



解析用データセットのあり方 - CDISCを意識して -

麒麟麦酒株式会社 医薬カンパニー

開発本部 開発推進部

長谷川要、本山佳代子、小崎昌昭、外城靖子



おわび

- ◆ SASユーザー会でありながら、SASに関する話題が少ないです。
- ◆ もっぱら社内における解析用データセットの現状の問題点の改善提案を中心に話します。従って、既に標準化を行っている会社の方にはあまり役立ちません。
- ◆ CDISCの活動成果を参考にした発表内容であるため、新規性のある話ではありません。



本日のメニュー

- ◆ 解析用SASデータセットの標準化の必要性
- ◆ 社内における現状の問題点の分析
- ◆ 海外における解析用データセット標準化に向けた動き
 - FDAの電子申請受け入れ態勢
 - CDISCの活動内容の紹介と成果
- ◆ 解析用データセットの標準化に向けた社内の改善策の紹介
- ◆ 今後の検討課題



本日のメニュー

- ◆ 解析用SASデータセットの標準化の必要性
- ◆ 社内における現状の問題点の分析
- ◆ 海外における解析用データセット標準化に向けた動き
 - FDAの電子申請受け入れ態勢
 - CDISCの活動内容の紹介と成果
- ◆ 解析用データセットの標準化に向けた社内の改善策の紹介
- ◆ 今後の検討課題



解析用データセットの標準化の 必要性

- ◆ ICH-E9、ICH-M4ガイドライン
複数の臨床試験成績を併合した解析の必要性に言及されている
- ◆ ICH-E5ガイドライン
海外臨床試験成績との類似性の評価に関して言及されている



解析用データセットの標準化によるメリット

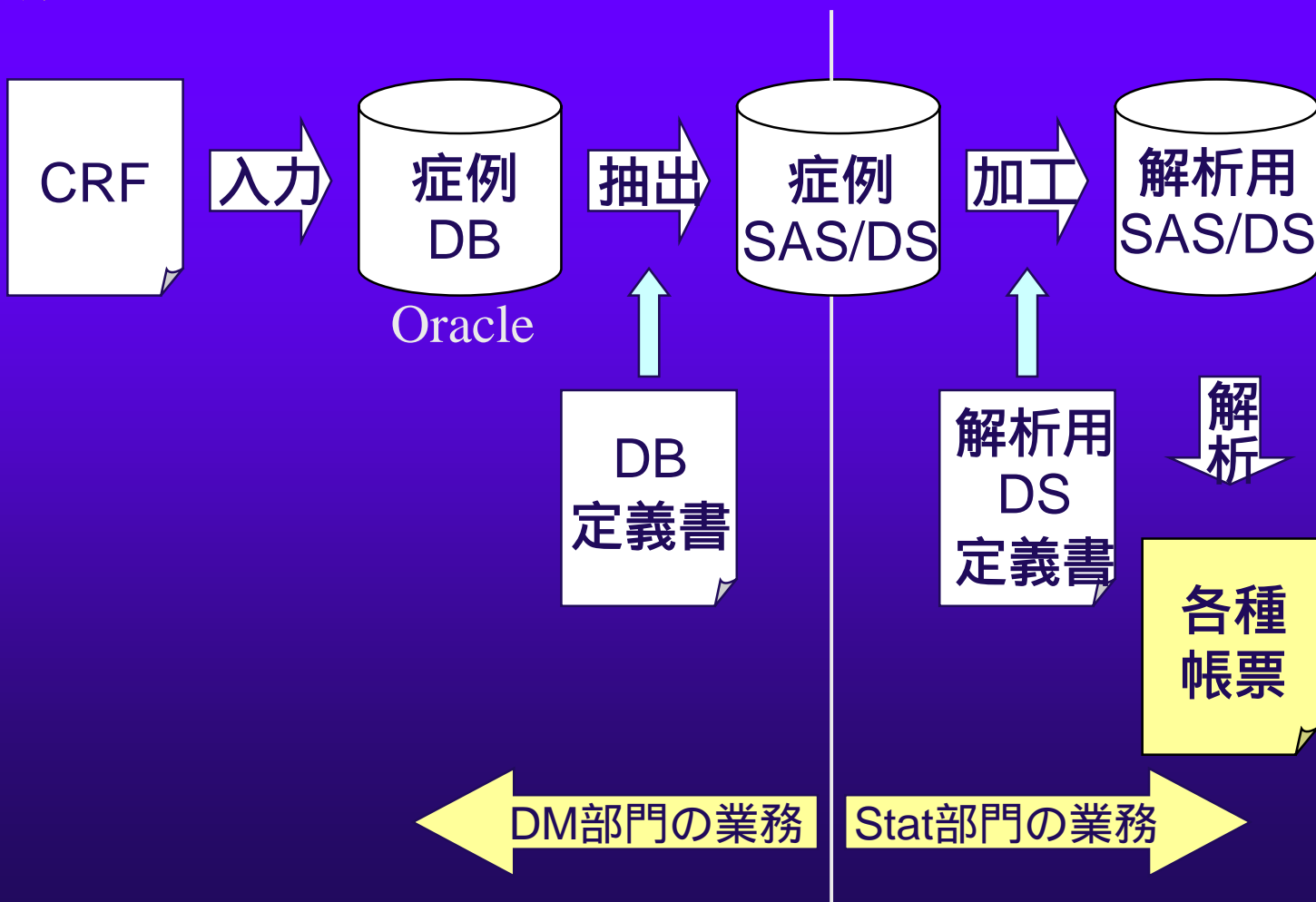
- ◆ 解析プログラムの汎用性の向上
- ◆ プログラミングのためのリソースの節約
- ◆ 解析結果の品質向上
- ◆ データの再利用が容易
- ◆ データの財産価値を高める



