

データマネジメント

CBI公開講座「臨床試験方法論」
東京大学医学系研究科
Clinical Bioinformatics 研究ユニット
大津 洋



(再)臨床試験を学ぶ

- 方法論

- 生物統計学 1/21 大橋
- データマネジメント 2/4 大津

- 実施

- 資金・組織・運営 2/18 甘利 1/7 荒川
- プラニング
- デザイン 1/14 大橋・木村(プロトコル)
- 実施 2/18 高木(安全性)・2/25坂口(QC/QA)
- 解析 1/28 松山
- 報告 2/25 木村

- 薬事・行政

- 倫理

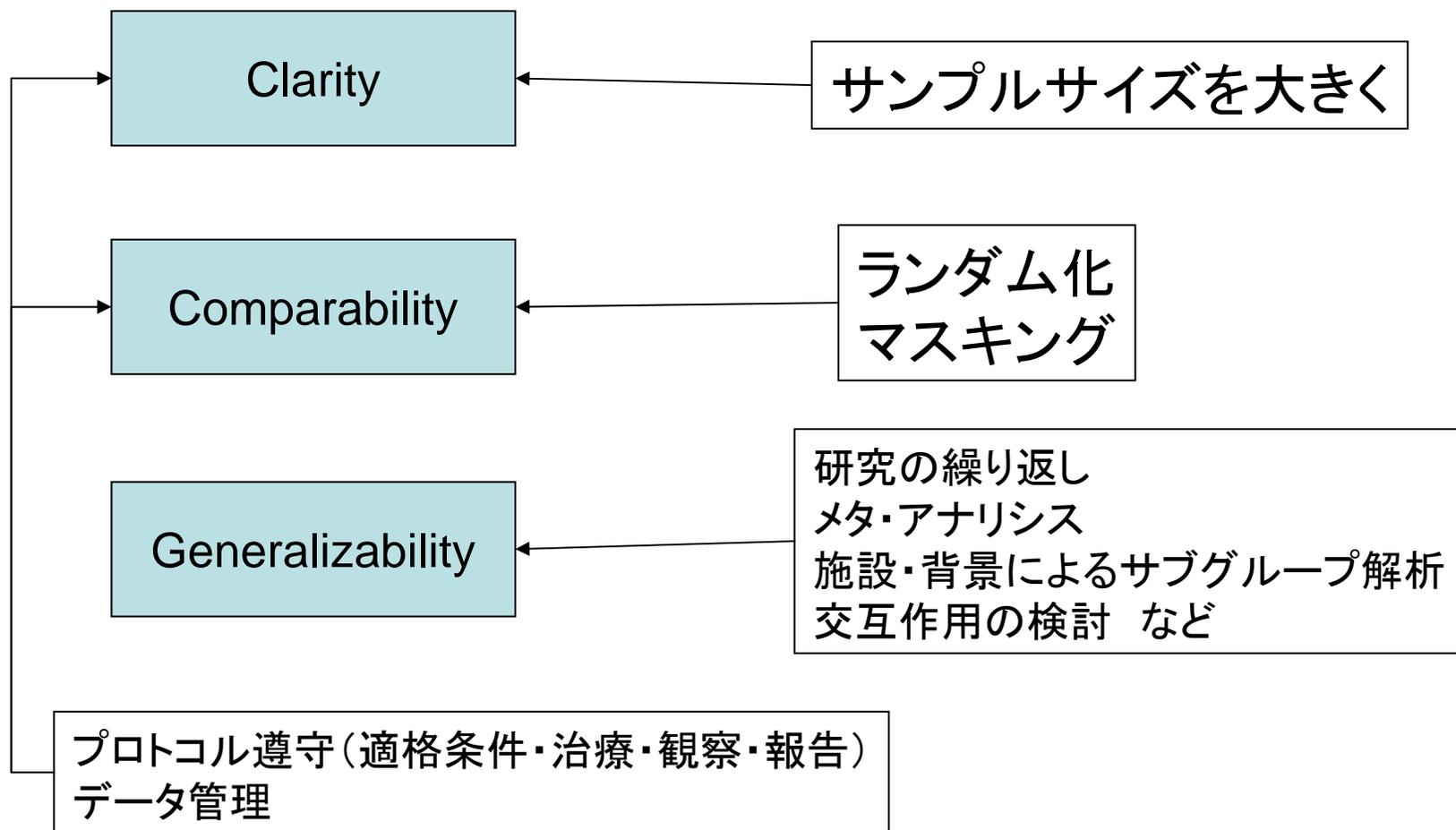


Issue

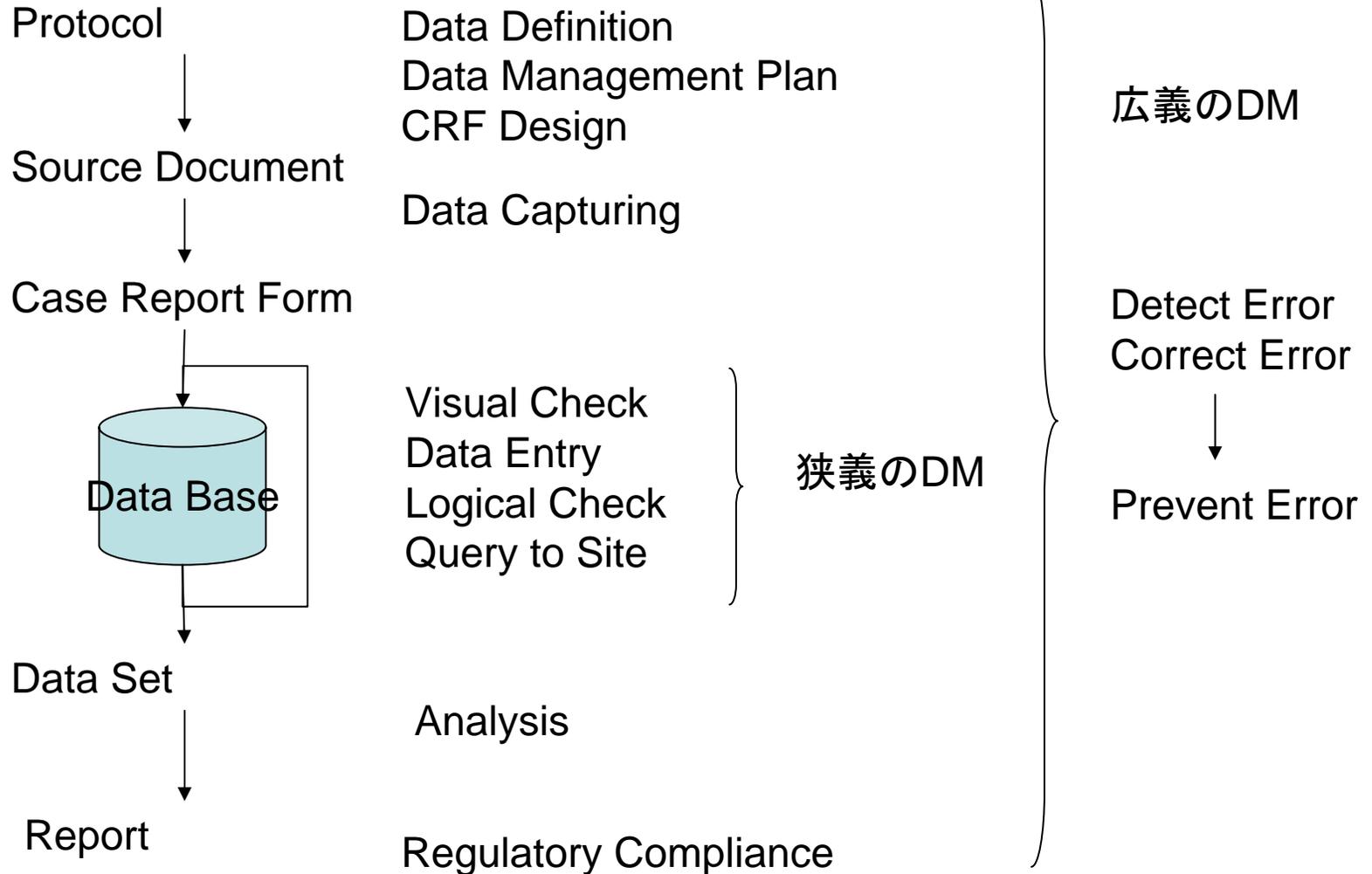
- Introduction
- ICH-GCP E9を眺める
- DM
- 統計解析
- 終わりに



研究(計画・実施)に要求される条件



(広義の) データマネジメントとは？



データマネジメントに対する認識

- CRF(Case Report Form)の情報をコンピュータに入力しておけばいいんじゃないの？
 - 入力するのは誰でも構わない。
 - 必要だと思われるデータはすべて取っておきたい。



さらに困ったことに

- 臨床研究のプロセスの中で、重要だけど脚光があたりにくい仕事。
 - データマネージャ(DM)の仕事が未定義。
 - DM = データ入力者
 - DM = 統計家(広義で言えば入るけど....)
 - DM = コンピュータに強い人
 - DM = うるさい人?
 - 実担当者のモチベーションが上がらない。
 - QOLの低下
 - 医学研究に対する評価が高くない。



統計解析に対する認識

- ・ 統計担当者は計算屋さんですか？
- ・ 統計担当者は最終判定者ですか？
- ・ 統計処理は論文の飾りですか(多くの医学研究者の本心？)
