

ADaMデータ作成時の懸念事項

三沢 喬

CDISC Japan Users Group SDTMチーム

Concerns of generate ADaM data

Takashi Misawa

CDISC Japan Users Group

SDTM Team

要旨:

Analysis Data Model (ADaM) は解析用データ標準であるが、非常に自由度の高いデータモデルであり、実装する際に様々な問題が発生する。

実際、ADaMデータ作成の仕様は各社取り決めが異なることが多く、担当者で仕様が異なることも少なくない。

本発表では、導出した値、採否フラグの持たせ方やそれらの情報の導出をSDTM/ADaMのどちらで行うかなど、CJUG SDTMガイドラインタスクで得られた各社SDTM仕様を基に、ADaM作成時の懸念事項をADaMデータ作成の実例やレビューアの視点を踏まえて紹介する。

キーワード: CDISC, ADaM, SDTM

NOTE

本発表は、今後社内でCDISCプロセスを導入する上で検討しなければならない事象を示すことを目的としております。

本資料はCDISCから発行されたドキュメントやCJUG SDTMガイドラインチームに寄せられた疑問を基に構成しております。

ADaMデータ作成時の懸念事項

SDTM/ADaMデータ間

採否フラグをどちらに持たせるのか

導出データをどちらに持たせるのか

ADaMデータ内

レコードを追加するか又は変数追加するか

採否フラグの使用方法

(Recorded, Parameter-Level)

ADaMデータ作成時の懸念事項

SDTM/ADaMデータ間

採否フラグをどちらに持たせるのか

導出データをどちらに持たせるのか

ADaMデータ内

レコードを追加するか又は変数追加するか

採否フラグの使用方法

(Recorded, Parameter-Level)

今までの採否フラグ

- 多くの会社で症例、時点採否が存在
- 各社で採否を作成している部署は様々
- ...ではCDISCに従った場合、どう対応すればよい？

- ①SDTM/ADaMそれぞれに症例採否変数がある
- ②時点採否変数の定義はADaMにしか存在しない

*SDTM症例採否はSDTM IGv3.2 Appendix Cに定義があるが必須として定義されているわけではない



- ①SDTM/ADaMどちらにデータを格納すればいいのか...
- ②DMが時点採否作成担当なのでSDTMに格納したい...

SDTMとADaMの整合性

- 症例採否変数名がそれぞれ定義されている理由

採否に修正が発生した場合にSDTMに戻らずに
修正対応できるため

3.4 Additional ADaM Variables

3.4.1 Differences between SDTM and ADaM Population and Baseline Flags

Table 3.4.1.1 ADaM Subject-Level Population Flags Corresponding to SDTM Supplemental Qualifiers

SDTM QNAM	SDTM QLABEL	ADaM Subject-Level Population Flags
COMPLT	Completers Population Flag	COMPLFL
FULLSET	Full Analysis Set Flag	FASFL
ITT	Intent to Treat Population Flag	ITTFL
PPROT	Per Protocol Set Flag	PPROTFL
SAFETY	Safety Population Flag	SAFFL



SDTMとADaMの整合性

- SDTMにParameter, Record-Levelがない理由

症例採否と同様の理由 + ADaMで導出するレコードに対して付与する際二度手間になるため

3.4 Additional ADaM Variables

3.4.1 Differences between SDTM and ADaM Population and Baseline Flags

ADaM also supports parameter-level and record-level population flags, which do not exist in SDTM.

Similarly, a baseline record identified in SDTM may not be the record identified in an ADaM dataset and there are many reasons why this may occur. There are ADaM parameters that are highly derived and do not have simple counterparts in a findings domain. An ADaM parameter may be derived from SDTM data spanning multiple domains and classes. Such a parameter would not exist in the SDTM and so its baseline could only exist in the ADaM dataset. Also, it may be necessary to have separate baselines for different periods within the study, for example to support analyses of change from screening baseline, double-blind treatment baseline, and open-label extension baseline (see Section 4.2, Rule 6). When there is record-level population flagging, it may be necessary to have different baselines for two different analysis populations. Lastly, it may be desired to conduct analyses for different definitions of baseline. The ADaM baseline flag ABLFL, coupled with the BASE and BASETYPE columns, plus population flags, can handle all of these practical scenarios.

