

- 実施計画書のレビュー
- データクリーニング(1)～(6)
- データのコード化
- SAE報告書とのリコンシレーション
- データの固定

お腹いっぱいです！！  
でも、頑張って聞きます



# 実施計画書のレビューは どのような観点で行うのか

実施計画書に収集する旨の記載がないデータは収集することができないため、有効性・安全性に関するデータなど解析に必要なデータがEDCに収集できるか、という観点で確認します。

- 例1 解析すると記載されているが、治験スケジュールや収集項目に記載がない。
- 例2 治験スケジュールと、収集項目に齟齬がある。
- 例3 収集したいデータの詳細が記載されていない。(例えば小数点第●位まで収集したい)
- 例4 プロトコル内容からどのようにデータ収集する想定か (EDC収集、外部データ? など) 推察できない。
- 例5 secondary Endpointの例えば治療完遂割合など用いる完遂の定義に漏れていないか。  
(〇〇の事例が発生した場合、どのデータを収集しておけば大丈夫か、そのデータは収集項目に記載されているかなど。)

実施計画書に記載のないデータは  
収集できませんね。  
漏れないよう確認なくてはですね。



# データクリーニングはどのような基準で実施しているのか(1)

DMは、カルテを見ることはできません。カルテを見ると先入観でEDCを見てしまうため、独立しています。

その為、主にEDCデータ内で確認しているのですが、下記のような例を紹介できればと思います。

- (1) 選択、除外基準に抵触するデータの確認
- (2) EDCデータ内の整合性を確認
- (3) 欠損？の確認
- (4) コード化できないデータの確認
- (5) 施設間でのデータの差異

カルテを見れないのはわかりましたけど  
DMの人がどのような観点でEDCデータを見て  
いるのか  
いまいちわからないのよね・・・



# データクリーニングはどのような基準で実施しているのか(2)

(1) 選択、除外基準に抵触するデータの確認をしています。

無事登録されたけれど、確認したら選択、除外基準に抵触しているデータが...

① 疾患に選択・除外基準に抵触する記載が...(高血圧)

既往症又は合併症がありますか？		
#	既往症・合併症名	継続していますか？
1	高血圧	はい

コントロール可能な高血圧かどうかは、表記のデータからだけではわからないため、クエリー確認？

② ベースラインに除外基準に抵触する異常値が...

データ	
Hb	7.9
血小板数	74000
白血球数	8000
好中球数	1700

登録時は問題なかったのか、EDCデータからはわからないため、クエリー確認？

③ ベースラインに、除外基準に抵触する異常値が...

項目名	測定実施状況	測定結果
体温	<input checked="" type="radio"/> 未実施	37.1
脈拍数	<input type="radio"/> 未実施	90
収縮期血圧	<input type="radio"/> 未実施	150
拡張期血圧	<input type="radio"/> 未実施	100

登録時は問題なかったのか、EDCデータからはわからないため、クエリー確認？

なるほど...ね



注：該当箇所のみ情報だとクエリーとなりますが、他のデータも参照して確認するため、事例として用いた内容は、クエリーとはならないケースが多いです。

# データクリーニングはどのような基準で実施しているのか(3)

(2) EDCデータ内の整合性を確認しています。

## 薬剤の例

薬剤/療法名	AAA薬剤
投与経路	経口
試験開始前に投薬/治療を開始していた場合にチェック	<input type="radio"/> はい
開始日	2024 3 24
試験終了後も投薬/治療を継続している場合にチェック	<input checked="" type="radio"/> はい
終了日	
併用理由	<input checked="" type="radio"/> 有害事象 <input type="radio"/> その他

使用理由に有害事象が選ばれているものの、添付文書内の有害事象の記載がなく、有害事象を特定できない。

開始日が実は「試験開始前に投与」に該当する日付だった。

## 中止記録の例

中止判定日または完遂日	2024 5 24
中止理由	<input type="radio"/> 完遂 <input type="radio"/> 原病の増悪 <input checked="" type="radio"/> 有害事象 <input type="radio"/> 患者希望 <input type="radio"/> プロトコール治療中の死亡 <input type="radio"/> その他、治験担当医師が治験の中止を必要と判断
その他詳細	