- ・ 大抵の統計学者は難しいと感じ、多くの臨床家は簡単だと考えている学問。足したり、時には2乗したりすることからなる簡単な数学。
[Senn, S(1997), *Statistical Issues in Drug Development*, Wiley]
- ・ (臨床)研究の産婆さん[佐久間 昭先生]



疫学研究では...

ある統計家との話で

- 「疫学研究の統計担当者になっても、実施する仕事の9割がデータを精査・クリーニング。データを見直すことに忙殺される。」
 - データ解析実施の遅れ。
 - 解析までに膨大なやりとり。
 - 論文にできない。 <= 統計家に対する不信。
 - データマネジメントに対する意識が低い。
 - データを集めれば、それなりの結果が得られるという幻想。
 - データが入っていればよい。



疫学研究では...

A)先生、〇〇の解析について相談があるのですが..

B)(話を聞く)

分かりました。じゃあ、△の件で検討しましょう。データを提供してください。

A)分かりました。

.....後日.....

B)先日、話を伺った時に聞いた背景データの集計結果と手元の結果が異なるのですが...

A)あ、先生。忘れてました。このデータは...(長々)

B)(心の中で)じゃあ、どのデータが信頼できるんだ!



臨床研究では(2)

- 例) 倦怠感をあらわす言葉
 - 倦怠感
 - だるさ
 - 体がだるい
- 入力したままだと「倦怠感(1),だるさ(1),体がだるい(1)」となる。
 - 折角集計しても、頭で再集計しなければならない。
 - 集計することが目的になり、うまく整理できていない



データマネジメントが うまく機能していない

- 「この項目を追加してください。」
 - 集計するかどうか分からないけど。
- 参照できるマニュアル類がなく、DM担当と言われる人々の間でコンセンサスができてない。
 - 読み替える人、そのまま入力する人
- などなど。



データ管理：日米比較 (製薬会社の場合)

