

ダブルプログラミングによる 統計解析の品質管理

興和株式会社

臨床解析部

* 菅波秀規 益田隆史

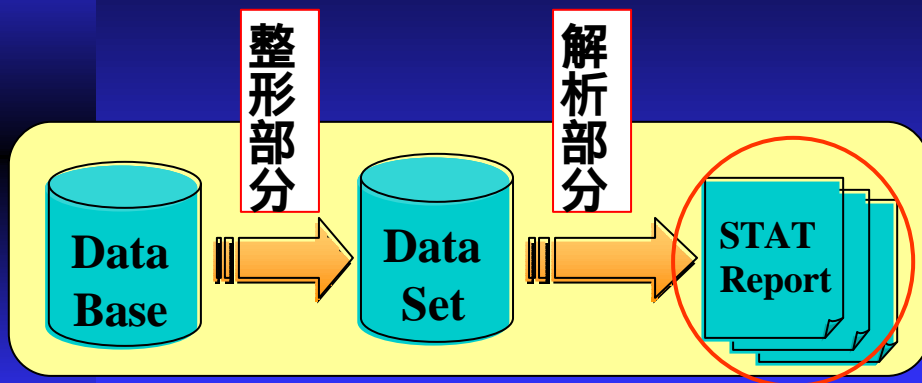
はじめに

- データの発生源に対する品質管理が注目されているとともに、電子データ処理に関する品質管理に対しても注目されてきた
- データマネジメントに関連する話題の整備は進んでいるが、統計解析に対する品質管理では決定的な方法論はいまだない

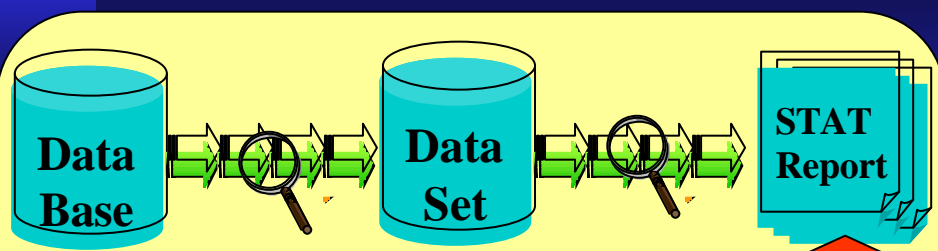
はじめに

- 解析結果に対する品質管理のひとつの方法として「ダブルプログラミング」を採用した
- 今回の報告では
 - ◆ ダブルプログラミングの採用理由
 - ◆ 解析業務システム
 - ◆ エラー率
 - ◆ 比較における注意点を例示する