有害事象にチェックされているが、有害事象欄に「被験薬投与に関する処置」が中止の有害事象がない。該当する有害事象の発現日が中止日より後ろの日付だった。

## 有害事象の例

有害事象名	白血球減少
-------	-------

EDCで収集している検査項目だけれど、EDCデータで該当するデータの確認ができない。誤記？

日付だけじゃないのね



注: 該当箇所のみ情報だとクエリーとなりますが、他の情報があれば、その情報も確認するため、事例として用いた内容は、クエリーとはならないケースもあります。(例 SAE報告書や事前に展開されたメールなど)

# データクリーニングはどのような基準で実施しているのか(4)

(3) 欠損？の確認もしています。

例

各VISITにおけるEDCデータの有無										
ケース		ベースライン	VISIT1	VISIT2	VISIT3	VISIT4	中止時or治験完遂	追跡1	追跡2	状況
A症例	問題ないケース	あり	あり	あり	あり	あり	治験完遂	あり	あり	生存
B症例	[REDACTED]	あり	あり	あり	あり	あり	治験完遂	なし	あり	生存
C症例	[REDACTED]	あり	あり	あり	なし	あり	治験完遂	あり	あり	生存
D症例	[REDACTED]	あり	あり	なし	—	—	VISIT2より後で中止	あり	あり	生存
E症例	[REDACTED]?	あり	あり	あり	—	—	VISIT2の時点で中止	なし	なし	生存

A症例: VISITデータの欠損はない

B症例: 追跡2のデータがあるので、追跡1のデータもあるはず

C症例: VISIT3のデータはあるはず

D症例: VISIT2のデータはあるはず

E症例: 生存しているので、追跡1、2のデータはあるはず

※DM担当者もEDCをぱっと見ただけでは、クエリーで問合せるべきか、判断できないことが多いため、他データとの関係性、VISITのアローワンス、例えばメールなどで事前情報ないか、など調べます。

なぜEDCに入力されていないかEDCデータだけではわからないのですね  
確かに、他に情報がないとですね



# データクリーニングはどのような基準で実施しているのか(5)

(4)コード化できないデータの確認をします。

コード化するため、クエリーにより詳細確認する

- 例
- 投与経路が不明の場合 : アスピリン、ロキソプロフェンNa(内服 or 注射) → クエリーで投与経路や用途など確認
  - 配合剤の場合 → 状況によりクエリーで分割入力など確認
  - 総称などが用いられている場合 : カテコラミン(実際は、どの薬剤を使用している?) → クエリーで詳細確認
  - 誤記などより明確に●●薬剤と特定できない場合 → クエリーで詳細確認
  - 記載名称では特定できない場合 : しびれ → 部位などクエリーで確認  
: 胃切迫穿通 → コード辞書にある似た用語よりクエリー確認

コード化をあきらめる例

- OTC: 例 パブロン 熊の胆 など
- 治験薬(コード辞書に登録されていない)
- 販売直後の薬剤(コード辞書にまだ登録されていない)

集計時にどのように集計するか検討が必要  
集計が必要ないデータはコード化は不要



# データクリーニングはどのような基準で実施しているのか(6)

(5) 施設間でのデータの差異も確認します。

集計結果も、見た目も、大丈夫そうだけど・・・

中央値	最大	最小
50.45	51	50

施設	VISIT	データ
施設A	VISIT1	50.5
施設B	VISIT1	50
施設C	VISIT1	50.3
施設A	VISIT2	51
施設B	VISIT2	50
施設C	VISIT2	50.3
施設A	VISIT3	50.5
施設B	VISIT3	51
施設C	VISIT3	50.7
施設A	VISIT4	51
施設B	VISIT4	50
施設C	VISIT4	50.4

A施設  
入力が0.5刻みです

B施設  
小数点が入力されてないです

本当にそうかもしれませんし、実は小数点第1位で加工化されているデータが入力されているかもしれないです。

もし少数点第1位が入力されていないとしたら、特に有効性に用いる場合、評価に影響は・・・品質は大丈夫だろうか。

そんなことあるの!?