Melinda Tao(Roche)による

日本

欧米

組織化・責任分掌

未分化・無責任体制

プロトコル策定からの参加

試験開始直前から

標準化

試験毎

Visit-type CRF

データ固定直前の作業増大 e CRF

許容できる品質水準

Over Quality!! 許さない

Fax送付・入力・修正のプロセス管理

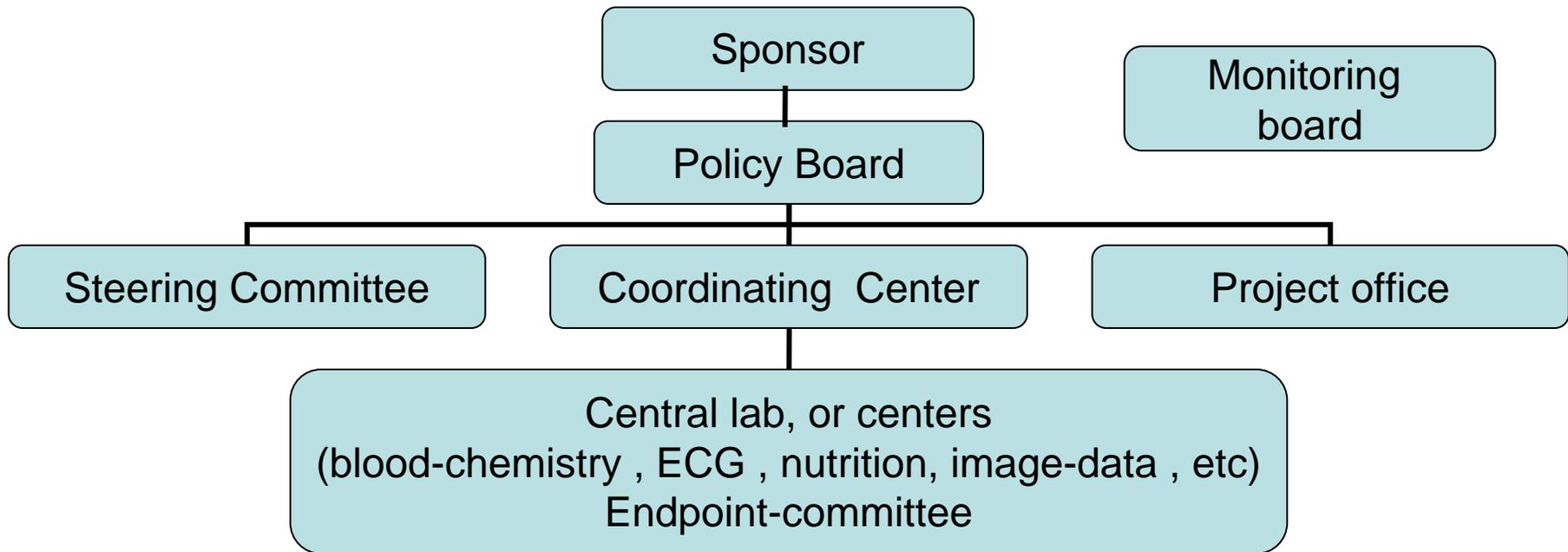
作業者のQOL低下！ 合わせ

過労死！

13



National Heart Lung Blood Institute (NHLBI) Model



3極比較

～データ(統計)センター～

JCOG EORTC SWOG NSABP

資金／年(¥)	0.5 億	5.6 億	2.5 億	?
うち人件費	0.35 億	3.2 億	2 億	?
資金／登録中試験	167 万	487 万	242 万	?

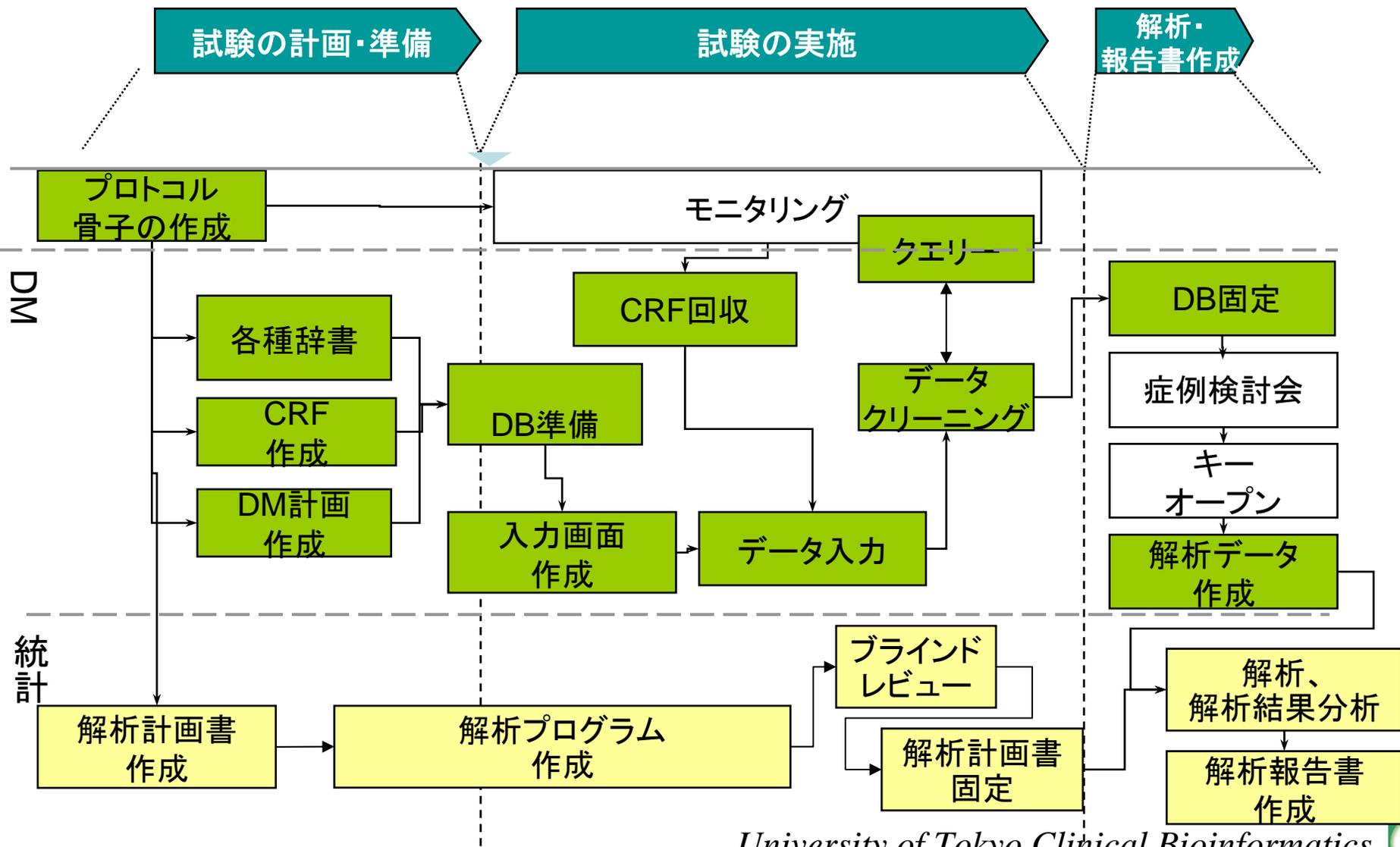
スタッフ

全体	29 (volunteer含む)	65	54	73
full-time	9	60+ α	54	70
Central Data Manager	3+3	22	15	10
統計家(解析担当)	+5	9	15	6
システム管理・サポート	1+3	5+2	10	8
秘書・アシスタント キーパンチャー	2	10	12	18
登録中試験数／FTE	3.2	1.8	1.9	0.1

15



臨床試験プロセスと各役割



医師、統計家、DMの関係

- 医師
 - 日常診療などから得られる視点
 - 「この薬(診療)は有効じゃないか？」
- 統計家
 - マクロな視点
 - 集団全体で効果はどうか？
 - 基本は、「プラセボ(既存治療)と変わらない」
- DM
 - 「臨床の視点」からマクロな視点へジョイント
 - データの標準化(データ収集からデータクリーニングまで)
 - 試験へ関与している方々の考え方の標準化



