

## 富士通電子カルテシステムから Reporter を通じてのデータ抽出検証方法について

(追補版)

### 1. 目的

本資料は、2014年10月19日作成の『厚生労働科学研究費補助金（医薬品・医療機器等レギュラトリーサイエンス総合研究事業）「治験活性化に資する GCP の運用等に関する研究」（研究代表者：渡邊裕司）「データ抽出に係わる WG」検討資料（たたき台）～富士通電子カルテシステムから Reporter を通じて EDC へデータ抽出を行う場合のチェック方法について』に関し補足解説を行うための資料である。

前提として、渡邊班 WG において取り上げられた下記宿題事項の検討結果を記載している。

- ・EDC への抽出システムを用いる場合ベンダー側のシステムに関わるチェック項目（ロジックを組んで自動抽出する部分）の論点をリストアップする
- ・富士通の持っている仕組みに沿った形で、ユーザーが検証を行う場合にはこうしてもらいたい、という項目を洗い出す

また、日本医療研究開発機構委託研究開発事業（臨床研究・治験推進研究事業）「臨床研究・治験の IT 化推進のための実施プラン策定に関する研究」（研究代表者：松村泰志）（以下、「松村班」とする。）で検討された内容に関連のある事項、ならびに国内・諸外国の関連状況等について、記載したものである。

### 2. 記載範囲

富士通の現時点における電子カルテ（レポーター）と EDC 連携のフロー（実証実験等での実績ベース）を前提としたデータ抽出のための検証ポイントの抽出検討を範囲とする。

下記は検討の範囲外とする。

- ① 各検討箇所に用いるチェック表などの書式や手順に関する具体的検討
- ② ユーザーによる検証を実施するフェーズに関する検討。CSV のアプローチを踏まえた検討（設計フェーズ + 受け入れテストフェーズ）
- ③ 複数候補データを人が選択する（画面機能を用いる）場合、または自動での抽出と混ざる場合の検証方法
- ④ テキスト型で保持している「単位」の変換や 2 バイト文字が混ざっている場合などの変換ロジックに関する検証方法
- ⑤ 薬事法や GCP ガイダンス等の各種規制や、倫理指針での要求事項などへの対応有無の比較検討

### 3. データ抽出を行う場合に必要となる検証方法の仮説検討

EDC への抽出システム（Reporter）を通じてデータ抽出を行うフローについては、

- ① CRF（症例報告書）の項目定義
- ② 電子カルテシステムの情報と CRF 項目のマッピング
- ③ 電子カルテからデータ抽出
- ④ Reporter から EDC 側への連携

の一連の流れを組み合わせたフローで構成されると考えられる。このため、上記組み合わせフローの最初から最後までを一貫性を持った形で連携できるようにしシステムの検証を行う必要があると考える。

従って、一連のフロー上で定義可能な情報の INPUT/OUTPUT に関わる資料やマッピング表など、システム構築作業で作成される「成果物」を明確にし、ポイント間の「トレーサビリティ」をシステム構築時に確認する方法を基本とすることが望ましいと考え、その方法を検討した。

#### 4. 検証箇所の抽出について

##### (ア) 連携全体のシステムフローの洗い出し

EDC への連携全体の流れについて、プロトコル（連携対象となる臨床試験）のセットアップから、試験実施、データの連携にいたる全体図を富士通方式の実績をベースにして作成した。

(図 1)

