

JAPAN CRO ASSOCIATION

ワークショップ：
オープンソースソフトウェア「R」の承認申請業務での利用実態と今後の展望

CROの「R」に関する取組みと受託事例

室永 遼太郎、片倉 大昌
2025年4月2日（水） 10：00～12：30

製薬協 DS部会 KT11から頂いた課題

「RについてのCRO側での現状の対応と今後の展望について」

- ・日本のCROのRへの取り組みの現状の紹介（グローバルCROの取り組み含む）
- ・日本におけるRの受注事例の経験談

日本CRO協会 DS-WGの取組み

DS-WG所属会社から本課題に対する対応メンバー（R検討チーム）を招集し、検討チームメンバーの所属会社での「R」に関する情報をメンバー個々人が収集した。

本日の発表

日本CRO協会/DS-WG/R検討チームメンバーで収集した情報について本日発表致します。

日本CRO協会 DS-WG R検討チームメンバー



会社名	氏名
ClinChoice株式会社	乙黒 俊也
	矢島 勉
シミック株式会社	室永 遼太郎
	隈本 滉一朗
	中西 豊支
イーピーエス株式会社	大場 紀彰
	片倉 大昌
	吉良 織江
	中易 (支援)
エイツーヘルスケア株式会社	平野 真太郎
	白石, 湯川 (支援)
IQVIA JAPAN	設楽 俊也
Fortrea Japan株式会社	Shuang Li
ICONクリニカルリサーチ合同会社	岡 昌子
パレクセル・インターナショナル株式会社	田野 貴洋
日本CRO協会 アドバイザー	渡辺 敏彦
日本CRO協会 事務局長	山田 章二

日本 CRO協会 データサイエンスワーキンググループ (DS-WG)

<活動目的>

臨床試験および臨床研究に関する試験データの科学的な評価および品質保証に関するものをテーマとして情報収集を行う。

<参加メンバー>

13社, 38名 (CRO, ベンター)

<検討テーマ>

- ・データサイエンティスト育成
- ・Estimand
- ・DCT
- ・CDISC
- ・RWD
- ・DDF
- ・オープンソース

健やかな未来を
共に実現するパートナーとして

- 1) CROのRに対する取り組みの紹介
- 2) 日本におけるRの受注事例の経験談
- 3) まとめ



JAPAN CRO ASSOCIATION

1) CROのRに対する取り組みの紹介

1. 解析の幅が広がる

- オープンソースのため開発が早く、Rの方がSASより様々な解析手法に対応できることがある。
また、機械学習等への応用を踏まえると、Rの方が有利な場合がある
- Rmarkdownを活用すると、コードや図表と文章を一本化したドキュメントが作成しやすく、
仮にCSRの文章まで含めて出力可能になれば、SAS以上の再現可能性が担保できるのではないか。
- R shiny使用してというWebアプリケーションを作成することで、
動的にリアルタイムで分析結果を提供できることが期待される。
- バリデーションや妥当性を考えると、Rの方が取捨選択は難しい。

2. コストの観点

- SASユーザーのクライアント様にも対応するため全面的な置き換えは難しく、SASで可能な解析をより低コスト化する見通しは立て難い。
しかし、インタラクティブツールや機械学習的手法の応用を考えれば、SASより低コストで提供が可能である。

3. 人材について

- 近年、SASよりもRの経験がある学生が多いため、若手層に対しては柔軟な対応が可能である。
- 中堅層・ベテラン層のR未経験者に対する教育などに関しては課題がある。

