

# 産官学でのCDISC標準利用に向けた取り組み：実務担当者のために

## 承認申請のためのCDISC実装とメタデータ作成

氏名

浅見由美子（第一三共）、奥田恭行（第一三共）、Tony Chang（Amgen Inc.）

Efforts toward the utilization of CDISC standards in industry, government, and academia

Implementation of CDISC Standards and the Metadata for Regulatory Submission

Name

Yumiko Asami<sup>1)</sup>, Yasuyuki Okuda<sup>1)</sup>, Tony Chang, Amgen Inc.<sup>2)</sup>

1) Daiichi Sankyo Co., Ltd., 2) Amgen Inc.

## 要旨

Clinical Data Interchange Standards Consortium (CDISC) とは、医薬品に関する臨床・非臨床データの標準を推進する非営利団体である。2016 年（予定）より、PMDA 承認申請の際に CDISC のデータ標準である SDTM（申請臨床試験データモデル：Study Data Tabulation Model）/ ADaM（統計解析データモデル：Analysis Data Model）に基づく電子データセットの提出が義務化されることが公表されている。CDISC 標準に基づくデータセット（以下、CDISC データセット）の作成には、CDISC ガイダンスに基づいたメタデータ（データセットの仕様）が必要となる。メタデータには CDISC ガイダンスをそのまま適用すれば作成できる部分もあるが、承認申請のスポンサー（以下、スポンサー）が作成しなければならない部分も多く含まれる。承認申請用の併合解析や、SDTM/ADaM 作成用の SAS プログラムを効率よく構築するためには、臨床試験単位ではなく、プロダクトもしくはスポンサー内で共通で使用できるメタデータを作成することが重要となってくる。本稿では、グローバル承認申請における CDISC 実装の経験をもとに、プロダクト内で適切に効率よくメタデータを構築する方法を提案する。あわせて、メタデータ作成における課題についてもまとめる。

キーワード：CDISC、SDTM、ADaM、メタデータ、SAS、承認申請

## 1 背景

### 1.1 CDISC

Clinical Data Interchange Standards Consortium（以下、CDISC）とは、医薬品に関する臨床・非臨床データの標準を推進する非営利団体である。<sup>1)</sup> CDISC が作成した標準（以下、CDISC 標準）には、以下のようなモデルが含まれる。

- Study Data Tabulation Model (SDTM)：医薬品の規制当局の申請臨床試験データ形式
- Analysis Data Model (ADaM)：医薬品の規制当局の承認申請統計解析データ形式
- その他、CDASH、SEND、PR 等

代表的な例として、Electric Data Capture (EDC)等で収集された臨床試験データから SDTM を経て、SAS 等を用いて ADaM データセットが作成される。そして ADaM データセットを基に、SAS 等を用いて統計解析結果が生成される。図 1 に、SAS と SDTM/ADaM を用いた統計解析プロセス例（模式図）を示す。

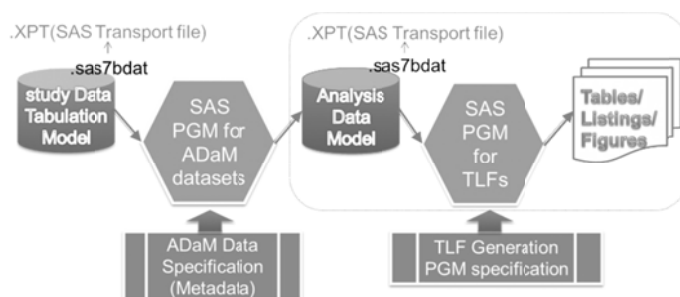


図 1：SAS と SDTM/ADaM を用いた統計解析プロセス例（模式図）

## 1.2 国際共同開発

「国際共同治験に関する基本的考え方」<sup>2)</sup>の発行により、日本が参画する国際共同試験の数が急激に増加し、複数のスポンサー（製薬企業またはアカデミア）や複数の地域（日本、アメリカ、ヨーロッパ等）が開発計画に関与し、同じ時期に世界の各規制当局に対する承認申請を実施する場合もある。その結果、日本の製薬企業、アカデミアおよびCRO等の組織においても、FDA 承認申請のために、日本で実施した臨床試験の CDISC データセットを作成したり、アメリカで実施された臨床試験の CDISC データセットを用いて、日本における承認申請のための解析を実施したりする場面も増加してきた。