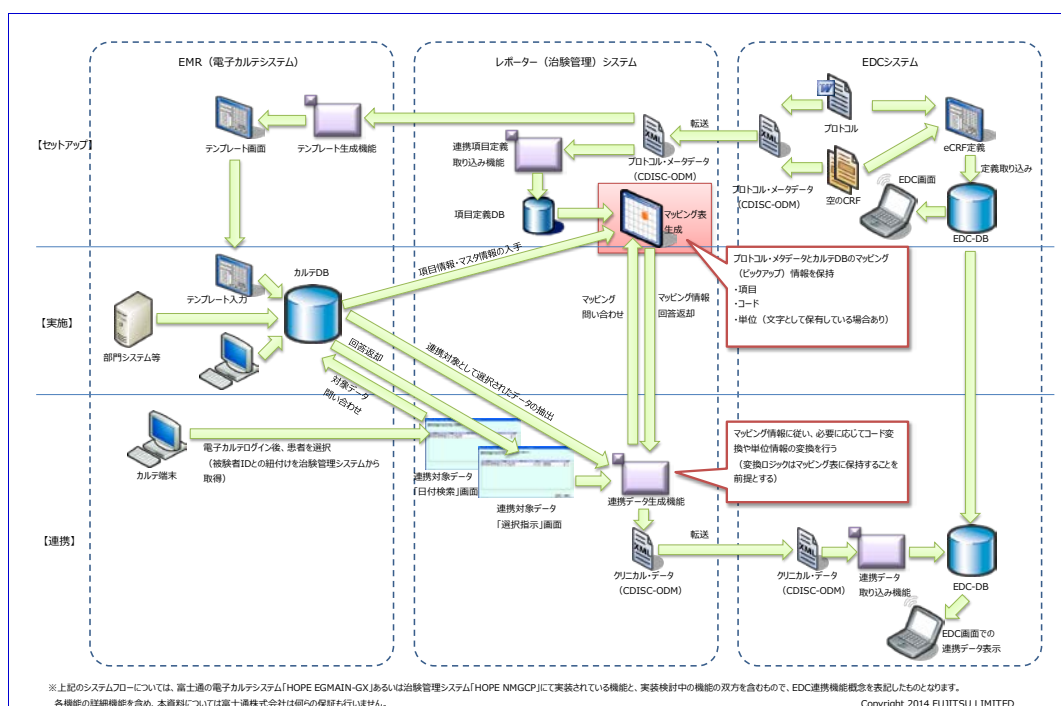


図 1：システム全体フロー

図の上段には「セットアップ」の工程における連携フローをあらわしている。図の中段には「試験の実施」時の工程におけるシステム連携の流れ、同様に下段には「データ抽出連携」時の流れを記載した。

セットアップ時は主に、EDC システムで定義した CRF 項目定義を CDISC-ODM の Study データ（メタデータ）で記述し、CRF の項目定義情報とレポーター側、さらには電子カルテ側で保持する項目情報との「マッピング表」を定義、さらには電子カルテ側で利用する入力用テンプレートの定義を行う

ことが主な作業となる。

実施時には主に電子カルテシステムにおけるデータ入力の作業が中心となる。

連携時はレポーターが連携用データを患者ごとに抽出（たとえば、抽出期間などの条件を与え）し連携データを生成、EDCシステム側に実際のデータ連携を行う部分の作業となる。

このような一連のフローを富士通においては電子カルテからレポーターを通じて EDC ヘデータ抽出連携を行う流れとして定義可能である。

(イ) フロー上のポイントとなる箇所の成果物等の抽出

次に、上記（ア） ので定義したフロー上で発生しうる、項目の定義やレイアウト、マッピング表、生成データなどのシステム上作成される情報（データや設計資料などを含む） 成果物の所在ポイントに番号を付けてマッピングを行ってみたのが図 2 となる。