本日のメニュー

- ◆ 解析用SASデータセットの標準化の必要性
- ◆ 社内における現状の問題点の分析
- ◆ 海外における解析用データセット標準化に向けた動き
 - FDAの電子申請受け入れ態勢
 - CDISCの活動内容の紹介と成果
- ◆ 解析用データセットの標準化に向けた社内の改善策の紹介
- ◆ 今後の検討課題

解析業務における電子データの 流れ



問題点の箇

変数名やデータ
セットの構造が不
統一！

解析担当者各人が個別
に定義書を作成し、標準
化されていない。

定義書

プログラムの使
いまわしが出来
ない。

症例
/DS

解析用
DS

解析用
SAS/DS

解析

各種
帳票

Div. of Statistics 部門の業務



現状の問題点の把握

◆ 解析用データセット作成時の問題点

- 解析用データセット作成時に必要な定義書が、担当者間で不統一
- 解析用データセットの形式が不統一
 - データセット名、変数名の不統一
 - データセット構造の不統一

◆ 解析用データセット利用時の問題点

- 併合した解析が困難
- プログラムの再利用が困難



本日のメニュー

- ◆ 解析用SASデータセットの標準化の必要性
- ◆ 社内における現状の問題点の分析
- ◆ 海外における解析用データセット標準化に向けた動き
 - FDAの電子申請受け入れ態勢
 - CDISCの活動内容の紹介と成果
- ◆ 解析用データセットの標準化に向けた社内の改善策の紹介
- ◆ 今後の検討課題



電子申請に関するガイダンス

- ◆ Guidance for Industry, Providing Regulatory Submissions in Electronic Format – General Considerations. IT2.
- ◆ Guidance for Industry, Providing Regulatory Submissions in Electronic Format – NDAs. IT3.
- ◆ Guidance for industry, Providing Regulatory Submissions to the Center for Biologics Evaluation in Electronic Format – Biologics Marketing Applications.