# データのコード化はなぜ行うのか さっきも出てきたけど、コード化って何！？

## 有害事象 コード化の例

### 入力例

no	事象名
1	発熱
2	発熱 (38.2)
3	発熱
4	発熱
5	発熱
6	発熱
7	発熱
8	発熱 (37.8)
9	発熱
10	発熱

### プログラム集計

事象名	件数
発熱	7
発熱	1
発熱 (38.2)	1
発熱 (37.8)	1

同じ発熱なのに別で集計されます。  
これぐらいなら、目で見ればわかるので、手カウントも可能ですが、100事象もあると…



### コード値の付与

no	事象名	コード
1	発熱	発熱
2	発熱 (38.2)	発熱
3	発熱	発熱
4	発熱	発熱
5	発熱	発熱
6	発熱	発熱
7	発熱	発熱
8	発熱 (37.8)	発熱
9	発熱	発熱
10	発熱	発熱

### コード値を使った プログラム集計

事象名	件数
発熱	10

正しく集計されました。

ポッポポ  
(プログラムで読める  
言葉に変更する必要  
があるのですね)



有害事象・既往症/合併症にはMedDRAというコード辞書が用いられます。承認申請時にMedDRAを用いたコード化が必須となります。

また、コードを用いる事により集計は問題なくできるようになりますが、誤記の場合、CSRに誤記のまま記載されるので、悩ましいところです。

# データのコード化はなぜ行うのか さっきも出てきたけど、コード化って何！？

## 薬剤 コード化の例

### 入力例

no	薬剤名
1	アスピリン
2	ロキソプロフェンN a
3	アスピリン注
4	ロキソプロフェンナトリウム
5	アスピリン
6	ロキソプロフェンN a
7	アスピリン

### コード値の付与

no	薬剤名	コード
1	アスピリン	アスピリン
2	ロキソプロフェンN a	ロキソプロフェンN a
3	アスピリン注	アスピリン
4	ロキソプロフェンナトリウム	ロキソプロフェンN a
5	アスピリン	アスピリン
6	ロキソプロフェンN a	ロキソプロフェンN a
7	アスピリン	アスピリン

DMは集計を意識してコードを付与します。  
実際のコードは数値とそれに紐づく薬剤名を持っています。  
また、コード化することでプログラムチェックにも活用可能となります。

ポッポポポ  
(ロジカルチェックに  
用いられることもある  
のか)

### プログラム集計

薬剤名	件数
アスピリン	3
アスピリン注	1
ロキソプロフェンN a	2
ロキソプロフェンナトリウム	1

### コード値を使ったプログラム集計

薬剤名	件数
アスピリン	4
ロキソプロフェンN a	3

薬剤ではJDRUG、承認申請時にはWHODDというコード辞書が用いられます。

コードを用いる事により集計は問題なくできるようになりますが、誤記の場合、CSRに誤記のまま記載されるので、悩ましいところです。



# SAE報告書とのリコンシレーションはなぜ行うのか

SAE報告書とEDCデータの収集はタイミングが異なります。また、必要なデータも異なります。そのため、不整合が起こりやすいデータとなります。

症例	有害事象名	発現日	転帰確認日	転帰	因果関係	措置	データ
...	食欲不振	2020/06/04	2020/07/09	回復	関連あり	減量	EDC
	食欲不振	2020/06/04	2020/07/06	軽快	関連あり	減量	SAE報告書
...	細菌性肺炎	2021/06/19	2021/07/13	軽快	関連なし	休薬	EDC
	細菌性肺炎	2021/06/19	2021/07/13	軽快	関連なし	該当せず	SAE報告書
...	気胸	2021/07/18	2021/08/26	軽快	関連なし	休薬	EDC
	右気胸	2021/07/18	2021/08/26	軽快	関連なし	中止	SAE報告書

どちらのデータが間違っているのか、または、どちらも正解か？EDCに必要なのは、治験実施計画に定めてあるデータとなり、解析に利用するデータとして問題がないのか、確認が必要になります。

総括報告書にどちらも記載されるので、SAE報告書とEDCデータで不整合がある場合、状況により監査指摘になります。



# データの固定はなぜ行うのか

すべてのデータが収集・クリーニングされた後や特定のポイントでデータの固定を行います。

データ固定中は、そのデータの追記・修正など実施できないため、データを追記・修正などしてほしくない際にデータの固定をします。また、●●時点のデータと明示するためにデータの固定をする場合もあります。