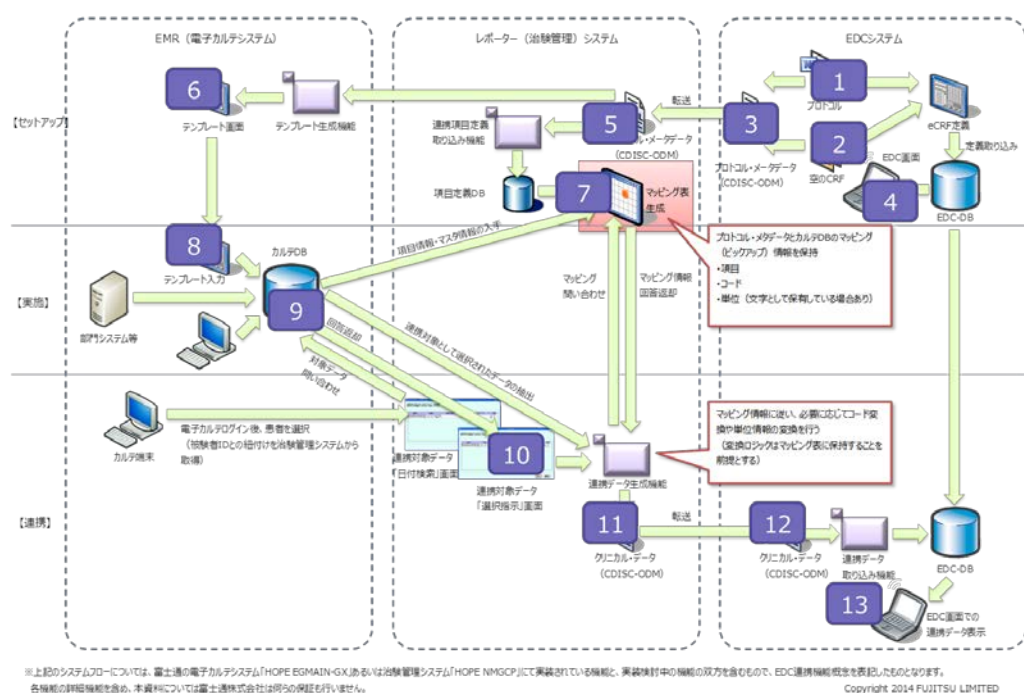


図 2：ポイントのマッピング図

各ポイントをリストアップすると下記のとおりとなる。

1. プロトコル（治験実施計画書）
2. 空の CRF
3. プロトコル・メタデータ＜EDC システム側＞（CDISC-ODM）
4. EDC 画面＜構築時＞
5. プロトコル・メタデータ＜レポーターシステム側＞（CDIS-ODM）
6. テンプレート画面＜電子カルテシステム構築時＞
7. プロトコル・メタデータと電子カルテのマッピング表
8. テンプレート画面＜入力時＞
9. 電子カルテ DB
10. 連携対象データ「選択指示」画面

- 1.1. クリニカルデータ<レポーターシステム側>
- 1.2. クリニカルデータ<EDCシステム側>
- 1.3. EDC画面<連携時>

(ウ) 検証ポイントの明確化と検証方法の抽出

システム構築時には、前述の情報成果物を単体で設計書としたり、テスト仕様書の元データとして活用することが可能である。

しかし、今回はEDCへの抽出システムの連携されるデータが妥当であるかどうかをシステム構築時等に事前検証にて明確化することが必要であることから、一連の連携フローに則って定義されたデータ項目が正しく連携されていることを「トレース」(=追跡)可能となっていることが必要であると考える。

今回、その「トレース」を行う方法として、前述の各ポイントにおける成果物と別のポイントの成果物の間で正しく意図された項目どおりのデータ連携が行われていることを検証し、それらをいくつかのポイントで検証を行って、最終的にはトレースの結果をすべてつなげることで一連の連携が問題ないことを証明する方法が良いと考えた。

そのために各ポイント成果物間の検証実施項目と内容について、下記直交表に洗い出しを行った(表1)。

項目	チェックポイント名	項目														
		1	2	3	4	5	6	7	8	9	10	11	12	13		
1	サイトごとのCRF (最終実施計画書)															
2	元のCRF															
3	サイトごとのデータ <EDCシステム側> (CDISC-ODM)	10のサイトごとのデータで必要な項目が正しくODMファイルに記述されていること。	20のCRFレイアウトが正しくODMファイルに記述されていること。													
4	EDC画面<構築時>	10のサイトごとのデータで必要な項目が正しくEDC画面で表示されていること。	20のCRFレイアウトが正しくEDC画面で表示されていること。													
5	サイトごとのデータ <レポーターシステム側> (CDISC-ODM)			抽出されたODMファイルの同一性が確認できること。												
6	ランプレート画面<電子カルテシステム構築時>						データがランプレート画面で正しく生成されていること。									
7	サイトごとのデータ<電子カルテのシステム側>						データがランプレート画面で正しく生成されていること。									
8	ランプレート画面<入力時>															
9	電子カルテDB							ランプレート画面で電子カルテの抽出結果が正しく生成されていること。								
10	連携対象データ一覧選択画面															
11	サイトごとのデータ<レポーターシステム側> (CDISC-ODM)							70の11のデータ一覧選択画面でODM出力されていること。	80の11のデータ一覧選択画面でODM出力されていること。	90の11のデータ一覧選択画面でODM出力されていること。						
12	EDC画面<連携時>															
13	EDC画面<連携時>															