データマネジメントに必要な知識

CRF設計の場合

	必要となる基礎知識			
	DM実務	IT	統計	臨床
プロトコルとの整合性	●			
記入しやすく、処理しやすく	●	●		
項目配置とデータベース構造の整合性	●	●		
範囲チェック・論理チェック (分布の範囲、精度)				●
品質管理レベル (解析上の必要性)			●	
そもそも必要な項目か？ (必要なら測定頻度は？)			●	●
発生源の信頼性は？	●			●



ICH-GCP E9

「臨床試験のための統計的原則」



ICH-GCP E9とは？

医薬審第1047号より

別添の「臨床試験のための統計的原則」(以下「本ガイドライン」という。)は、ICHにおける合意に基づき、臨床試験における統計的原則について記載されたものであり、臨床試験から得られる結果の偏りを最小にし、精度を最大にすることを目標としている。

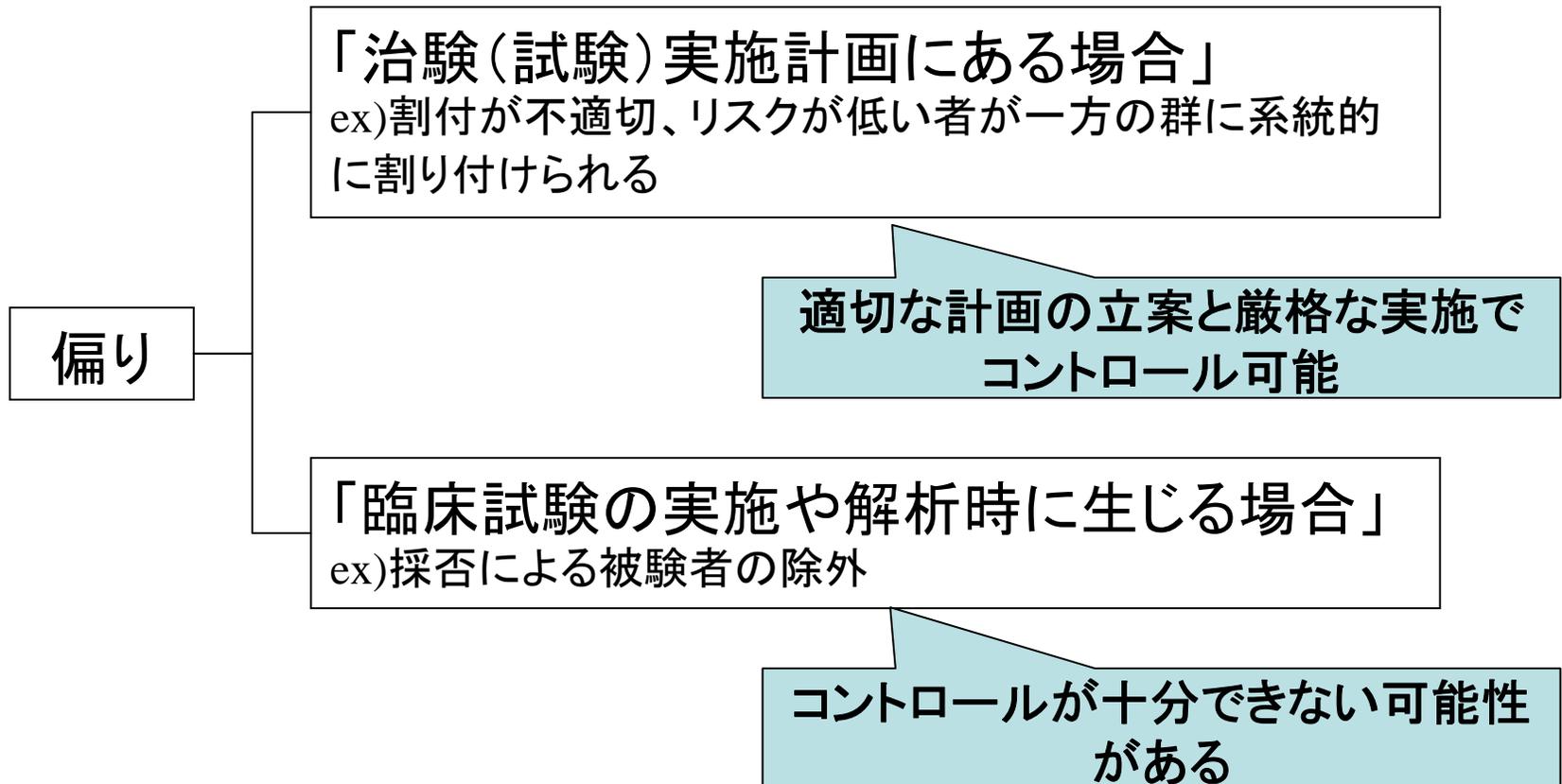


偏り(BioS)

- 偏り(Bias)とは？
 - 臨床試験の計画、実施、解析及び結果の解釈と関連した因子の影響により、試験治療の効果の推定値と真の値に系統的な差が生じること。
- 偏りがあると
 - 臨床試験から妥当性のある結論を導くことが困難になる可能性がある。
- 偏りを低く押さえるためには...
 - 偏りの潜在的な原因を可能な限り明らかにすることが重要。



偏り



適切な計画の立案と厳格な実施 のために

- 適切な試験計画書(プロトコル)の立案
- プロトコルに基づいたDMおよび統計解析計画書の作成
- マニュアルの整備(入力マニュアルなど)
- 検証可能なDB,プログラムの作成
 - 履歴管理
 - DB構造、定義書
 - システム・バリデーション
- 報告書の作成(DM報告書、解析報告書)



(参考)統計ガイドライン

- I.はじめに
 - 1.1 背景と目的
 - 1.2 適用範囲と方向性
- II. 臨床開発全体を通して考慮すべきこと
 - 2.1 試験の性格
 - 2.1.1 開発計画
 - 2.1.2 検証的試験
 - 2.1.3 探索的試験
 - 2.2 試験で扱う範囲
 - 2.2.1 対象集団
 - 2.2.2 主要変数と副次変数
 - 2.2.3 合成変数
 - 2.2.4 総合評価変数
 - 2.2.5 複数の主要変数
 - 2.2.6 代替変数
 - 2.2.7 カテゴリ化した変数
 - 2.3 偏りを回避するための計画上の技法
 - 2.3.1 盲検化
 - 2.3.2 ランダム化(無作為化)