統計解析業務



二つの品質管理法



適切性の確認されたステップを組み合わせる

選択の理由

適切性の確認されたプログラム

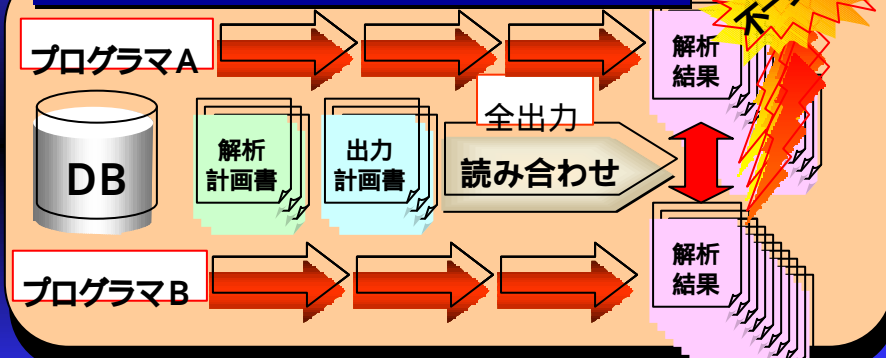
目の前の解析結果だけを
保証する品質管理方法

試験運営期間中のDBがロックされる
例外処理の可能性もある

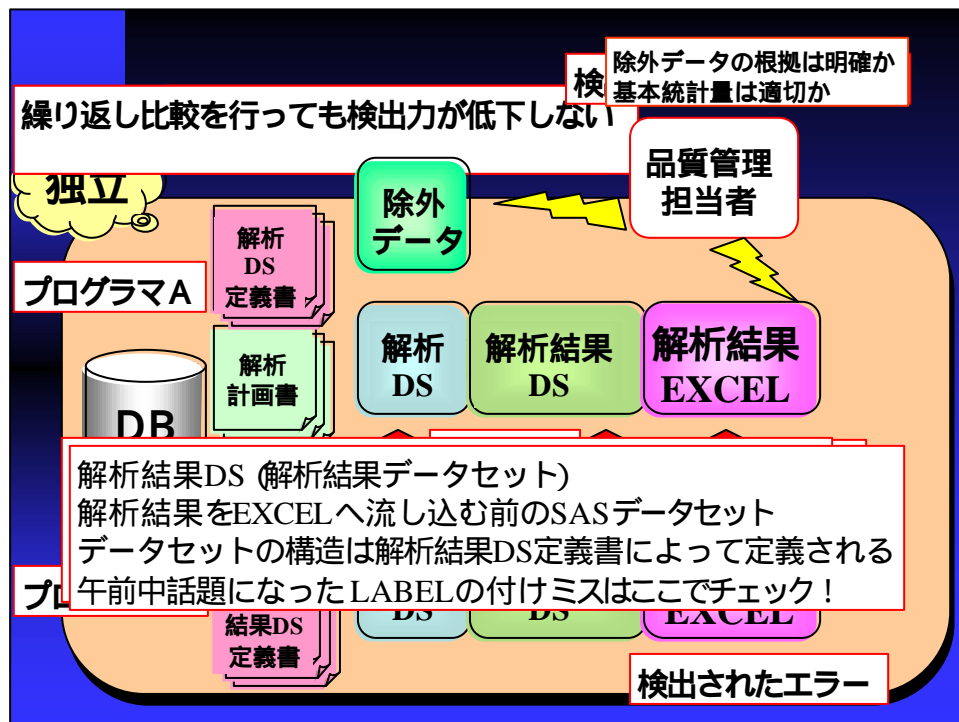
ダブルプログラミングによる
統計解析の品質管理

旧バージョンのマクロは再バリデーション

三つの問題点



3. 繰り返しによるエラー検出力低下



検討した臨床試験

試験	群数	総症例数	出力ページ数
A	1	34	376
B	3	35	912
C	2	240	501
D	3	317	1049
E	1	36	798
F	3	42	231
計	-	704	3867

検討結果

結果

- 検出されたエラー数
 - ◆ 多数に及ぶため特定していない
- クリーニングサイクル
 - ◆ 平均2.3回 (1回 ~ 3回)
- 検出されなかったエラー率
 - ◆ 0.15% (6/3867page)
 - ◆ 1errorが複数項目に影響を及ぼすためページ単位での特定

検出できなかったエラー

- 変数取得ミス
- 悪化率の方向選択ミス

変数取得ミス

家族性疾患の集計



~~FAMILY~~

FAMILY (家族歴)

DIAG(診断名)