## 1.3 医薬品承認申請における電子データ提出に関する各規制当局の対応

以下に各規制当局における医薬品承認申請における電子データ提出に関する対応を簡単に述べる。

FDA は 1980 年代から SAS データセット（その後、SAS Transport v5 format）を移送形式とした電子データ提出をスポンサーに要求し、さらに 2004 年からは、データセットの構造、変数の属性等の標準として CDISC 標準を推奨してきた。そして、2014 年に「Draft Guidance for Industry Providing Regulatory Submissions in Electronic Format Standardized Study Data」ならびに「Draft Study Data Technical Conformance Guide」を発行した。それにより、FDA の医薬品承認申請における CDISC 標準に基づく電子データ提出の強化が予想される。<sup>3)</sup>

一方、PMDA（医薬品医療機器総合機構）は臨床試験の電子データ提出を従来スポンサーに要求していなかったが、最近 CDISC SDTM/ADaM 標準に基づく電子データ提出を 2016 年（予定）より義務化することを公表した。これにより、日本の製薬企業、アカデミアおよびCRO等の組織においても、PMDA の承認申請のために CDISC データセットを作成し、提出することが必要となる。<sup>4)</sup>

# 2 CDISC データセットのメタデータ作成における問題点

## 2.1 CDISC データセット作成におけるメタデータ

メタデータ（metadata）とは、一般的には「データに関するデータまたは情報」を意味する。CDISC 標準の実装においては、データセット、変数等に関する定義、仕様書にあたり、SAS 等を使用して SDTM/ADaM データセットを作成（変換）する際のベースとなる。メタデータは、SDTM Implementation Guide（以下、SDTM IG）、ADaM Implementation Guide（以下、ADaM IG）等の CDISC 標準に関するガイド類<sup>1)</sup>に従う必要がある。承認申請時には「Define file」に含めて、規制当局へ提出される。なお、本稿においては、メタデータは SDTM/ADaM データ作成（変換）の元になる仕様書を意味し、Define file の意味を含めないこととする。

図2にSASを用いて作成した場合のSDTM AEドメインの一部（見本）を示す。メタデータの検討の際には、データセット（ドメイン）レベル、および、変数レベルで検討する必要がある。



図2：SASを用いて作成した場合のSDTM AEドメインの一部（見本）

## 2.2 メタデータ作成における問題点

日本製薬工業協会（以下、製薬協）が2013年に実施したアンケート<sup>5)</sup>によると、製薬協加盟企業の約半数がSDTM実装の経験があるが、そのうち、その約3/4はSDTM標準に適合しているとは言えない（いわゆる「SDTM Like」）であるとのことであった。それらの企業では、SDTMデータセット作成の基となるメタデータがSDTM IGに適合していない可能性がある。

EDC等で収集された臨床試験データ（以下、CRFデータ）を基にSDTMデータセットを作成するためには、SAS等を用いたデータ加工（複数データセットのマージや転置等の構造変換、変数の属性変更、単数もしくは複数の変数の値を代入した四則演算や関数等を用いた変換など）が必要となる場合がある。CRFデータの構造によっては、データ加工の仕様が複雑になることもあるため、SDTMメタデータ作成のためにはSAS等によるデータ加工のプログラミングの知識も必要となる。SDTM実装の経験が少ない組織では、SDTMメタデータ作成におけるSASプログラミング知識の必要性が認識されていない場合もあり、SDTM作成を担うファンクションにSASプログラミングの知識のある人材が不足している可能性もある。

また、SDTM IGやADaM IGには画一的ではなく、複数のオプションを取りうる事項が含まれている。そのことが、メタデータ作成を困難にしている可能性がある。

## 3 医薬品の承認申請における統計解析とデータ標準化

図3にSASとADaMを用いた承認申請用解析のプロセス例（模式図）を示す。使用する臨床試験数で出力する解析帳票を分類すると以下ようになる。

- 1) 単一の臨床試験のデータを用いる解析帳票（主に総括報告書用）
- 2) 複数の臨床試験のデータを用いる解析帳票（主に承認申請資料用）

複数の臨床試験データを用いる解析帳票には、例として併合解析や複数試験の「横並び」の解析（臨床試験、投与群を横に並べたような解析）がある。

臨床試験間で解析データの仕様が統一されていない場合、1)において解析プログラムを再利用することが難しくなり、臨床試験毎に解析帳票作成プログラムを作り直すことが必要となる。また2)において解析を実施する前に、承認申請資料に含まれる複数の臨床試験の解析データの仕様をそろえるための変換が必要となる。その変換作業は煩雑になることも多く、承認申請タイムラインに影響を与える場合もある。