(参考)統計ガイドライン

- III. 試験計画上で考慮すべきこと
 - 3.1 試験計画の構成
 - 3.1.1 並行群間比較計画
 - 3.1.2 クロスオーバー計画
 - 3.1.3 要因計画
 - 3.2 多施設共同治験
 - 3.3 比較の型式
 - 3.3.1 優越性を示すための試験
 - 3.3.2 同等性又は非劣性を示すための試験
 - 3.3.3 用量-反応関係を示すための試験
 - 3.4 逐次群計画
 - 3.5 必要な被験者数
 - 3.6 データの獲得と処理



(参考)統計ガイドライン

- IV. 試験実施上で考慮すべきこと
 - 4.1 治験モニタリングと中間解析
 - 4.2 選択基準と除外基準の変更
 - 4.3 集積率
 - 4.4 必要な被験者数の調整
 - 4.5 中間解析と早期中止
 - 4.6 独立データモニタリング委員会の役割



(参考)統計ガイドライン

- V. データ解析上で考慮すべきこと
 - 5.1 解析の事前明記
 - 5.2 解析対象集団
 - 5.2.1 最大の解析対象集団
 - 5.2.2 治験実施計画書に適合した対象集団
 - 5.2.3 二つの異なる解析対象集団の役割
 - 5.3 欠測値と外れ値
 - 5.4 データ変換
 - 5.5 推定、信頼区間及び仮説検定
 - 5.6 有意水準と信頼水準の調整
 - 5.7 部分集団、交互作用及び共変量
 - 5.8 データの完全性の維持とコンピュータソフトウェアの妥当性



(参考)統計ガイドライン

- VI. 安全性及び忍容性評価
 - 6.1 評価の範囲
 - 6.2 変数の選択とデータ収集
 - 6.3 評価される被験者集団とデータの提示
 - 6.4 統計的評価
 - 6.5 統合した要約
- VII. 報告
 - 7.1 評価と報告
 - 7.2 臨床データベースの要約
 - 7.2.1 有効性データ
 - 7.2.2 安全性データ



データマネジメント



役割

臨床試験データを取り扱うための品質管理を実施し、最終的にその品質を保証すること。

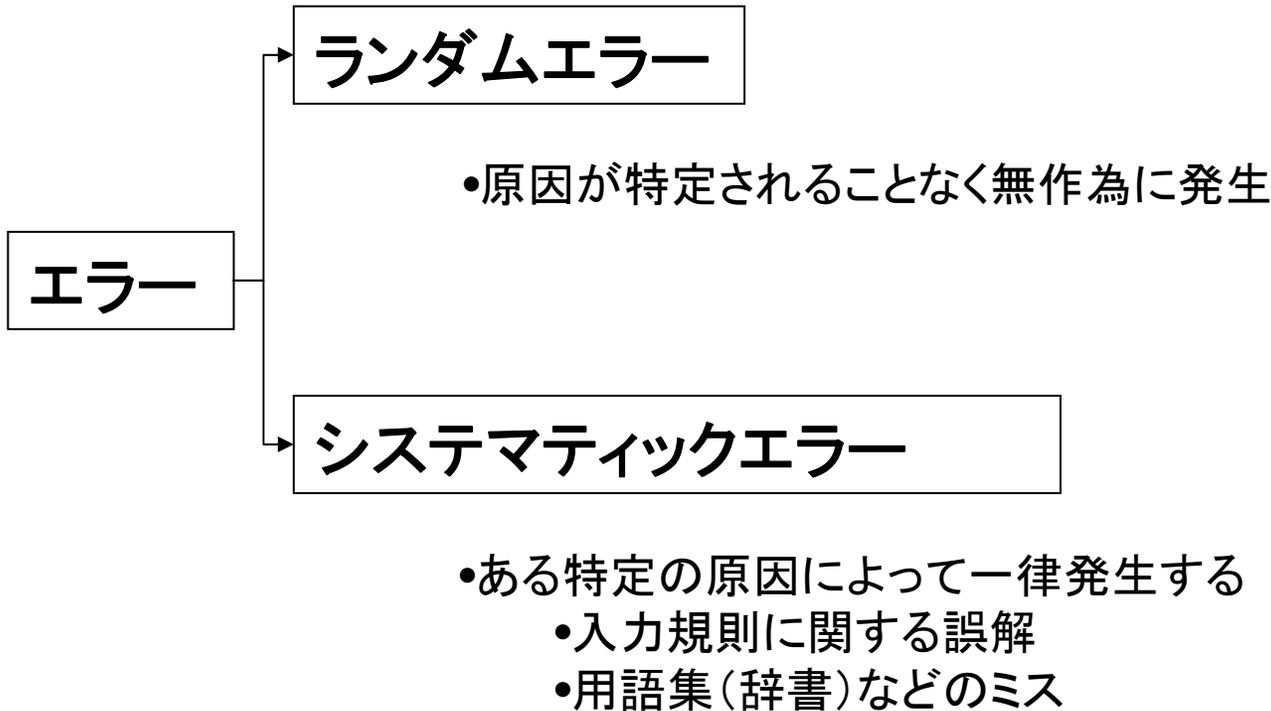


臨床試験では

- 製薬会社
 - ダブルエントリー・トリプルエントリー
 - 読み合わせ
- そもそも臨床試験の質を一定に保つとは？
 - エラーをなくす？
 - エラーを一定以下に抑える？



エラーとは？



エラーの制御は可能か？

- ランダムエラー
 - 100%制御することは難しい。
 - 回避策
 - 書きやすい／入力しやすいCRFデザイン・画面が重要
 - エラーを抽出する方法が重要
 - ダブルエントリー
 - 論理チェック(コンピュータなどの利用)



エラーの制御は可能か？

- システムマティックエラー
 - 担当者の誤解・勘違いによるもの
 - データハンドリングプロセス全体での制御
 - プロトコルからDBまで一貫した整合性の担保
 - 各プロセスにおいて発生するエラーを想定したトラップ
 - プロセスバリデーション
 - コードリスト・マニュアル管理の統一
 - データ入力、チェック担当者への指導
 - 入力・訂正履歴管理をうまく活用する。



エラーの制御のために

- DM計画を立てる。
 - プロセスの検討
 - どんなエラーが発生するのかを想定して、何を制御するのかを基本におく。
 - プロトコルのチェック
 - CRFレイアウト
 - DB設計
 - 入力マニュアル
 - 入力データのチェック