...であれば全てADaMで導出すれば？



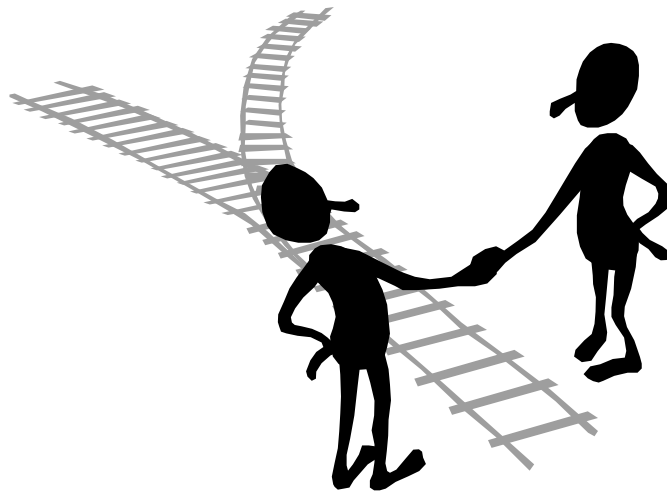
- あとは運用上の問題になります。

図式化ツールで集計するならSDTMにあった方がいいかも？

DM側で基準に抵触しているかを判断しているのでSDTMに持たせる方が業務の切り分けになるかも？

採否フラグをどちらに持たせるのか

- CDISCドキュメント及び社内体制を考慮して採否フラグの格納方法を検討しよう！



ADaMデータ作成時の懸念事項

SDTM/ADaMデータ間

採否フラグをどちらに持たせるのか

導出データをどちらに持たせるのか

ADaMデータ内

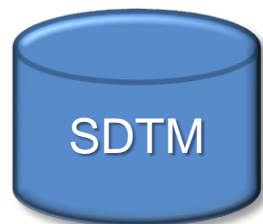
レコードを追加するか又は変数追加するか

採否フラグの使用方法

(Recorded, Parameter-Level)

導出データの格納先

- 評価パラメータをある情報から導出する場合、SDTM/ADaMどちらで導出すればいいの？
- IG上どちらでも導出することは可能！



SDTM

Value Level Metadata - EG [EGSTRESC]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
EGSTRESC	EGTESTCD EQ QTcB (QTcB - Bazett's Correction Formula)	float	5.1		Derived	QTcB = QT interval / square root of (60 / heart rate). For the complete algorithm see the referenced external document. Complex Algorithms (complexalgorithms.pdf)
EGSTRESC	EGTESTCD EQ QTcF (QTcF - Fridericia's Correction Formula)	float	5.1		Derived	QTcF = QT interval / cubic root of (60 / heart rate). For the complete algorithm see the referenced external document. Complex Algorithms (complexalgorithms.pdf)