4. 規制当局関連

- FDAはSASのみに限定してはいないが、R consortiumが立ち上がったことでRを用いた申請が活発化した。

「R」に関する取組み

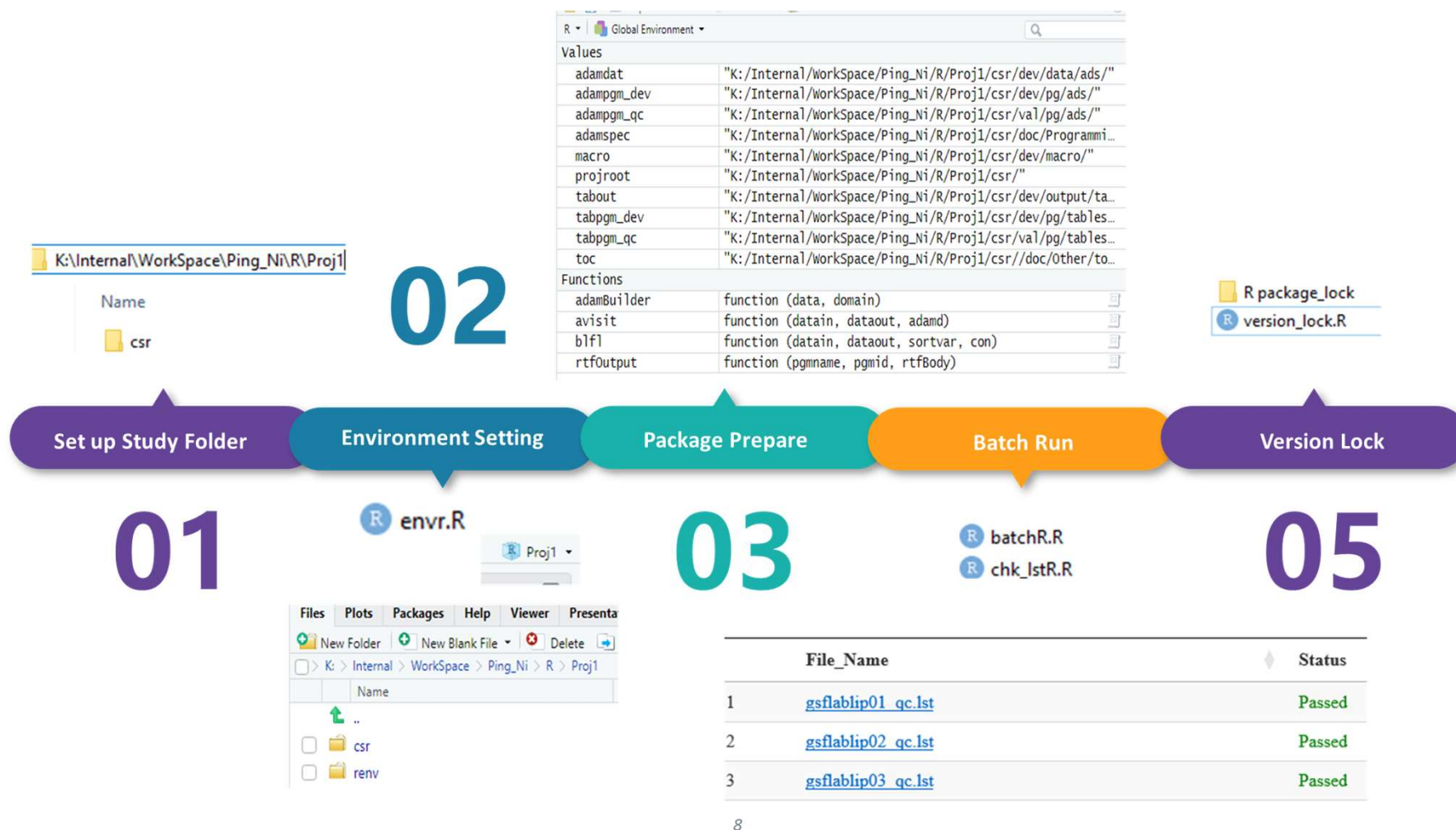
1. システム・プロセスに関する取組み

- R Serverを使用可能
- Figure Designer を使用可能
- RをベースとしたGUI型のツール(R Studioなど)の使用

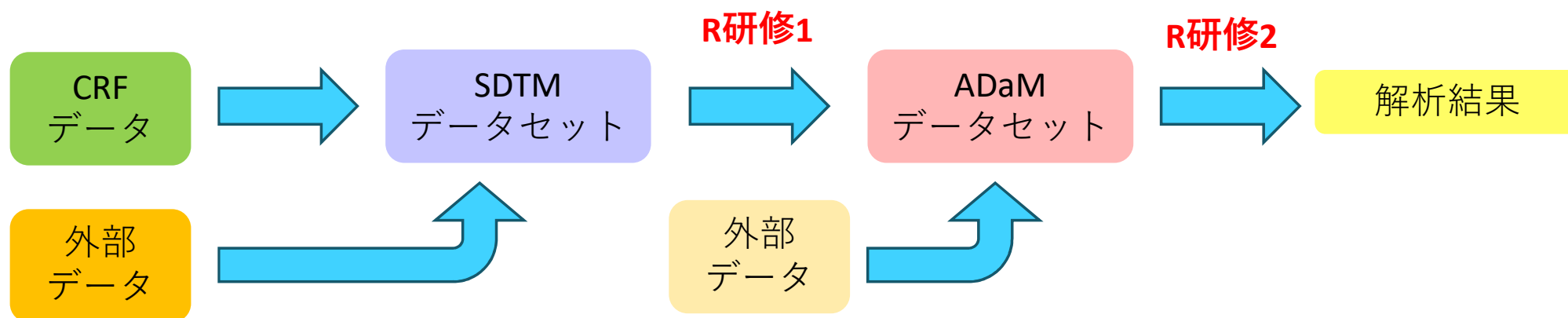
2. 人材育成に関する取組み

- Rの学習
- 一般的な解析プロセスにRを利用

システム・プロセス（事例）



B社では以下の治験業務を想定し、R研修(社内研修)を構築中です。



Rを使用し、