DM計画書

記載例

- 概要(業務範囲)
- 実施体制(責任者、担当者)
- 業務手順
 - 準備までの手順
 - 利用するシステム
 - メタデータ
 - システムバリデーションのポリシー
 - 用意するマニュアル類
 - データ入力の流れ
 - コーディング
 - 有害事象の報告
 - 終了報告
- 資料の保管方法



DM計画書

- 変更不可能文書ではない。
 - 必要に応じて改訂
 - 関係者に周知徹底
 - 必要があれば、マニュアル改訂も実施



Data Cleaning

- 臨床情報システムでは、データ洗淨（データクレンジング）とも言う。
 - データクレンジング
 - データの情報の質を向上させるための一連の作業のこと。
(by Matsuya, 2004 Feb 3 臨床情報システム 講義より)
- ⇒ Data Cleaning と作業内容は同じ！



データ洗淨の手順

(2004.2.3医療情報システム工学講義より)

- Parsing (解析)
- Correcting (修正)
- Standardizing (標準化)
- Matching (マッチング)
- Consolidating (集約)



Parsingの例

(2004.2.3医療情報システム工学講義より)

入力データ

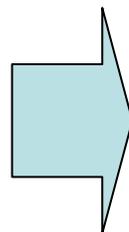
東大 太郎、研究担当部長

XXX製薬(株)

赤門bldg.

本郷一丁目二番三号

文京、東京



解析済データ

姓: 東大

名: 太郎

役職: 研究担当部長

会社: XXX製薬(株)

ビル: 赤門bldg.

番地: 本郷一丁目二番三号

市: 文京

県: 東京都

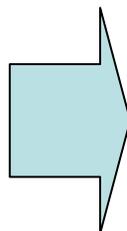


Correctingの例

(2004.2.3医療情報システム工学講義より)

解析済データ

姓： 東大
名： 太郎
役職： 研究担当部長
会社： XXX製薬(株)
ビル： 赤門bldg.
番地： 本郷一丁目二番三号
市： 文京
県： 東京都



修正済データ

姓： 東大
名： 太郎
役職： 研究担当部長
会社： XXX製薬(株)
ビル： 赤門bldg.
番地： 本郷一丁目二番三号
市： 文京区
県： 東京都

郵便番号： 113-8888 41

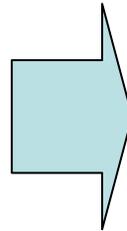


Standardizingの例

(2004.2.3医療情報システム工学講義より)

修正済データ

姓： 東大
名： 太郎
役職： 研究担当部長
会社： XXX製薬(株)
ビル： 赤門bldg.
番地： 本郷一丁目二番三号
市： 文京
県： 東京都
郵便番号：113-8888



標準化済データ

姓： 東大
名： 太郎
役職： 研究担当部長
会社： **XXX製薬株式会社**
ビル： **赤門ビル**
番地： **本郷1-2-3**
市： **文京区**
県： 東京都
郵便番号：113-8888⁴²



Data Cleaning(1)

再登場

- 例) 倦怠感をあらわす言葉
 - 倦怠感
 - だるさ
 - 体がだるい
- 入力したままだと「倦怠感(1),だるさ(1),体がだるい(1)」となる。
 - 折角集計しても、頭で再集計しなければならない。
 - 集計することが目的になり、うまく整理できていない



Data Cleaning(1)

- 「体がだるい」「だるさ」「倦怠感」といった記載用語を
 - どの用語でまとめると理解しやすいか？
 - いつまとめるのか？
 - 入力時なのか？
 - データ入力完了時なのか？
 - On time でよいのか？



Data Cleaning(2)

- 例

- 既往歴10年前に梗塞で倒れた。1994年の12月ごろだったと思う。

- どう入力させるの？[日付型?/テキスト型?]

- 「頃」っていつなのか？

- 医師に問い合わせる

- 患者に問い合わせる ⇒ 大体の人が忘れている



Data Cleaning(2)

- 曖昧な情報をどこまで問い合わせるのか？
 - ルールのない問い合わせは労力だけがかかる。
 - 例えば、1日の違いがどこまで影響を与えるのかをCRF作成段階で考えておく。
- ルールを決める。
 - 上旬 -> 1日と読み替える
 - 読み替えは、常識の範囲で...
 - コンピュータ処理で自動化。
 - 定型処理はコンピュータの最も得意なところ。



Time Window

- 臨床検査値データの取り扱い
 - 2週 -> 14日目に必ず来院するとは限らない。
 - 被験者ごとに採用するのは好ましくない。
 - 期間を定めて、その期間に来院したデータを採用すればよい。
 - 考慮のポイント
 - 妥当な期間は？
 - 解析に与える影響は？



その他考慮すること

- 欠測値の扱い
 - 空白は、「未検査」と「検査したが、入力に値するデータではない」の2つの意味がある。
 - 解析を実施する上で、これら2つの情報は分けて考える必要がある。



ITの活用

- 効率的にDM業務を行うためには、ITを利用する必要性がある。
- では、そのシステムは？
 - In-houseのシステム
 - DM担当者が必要なことはすべて可能
 - 開発に時間がかかる。開発終了時には、新たなテクノロジーが生まれる。
 - パッケージ
 - 製品について理解がないと使いこなせない。
 - システムバリデーションが取られている。
 - ある程度の運用は、比較的短期に可能。



DM System

- 日本の製薬企業で採用されているシステム
 - Global
 - Phaseforward社
 - Clintrial (DM)
 - Inform (EDC) : アカデミックでも利用実績あり。
 - » Children's Hospital of Philadelphia
 - » Duke Clinical Research Inst
 - » Harvard Clinical Research Inst
 - » Mayo Clinic
 - Oracle/CTCLS
 - Oracle Clinical (DM/EDC機能あり)
 - Domestic
 - 日立？富士通？
 - 自作システム



メタデータとは

(2004.2.3医療情報システム工学講義より)

- メタデータとは、「データに関する記述データ（データ辞書）」です。データ構造、データ定義、データの意味、データの出所、発生場所、保管場所、入手時期、データ間の関連性等の情報を保持。
- データ・ウェアハウスの管理やアプリケーション開発に必要な“データに関する技術的な青写真”を提供するもので、データベース内の「リポジトリ」と呼ばれるテーブルに保存。