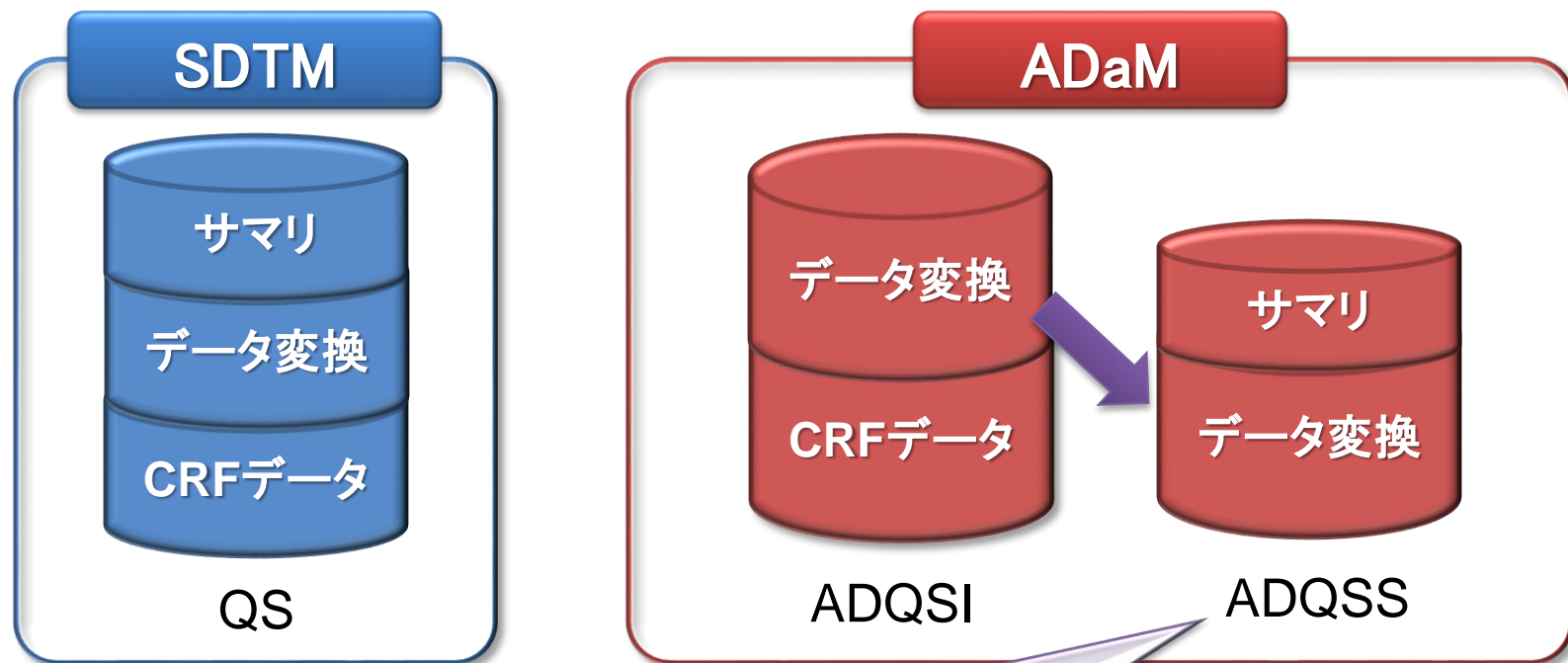
ADaM

Parameter Value List - ADQSADAS [AVAL]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
AVAL	PARAMCD EQ ACTOT (Adas-Cog(11) Subscore)	integer	8		Derived	Sum of ADAS scores for items 1, 2, 4, 5, 6, 7, 8, 11, 12, 13, and 14, see Analysis Data Reviewers Guide (Page 3) for details on adjusting for missing values. Analysis Data Reviewer's Guide (analysis-data-reviewers-guide.pdf)

導出データの格納先

- 数十個の項目から評価指標を算出する場合(SF-36等)



Traceabilityを確保するために
データセットを分けることが可能！

Reviewerの立場になって考えると...

SDTM

STUDYID	DOMAIN	USUBJID	QSTESTCD	QSCAT	QSORRES	QSSTRESN	QSDRVFL
STUDY-A	QS	101	T01	T-SOCRE	3	3	
STUDY-A	QS	101	T02	T-SOCRE	2	2	
STUDY-A	QS	101	T03	T-SOCRE	2	2	
STUDY-A	QS	101	T01	T-S_INTERNAL		-3	Y
STUDY-A	QS	101	T02	T-S_INTERNAL		-2	Y
STUDY-A	QS	101	T03	T-S_INTERNAL		-2	Y
STUDY-A	QS	101	SUMMARY	T-S_SUMMARY		30	Y
STUDY-A	QS	102	T01	T-SOCRE	2	2	
STUDY-A	QS	102	T02	T-SOCRE	1	1	
STUDY-A	QS	102	T03	T-SOCRE	2	2	
STUDY-A	QS	102	T01	T-S_INTERNAL		-2	Y
STUDY-A	QS	102	T02	T-S_INTERNAL		-1	Y
STUDY-A	QS	102	T03	T-S_INTERNAL		-2	Y
STUDY-A	QS	102	SUMMARY	T-S_SUMMARY		20	Y

--DRVFL="Y"は導出レコードを示し、--CAT等の
カテゴリ変数で導出された項目を区分していることが分かる

Reviewerの立場になって考えると...