R研修1：SDTMデータセットからADaMデータセットの作成

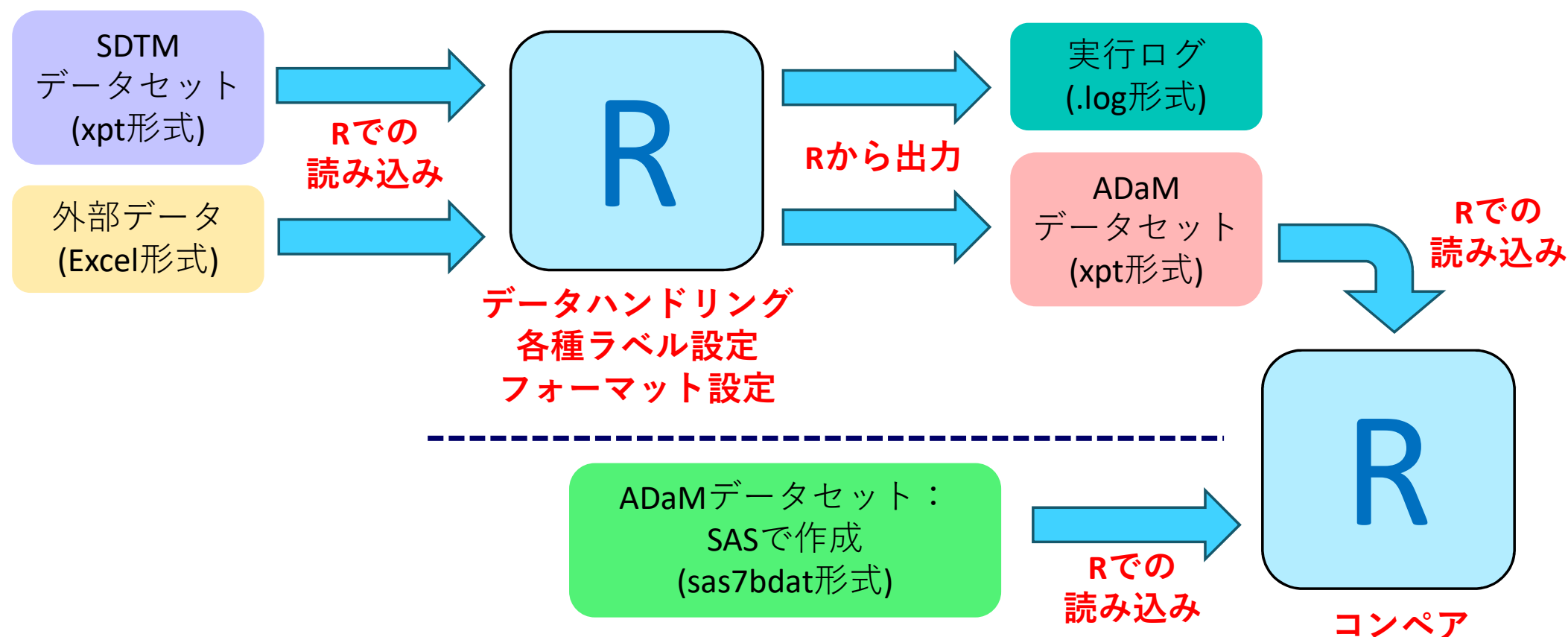
R研修2：ADaMデータセットから解析結果の作成

を実施できるよう、研修資料の準備を進めております。

まだ課題は残っておりますが、R研修の概要を紹介します。

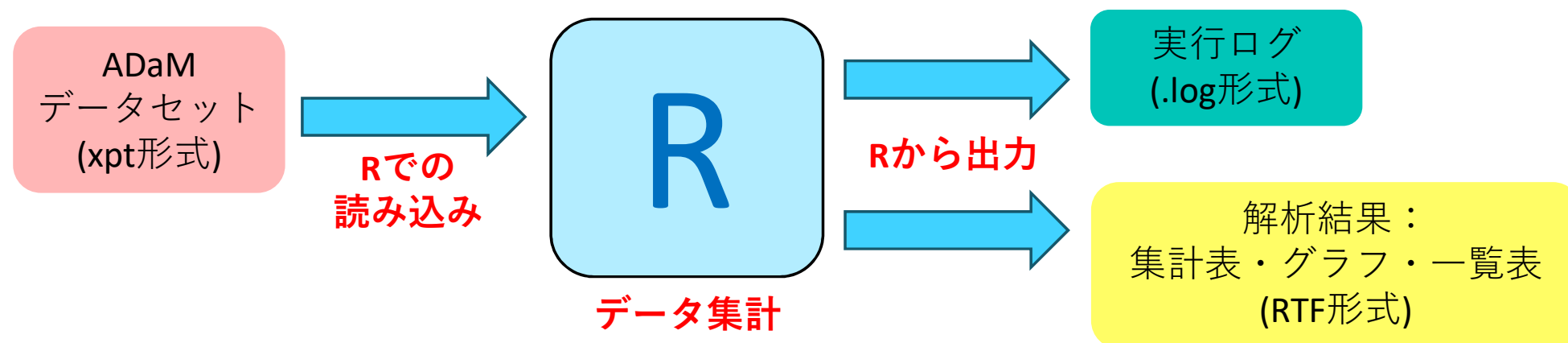
<R研修1: ADaMデータセット トレーニング>

ADaMデータセットを作成する研修では、以下の実施項目(赤字)を想定した準備を進めております。



<R研修2: 解析結果 トレーニング>

解析結果を作成する研修では、以下の実施項目(赤字)を想定した準備を進めております。



※データ集計にはパッケージを使用するため、
ADaMデータセット トレーニングのような、
SASデータセットとのコンペアは実施していない

- ◆ **タスクチームレベル**でRによる臨床試験データ解析方法を調査し、トレーニング資料を作成

<内容に関する検討>

- ・ SASによる導入研修の内容を、R用にアレンジ
- ・ 現時点ではternパッケージを用いる場合に限定
- ・ “TLGカタログ”の参照のみではなく、SASとの差異を認識できるよう工夫

<運用に関する検討>

- ・ SASによるデータハンドリングトレーニングも必須とし、SASを習得したうえでのRリソース展開を検討
- ↑ Rには暗黙の処理が多いことを考慮

- ◆ **部レベル**へのトレーニング展開

- ・ 今後のニーズに応じ、既に経験のあるもののみ（②まで）or 未経験リソース含め展開が必要か検討

タスク取り組み前のRに関する社内調査 (主にプログラミング・統計チーム対象)