そこで解析データ仕様の標準化が重要となってくる。CDISC標準は規制要件としてだけでなく、上記のような問題を回避するためにもメリットがあると考えられる。更に、レビューアー（社内の関係者、規制当局

審査官等）がデータをレビューしやすくなるというメリットもあると考えられる。

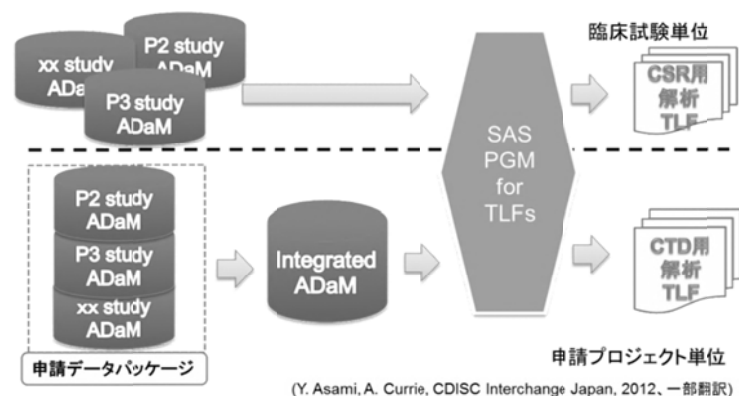


図3：SASとADaMを用いた承認申請用解析のプロセス例（模式図）

## 4 メタデータの作成手順と管理項目

### 4.1 メタデータの記述パターン

SDTMとADaMでは、より自由度の低いSDTMメタデータのIGへの適合性がより重要であると考えられるため、SDTMメタデータを中心に検討を進める。実際のSDTM実装の経験を元に、表1にSDTMメタデータにおける記述のパターンの分類を示す。本稿においては4つの分類とした。CRFデータの構造に依存し1)に該当する割合が高くなり、メタデータの作成、及び、その後のプログラミング作業の負担が大きくなる。また、SDTMの経験が少ない組織の場合は、2)、3)に該当するメタデータ作成、および、そのSDTM IGへの適合性の確認に時間が掛かる可能性がある。

表1：SDTMメタデータの記述パターンと代表的なデータ加工例

メタデータの記述パターン	代表的なデータ加工例 (データ構造変換、属性の変更、派生等)
0) CRFデータのCopy	—
1) SDTM IGで決定されている属性とCRFデータのギャップの穴埋め	Dataset Name、Structure、Key variableの変更 Variable Name、label、Typeの変更 Controlled Terminologyの変更、等
2) 一般的にCRFデータとして収集されておらず、他のマテリアルの記述に依存するもの	Trial Design Terminologyの作成 Baseline Flagの作成 Population Flagの作成、 Extensible Terminologyの追加 Sponsor Specified Terminologyの追加、等
3) その他、一般的にスポンサーが決定する属性	Custom Domain Name、Structureの設定（Classの選択） Variable Lengthの調整 Supplemental Qualifierの作成、等

\*注：2)と3)は重複もありうる。

## 4.2 医薬品承認申請を考慮したメタデータ作成手順

表 1 のメタデータの記述パターンを基に、医薬品の承認申請時の工数軽減も考慮した、メタデータ作成の負担軽減策を表 2 にまとめる。

表 2：SDTM メタデータの記述パターンと負担軽減策

メタデータの記述パターン	負担軽減策
0) CRF データの Copy	ー
1) SDTM IG で決定されている属性と CRF データのギャップの穴埋め	・ CDASH の活用、プロジェクトまたは組織内の CRF データ標準化、英語化（日本ローカル試験の場合）
2) 一般的に CRF データとして収集されていない項目	<ul style="list-style-type: none"> <li>・ SDTM を考慮したプロトコル標準化、マシリーダブル形式の採用、英語化（日本ローカル試験の場合）</li> <li>・ SDTM を考慮した統計解析早期策定（ベースラインの定義、解析集団の定義、等）等</li> <li>・ 図 4 及び表 3 参照</li> </ul>
3) 一般的にスポンサーが決定する属性	・ 図 4 及び表 3 参照

