

Associated Person Data 拡張解説

注意：本文書はとあるサークルが作成した私的なメモです。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 イントロダクション

1-1. Associated Person Data とは何か

Associated Person Data Supplement は、AP 標準とも呼ばれる SDTM の拡張標準の一つです。この文書は何を記述しているのでしょうか? Associated Person Data とは何でしょう?

臨床試験においては、被験者ではない“別の人物”のデータが収集される場合があります。

「Associated Person Data」とは、被験者ではない別人物のデータを指したものです。そして、AP 標準は Associated Person Data を SDTM の一環として実装するためのルールを示したガイドラインです。このガイドラインでは、AP 標準の基本的なコンセプト、重要な原則、そしていくつかの追加要素が記載されています。以下、これらの要素を順に見ていきます^{+1,2}。

+1: もしあなたが、SDTM を一度も読んだことがないのなら、ここから先を読み進めることは完全に無意味です。今すぐに SDTM を(可能ならば SDTM IG も)読んでください。また、そのバージョンは **3.1.4 以降** でなければなりません。なぜならば、AP 標準は SDTM の拡張であり、SDTM の知識を前提としているからです。SDTM を理解せずに AP 標準を読むことは、読めない字で書かれた本を眺めるようなものです。

+2: AP 標準は、かつて NSIG (Non Subject Data Implementation Guide) と呼ばれていました。NSIG と AP 標準には微妙な違いがあるものの、基本的なアプローチは同一です。

1-2. Associated Person Data の事例

新しいものを知る際に、その背景を確認しておくことは大切でかつ有用です。AP 標準を読み進める前に、Associated Person Data とはどのようなものかを考えてみましょう。ここでは具体的な事例を想定することでイメージをより確実なものにします。Associated Person Data の例として、次のシナリオを想定できるでしょう。

1. 被験者本人のリスクファクターを知るために、親族の病歴を調査する
2. ドナーの情報を収集する
3. 性的パートナーの状態を把握する
4. 被験者の世話をしている人に対して質問する
5. "被験者の環境"の記述としての他者
6. オペレーター

各シナリオについて、詳しく考察してみます。

(1) 被験者本人のリスクファクターを知るために、親族の病歴を調査する

親族の病気の情報は、遺伝的な疾患を対象にした臨床試験で重要です。例えば、1型糖尿病は遺伝性の病気です。被験者の親族に糖尿病患者がいるか調べる必要があるかもしれません。このとき「親族が糖尿病である」という情報は Associated Person Data になります。なぜならば、このコンテキストにおいて「糖尿病の罹患状態」は被験者本人の病気ではないからです。

(2) ドナーの情報を収集する

これは分かりやすい事例です。血液や臓器提供者の「性別」や「人種」といった背景情報を収集します。これは明らかに被験者本人の情報とは異なります。被験者が女性の場合、血液のドナーは女性でしょうか?男性であるかもしれませんし、複数のドナーがいるかもしれません。

(3) 性的パートナーの状態を把握する

これは意味深長な事例です。単純な例として「臨床試験中に性的パートナーが妊娠しましたか?」という設問が考えられます。この質問をするのは男性の被験者に限られます。そして男性は妊娠できません。したがって、妊娠の有無は Associated Person Data と考えることができます³。

(4) 被験者の世話をしている人に対して質問する

アルツハイマー病の臨床試験では、患者の世話をしている人に質問をすることがあります。質問に対する回答は、被験者のデータでしょうか?残念ながら、被験者は質問に回答していません。これも Associated Person Data にカテゴライズされるべきです。

(5) 被験者の置かれた環境の記述としての他者

ここに「ぜんそく」の臨床試験があると考えてください。ぜんそくをもつ被験者がホコリの多い環境（工業地帯に住んでいる）に暮らしているならば、それは考慮すべき情報かもしれません。では、同居人がヘビースモーカーである場合はどうでしょうか?これも同様に考慮に値する問題です。しかし、同居人の喫煙状況は『環境データ』に過ぎないのでしょうか。それとも、被験者以外の誰かの嗜好品 (Substance Use) のデータなのでしょう。後者と考えるなら、同居人の喫煙状況は Associated Person Data です。

(6) オペレーター

デバイスの試験では、測定機器のオペレーターが事故に巻き込まれる可能性があります。オペレーターは被験者ではありません。試験計画書にて、オペレーターに起きた事故が調査対象とされているなら、情報が収集されます。そして、それは Associated Person Data です。

ここで取り上げたシナリオは Associated Person Data の一部に過ぎません。この世界には、様々なデザインの臨床試験があります。その中には上記の例ではカバー仕切れない Associated Person Data が存在するでしょう。