## 例

①データ収集・クリーニングされ、データの変更が不要と判断された際、そのタイミング以降、データの変更ができない様に、データの固定をします。

最終的なデータを指し、解析やCSRに用いられます。

②プロトコールに定められている解析に使用するデータを特定するために、データの固定をします。

解析で使用したデータがどれであるか、特定できるようデータの固定をします。

③試験途中で何かしらの要因でデータの修正をしてほしくない際にデータの固定をします。

抄録や論文用にデータを利用する際などデータ固定する場合があります。

注: 固定データの中に間違いを見つけた際は、状況により固定解除しデータ修正します。

最終的なデータの場合は、EDCによる修正が困難な場合があるため、EDCデータを修正せず、DCFでデータを収集し、DCFの内容を基に、解析を進めていきます。

メモ: データの固定をしない場合、解析に●●時点のデータを使いました。とはいえものの、確認したデータが修正されている可能性や、追記されている可能性もあり、予期せぬデータを解析に使用してしまう恐れが。

何かに利用するデータを特定するため、データの固定を  
するのです。  
解析時にどのデータを用い  
たかがわからなくなったら、  
再現できませんからね



# 記録の必要性

## DMは何故記録にこだわるのか

質問例	回答
EDCにはどのようなルールで入力されましたか	入力マニュアルに記載されています。また、入力マニュアルは、●●によって承認され、版管理されています。
解析にはどのデータを使用しましたか	●●日にデータカットオフし、●●日に固定したデータを使用しました。データカットオフした日付、固定した日付は、●●の書類で確認できます。 また、データは、DVD-Rに保存されているデータを用いており、移管・提供した記録はこちらになります。
システムは正しく動いていますか	要求仕様に合わせて構築されていることを、UATで確認しました。UATの計画や手順、記録は、こちらになります。
データの整合性は確認されていますか	チェック仕様を基にデータの整合性確認を行い、必要に応じてクエリーなどで確認しています。チェック仕様とクエリー確認の記録は、こちらになります。
そのデータは正しいですか	システムの確認はUATで、データの確認は、チェック仕様を基に実施しております。DM観点での問題点は解決したデータとなっております。
再現性はありますか	DMPや入力マニュアル、EDCの構築ドキュメントなどありますので、それらを基に再度収集すれば、再現性はあります。(実際に再現できるかは、不明だが、各ドキュメントがないと不可能)
どのような作業してデータの品質を担保しましたか	DMPで定めた書類やそれに記載されている手順を実施しました。また、イレギュラーな対応はDMRなどに記載しており、それらのドキュメントに従った担保をしております。
同じ電子データ(DVD-R 2枚)がありますが、どちらが原資料ですか。	こちらのDVD-R(A)がマスターで、こちらのDVD-R(B)がその写しです。そのため、DVD-R(A)が原資料ですDVD-R(B)がその写しです。

様々な内容や状況を説明する際の根拠資料として記録が大事なのですね。



上記のようにデータの品質を保証するためには根拠となる記録がないと説明ができない。

# 概要

- よくいだかれる疑問
- データマネジメントがしている事
- データマネジメント業務を支える資料、プログラムなど
- データマネジメントは何を目的に実施しているのか
- なぜ、〇〇ドキュメントを作成するのか
- **ドキュメントを基にした作業の紹介(一部)**
- 掛かる期間のお話
- 求めたい知識やスキルのお話
- まとめ

赤字のところが  
終わりました



# 掛かる期間のお話

## 例

- EDC構築でき、入力できるまでの期間は
- データクリーニング完了までの期間は

どれどれ



# 掛かる期間のお話(1)

## 例 EDC構築がされ、入力開始できるまでのざっくりスケジュール

ヶ月	0	1	2	3	4	5	補足
治験実施計画書レビュー	■						EDCに収集する項目の確認、内容の理解を深める
治験実施計画書FIX	■						
EDC構築ドキュメント作成		■	■				原則は、プロトコルがFIXしてから作成を開始する プロトコルが大幅に変更になると、再作成となるため入力項目の設定とEDCに設定するチェック基準の設定が必要(入力400項目、チェック500項目など)
EDC構築			■	■			EDC構築ドキュメントに沿ってEDCを構築する 一部分ですがEDCドキュメント作成と並行して作成可能である
EDC動作確認(UAT)					■		構築されたEDCが正しく動作するのか、確認する 構築途中で実施すると、実施した部分が修正される可能性もあるため、基本は纏めて実施する
EDCFIX					■		
入力マニュアル作成						■	EDC画面を変更すると再作成の可能性があるため、EDCがFIXしてから作成する
トレーニング・ユーザーID発行						■	入力マニュアルを基にトレーニングなどの後に、ユーザーID発行可能となる 入力開始後にユーザー追加・削除などあれば逐次実施する
入力開始						■	