システム・バリデーション？

ICH-GCP E9 5.8

データの完全性の維持とコンピュータソフトウェアの妥当性について より

治験データを行うコンピュータシステムは十分な検証が行われ、その妥当性が示される必要がある。

- 現在、Windows版のソフトウェアが沢山あるが、きちんと検証されていないものは駄目ですということです。

①信頼性が高い統計解析を行うためには、質の高いデータが必要。

- 信頼のおけるコンピュータシステムを適切に用いたデータマネジメントが業務のすべてにわたってなされ、その結果として、**データの誤りはある限られた範囲内に保証されなければならない。**

② ①で得られたデータについて、信頼度の高い統計解析ソフトウェアを適用することにより、初めて信頼の置ける結果を得ることができる。



品質管理

- 100%ミスはないと言えない。
 - 人間の作業によるため
 - ダブル・トリプルエントリー、読み合わせ
 - 何回チェックしてもよいが、チェックする人のモチベーションの低下によりチェックがうまく機能しないこともある。
- 品質管理の考え方をを用いる。
 - 純度を示すこともよいかもしれない。



EDC



Web-based Trialは救世主か？

- インターネット環境の普及
- すでに数百の実施例
- わが国でも実施例
 - JALSG、京大EBMセンター、HOME-BPなど、一部治験も
- 規制 21CFR Part11 電子申請の実現
- ある外資系大企業はすべてを数年以内にWeb-basedに
- 商品化されたソフトウェアの日本上陸、多くのベンダー
 - DIA/DM 2002JAN24-25、 JUSE CDMセミナー 2002JAN31



Web-based Trialは救世主か？

セントラルでのメリット

ローカルでのメリット

ヒューマンエラーの減少

転記作業削減

問い合わせ減少、モニタリング簡素化

迅速な進捗管理

入力時間と資源の節減

ペーパーレス

ペーパーレスと一元化



EDCを用いたDMCの現状

規模／種類

(2002年末調査時点/東京大学卒業論文より)

	国立大学		財団法人					民間		F	
	A		B	C	D			E			
研究	A-1	A-2#			D-1	D-2	D-3	D-4	E-1	E-2	
試験の種類	臨床試験	治験	臨床試験	治験							
疫学研究											
参加施設(数)	589	4		2	50	80	10	10	50	80	
症例数	4746	10		4※	300	300	100	300	400	800	22
医師主導型	○	○	○	○	○	○	○	○		○	

※目標症例数は、2年間で1800例

#抗がん剤の第1相試験



EDCを用いたDMCの現状

構成要員

(2002年末調査時点/東京大学卒業論文より)

	国立大学		財団法人		民間	
	A	B	C	D	E	F
実務担当医		2				
データマネージャ	2	2	1	(5)	7(1)	8
SE		1	(1)	(2)	8	(2)
プログラマ		(1)		(2)	8	
生物統計家	1	1			1(1)	10
庶務	3	1	1	(1)		1
その他						

()内の数字は、非常勤での担当者



EDCを用いたDMCの現状

～データ入力者～

(2002年末調査時点/東京大学卒業論文より)

	国立大学		財団法人					民間		
	A		B	C	D			E		F
研究	A-1	A-2			D-1	D-2	D-3	D-4	E-1	E-2
医師	○				○	○	○	○	○	○
施設専 属 CRC										○
派遣CRC	○	○		○					○	
その他			○							



日本におけるElectronic Data Capture systemの 現状、開発・導入の留意点及びあるべき姿について 臨床DM研究会 EDC検討班

参加メンバー(製薬会社 7、CRO 3、システム開発 2、大学 1)

(DM以外の経験: システム、他業種のシステム、解析、臨床開発)

伊藤 信行、上原 麻子、卯城 靖子、奥野 潤、白井利明、田村 洋介、友田 裕子、
西川 明宏、畑中 雄介、深澤 一郎、福田 紋子、古幸 美保、水野 享英

(株)シーエーシー、(株)ベルシステム24、キリンビール(株)、日本ワイスレダリー(株)、
三共(株)、アストラゼネカ(株)、第一製薬(株)、(有)電助システムズ、CRCソリューションズ、
武田薬品工業(株)、東京大学大学院医学系研究科 薬剤疫学講座、
日本ベーリンガーインゲルハイム(株)、イーピーエス(株)

2003年5月23日 於 日科学技術連盟

臨床医薬にて出版



日本の薬事規制上の問題点

1. pCRF

- ・当局の指導はpCRF：紙に印刷し記名・捺印(署名)(2003.4現在)
- ・GCPの一部資料に電子媒体の使用を認めた改正通知
「厚生労働省令第36号(H13.3.26)」でもCRFは盛り込まれていない。

2. CRF作成者は医師

- ・J-GCPでCRFは医師が作成(ICH-GCPでは保証)。
- ・CRCは原資料からの転記のみ許される。



薬事行政への働きかけが重要

1. pCRFからeCRF

- ・GLP、GMPでは、既に電子記録、電子署名を受け入れ
 - ・GCP課長通知等では、既にEDCを含む電子記録、電子署名が明記
- 信頼性を保証できるシステムと体制を構築し、業界から働きかけ

2. CRCのCRF記入

- ・Q&Aハンドブック等運用の規制緩和
- ・J-GCPの改定：ICH-GCPとの齟齬を埋め、時代に追いつく



EDCの利用について

- 良い点

- データの信頼性向上。
 - 簡単な問い合わせは、瞬時に可能。

- 悪い点

- 入力担当者は、紙の時よりも1枚のCRFを書き上げる時間が増える。
 - 多くの場合、担当医師の負担増。
- 運営するセンター側は、ヘルプデスクを開設する必要がある。



新たなテクノロジーの利用

- *Teleform* (ハンモック株式会社)
 - Faxサーバー + OCR
 - ベリファイ実行後、DBに格納可能
- PDAを用いたシステム (欧米)
 - データの通信は、電話回線

いずれも一長一短あり。



EDCを行う上での留意点

- ユーザー教育が重要
 - トレーニング
 - マニュアル
 - 入力初期段階のフォロー



統計解析



統計

→ 少数の情報から全体を推論する。

母集団から標本を取り出すことで、少数の情報から効率良く全体を推論することが可能。

検定: 仮説を設けて、その妥当性を確率論的に検証する方法。

推定: 標本の特徴から、母集団の特徴を推論する方法。

→ 情報の基本構造を明らかにする

標本から、得られた情報を数量的に要約したり分類したりすること。

平均・標準偏差を求めてデータを集約することから、様々な方法がある。

Ex) ヒストグラムを作成する。



統計解析

- プロトコル作成時

- 統計解析部分の執筆

- 臨床研究におけるおおまかな統計解析方針の記述
主に、主要評価項目
 - 症例数の設定(症例数とその根拠)