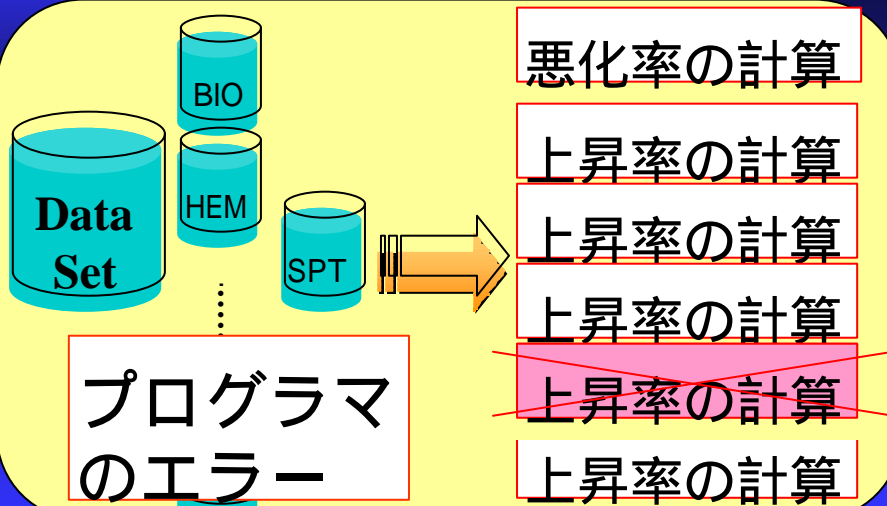
原因

紛らわしい変数命名
誤解を招く指示
を避ける

？ たか？

家族歴の集計 診断名の集計
FAMILY F_HIST

悪化率の方向選択ミス



原因

- プログラムのエラー
- 当時の解析計画書では
悪化率とは対象例に対する悪化例の割合である」
という定義であった
- 悪化の方向性を定義する必要があった

指示文書の記載不十分

検出されなかったエラー

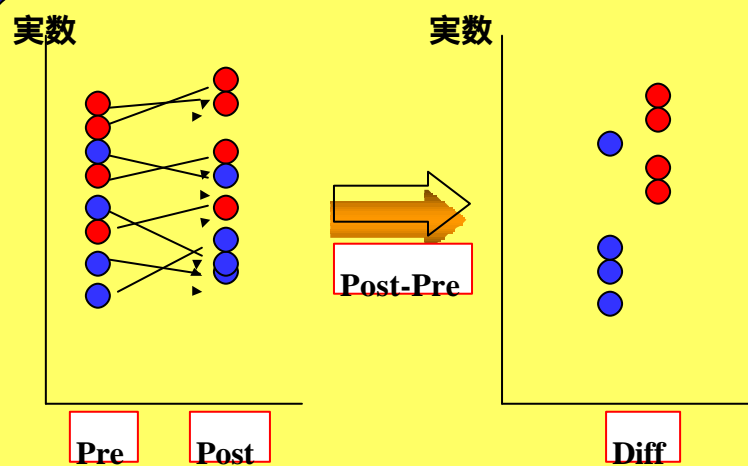
- プログラムのエラー
- エラーを誘引する原因の存在
- 指示文書 (解析計画書) を改善
 - ◆ 指示の方法を明瞭
 - ◆ 詳細な定義
- 変数命名法を改善
 - ◆ 紛らわしい変数命名の排除

検出された注目すべきエラー

- 変化量を用いた順位和検定
- 表示上の丸めの問題
- 重複処理のタイミング

論文参照

変化量を用いた順位和検定



変化量を用いた順位和検定

プログラムA

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 13408.0 Z = -1.95529 Prob > |Z| = **0.0505**

T-Test Approx. Significance = 0.0517

Kruskal-Wallis Test (Chi-Square Approximation)

CHISQ = 3.8268 DF = 1 Prob > CHISQ = **0.0504**

プログラムB

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 13405.5 Z = -1.95998 Prob > |Z| = **0.0500**

T-Test Approx. Significance = 0.0512

Kruskal-Wallis Test (Chi-Square Approximation)

CHISQ = 3.8452 DF = 1 Prob > CHISQ = **0.0499**

桁落ちによる同順位判定

	Pre	Post	DIFF
A:	-0.308	0.419	0.7270.7270000000000000
B:	0.526	1.253	0.7270.7269999999999990

Under Flow (桁落ち)

浮動小数点表現 計算機は内部的に2進数であるが

検定前に数値を丸める必要がある

$X = m \times 2^n$ 2進数の浮動小数点表示はつぎの
(mは固定小数点:2進数) かないことがある

比較における注意点

エラーではない！！

- 等分散を仮定しない t 検定
- リサンプリングによる並び替え検定

コンピュータを使用して
結果を比較する際には気をつけなければならない

等分散を仮定しない 検定

- TTEST procedure
- GLM procedure
 - ◆ means/welch
- MIXED procedure
 - ◆ repeated/group ddfm

等分散を仮定しない検定

GLM

General Linear Models Procedure

Welch's ANOVA for Y			
Source	DF	F Value	Pr > F
G	1.0000	1.4524	0.2568
Error	9.6782		

MIXED

Tests of Fixed Effects

Source	NDF	DDF	type III F	Pr > F
G	1	9.68	1.45	0.2568

TTEST

TTEST PROCEDURE

Variable: Y

G	N	Mean	Std Dev	Std Error	Variances	T	DF	Prob> T
1	7	1.51196451	1.44655515	0.54674646	Unequal	-1.2051	9.7	0.2569
2	7	2.27541121	0.84653299	0.31995939	Equal	-1.2051	12.0	0.2514