ADaM

STUDYID	USUBJID	PARAMCD	PARAMTYP	PARCAT1	AVAL
STUDY-A	101	T01		T-SOCRE	3
STUDY-A	101	D01	DERIVED	T-S_INTERNAL	3.4
STUDY-A	101	T03		T-SOCRE	5
STUDY-A	101	D03	DERIVED	T-S_INTERNAL	2.2
STUDY-A	101	T04		T-SOCRE	3
STUDY-A	101	D04	DERIVED	T-S_INTERNAL	3
STUDY-A	102	T01		T-SOCRE	2
STUDY-A	102	D01	DERIVED	T-S_INTERNAL	4.4
STUDY-A	102	T03		T-SOCRE	2
STUDY-A	102	D03	DERIVED	T-S_INTERNAL	5.4
STUDY-A	102	T04		T-SOCRE	1
STUDY-A	102	D04	DERIVED	T-S_INTERNAL	5

PARAMTYP=“DERIVED”は導出レコードを示し、
PARCAT1というカテゴリ変数で導出された項目を区分して
いることが分かる

Reviewerの立場になって考えると...

SDTMで導出した際のADaM Define+xml

Parameter Value List - ADQSADAS [AVAL]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
AVAL	PARAMCD EQ ACTOT (Adas-Cog (11) Subscore)	integer	8		Derived	QS.QSSTRESN where QSTESTCD=PARAMCD

ADaMで導出した際のADaM Define+xml

Parameter Value List - ADQSADAS [AVAL]

Variable	Where	Type	Length / Display Format	Controlled Terms or Format	Origin	Derivation/Comment
AVAL	PARAMCD EQ ACTOT (Adas-Cog(11) Subscore)	integer	8		Derived	Sum of ADAS scores for items 1, 2, 4, 5, 6, 7, 8, 11, 12, 13, and 14, see Analysis Data Reviewers Guide (Page 3) for details on adjusting for missing values. Analysis Data Reviewer's Guide (analysis-data-reviewers-guide.pdf)

SDTMで導出
ルールを確認

どちらがデータ把握に有益かを考えよう！

SDTMでは評価項目
を確認できない

導出データの格納先

- 複数レコードを用いた導出はADaMで行う
→ 採否情報を考慮する必要が発生
- 1レコード内での導出又は明確な仕様の場合、SDTMで導出してもTraceabilityに支障をきたさない
→ BMI, QTCB, QTCFなど

…ではAVISITはSDTMで導出してもいいの！？

ADaM IGで定義がある変数は、ADaMで作成するのが適切

ADaMデータ作成時の懸念事項

SDTM/ADaMデータ間

採否フラグをどちらに持たせるのか

導出データをどちらに持たせるのか

ADaMデータ内

レコードを追加するか又は変数追加するか

採否フラグの使用方法

(Recorded, Parameter-Level)

レコード追加するのか又は変数追加するのか

Row	USUBJID	VISIT	AVISIT	ADY	PARAM	AVAL	DTYPE	VSSEQ
1	1001	Screening	Screening	-12	SUPINE SYSBP (mm Hg)	144		1
2	1001	Baseline	Baseline	1	SUPINE SYSBP (mm Hg)	145		2
3	1001	Week 1	Week 1	6	SUPINE SYSBP (mm Hg)	130		3
4	1001	Week 2	Week 2	12	SUPINE SYSBP (mm Hg)	133		4
5	1001	Week 1	Post-Baseline Minimum	6	SUPINE SYSBP (mm Hg)	130	Minimum	3
6	1001	Week 2	Post-Baseline Maximum	12	SUPINE SYSBP (mm Hg)	133	Maximum	4
7	1001		Post-Baseline Average		SUPINE SYSBP (mm Hg)	131.5	Average	
8	1001	Week 2	Endpoint	12	SUPINE SYSBP (mm Hg)	133	LOV	4
9	1002	Screening	Screening	-14	SUPINE SYSBP (mm Hg)	144		22
10	1002	Baseline	Baseline	-1	SUPINE SYSBP (mm Hg)	144		23

最大値

最小値

Row	USUBJID	VISIT	AVISIT	ADY	PARAM	AVAL	ANL01FL	ANL02FL
1	1001	Screening	Screening	-12	SUPINE SYSBP (mm Hg)	144		
2	1001	Baseline	Baseline	1	SUPINE SYSBP (mm Hg)	145		
3	1001	Week 1	Week 1	6	SUPINE SYSBP (mm Hg)	130	Y	
4	1001	Week 2	Week 2	12	SUPINE SYSBP (mm Hg)	133		Y
9	1002	Screening	Screening	-14	SUPINE SYSBP (mm Hg)	144		
10	1002	Baseline	Baseline	-1	SUPINE SYSBP (mm Hg)	144		

