

SUGI-J2000 平成12年8月31日

臨床試験の解析計画書から SASプログラム自動生成の試み

持田製薬株式会社

萩野 篤司

水留 稔

矢島 勉

舟喜 光一

1

目次

- はじめに
- 解析業務の流れ
- 現状の問題点
- 解析プログラム自動生成 (概要)
- プログラムジェネレータ
- 解析プログラム自動生成の手順
- 解析プログラム自動生成の利点
- 今後の課題

2

はじめに

【SUGI - J' 98】

『SAS / IntrNet を活用した臨床試験解析報告書
作成の試み』 矢島 勉、水留 稔、萩野 篤司、舟喜 光一

解析報告書を電子化することにより

- 資料の整理
- ブラウザを用い資料閲覧が容易
- 使いまわしによる業務の効率化

さらに、業務の効率化を図るため

プログラム自動生成を検討

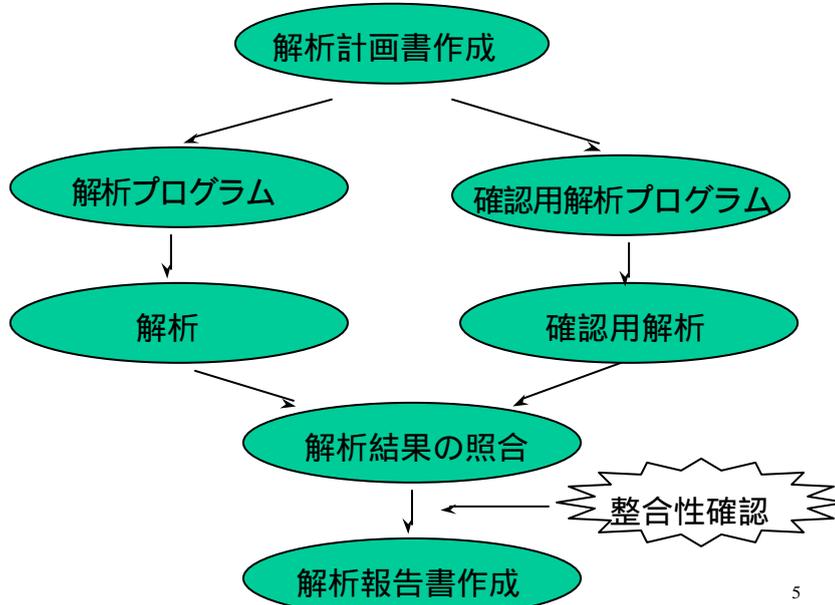
3

解析業務の流れ

- (1) 解析計画書作成 (Excel)
- (2) 解析プログラムを作成
(解析計画書をもとに2名の解析担当者)
- (3) 解析
- (4) 解析結果の照合 (整合性確認)
- (5) 解析報告書作成

4

【解析業務の流れ】



5

現状の問題点

- プログラム作成に膨大な作業量、時間を要する。
- 解析計画書変更に伴う プログラム修正に手間を要する。

6

現状の問題点 (つづき)

解析計画書と解析プログラムを
リンクさせることはできないか？



7

解析プログラム自動生成 (概要)

手順1 解析計画書を Excel で作成



手順2 DS名、変数、手法等を解析計画書に記載



手順3 上記変数をDDEで読み込みデータセット化



手順4 上記データセットを読み込みジェネレータにより解析プログラムを自動生成

8

プログラムジェネレータ

【パラメータセット：temp】

```
OBS  DS  ID
1   work1  id
2   work2  id
```

【プログラムジェネレータ】

```
data _null_ ; set temp;
file 'c:¥SAS¥テスト.sas';
if _N_ > 0 then do;
  put 'proc print data=' ds ' ;var ' id ' ;run ;' ;
end;run;
```

9

プログラムジェネレータ

【生成されたファイル：c:¥SAS¥テスト.sas】

```
proc print data=work1 ;var id ;run;
proc print data=work2 ;var id ;run;
```

この方法を解析計画書から、解析プログラムの作成に適用してみる。

10

解析プログラム自動生成の手順

手順1 解析計画書を Excel で作成する。

- 11 有効性の評価
 11.2 人口統計学的及び他の基準値の特性
 (1) 患者背景 (FAS解析対象例)
 背景因子を群別に集計する。
 また、比較可能性の確認のために、Fisherの直接確率計算法を行い、有意水準は両側 %とする。

項目		A群	B群	計
性別	男性 (%)			
	女性 (%)			
年齢	20=<.<30 (%)			
	30=<.<40 (%)			
	40=< (%)			
重症度	軽度 (%)			
	中等度 (%)			
	重度 (%)			

11

解析プログラム自動生成の手順

手順2 解析集合名、変数等を解析計画書に記載する。

- 11 有効性の評価
 11.2 人口統計学的及び他の基準値の特性
 (1) 患者背景 (FAS解析対象例)
 背景因子を群別に集計する。
 また、比較可能性の確認のために、Fisherの直接確率計算法を行い、有意水準は両側 %とする。

解析の内容 : 性別
 解析手法 : クロス集計
 解析対象集合 : FAS
 解析項目 (縦変数) : x50
 分類項目 (横変数) : x2
 その他 : なし

解析の内容 : 年齢
 解析手法 : クロス集計
 解析対象集合 : FAS
 解析項目 (縦変数) : x55
 分類項目 (横変数) : x2
 その他 : なし

解析の内容 : 重症度
 解析手法 : クロス集計
 解析対象集合 : FAS
 解析項目 (縦変数) : x60
 分類項目 (横変数) : x2
 その他 : なし