米国における電子承認申請の 申請件数の推移



出典: www.fda.gov/cder/regulatory/ersr/levin_electronic/sld004.htm



CDISCの紹介

◆ CDISCとは？

- Clinical Data Interchange Standards Consortiumの略名
- 米国にある非営利団体組織
- 活動内容は、臨床試験全般に関わるデータの標準化の活動を行っている
- 国内では製薬協が主体となり活動の準備が開始されている模様



CDISCの活動内容

Working Team	活動内容
ODM	XMLを用いたデータの収集、変換などの標準化
SDS	安全性データセット(SDM)の標準化
ADaM	有効性データセット(ADaM)の標準化
LAB	Central Lab.データの受け渡し方法の標準化
TAP	標準化されたデータモデルの確認作業
EDU	教育カリキュラムの作成、CDISCの普及活動
OIS	CDISCのインフラに関わる活動を行う




SDMのデータセットドメイン

Dataset Name	Description
DEMO	Demographics and subject characteristics
AE	Adverse Event
CONMEDS	Concomitant Medication
DISPOSIT	Disposition
ECG	ECG
EXPOSE	Drug Exposure
CHEM	Labs-Chemistry
HEMAT	Labs-Hematology
URINE	Labs-Urinalysis
MEDHIST	Medical history
PE	Physical Examination
VITALS	Vital Signs

DEMOデータセットの変数一覧

DEMO.xpt, Demographics, Version 2.0, October 22, 2001, One record per subject, CRT						
Variable Name	Variable Label	Type	Decodes/ Format	Origin	Role	Comments
STUDYID	Study ID	Char		Demographics CRF Page		
SITEID	Center or Site ID	Char		Demographics CRF Page		
INVID	Investigator ID	Char		Demographics CRF Page		
INVNAME	Investigator Name	Char		Demographics CRF Page or Derived		
USUBJID	Unique Subject ID	Char		Sponsor Defined		
SUBJID	Subject ID	Char		Demographics CRF Page		
AGE	Age in Years at Baseline	Num		Demographics CRF Page or Derived		
SEX	Sex	Char	M, F	Demographics CRF Page		
RACE	Race	Char		Demographics CRF Page		
TRTCD	Treatment Code	Num		EXPOSE TRTCD		
TRTGRP	Treatment Group	Char		EXPOSE TRTGRP		
COUNTRY	Country	Char	ISO 3165	Demographics CRF Page or Derived		
VISITNUM	Visit Number	Num		Demographics CRF Page or Derived		
VISITDY	Visit Day	Num		Demographics CRF Page or Derived		
VISIT	Visit Name	Char		Demographics CRF Page or Derived		
DMACTDY	Actual Study Day of Visit	Num		Demographics CRF Page or Derived		
DMDT	Date of Visit	Num	ISO 8601 YYYY-MM-DD	Demographics CRF Page or Derived		



ONE PROC AWAY

```
proc phreg
```

```
    data=s.survival(where=(s_pp  
    eq  ));
```

```
    model ttp*ttpcnsr(1)=trtcd/r1;
```

```
run;
```



本日のメニュー

- ◆ 解析用SASデータセットの標準化の必要性
- ◆ 社内における現状の問題点の分析
- ◆ 海外における解析用データセット標準化に向けた動き
 - FDAの電子申請受け入れ態勢
 - CDISCの活動内容の紹介と成果
- ◆ 解析用データセットの標準化に向けた社内の改善策の紹介
- ◆ 今後の検討課題



解析用データセット標準化の利点

- ◆ DS定義書が標準化されており、解析目的とするデータがどのDSに含まれているのかが容易に識別できる
- ◆ SASプロシジャで集計を行う際に、データ加工を行わずに集計処理が可能
- ◆ 複数の試験データを併合した解析が容易
- ◆ プログラムの再利用が可能



