CDER Common Data Standards Issues Document に関するメモ（私家版）

内容

1 この文書は何を記述しているのか？ 3
2 データ標準に見受けられる諸問題 4
2-1 一般的事項 4
2-2 SDTM に関する事項 4
2-3 用語に関する事項 5
2-4 SDTM のドメインに関する事項 6
2-5 ADaM に関する事項 7
2-6 変数に関する事項 7

注意：本文書はとあるサークルが作成した私家的なメモです。内容には意図しない誤りが存在する可能性があります。本文書の使用は自己責任にて行って下さい。本文書の使用にて発生したあらゆる損害に対して、当サークルは一切の責任を負いません。また、本文書に関する質問に対して当サークルは基本的に回答いたしません。しかしながら、コメントや提案を歓迎しております。

現バージョンの公開は以下のサイトに限定しています。
http://doubledealer989.blog74.fc2.com/blog-category-14.html

1 この文書は何を記述しているのか？

CDER は、申請業務の品質向上と効率アップのために、標準的なデータ構造を用いたデータ提出を求めています。予算削減の一方で、医薬品の安全性向上が求められており、CDER は『より少ないもので、より多くを』こなすという使命に立ち向かっています。この目的のためには、統一的な構造でデータを集める事が必要です。「標準的なデータ構造」は唯一無二に決まる物ではありません。これは決断の問題次第ですが、CDISC という団体が作成している SDTM、ADaM、SEND といったデータ定義が採用されています。

CDER は 2004 年から CDISC 標準を用いたデータ提出を受け入れてきました。しかし、データ提出者側（典型的には製薬企業です）に様々な誤解があり、かつ実装の個性が見受けられています。これでは『より少ないもので、より多くを』という究極的な目標を達成することはできません。そこで、CDER はどのような SDTM を求めているか、ガイダンスを出しました。それが CDER Common Data Standards Issues Document になります。しかし、仕様の詳細を全て記述することは不可能です。したがって、本文書の参照だけでは不十分であり、実務レベルでは当局に相談することが非常に重要です。

元文書[†]は定期的にアップデートされる予定です。ここでは Version1（2011 年 5 月）を参照しています。

†:http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM254113.pdf

2 データ標準に見受けられる諸問題

2-1 一般的事項

2-1-1 最新の規格に対応する

データ提出者は最新版のスペックを参照するべきです。新しい規格は、以前のバージョンでカバーされていなかった領域を考慮しているかもしれません。また、新しいバージョンには、レビューを容易にするための改良が期待できます。これはレビューワーにとって望ましいものであり、同時にデータ提出者側に最新規格への対応を十分に考慮することを求める理由になります。

2-1-2 正しい Define.xml を作成する

Define.xml（xml 形式が望ましい）は重要なコンポーネントです。コードリストの内容等をきちんと記述することが求められます。

2-1-3 時点コードの有用性

EPOCH や ELEMENT の情報は有用です。どの時期に何が発生したか把握できます。例えば、有害事象のおおよその発現時期や薬剤師用のタイミングを知ることができるでしょう。そして、時期情報は”発生時点”を反映しなければなりません。すなわち、有害事象であれば発現時点の EPOCH や ELEMENT をデータとして採用します。

2-2 SDTM に関する事項

2-2-1 事前に準備する

最初から SDTM を想定することが肝要です。具体的には、CDASH 等を利用してデータ収集段階から SDTM に備えることです。後付けで臨床データベースを SDTM へ変換することは、賢い選択ではありません。実例を見る限り、後付けの変換は複雑で難解となり、レビューワーにとって全く有用ではありません。

2-2-2 Traceability

SDTM のデータの出自が辿れるようにしておかなければなりません。

2-2-3 解析データセットとの関係

臨床データをデータベースに保管し、そこから SDTM を作成することがあります。このシナリオでは、解析データセットをデータベースから作成できます。これはスポンサーが決定するべき手順です。しかし、スポンサーの手順がどうあれ、CDER に提出する SDTM と ADaM データセットは本質的に整合性が確保されているべきです。解析データセットと SDTM が独立して作成されたために、両者が矛盾している例があります。これはもちろん良くない例です。