- 統計解析計画書の作成

- 主要評価項目については現段階での計画詳細
 - その他項目については、評価項目と概略



統計解析

- 試験中

- 統計解析計画書の改訂

- ⇒ 試験終了直前に解析計画書の固定

- 解析プログラムの作成。

- プレデータでの解析（解析プログラムのチェック）

- 試験終了

- DMが固定したデータでの統計解析・報告

- プロセス・バリデーション



解析計画書

最低限の記載内容

- 概要
- 実施体制
- 解析対象集団について
 - ITT/FAS/PPS
- 有意水準について
- 解析項目と計画
 - 患者背景解析
 - 主解析項目
 - 副次解析項目
 - 安全性解析
 - その他解析項目
- 解析ソフトウェア

- 解析用データ変換項目とその概略
- 解析帳票レイアウト



解析計画書を作る意義

- 解析方針の明確化
 - どの項目をどのような解析手法を用いるかを明記することにより、解析担当者の視点を明確にする。



解析計画書を作る意義

- 作業の明確化
 - 項目にもれがないことを確認する。
 - 自分のためにも、すべての研究者に対しても。



解析プログラムの作成

- 市販パッケージの場合
 - 本解析前までに操作に慣れておく。
 - 設定が保存できる場合は、設定を済ましておく。
 - 解析計画書通りに設定がされているかチェックを完了しておくこと。
- SASなどでプログラミングする場合
 - プログラム仕様を作成しておく。
 - 本解析前までに、解析計画書に記載されている解析プログラムが作成・チェック完了しておくことが望ましい。



解析ソフトウェア

- 商用ソフトウェア
 - SAS, JMP
 - SPSS
 - Splus
- 自作など
 - Excelマクロ
 - On lineソフトウェア
- 統計解析上考慮しなければならないこと
 - 有効数字
 - 解析名称とロジックの整合性
- 解析精度は、
 - ソフトウェア
 - 実行PCに依存。



再)E9 5.8

ICH-GCP E9 5.8

データの完全性の維持とコンピュータソフトウェアの妥当性について より

治験データを行うコンピュータシステムは十分な検証が行われ、その妥当性が示される必要がある。

- 現在、Windows版のソフトウェアが沢山あるが、きちんと検証されていないものは駄目ですということです。

①信頼性が高い統計解析を行うためには、質の高いデータが必要。

- 信頼のおけるコンピュータシステムを適切に用いたデータマネジメントが業務のすべてにわたってなされ、その結果として、データの誤りはある限られた範囲内に保証されなければならない。

② ①で得られたデータについて、**信頼度の高い**統計解析ソフトウェアを適用することにより、初めて信頼の置ける結果を得ることができる。



解析のプロセスバリデーション

- 項目
 - 解析計画書通りの項目を指定しているか？
 - 統計解析手法に誤りはないか？
 - 適切な位置に出力されているか？
- ダブルプログラミングによるチェックが必要か？
 - サイトポリシーによる。
 - 解析の重要性（主解析かどうか？）



ひとつのアイデア

- 汎用性のあるマクロの利用
 - 患者背景データの解析などに有効。設定を変更するとプロセスバリデーションの再実行のため。流用できる解析用データベースの設定が必要。
- 解析に応じて、レベル設定
 - ダブルプログラミング
 - 目視確認(作成者／確認者)



解析報告書

- 統計解析の観点から、結果（帳票）および考察。
 - 臨床試験では、総括報告書と統合していることも。
 - ただし、総括報告書には解析者のコメントをすべて網羅することができない。
 - ここには、臨床的見地は混入しないように、客観的に事実を述べておく。



最後に



(広義の)データマネジメントのために

- DM教育
 - テキスト
 - 研究会(日本科学技術連盟, SoCRAなど)
 - CBI公開講座？
- 専門家の活用
 - 生物統計／バイオ統計学
- インフラの整備



環境の変化

- CDISC
 - Clinical Data Interchange Standards Consortium
 - 1997年から活動(<http://www.cdisc.org>)
 - 日本支部は、2003年から
 - メタデータの定義[XMLによる]
 - Operational Data Model (ODM)
 - Submissions Data Model (SDM)
 - Analysis Dataset Model (ADaM)
 - 将来は、HL7との融合も予定



環境の変化

- 電子カルテ
 - 院内情報のIT化
 - DWHは浸透するか？

- 治験ビジネス？
 - 治験推進センター？
 - 内容は？



CDMC

(Clinical Data Management Center)

- DM教育、訓練、人材育成
- 標準化されたデータベース、作業(SOP)、システムを利用する
 - テンプレートの整備
 - 標準的なCRF, Data base ,Check内容
 - 解析プログラム
 - 試験のセットアップの円滑化
 - 1試験あたりのコストを軽減



研究者主導臨床試験を推進するグループ

研究者主導臨床試験のあり方を考える会への参加グループ

2001DEC14 (第1回の講義より)

JCOG	日本臨床腫瘍研究グループ
JMTO	日本・多国間臨床試験機構
WJTOG	西日本胸部腫瘍臨床研究機構
JALSG	日本成人白血病研究グループ
TCOG	東京がん化学療法研究会
CSPOR-BC	乳がん臨床研究支援事業
EBMセンター	京大
JCRAC	臨床研究支援センター (メディカルフロンティア:がん以外)

その他、その後	ATOP	骨粗鬆症至適治療研究会
	JACCRO	日本がん臨床試験推進機構
	JGOG	婦人科腫瘍研究機構
	KCOG	関西臨床腫瘍研究グループ
	TRI	臨床研究情報センター(神戸) . . . 84



質の高い臨床研究のために

- 日常のデータをいかに活用するか？
 - 海外論文から引用するだけの国内臨床試験 (RCT) では、定義できないことが多すぎる。
 - 日本での疫学データの利用
 - 臨床DBの逆引きなど (Data Mining)
 - 日常から、「データをきれいにする」という心がけ
 - 診療DBにおいても、「入カールール」などの策定
 - 診療カンファレンスなどでサマリーされた情報をいかに保管しておくか？



臨床研究とは



医師による臨床評価



生物統計家による
適切な試験デザイン

協調
コミュニケーション



メディカルライターによる報告書作成



DMIによる標準化、品質管理



参考文献

DM業務の詳細

辻井 敦著

臨床試験データマネジメント(近刊予定:医学書院)

EDCの現状

臨床DM研究会 EDC分科会

日本におけるElectronic Data Capture systemの現状、開発・導入の留意点及びあるべき姿について(臨床医薬 第19巻 10号)