①触れたことがない: 約50%

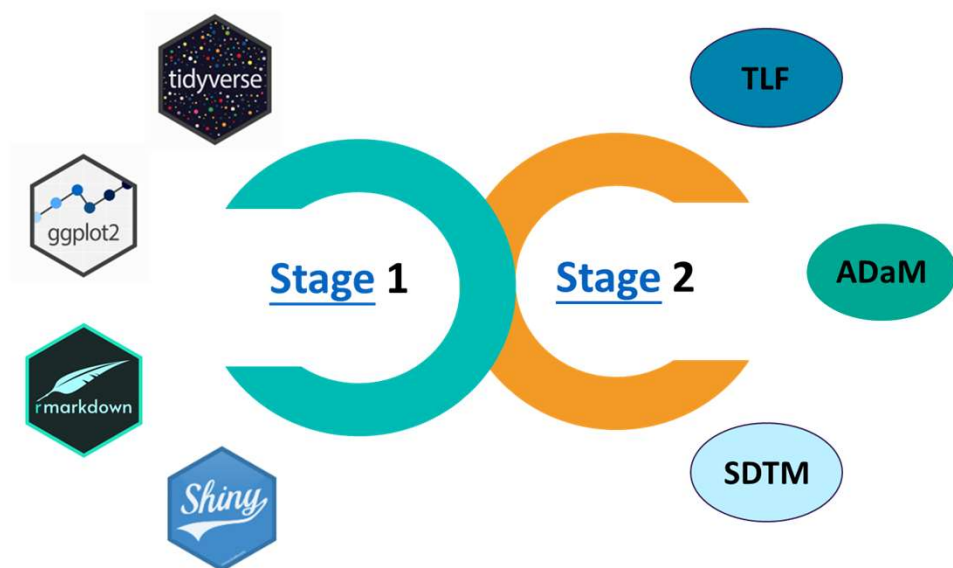
③ハンドリング
可能: 約15%

②定型プログラム
実行程度:
約30%

④習熟: 約5%

主にこちらに
該当する人員の
一部でタスク活動

Shinyアプリや
パッケージ自体
の作成可能な
リソースの確保



R Training Sources - Stage 1

Slides Recordings References

- [Overview of R & R Markdown](#)
- [Basics of R programming](#)
- [Data manipulation via tidyverse \(focus on readr, dplyr\)](#)
- [Data manipulation via tidyverse \(focus on tidyr, stringr\)](#)
- [Writing functions in R](#)
- [Basic R graphics](#)
- [Advanced R graphics](#)
- [Statistical Analysis in R](#)
- [Survival Analysis in R](#)
- [R packages](#)
- [R shiny](#)

R Training Sources - Stage 2

TLF ADaM SDTM

- [How to create demographics table](#)
- [How to create disposition table](#)
- [How to create LB summary tables](#)
- [How to create AE summary and other AE tables](#)
- [How to create efficacy table](#)
- [How to create figure, such as KM plot](#)

JAPAN CRO ASSOCIATION

2) 日本におけるRの受注事例の経験談

「R」の利用理由

①クライアントの要望

- ・海外臨床試験で用いられた「R」を再利用したい（海外）
- ・論文に記載されている「R」のコードを利用したい（論文）
- ・オープンソース（「R」）の可能性を試したい（試用）
- ・クライアントの標準（標準）

②最適なツールとして（※SASが得意でない・未実装の解析内容）

- ・複雑な解析（解析）
- ・複雑な図表（図表）

「R」の利用経験

No	分類	「R」の利用内容
1	図表	サンキーダイアグラム、ヒートマップの作成
2	解析	時系列クラスタリング(Ward法)によるカテゴリー分類
3	論文	Lasso回帰
4	解析	ベイズ流解析による指数変換や95%信用区間の算出
5	解析	MCP-Modによる例数設計
6	標準	医療データベースを使った医薬品別、疾患別集計
7	解析	BOIN(Bayesian Optimal INterval designs)デザインのシミュレーション
8	解析	Lasso回帰
9	試用	CDISC対応を含めた一般的な解析プロセスをRで実施（海外）
10	図表	ggplotを中心としたグラフ作成
11	解析	機械学習的手法の実施
12	解析	PKシミュレーション解析

「R」の利用場面

- ◆臨床試験におけるCSR解析全体の解析業務（臨床薬理分野除く）⇒ 受託経験なし
- ◆臨床研究，論文，学会用資料，クライアント社内用資料 ⇒ 外資系製薬会社からの依頼が多い

分類	内容
統計解析	・ EX-NEXモデルの解析 ・ クラスター分析 ・ ベイズ流解析 ・ Lasso回帰 ・ ロジスティック回帰分析 ・ 逐次検定 ・ 時系列クラスタリングによるカテゴリー分類 ・ 例数設計のシミュレーションデータ生成 ・ 階層的クラスタリング ・ 決定木における特徴量選択
作図	・ サンキーダイアグラム ・ ヒートマップ ・ デンドログラム ・ バブルプロット ・ ペイシェントジャーニー
その他	・ html出力