最後に、2つのタイプの Associated Person があることを心に止めておいて下さい。1つは被験者と生物学的な繋がりがあるもの。もう一つは、生物学的な繋がりがありません。薬剤の評価という観点から考えて、被験者と意味のある関係者（血縁関係や同居人）と無関係の第三者（オペレーター・世話人）を区別することは重要です。

†3：この質問は被験者に対する Questionnaire であると考えられ、その結果として QS ドメインに配置できそうです。しかし、QS ドメインは確立された質問票（SF-36 など）に使用することが推奨されています。そのため、もっともありそうな選択肢は RP ドメインになるでしょう。

2 AP 拡張標準のコア

AP 標準がどんなものであるか、前節の例でイメージを掴みました。この節では AP 標準の基本コンセプトと原則について見ていくことにしましょう。基本コンセプトでは AP 標準の構造を、原則では AP 標準を理解するためのポイントを解説していきます。

2-1 基本コンセプト

2-1-1 データの管理方法

もしあなたが Associated Person Data の保存形式を考える立場であったなら、どんなデータベースに Associated Person Data を格納するでしょうか？この答えは人により異なるはずですが、データの構造化として、複数の異なるアプローチが成立します。AP 標準ではどのようにデータ構造を設計しているのでしょうか？

最もシンプルな方法は、SDTM のデータ規格を厳密にそのまま適用して Associated Person Data を構造化することです。既存の規格を用いるという点において、このアプローチは明解で強力です。都合の良いことに SDTM は自由度が高い規格です。独自ドメインを作成したり、SUPPxx データセットを作成したり、様々な方法で Associated Person Data を保存することができるでしょう。しかし、データ利用⁴の観点から考えると、このアプローチには大きな問題があります。

致命的な点は Associated Person Data が散逸し、情報の再構成が難しくなることです。ここで、血液のドナーの情報を集めた例を取り上げましょう。ある被験者に対して輸血がなされました。輸血に用いられた貴重な血液のドナーは以下の 2 名です。

ドナー A 男性 西部地方 2011-06-11 採血 A Rh+

ドナー B 女性 中部地方 2011-06-11 採血 A Rh+

そして幸運にも血液の提供を受けた被験者は次の通りの背景です。

被験者 X 男性 北部地方 A Rh+

この情報を SDTM の規格に従って構造化してみましょう。血液のドナーの情報を独自ドメインを利用して格納しましょう。独自ドメインは SDTM のデータモデルのいずれかをベースに作成します。従って、この事例ではドナーの背景情報と採血日は別ドメインに格納される公算が高くなります。一方で、ドナーの背景情報と採血日データは意味的に結合しています。分離したデータを結びつきをどのように保証できるでしょうか？SDTM の規則に従うなら、この問題は RELREC を利用することで解決できます。しかし、レビュワーは複数のデータテーブルを参照しなければならず、見やすい状態ではあり

ません。

では、SUPPxx を用いるとどうなるでしょうか？残念なことに事態は更に深刻です。第一にどのドメインの SUPP を選択すればよいのでしょうか？明確なガイドラインは存在しません。仮にドナーの背景情報を SUPPDM に、採血日や採血場所を SUPPLB に格納すと考えましょう。ここでもやはり、RELREC を利用したデータ連結情報が必要になります。そして Associated Person Data 以外にも、様々な余剰データが SUPPxx に直積されることにも注意が必要です。Associated Person Data を探すためには、大量の余剰データから必要な情報を引き出さねばなりません。

この問題を踏まえ、AP 標準では別のアプローチを選択しています。第一に Associated Person Data を SDTM ドメインから明確に切り出しました。第二にデータの連結情報を保つための変数を追加しています。この決断により、Associated Person Data の所在が明らかになり、データの利用効率が高められています。次の節ではより詳しく AP 標準が選択したデータ構造を見てみましょう。

†4：ここではデータレビューや数年前のデータを再利用する例を暗黙的に想定しています

2-1-2. AP 標準が採用したデータモデル

AP 標準における Associated Person Data のモデリングを以下の図に示します。



まず最初に、上の図が極端な単純化をした模式図であることを断っておいてから、Associated Person Data の構造の特徴に触れましょう。第一のポイントは Associated Person Data を被験者データから分離したことです。この図では SDTM に定義されたドメインが図の上部にあります。そして SDTM のドメインと対応するように Associated Person Data のデータセットが存在します（このテキストでは Associated Person Data を格納するドメインの集合を **AP ドメイン** と呼ぶことにします）。これは、Associated Person Data のために別のドメインを作成することを意味しています。ドメインの数は倍増^{†5}していますが、この犠牲を元に Associated Person Data をデータテーブルのレベルで分離することに成功しています。