標準的なデータセットのあり方

◆ 理想

- DM部門のDMシステムから直接標準的な解析用データセットが作成されることを望んだ。

◆ 現実

- DMシステムの制限から、直接標準的な解析用データセットの出力は不可能。

◆ 解決策

- DM部門が作成する症例SASデータセットから標準的な解析用データセットを作成する際の、定義書の標準化を目指すこととした。



解析用DS定義書の標準化

- ◆ 表紙
- ◆ 履歴管理
- ◆ 解析用データセット一覧表
- ◆ 解析用データセット変数一覧表

解析用データセット一覧表

Protocol: APC0015 第1 相臨床試験 - 前立腺癌を対象とした探索的試験 -
 Protocol #: APC0015.001-A01
 Date: 2002/7/29

解析用データセット名一覧表

ID	データセット名	内容	構造	キー変数
1	BS_DEMOG	患者背景	1 record per patients	CSENO
2	BS_OCOMP	合併症	multiple records per patients	CSENO
3	BS_HST	既往歴	multiple records per patients	CSENO
4	BS_VITAL	バイタルサイン	multiple records per patients	CSENO, DAY
5	BS_HEMA	血液学的検査	multiple records per patients	CSENO, DAY
6	BS_LAB	血液生化学的検査	multiple records per patients	CSENO, DAY
7	BS_URC	尿検査	multiple records per patients	CSENO, DAY
8	BS_AE	有害事象	multiple records per patients	CSENO, DAY
9	BS_DRUG	投与薬剤	multiple records per patients	CSENO, DAY
10	BS_ABN	臨床検査値変動	1 record per patients	CSENO, DAY
11	BS_ECG	心電図検査	multiple records per patients	CSENO, DAY
12	BS_PSA	PSA濃度	multiple records per patients	CSENO, DAY
13	BS_CANC	腫瘍病変	multiple records per patients	CSENO, DAY
14	BS_FLAG	症例評価	1 record per patients	CSENO

A large, dark metal key with a heart-shaped bow and a notched bit, resting on a textured, yellowish-brown surface.

データセット名	SS_DEMO5
内容	患者健康データ
入力データセット	YMO1
構造	1 record per patient
キー変数	CEDNO
変換プログラム名	SS_DEMO5.SAS

Variable	Label	Type	Format	Variant	Comment	Origin
EXPNO	実験番号	IS		40001~の連番		PROT_EXPNO
EXP1	カルナ重荷	IS				PROT_EXP1
EXP2	イオンナル	IS				PROT_EXP2
EXP3	中核(重)	IS				PROT_EXP3
EXP4	中核(軽)	IS				PROT_EXP4
EXP5	中核(重)コード	IS		20000000-99999999		PROT_EXP5
EXP6	中核(軽)コード	IS		20000000-99999999		PROT_EXP6
EXP7	中核(重)レベル	IS	EXP7L	1:Level 1 2:Level 2 3:Level 3 4:投影線有り 5:投影線無し		PROT_EXP7
EXP8	中核(軽)レベル	IS	EXP8L			PROT_EXP8
EXP9	基準個体名	IS	EXP9P			PROT_EXP9
EXP10	当核重核基準名	IS	EXP10P			PROT_EXP10
EXP11	当核分相基準名1	IS	EXP11P			PROT_EXP11
EXP12	当核分相基準名2	IS	EXP12P			PROT_EXP12
EXP13	同位体	IS	EXP13P	1:本人		PROT_EXP13
EXP14	同位体年月日	IS	EXP14P			PROT_EXP14
EXP15	生年 月日	IS	EXP15P			PROT_EXP15
EXP16	身長(Cm)	IS	EXP16P			PROT_EXP16
EXP17	体重(Kg)	IS	EXP17P			PROT_EXP17
EXP18	体表面積 (m ²)	IS	EXP18P		SURFACE=(EXP18P/100) ^{0.725} × (EXP16/100) ^{0.725} × EXP17によって算出する。	PROT_EXP18
EXP19	経緯度	IS	EXP19P			PROT_EXP19
EXP20	経緯度年月日	IS	EXP20P			PROT_EXP20