項目		A群	B群	計
性別	男性 (%)			
	女性 (%)			
年齢	20=<.<30 (%)			
	30=<.<40 (%)			
	40=< (%)			
重症度	軽度 (%)			
	中等度 (%)			
	重度 (%)			

12

---『以下、省略』---

解析プログラム自動生成の手順

手順3 :プログラムに使用する変数をDDEで読み込む。

11.1 有効性の評価
 11.2 人口統計学的及び他の基準値の特性
 (1) 患者背景 (FAS解析対象例) 背景因子を群別に集計する。
 また、比較可能性の確認のために、Fisherの直接確率計算法を行い、有意水準は両側 %とする。

解析の内容 : 性別
 解析手法 : クロス集計
 解析対象集合 : FAS
 解析項目 (縦変数) : x50
 分類項目 (横変数) : x2
 その他 : なし

解析の内容 : 年齢
 解析手法 : クロス集計
 解析対象集合 : FAS
 解析項目 (縦変数) : x55
 分類項目 (横変数) : x2
 その他 : なし

解析の内容 : 重症度
 解析手法 : クロス集計
 解析対象集合 : FAS
 解析項目 (縦変数) : x60
 分類項目 (横変数) : x2
 その他 : なし

... 『以下、省略』 ...

DDEで読み込む

項目	A群	B群	計
性別	男性 (%)		
	女性 (%)		
年齢	20=<_<30 (%)		
	30=<_<40 (%)		
	40=<_< (%)		
重症度	軽度 (%)		
	中等度 (%)		
	重度 (%)		

13

解析プログラム自動生成の手順

手順3 :DDEで読み込み、パラメータセットを作成する。

```
filename xls1 dde "excel|解析計画書&k!R7C4:R100C4";
data pms_&k; /* k=1,2,... */
  infile xls1; /* 例) k=1の時、pms_1(解析1のパラメータセット) */
  length p $20;
  input p $ ;
run;
```

```
data pms_&k;set pms_&k;
  if compress(p)='クロス集計' then proc='FREQ';
  if compress(p)='要約統計量' then proc='UNIVARIATE';
run;
```

14

【作成されたパラメータセット】

【解析1に対するパラメータセット】

OBS	PROC	DATA	ID	FMTID	GR	FMTGR	COMMENT	CODE	CELL
1	FREQ	FAS	x50	x50c.	x2	x2c.	性別	1	R3C3R4C6
2	FREQ	FAS	x50	x50c.	x2	x2c.	性別	2	R5C3R6C6
3	FREQ	FAS	x55	x55c.	x2	x2c.	年齢	1	R7C3R8C6
4	FREQ	FAS	x55	x55c.	x2	x2c.	年齢	2	R9C3R10C6
5	FREQ	FAS	x55	x55c.	x2	x2c.	年齢	3	R11C3R12C6
6	FREQ	FAS	x60	x60c.	x2	x2c.	重症度	1	R13C3R14C6
7	FREQ	FAS	x60	x60c.	x2	x2c.	重症度	2	R15C3R16C6
8	FREQ	FAS	x60	x60c.	x2	x2c.	重症度	3	R17C3R18C6

【解析2に対するパラメータセット】

OBS	PROC	DATA	ID	FMTID	GR	FMTGR	COMMENT	CODE	CELL
1	FREQ	FAS	x2	x2c.	x300	x300c.	有効性	1	R3C3R4C9
2	FREQ	FAS	x2	x2c.	x300	x300c.	有効性	2	R5C3R6C9

パラメータセットを読み込みプログラムジェネレータにて
SASプログラムを生成する。

15

解析プログラム自動生成の手順

手順4:作成されたパラメータセットを読み込み、プログラムジェネレータにて、解析プログラムを生成する。

```
data _null_;
  set pms_&k end=lastrec;
  file " C:\%USER%\program_&k..sas";
  if _n_>0 then do;          /* 例) k=1の時、program_1(解析1のプログラム) */
  /* FREQプロシジャによる解析 */
    if proc=' FREQ' then do;
      put ' title " ' comment' ";
    ;
      put ' data temp; set ' data'; if ' id' = ' code'; run;
    ;
      put ' proc freq data = temp ;
    ;
  /* 分類変数がある場合 */
  ;
  ;
  ;
```

16

解析プログラム自動生成の手順

手順5 生成された解析プログラムを読み込み実行

```
title " 性別 ";
data temp;set FAS ;if x50 = 1 ;run;
proc freq data=temp;
  tables x2 * x50 / nopercnt nocol out=out;
  format x2 x2c. ; format x50 x50c. ;
run;
```

----- 『 以下、省略 』 ----- 赤文字は、パラメータ

```
%inc " C:\USERS\program_1.sas " ; /* program_1(解析1のプログラム)*/
%inc " C:\USERS\program_2.sas " ; /* program_2(解析2のプログラム)*/
----- 『 以下、省略 』 -----
```

17

解析プログラム自動生成の利点

- (1)解析プログラム作成の作業量および時間を削減できる。
- (2)解析計画書と解析プログラムが1対1対応となる。
- (3)解析計画書の表現に多義性、曖昧さがなくなる。

18

解析プログラム自動生成の利点

- (4) 解析計画書の信頼性を保証すると、自動的に解析プログラムの信頼性も保証される。
- (5) 解析計画書の追加、変更に伴い、解析プログラムも自動的に更新される。
- (6) 実際に使用した解析プログラムを残しておくことができる。

19

今後の課題

解析計画書の保証方法の検討する。

解析計画書の標準化を行う。

出力帳票を柔軟に調整できるようにする。

解析結果の出力部分については、Ver8のODSも考慮に入れ再検討する。

20