For H0: Variances are equal, F' = 2.92 DF = (6,6) Prob>F' = 0.2179

原因

昨日話題に
なった

■ TTEST

- ◆ P値の計算に線形補間を使用
永田靖 統計的方法のしくみ 日科技連

■ GLM・MIXED

- ◆ 小数点自由度のt分布を利用

使用するProcedureを特定しなければならない

リサンプリングによる並び替え検定

■ MULTTEST プロシジャ

◆ Permutation option

SEED

SIMULATION

プログラマ A		
Contrast	Raw_p	Perm_p
LINEAR	0.0102	<u>0.0192</u>
M.SATU	0.0954	<u>0.1706</u>

プログラマ B		
Contrast	Raw_p	Perm_p
LINEAR	0.0102	<u>0.0187</u>
M.SATU	0.0954	<u>0.1768</u>

リサンプリングによる並び替え検定

■ SORT前のデータセットの並びによって生成される検定統計量の並び替え分布が異なる

◆ OUTSAMP Optionで確認

使用するProcedureとOptionだけでなく
データセットのソート順まで特定しなければならない！

コンピュータによる比較のために

- 解析計画書に解析方法を明確に定義する
 - ◆ 独立でプログラミングを行っているため定義があいまいでは一致しない
 - ◆ 使用するprocedureやoptionを特定する

コンピュータによる比較のために

- SASやコンピュータの性質を良く知る
 - ◆ Under Flow
 - ◆ MULTTESTなど解析に使用するデータセットの状態が結果に影響するProcedure
 - ◆ if A le 10 then ... 欠測値も選択される

プログラミング担当部門の
マニュアルとして整理する

SAS Ver8e TTEST procedure

The TTEST Procedure

Variable	G	N	Statistics								Minimum	Maximum
			Lower CL	Mean	Upper CL	Mean	Lower CL	Std Dev	Upper CL	Std Dev	Std Err	
Y	1	7	0.1741	1.512	2.8498	0.9322	1.4466	3.1854	0.5467	0.2884	4.4058	
Y	2	7	1.4925	2.2754	3.0583	0.5455	0.8465	1.8641	0.32	0.6753	3.4828	
Y	Diff (1-2)		-2.144	-0.763	0.6168	0.8499	1.1851	1.9564	0.6335			

Variable	Method	T-Tests		DF	t Value	Pr > t
		Variances				
Y	Pooled	Equal		12	-1.21	0.2514
Y	Satterthwaite	Unequal		9.68	-1.21	0.2568

Variable	Method	Equality of Variances		F Value	Pr > F
		Num DF	Den DF		
Y	Folded F	6	6	2.92	0.2179

SAS Ver 8e 変数の制限

The FREQ Procedure

Table of HYPERTENSION_DIAGNOSIS by Y

HYPERTENSION_DIAGNOSIS(高血圧(診断名))
Y(反応)

Frequency				Total
Row	Pct	1	2	
1		9	28	37
		24.32	75.68	
2		8	29	37
		21.62	78.38	
Total		17	57	74

まとめ

- 解析結果の品質管理の基礎にダブルプログラミングを採用した
- 解析結果の比較を読み合わせからコンピュータによる比較へ移行

エラー率低下の可能性

- エラー率は0.15% (6/3867page)
- 改善点
 - ◆ 指示文書
 - ◆ 変数命名 (DB構造)
 - ◆ SASに関するマニュアルの整備

最後に

- ダブルプログラミングによって保証されるのは一致していることまで
- 品質管理の方法論に関する議論はいろいろある
- 今回はダブルプログラミングによる品質管理について評価を行った

余談

- 解析結果をSASデータセットで保存
症例取り扱いを変更した後に再解析を実施した場合影響があった部分を特定することが容易
P値の比較や要約統計量の比較など
- Compare Procedureの出力
読み合わせでは一致していたという報告
Compare Procedureではデータセット名・実施時刻が自動的に記録される

解析業務システムの検討

- Output Delivery System
解析から結果の報告までをSASだけで行うことができる可能性
- バリデーションの可能性



指示文書

■ 解析計画書

- ◆ ページ数 平均64.3 (min:53 ~ max:75)
 - ◆ 用語の定義、症例取り扱い、解析対象集団の定義、解析方法の詳細

■ 出力計画書

- ◆ ページ数 86.0 (min:42 ~ max:138)
 - ◆ 集計表イメージ、数値の性質を示すために数値 (9)や符号(-)を表示した