#1 BS DEMOG-1



変数一覧表に含めた項目

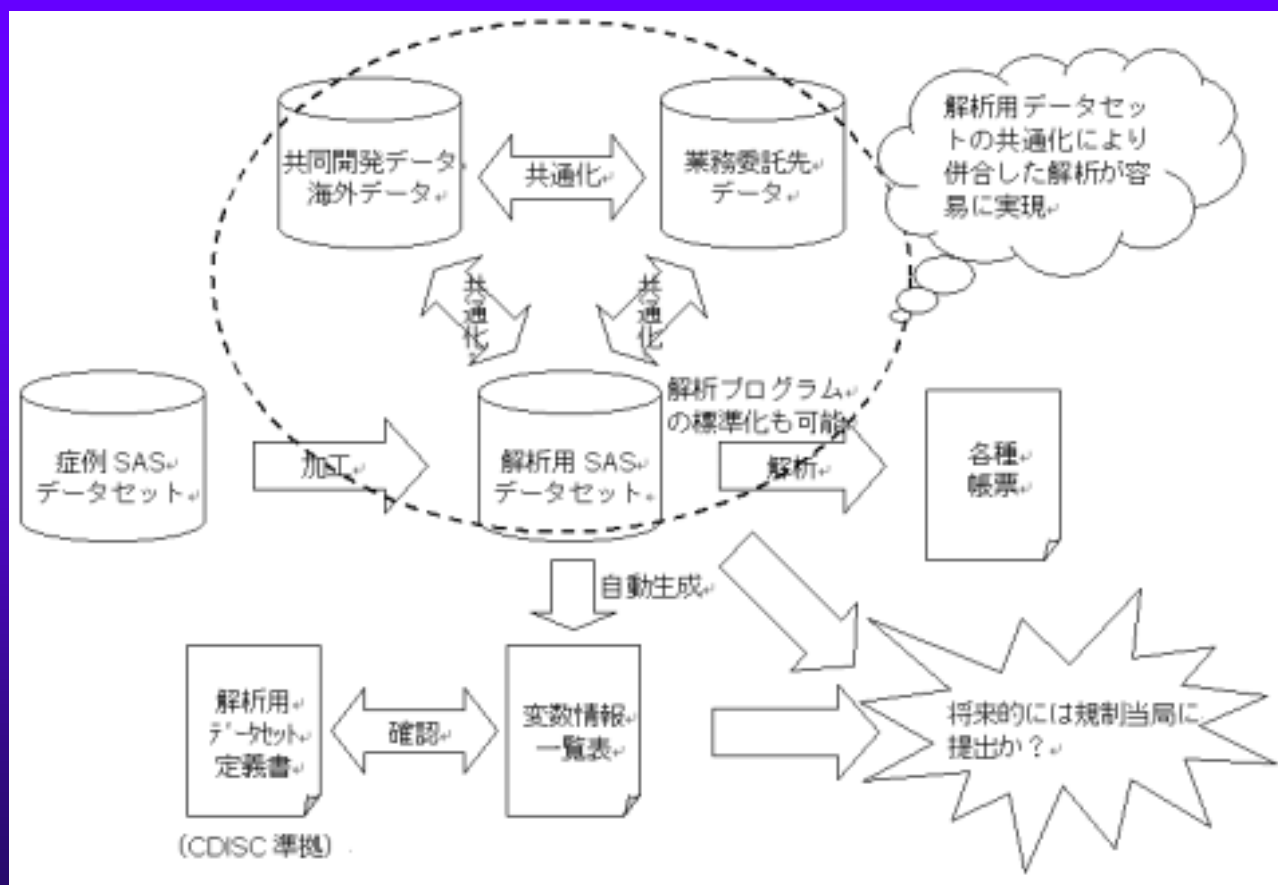
◆ ヘッダー部分

- データセット情報 (DS名、内容、入力データ、構造、キー変数、変換プログラム)