表1：直交表

直交表で表しうるすべての組み合わせで検証が必要あるいは可能というわけではないため、必要な箇所のみを記載(表中の黄色で網掛けされているところ)している。また、検証の実施が必要な箇所ならびに検証方法については、さらに詳細な検討が必要な状況と考える。

下記に、直交表（表 1）に記載した検証項目をリストアップする（表 2）。

4 (イ) で定義した ポイント番号の組 み合わせ		検証内容
1	3	1 のプロトコル情報で必要なものが正しく ODM ファイルに記述されていること
1	4	1 のプロトコル情報で必要なものが正しく EDC 画面で表示されていること
2	3	2 の CRF レイアウトが正しく ODM ファイルに記述されていること
2	4	2 の CRF レイアウトが正しく EDC 画面で表示されていること
3	5	転送された ODM ファイルの同一性が確認できること
4	1 3	4 の画面表示が 13 においても正しく表示され、連携データが表示されていること
5	6	メタデータからテンプレート画面が正しく生成されていること
5	7	メタデータからマッピング表に正しくデータが転記されていること
7	9	7 のメタデータと電子カルテの項目情報やその他情報と正しくマッピングされていること
7	1 1	7 から 11 へ実データがマッピングどおりに ODM 出力されていること
8	1 1	8 から 11 へ実データがマッピングどおりに ODM 出力されていること
9	1 1	9 から 11 へ実データがマッピングどおりに ODM 出力されていること
1 0	1 1	10 にて指定された抽出対象が ODM に出力されていること
1 1	1 2	転送された ODM ファイルの同一性が確認できること
1 2	1 3	転送された ODM が正しく EDC 画面で表示されること

表 2：直交表記載の検証項目リスト

5. テストの計画ならびに実施方法について（元資料には記載なし）

前章で洗い出しを行った検証ポイントと方法について、システム構築時等に行うテストへの計画化と実施方法についての詳細な検討が必要である。

このうち、テストの計画化については、どの段階でどの単位でテストを実施すべきであるかなどについて、検討が必要であると考え。たとえば、

- A) 抽出システム（レポーターシステム）の実施施設導入時に 1 度行うテストとして
- B) EDC ベンダとの初回接続時に一度行うテストとして
- C) ある治験依頼者との接続時に一度行うテストとして
- D) 試験プロトコルごとに行うテストとして

上記に挙げられるようなテストとタイミングについて、前章記載の検証方法がどこで行われるべきか検討しておくべきと考える。

さらに、テストを誰が実施すべきかについても明確化が必要であり、施設あるいは依頼者、さらには構築ベンダのいずれであるべきかについて、その妥当性を含めた検証が必要であると考え。

テストの仕様書に関する検討事項については、詳細の検討が必要なこと、全体としての「トレーサビリティ」を確認可能にする方法等、さらに詳細な検討が必要と考える。

## 6. 今後検討が必要な課題と参考情報

本書の第 5 章までに記載した事項について、検討課題化したものや「範囲外」としたものについては下記のとおりであり、今後更なる検討の継続が必要と考える。

- 第 2 章で検討範囲外とした①～⑤
- 第 5 章で詳細な検討が必要とした項目
  - テストをどこで実施するか？のタイミング
  - テストを誰が実施すべきか？  
(施設あるいは依頼者、さらには構築ベンダのいずれか？)
  - テストをどこまで実施すべきか？※第 5 章には未記載で新たに追加
  - テスト仕様書や記録に関する詳細の検討
  - 全体としての「トレーサビリティ」を確認可能にする方法

また、下記 (ア) ～ (エ) について、本書に関連のある外部動向について、最新状況を記載する。

### (ア) 松村班の関連検討事項について

松村班において治験レベルでの電子カルテから EDC へのデータ抽出連携のためには、バリデーション (CSV) 対応が必要という共通認識である。

松村班のモデルは、データ転送フォーマットに米国 Part11 (電子記録・電子署名に関する規定) 対応を考慮し、データの真正性担保を可能とする CDISC-ODM を利用することを前提としている。