どちらでも問題はないが、
社内で統一されていた方が利用しやすい

レコード追加するのか又は変数追加するのか

- 疾患ガイドラインで低下量が指標になっている場合
→ CHGは変化量であり、低下量ではない

PARAM	AVISIT	VSSEQ	ABLFL	AVAL	BASE	CHG	PARAMTYP	AMDEC
Weight (kg)	Screening	101		99	100	.		
Weight (kg)	Baseline	102	Y	100	100	0		0
Weight (kg)	Week24	103		94	100	-6		6
Weight (kg)	Week48	104		92	100	-8		8
Weight (kg) AOD	Week24	103		6			DERIVED	
Weight (kg) AOD	Week48	104		8			DERIVED	

どちらの対応でもADaM IG上は問題なし

レコード追加するのか又は変数追加するのか

- レコード追加をすると…

標準変数であるため使い勝手が良い

レコード数が多くなり、
どのレコードから導出されたものかが判別しづらくなる

- 変数追加すると…

レコードが少なくて済む

新たな変数を探し理解する必要がある

ADaMデータ作成時の懸念事項

SDTM/ADaMデータ間

採否フラグをどちらに持たせるのか

導出データをどちらに持たせるのか

ADaMデータ内

レコードを追加するか又は変数追加するか

採否フラグの使用方法

(Recorded, Parameter-Level)

採否フラグの使用方法

Table 3.3.8.1 Flag Variables for BDS Datasets

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core
ANLzzFL	Analysis Flag zz	Char	Y	Cond

Table 3.3.8.2 BDS Population Indicator(s) Variables

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core
ITTRFL	Intent-To-Treat Record-Level Flag	Char	Y	Perm
SAFRFL	Safety Analysis Record-Level Flag	Char	Y	Perm
FASRFL	Full Analysis Set Record-Level Flag	Char	Y	Perm
PPROTFL	Per-Protocol Record-Level Flag	Char	Y	Perm
COMPLRFL	Completers	Char	Y	Perm

ITTPFL	Intent-To-Treat Parameter-Level Flag	Char	Y	Perm
SAFPFL	Safety Analysis Parameter-Level Flag	Char	Y	Perm
FASPFL	Full Analysis Set Parameter-Level Flag	Char	Y	Perm
PPROTPFL	Per-Protocol Parameter-Level Flag	Char	Y	Perm

◆解析レベル

解析集団にかかわらず同じ
取扱いをする場合

◆解析集団レコードレベル

各解析集団でレコードごとに
取扱いが異なる場合

◆解析集団パラメタレベル

各解析集団で複数レコード
を考慮しパラメータの取り扱い
を決める場合

採否フラグの使用方法

Analysis Flag zz

- 使用方法
 - Time Allowance内に複数データが存在した場合に、どちらを採用するのか
 - 最大値データを明らかにする等
- 01から始まる必要はないので社内で用途を決めておくとう便利（ANL10FLは最大値データフラグ等）
- 変数ラベルが自由（ADaM IG v1.1 Draftにおいて）
 - 用途がわかるラベルにすることで使用しやすくなる

解析集団レコードレベル

- 使用方法
 - 解析集団ごとに使用するレコードが異なる場合
- 変数名を見れば用途が明確

採否フラグの使用方法

解析集団パラメータレベル

- 使用方法
 - ベースライン値が欠測でなく、少なくとも投与後の値が存在する場合のパラメータを採用する場合に使用する
 - 感度分析等に使用できる
- 条件によってパラメータ or レコードレベルかの区別が可能

解析集団レコード/パラメータレベルの変数は用途が明確であるが、Analysis Flag zzは変数名だけでは用途を特定することは困難
適切な使い分けを心がけよう！

部門連携強により
効率の良い申請データ作成を目指しましょう！

ご清聴頂きありがとうございました



＜参考文献＞

- **Analysis Data Model Implementation Guide Version 1.1 (Draft)**
- **Study Data Tabulation Model Implementation Guide: Human Clinical Trials Version 3.2**
- **CDISC Define-XML Specification Version 2.0**
- **A Guide to the ADaM Basic Data Structure for Dataset Designers**
<http://www.pharmasug.org/proceedings/2014/DS/PharmaSUG-2014-DS11.pdf>
- **CDISC Q&A Forum**
<http://qa.okada.jp.org/>