◆ 変数定義部分

- Variable
- Label
- Type
- Format
- Values
- Comment
- Origin

新たな電子データの流れ





本日のメニュー

- ◆ 解析用SASデータセットの標準化の必要性
- ◆ 社内における現状の問題点の分析
- ◆ 海外における解析用データセット標準化に向けた動き
 - FDAの電子申請受け入れ態勢
 - CDISCの活動内容の紹介と成果
- ◆ 解析用データセットの標準化に向けた社内の改善策の紹介
- ◆ 今後の検討課題



今後の検討課題

◆ 解析用データセット作成業務手順書

- 一連の作業過程を誰でも行えるようにするために、業務手順書を作成する

◆ CRFの標準化

- データを入手する段階、すなわちCRFそのものから標準化する必要がある



Adverse Events		CDISC Clinical Data Connectathon																	
		PROTOCOL NUMBER										VISIT ID				CRF PAGE ID			
		1 4 3 - 0 2										A E				A E			
COMPOUND		THERAPEUTIC		COUNTRY		<input type="radio"/> Fred Hutchinson Cancer Research Center <input type="radio"/> Memorial Sloan Kettering Cancer Institute <input type="radio"/> Roswell Park Cancer Institute												SCREENING / RANDOMIZATION #	
S D P		O N C		U S A															
Instructions: List one continuous Adverse Event per box and complete all fields when one of the following occurs: the severity increases, the event resolves or the patient terminates from the study.																			
Adverse Event Description						AE Start Date (mm/dd/yyyy)						SEVERITY:							
						/ /						<input type="radio"/> Mild <input type="radio"/> Moderate <input type="radio"/> Severe <input type="radio"/> Life Threat							
ACTION TAKEN REGARDING STUDY DRUG:						AE Stop Date (mm/dd/yyyy)						RELATIONSHIP TO STUDY DRUG:							
<input type="radio"/> None <input type="radio"/> Discontinued Permanently <input type="radio"/> Reduced <input type="radio"/> Interrupted						/ /						<input type="radio"/> None <input type="radio"/> Unlikely <input type="radio"/> Possible <input type="radio"/> Probable							
ACTIONS TAKEN REGARDING PATIENT CARE:										ADVERSE EVENT OUTCOME:									
<input type="radio"/> None <input type="radio"/> Medication required <input type="radio"/> Hospitalization required or prolonged <input type="radio"/> Other										<input type="radio"/> Resolved, no residual effects <input type="radio"/> Continuing <input type="radio"/> Resolved, residual effects <input type="radio"/> Death									
Adverse Event Description						AE Start Date (mm/dd/yyyy)						SEVERITY:							
						/ /						<input type="radio"/> Mild <input type="radio"/> Moderate <input type="radio"/> Severe <input type="radio"/> Life Threat							
ACTION TAKEN REGARDING STUDY DRUG:						AE Stop Date (mm/dd/yyyy)						RELATIONSHIP TO STUDY DRUG:							
<input type="radio"/> None <input type="radio"/> Discontinued Permanently <input type="radio"/> Reduced <input type="radio"/> Interrupted						/ /						<input type="radio"/> None <input type="radio"/> Unlikely <input type="radio"/> Possible <input type="radio"/> Probable							
ACTIONS TAKEN REGARDING PATIENT CARE:										ADVERSE EVENT OUTCOME:									
<input type="radio"/> None <input type="radio"/> Medication required <input type="radio"/> Hospitalization required or prolonged <input type="radio"/> Other										<input type="radio"/> Resolved, no residual effects <input type="radio"/> Continuing <input type="radio"/> Resolved, residual effects <input type="radio"/> Death									
Comments:																			
CDISC AE		<div style="display: flex; justify-content: space-between;"> Investigator Name (PRINTED) Investigator Initials Date </div>																	



他のWorking Teamの活動にも 注目

- ◆ CDISCでは紙ベースのCRFを用いずに電子的にデータの受け渡しを可能としたXMLによる標準化の活動がOperational Data Modeling Teamで行われ、その成果も公表されている。今後DM部門を含め、こうした動向にも注目したいと考える。