分類	内容
パッケージ	・ Tidyverce ・ Pharmaverce ・ ggplot - など

「R」利用時のバリデーション方法

＜ダブルプログラム（Rのコンペア）＞

- ・ Rで作成した出力結果を目視確認
- ・ Rパッケージ使用結果とパッケージ使用しない結果の整合確認

＜ダブルプログラム（SAS取込コンペア）＞

- ・ 分類結果をSASに取り込み、SASコンペアー一致を確認
- ・ 作図用データをSASで作成しSASコンペアー一致を確認

＜他集計表と整合確認＞

- ・ Rで出力した結果と既に作成された集計結果の目視確認（グラフの確認や、受け入れ確認等で使用）

＜実施せず＞

- ・ シングルプログラムで作成（コンサル業務等で使用）

「R」のパッケージバリデーション

Rのシステムバリデーション・パッケージバリデーションまで実施していた案件は現状ほぼない

バリデーションを取った案件について：

システム・パッケージについて、クライアント指定のバージョンと一致しているか

※今後の標準的な対応は検討が必要

◆委託者から理解が得られれば、CROとしての作業は可能（CRO側も標準的な手順を準備する必要あり）

◆パッケージについて

- ・バージョン記録をどのように整備するか
- ・使用するところを想定してOQテスト(Operational Qualification; 稼働性能適格性確認)を検討
- ・あとは当局がどのように考えているのか…

◆Rを使うサーバのバリデーション

- ・バージョン記録、インストール時にエラーが発生しなかったことの記録
- ・他に必要なものは何か…

「「医薬品の臨床試験の実施の基準に関する省令」のガイダンスについて」の改正について
(薬生薬審発0730第3号令和3年7月30日)

第26条 記録の保存等 第1項

3 治験依頼者は、データの処理に電子データ処理システム（遠隔操作電子データシステムを含む。）を用いる場合には、次の事項を実施すること。

- (1) 電子データ処理システムが、完全性、正確性、信頼性及び意図された性能についての治験依頼者の要件を満たしていることを保証し、文書化すること（すなわちバリデーションされること。）なお、その際には、システムの用途並びにシステムが被験者保護及び治験結果の信頼性に影響を与える可能性を考慮したリスク評価に基づいて行うこと。
- (2) 当該システムを使用するための手順書を整備すること。当該手順書の対象範囲にはシステムのセットアップ、インストール及び使用方法を含むものとする。当該手順書には、システムバリデーション及び機能テスト、データの収集及び取扱い、システムの維持管理、システムの安全対策、変更管理、データのバックアップ、修復、危機管理計画並びにシステムの廃棄について記載すべきである。これらの電子データ処理システムの使用に関する治験依頼者、治験責任医師その他の当事者の責務を明確に示し、使用者に対し、システムの使用に関する教育・訓練を提供すべきである。

JAPAN CRO ASSOCIATION

まとめ

- ◆ パッケージに依存する処理(例. 表形式を整える機能)は、柔軟なアレンジは難しくなる
 - ・クライアント様によって使用するパッケージ要望が異なる場合、パッケージごとに学習すべき内容が異なる
 - ・ternパッケージを基本とする場合は、特に“TLGカタログ” が公開されており実装検討しやすい
- ◆ ダブルプログラミングを基本とするQC方法の再検討が必要
 - ・コンペアを実施できない場合のQC方法をどのようにするか
- ◆ パッケージのバージョン管理レベルの検討などが必要

本発表のまとめ

- ◆ CRO側でも、Rを用いた解析業務を受注できる体制構築を進めている
 - ・ 実行環境の整備、社内教育を進めており、今後も拡充を検討している
- ◆ 臨床研究や論文用解析などの用途では、CROでもRを用いた解析実績が蓄積しつつある
 - ・ SASでは簡単に実装できない解析手法やグラフ作成の依頼が多い
- ◆ Rシステム・解析結果のバリデーションには課題があり、Rを用いたCSR解析の実施には事前の検討が必須である

- ・ CROの役割は、単に受注した業務を既存のシステムに流し込んで遂行することではなく、クリニカルクエスチョンに基づいたデータエンジニアリングを行い、データインテグリティを担保することが役割の一つである
- ・ 近年、臨床試験のデザインが複雑化したり、RWD（リアルワールドデータ）を活用したエビデンス創出に注目があつまっており、シミュレーションを含むものやSASで解析方法を実装することが困難なケースも増えており、OSS（オープンソースソフトウェア）活用の重要性が増している
- ・ RやPythonといったOSSを用いたシミュレーション・解析業務は一部行われており、一定の成果を生んでいる
- ・ 一方で、CSV（Computer System Validation）とOSSに精通した人材確保・育成に課題もあり、特に承認申請用解析業務に耐えられるレベルのCSVについて検討が必要

クリニカルクエスチョンに基づいたデータエンジニアリング