表 2 における 2) 及び 3)、及び、ADaM メタデータは、スポンサーで決定できる属性を含むため、承認申請プロジェクト内の臨床試験間でばらついてしまう場合がある。そのことが承認申請用の解析実施時の工数増大につながる可能性がある。また、臨床試験ごとにメタデータを作成する場合、本来同じ派生ルールの場合でも、記述方法が異なってしまう可能性もある。

そこで、図 4 のイメージにあるような、承認申請プロジェクト内の臨床試験のメタデータを横断的に管理できるようなシートを用いることとした。一般的に、承認申請プロジェクトに含まれる臨床試験は実施時期が異なるため、新規の臨床試験時はカラムを追加する形式とした。また、そのシートを用いたメタデータの管理項目と作成手順を表 3 に示す。

	A	B	C	D	E	F	G	H	I	J	K
	Dataset	Variable Name	Variable Label	Variable Type	Display Format	Codelist/ Controlled Terms	General Source/Derivation	Core	Study x	Study y	Study z
1											
2	ADxx	aaa	zzzz	Char	\$zz	zzzz	zzzz	Req	x	x	if xxx=xxx then aaa="X"
3	ADxx	bbb	zzzz	Char	\$zz	zzzz	zzzz	Cond	NA	x	NA
4	ADxx	ccc	zzzz	Char	\$zz	zzzz	zzzz	Req	x	x	x
5	ADxx	ddd	zzzz	Num	zz	zzzz	zzzz	Cond	NA	NA	x

図 4：承認申請パッケージ内のメタデータ管理シート（イメージ）

表 3：承認申請パッケージ内のメタデータ管理項目と作成手順

順序	レベル	管理項目
1	データセット（ドメイン） レベル	Dataset Name, Structure, Key Variable(s) *必要な場合はカスタムドメインの設計
2	変数レベル	Variable Name, Label, Type, Display format, Derivation Rule *特に CRF データの構造が臨床試験間で異なる場合は変数 派生ルール(Derivation Rule)の確認が重要
3	ターミノロジー、辞書	Controlled Terminology (Extensible or not) Sponsor Specified (Trial design, etc.) *臨床試験間で別の意味のデータに、同じ Terminology を当 てはめないように注意

## 5 実際の承認申請における適用

図 4 のシートと表 3 に示す方法を用いて、実際のグローバル承認申請プロジェクトにおけるメタデータを作成、管理した。その結果をまとめると以下のとおりである。

- 既に存在する変数はそのまま使用することができた。また、既に存在する変数を誤って別の名前または属性で新規に作成することが防げた。
- 臨床試験間で同じ意味の場合、Terminology を共通して使用することができた。一方、別の意味の場合、誤って臨床試験間で同じ Terminology にならないように、排他的に作成できた。
- 変数派生ルールの記述を再利用することができた。また、それに基づく SAS コードを再利用することができた。
- 異なる会社、地域間でメタデータを見える化、共有化することができた。

## 6 まとめ

SDTM/ADaM メタデータを IG に適合する形で、効率よく作成するためには CDASH およびプロトコル標準を用いて、End-to-End の標準化をすることが望ましい。しかし、組織全体でそのようなプロセスに移行するにはある程度の時間が掛かる。また、医薬品の承認申請準備の効率化も考慮すると CDASH の適用のみでは解決できない問題点もある。そこで、本稿では実際の SDTM メタデータ作成の経験を基に、SDTM メタデータの記述パターンを分類することにより、CDASH の適用で負担を軽減できる部分と、その他の方策も必要な部分を明確にした。CDASH 以外のその他の方策の一つとして、医薬品承認申請を考慮したメタデータ作成手順と管理項目を設定した。そして、実際のグローバル承認申請でその方法を用いた結果をまとめた。

PMDA 承認申請における 2016 年（予定）からの CDISC データセット提出の実現のためには、End-to-End で標準化されたプロセスへの移行準備と、それが定着するまでの過渡期の対応を同時に進めていく必要があると考える。

< 参考 >

- 1) CDISC Website <http://www.cdisc.org/>
- 2) 薬食審査発第 0928010 号. 国際共同治験に関する基本的考え方 ; 2007
- 3) FDA Study Data Standards Resources  
<http://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/default.htm>
- 4) 医薬品医療機器総合機構 次世代審査・相談体制について（申請時電子データ提出）  
<http://www.pmda.go.jp/operations/shonin/info/iyaku/jisedai.html>
- 5) 製薬工業協会発表資料、CDISC Interchange Japan ; 2013