

# 解析システムの作成手順

バリデーションがしやすいシステム作成

稲葉 弥一郎  
(東京CRO 株式会社)

# 目次

- 解析対象の統計解析環境の作成
- 解析対象のDB調査
- テストデータの用意
- 解析プログラムの作成
- プログラムテスト
- 解析処理手順書の作成
- 実解析処理
- 実務上での注意点
- 解析パッケージの選択

# 解析対象の統計解析環境の作成

- **テスト環境**
  - **プログラムの作成及び動作検証を行う環境**
- **運用環境**
  - **実データで解析を行う環境**

# 解析対象のDB調査 1 / 3

- データレイアウト
  - 項目の並び順
  - 項目のデータ形式
    - 文字、数字(小数部も)、最大長、最小長
  - 全ての項目のデータ例
  - 項目の変数名の決定
    - 一番のお勧めは、入力システム作成時に決める
    - 全て英数字にする

# 解析対象のDB調査 2 / 3

- カテゴリカル・データ
  - 入力システムのマスターと関連がある
  - 測定値から作る場合もある
    - 解析計画書に書かれている(年齢、体重、BMI等)
- 測定値・データ
  - 臨床検査値の場合
    - 正常範囲を調べておく(グラフでは必須)
    - 男女で正常範囲が違うものがある
    - 施設別で正常範囲が違う

# 解析対象のDB調査 3 / 3

- PPSとFAS作成するためのDB作成
  - カットフラッグDBと呼ぶ
  - 基本的に、PPSとFAS用の2種類はある
  - 上記の、変則パターンが発生する場合もある
- 症例検討会で決まったことを、このDBに反映させる

# テストデータの用意

- プログラム作成者が作成
  - 特殊条件データを組み込む
- プログラム作成者以外が作成
  - プログラムを良く読み、データを作成する
- ノーマル条件のデータ作成
  - 収集中のデータでも良い(DMに作ってもらう)

# 解析プログラムの作成 1 / 6

- 入力システムから受け取ったCSVファイルを、入力基本DBに変換する
  - 入力システムから受け取ったCSVファイルが変更されない限り再処理の必要性は無い
- 入力基本DBからPPSとFASの基本解析DBを作成する
  - カットフラッグDBを使って基本解析DBを作成
  - カットフラッグDBが変更されたら再処理をする

# 解析プログラムの作成 2 / 6

- 解析DBを作成
  - PPSとFASの基本解析DBから解析DBを作成
- 解析DBの詳細編集処理
  - カテゴリーズ、計算項目など解析に必要な項目の追加や、計算しやすい形への編集
  - 解析途中で、追加項目が発生した場合は、このプログラム群の修正をする

# 解析プログラムの作成 3 / 6

- **データ抽出**
  - 施設別などの解析は、ここでデータ抽出をする
  - 全データを扱う場合も、データ抽出と考える
  - 層別解析はここで言うデータ抽出ではない

# 解析プログラムの作成 4 / 6

- 解析プログラム構造 1

- 背景因子関連

- クロス集計
- 基本統計量
- グラフ
- 必要に応じて検定

# 解析プログラムの作成 5 / 6

- 解析プログラム構造 2
  - 臨床検査値関連
    - クロス集計
    - 基本統計量
    - グラフ
    - 必要に応じて検定

# 解析プログラムの作成 6 / 6

- 解析プログラム構造3
  - 有害事象関連
    - クロス集計
    - 一覧表
    - 関連性有り無し等の集計
- エンドポイントの解析
  - 指定された解析
    - 特殊な解析の場合もある
    - グラフ(生存率などは必須)

# プログラムテスト

- プログラマー自身のテスト
  - テストデータを使ってのプログラム動作検証
- プログラム作成者以外のテスト
  - テストデータを使ってのプログラム動作検証
- バリデーション
  - 上記のテストをすることがバリデーションになる
- 運用環境への移行
  - テスト完了後、運用環境に必要なものをコピー

# 解析処理手順書の作成

- 全てが自動化出来るわけではないので、Windowsの操作ができる人にわかる、処理手順書を作成する
- 操作ミスが発生しないように、処理チェックシートを作成する
- 解析処理手順書と解析処理チェックシートは同じものでもかまわない
  - 作業報告書として使える

# 実解析処理

- データをDM部門より受け取る
- 解析依頼書に基づき解析をする
- 解析処理チェックシートを確認しながら処理を進める
- 解析処理完了後、ログの確認
- 解析結果をプリンターに出力する

# 実務上の注意点 1 / 2

- 解析条件(データ条件)をハッキリさせる
- データの変更があった場合
  - CSVファイルの取り込みから始める(初めから)
- 追加解析が発生した場合1
  - 追加解析依頼書を発行してもらう
  - テスト環境で修正テストを行うこと
    - ERRORが無くなるまで(当たり前だが)
    - バリデーションは必ず行うこと

# 実務上の注意点 2 / 2

- 追加解析が発生した場合2
  - Warningエラーは気を付けること
    - 出て仕方ない場合と、絶対出てはいけない場合があるので、注意すること
  - テスト環境から運用環境へのプログラム移行には細心の注意を払うこと

# 解析パッケージの選択

- 以上の留意点を踏まえて、S - PLUSを選択した。以下は私としてのS - PLUSの評価です。
  - スクリプトベースでの作成が容易
  - 関数化の容易性
  - 開発ステップ数が減る
  - グラフ作成の自由度が高い
    - ファイル化が簡単に出来る
  - バリデーションがしやすい等

# おわり

- **ご静聴有難う御座いました。**