ODM の利用方法については、データ送信単位や利用目的等、様々な記述が可能となっており、Part11 対応のスキーマタグを利用するかどうか、その詳細については「ODM ベストプラクティス」ドキュメントに記載を行っている。

また、松村班においては上記以外に、連携されるデータの真正性を担保するうえでユーザーを一意に特定するための方法について議論がなされている。CRF レポーターを使用して EDC にデータを送信する場合は EDC の利用者 ID でログインしデータを送ることで、EDC 側が送信者を特定できれば問題ないとの見解を打ち出している。

したがって、EDC に対しては電子カルテシステム内における部門システムを含むデータ抽出経路や履歴、カルテ側の利用者 ID 等を送信する必要はなく、電子カルテシステム内でトレーサビリティが担保されていれば良いとされる。

さらに ODM 上の記載場所にデータ作成者 (CRF 作成者またはオリジナルのデータ作成者) や施設情報、署名情報 (データ作成や変更理由等) をどのように記載すべきかについては、松村班の継続検討課題となっている。

以上のとおり、電子カルテシステムから抽出され CRF レポーターを通じて EDC に連携されるデータの

真正性を担保するための仕組みの利用を前提とし、さらに CRF レポーターそのものの機能要件に対する検証方法を含め、本書の目的であるシステムのチェック方法を確立していく必要がある。

#### (イ) CDISC の eSDI について

FDA が CDISC に依頼しソースデータの連携を検討した「eSDI」（2006年）にソースデータの収集（＝今回テーマのデータ抽出連携）に関する12の要件が記載されている（表2）。

要件 No.	要件の内容
1.	ソースデータの収集に利用される道具は治験実施計画書に規定されたデータの収集を保証するものであること
2.	ソースデータは正確で、判読可能性、同時発生性、オリジナル性、帰属性、完全性及び一貫性があること
3.	監査証跡は、ソースドキュメントの一部として、すべてのソースデータのオリジナルの生成物とその後の変更内容を維持すること
4.	保存されたソースドキュメントは即時検索可能であること
5.	治験責任医師はオリジナルのソースドキュメントや認証コピーを保持すること
6.	スポンサーはソースドキュメントの排他的制御をしないこと
7.	ソースデータは治験責任医師の了解や承認のある時のみ変更されること
8.	ソースドキュメントとソースデータは破壊から保護されていること
9.	ソースドキュメントは正確な写しが作成できるようになっていること
10.	ソースドキュメントは不正アクセスから保護されていること
11.	ソースドキュメントと関連ソースデータの所在地は、収集プロセスの中で全てのポイントで明確に特定できること
12.	ソースデータをコピーする際には、使用されたプロセスが全てのオリジナルのデータとメタデータを保持した正確なコピーができるものであることを保証すること

7. 表2：eSDIの12の要件

これらの要件については、CDISCにおける検討時にICH-GCP等の関連する規制要件への対応等、幅広く検討された経緯があり、今回のスコープであるデータ抽出の妥当性を検証するための検証ポイントとして取り込みを検討すべきであると考ええる。

ただし、4章で記載した検証ポイントの洗い出しと検証項目リストの検討にはeSDIの要件までは反映されていないので今後の課題として取り扱いたい。

#### (ウ) FDA の eSource ガイダンスについて

eSDI後の2013年9月にFDAが“Guidance for Industry Electronic Source Data in Clinical Investigations”を発行しており、その中でも臨床試験の原資料・原データからCRFへのデータ連携に必要な要件が記載されているので、当検討でも参考にすべきと考える。

当ガイダンスにおいては、概ね下記図3に相当するSourceDocumentsやSourceDataの定義と電子カルテシステムからeCRFへデータ連携するフローが記載されており、本書第5章までで検討したチェックポイントの見直しを含めた検討が必要と考える。