結構かかるな

※ スケジュールは試験の複雑さや収集項目数の影響を受けます。



# 掛かる期間のお話(2)

## 例 データカットオフからデータ固定されるまでのざっくりスケジュール

ヶ月	0	1	2	3	4	補足
データカットオフ	■					この日までのデータをEDCで収集しますという宣言をする
データ入力完了		■				データカットオフまでの入力状況で期間は伸び縮みする
データクリーニング	■	■	■			データカットオフまでの入力、クリーニング状況で期間は伸び縮みする
SDV	■	■	■			データカットオフまでの入力、SDV状況で期間は伸び縮みする
クエリー対応など	■	■	■			データカットオフまでのクリーニング、SDV状況で期間は伸び縮みする
症例固定			■			データの修正が必要がないと判断された際に、「症例固定」というステータスを宣言する
症例検討会				■		症例固定後に症例検討会の議事次第や資料を作成開始する プログラムなどは事前に準備できる範囲もあるが、データ変更されると、資料の再作成が発生するため、原則、症例固定後に資料作成を開始する
EDCデータ固定					■	症例検討会后、「データの修正の必要がない」と判断されたら、データが固定される

※ スケジュールは試験データの複雑さや症例数、データカットオフまでにどの程度データが入力・確認がされているかの影響を受けます。

確かにかかるかも・・・



# 概要

- よくいだかれる疑問
- データマネジメントがしている事
- データマネジメント業務を支える資料、プログラムなど
- データマネジメントは何を目的に実施しているのか
- なぜ、〇〇ドキュメントを作成するのか
- ドキュメントを基にした作業の紹介(一部)
- 掛かる期間のお話
- 求めたい知識やスキルのお話
- まとめ

赤字のところが  
終わりました  
あともう少しです。



# 求めたい知識やスキルのお話

## 疾患領域に関する内容

医学・薬学に関する知識  
癌治療に関する知識  
治験実施計画書で対象となっている疾患への知識(医療機器、再生医療などを含む)  
実診療に関する知識  
RECIST  
CTC-AE

## 臨床研究に関する内容

各省令・指針・通知  
ICH  
治験実施計画書  
総括報告書

## データに関連する内容

個人情報  
データ品質(トレサビリティ)  
データベース(リレーショナルデータベース)  
CDISC  
コーディングおよびコーディング辞書  
解析手法におけるデータの在り方  
網羅性、正確性、整合性、安全性  
データ分析  
データ洗練

記載を省いたものもあります

## システム・アプリ操作に関する内容

ER/ES  
コンピュータ化システムバリデーション(CSV)  
利用するシステムへの認識(何ができて何ができないか、DCT含む)  
監査証跡(どのように取れて、どのように見れるか)  
プログラミング  
EXCEL、WORD、POWERPOINTなど(各ドキュメント作成や、データレビューに必須)  
真正性、見読性、保存性  
ITに関する基礎知識  
セキュリティ

## その他のワード

PDCA、OODA  
品質マネジメント  
ヒヤリハット  
リスク回避  
問題点の分析・解決能力(なぜなぜ分析など)  
進捗、スケジュール管理・調整  
伝わるコミュニケーション  
ロジカルシンキング  
目的意識  
自己チェック(作業などの正確性)  
効率化  
契約  
イレギュラー対応力  
プロジェクトマネジメント  
想定力

思ったより  
多くない!?



# まとめ

医師主導治験におけるデータマネジメントはデータの確認だけではなく、色々なことをしています。

本資料では全部は紹介できておりませんがデータマネジメントを担っている方々がどのような業務をしているのか、少しでも伝われば幸いです。

なっ何とか走り切れましたけど、全部は覚えられません。

とにかく色々しているのはわかりました





ご清聴  
ありがとうございました