2-3 用語に関する事項

2-3-1 用語

用語集は既に定義済みですが、完璧ではありません。必要な用語があれば、適宜作成します。その際は、追加した用語の情報を併せて提供します。

2-3-2 コード辞書

薬剤のコードには一般名のコードを利用します。コード辞書には WHO DRUG などを利用します。一方、病名のコードには MedDRA を利用します。このときは、辞書にある通りのリテラルを用いなければなりません。しばしば、MedDRA 辞書と一致しないリテラルが見受けられます。また、辞書には「数値コード」がつきものですが、これは SDTM に含めてはいけません。

2-4 SDTM のドメインに関する事項

2-4-1 SUPP のゴミ箱化

SDTM では、SUPPーというドメインを作成し、そこに余り物のデータを格納します。これは仕様に従っていますが、評価の上で重要なデータは SUPP--ドメインに含めるべきではありません。

2-4-2 DM

DM ドメインは 1 レコード/1 症例です。当ドメインに関して Amendment がリリースされました。この Amendment に対応することが強力に推奨されています。

2-4-3 DS

DS ドメインに複数のレコードが発生する場合、時点データを付属します。EPOCH 変数を利用します。また、死亡（DEATH）のレコードは最後に配置します。

2-4-4 AE

データ提出者は「TEAE(trestment emergent AE)」のみに興味があるかもしれません。しかし、TEAE に関わらず、AE ドメインには報告された全ての事象を盛り込みます。Amendment 1 にて、TEAE 用のフラグが追加されています。区別にはこのフラグを使用します。

AEBODYSYS と AESOC という変数には MedDRA の SOC コードを格納します。
MedDRA の構造上、SOC には Primary と Secondary があります。CDER は AESOC 変数に Primary のデータを格納する手法を想定しています。Secondary が必要な場合には AEBODYSYS を利用します。これは Amendment1 にも明記されています。

2-4-5 LB

LB は巨大なファイルになることが多いものです。意味のある分割（例えば、検査のカテ

ゴリごとにファイルを分割）をして提出します。このときは「分割する前」の巨大なデータセットも提出します。400MB 程度のファイルサイズであれば受け入れられるでしょう。

2-4-6 カスタムドメイン

カスタムドメインの作成はしても構いません。しかし、その前に別のドメインでカバーできないか、最新のドラフトがリリースされていないか確認をしましょう。

2-5 ADaM に関する事項

2-5-1 参考資料

3 つの文書を参照して作成しましょう。ADaM、ADaM IG、FDA Study Data Specification Document がそれに当たります。

2-5-2 何を ADaM に含めるか

今のところ、必要な解析データセットの数は指定されていません（唯一の例外は ADSL です）。だからといって、大量の解析データセットを作ればよいという訳ではありません。また、ADSL だけで十分でもありません。主要評価項目と安全性評価の上で必要な解析データセットを ADaM パッケージとしてまとめます。これはすなわち、どのデータセットをつくるか十分に吟味する必要があるということです。

2-5-3 ADSL

ADSL は試験毎に必ず作成します。ADaM IG で指定した情報に追加して、被験者のベースライン時の背景を追加するのが好ましいです。

2-6 変数に関する事項

SDTM の仕様には「Reqired」、「Expected」、「Permissible」の 3 種類の要求レベル

が存在します。しかし、これは元データにあるものを SDTM 作成時に省略してよいという意味ではありません。存在するデータはすべて SDTM に含めます。同時に、変換データを含めないことも重要です。

CDER Common Data Standards Issues Document に関するメモ（私家版）

2011, Circle TKD+SMZ, all rights reserved.