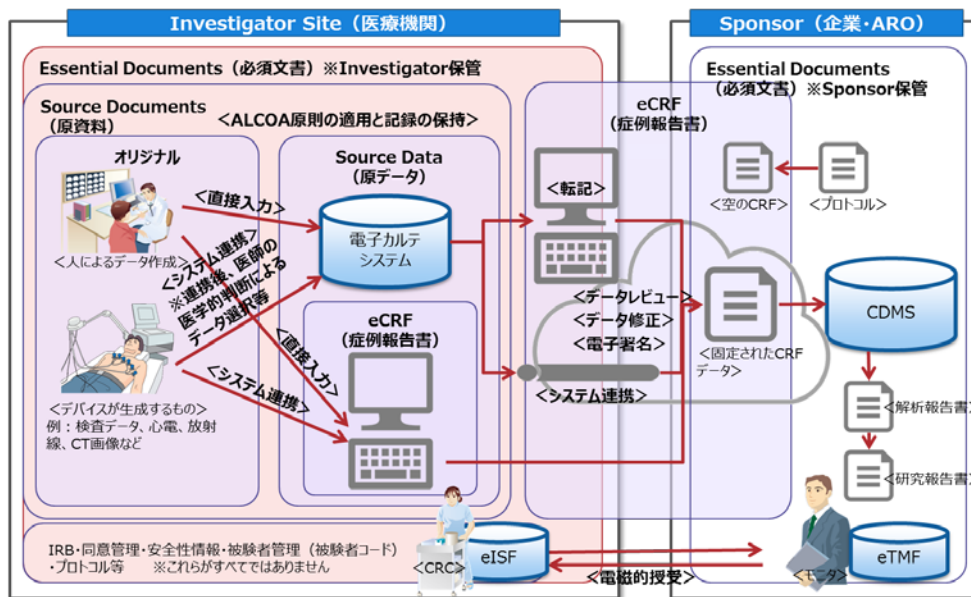


図3：eSourceの全体像

(工) FDAのFederal Register Noticeについて

FDAが2015年8月にFederal Register Noticeとして、Source Data Capture from EHRs: Using Standardized Clinical Dataと題して本書の検討範囲に関連する質問を発出している。

url:<https://www.federalregister.gov/articles/2015/06/26/2015-15644/source-data-capture-from-electronic-health-records-using-standardized-clinical-research-data>

下記は、そのうちの関連する質問の抜粋である。

- What are the gaps between the data collected in a healthcare setting by EHRs vs. clinical research data required for regulated drug development? (電子カルテの記録とCRFデータのギャップ)
- Are there any perceived regulatory obstacles to the implementation of a standards-based technology solution focusing on EHR and EDC integration (Examples include: Source data verification, remote monitoring, 21CFR Part11, patient privacy, access control and confidentiality safeguards)? If yes, what approach(es) would you recommend to overcome these obstacles? (規制上の障害について)

本書の範囲としているデータ抽出に掛かるシステムのチェック方法の実現がこれらの質問に対する回答となりうる。

米国内を含めた複数のベンダ、大学等が質問に対する回答もしくはアクションを提案しているので、これらの情報入手を含め、参考にすることができれば有用であると考えます。

(参考になる回答が含まれているかどうかは未確認である)

(オ) ICH-GCP 追補版 (E6R2) について



ICH GCP (E6) については R1 の追補版として R2 が 2016 年正式版になる予定である。

この中には、医療機関側の記録（原資料・原データ）の電子化を前提とした記述が複数追加されており、本書の検討範囲であるデータ抽出に掛かるチェック方法の検討に対して考慮が必要と考えられる。

一例を挙げると、下記の 4.9.0 章末尾にある、原データの変更に対するトレースの確保やオリジナルの記録の保持に関する事項の対応については、松村班の CRF レポートにおける実装が必要であると考えられ、実装に当たっては同時に本書記載範囲のチェック方法を明確にしておく必要があると思われる。

- 4.9.0 The investigator should maintain adequate and accurate source documents and trial records that include all pertinent observations on each of the site's trial subjects. Source data should be attributable, legible, contemporaneous, original, accurate, and complete. Changes to source data should be traceable, should not obscure the original entry and should be explained if necessary (e.g., via an audit trail).

以上