IIIT & HEXAWARE CHALLENGE – 2025

Hack the Clinical research challenges- an initiative of Hexaware and IIIT - Hyderabad

Challenge Concept: Agentic AI for End-to-End Clinical Trials

Theme: Reimagining Clinical Trials with Agentic AI in the Data Management, Data Analysis and CSR preparation.

Following are the agentic AI use cases for challenge

- ### CRA.AI

  A Clinical Research Associate (CRA) plays a pivotal role in the success of clinical trials. Their responsibilities span across multiple stages of a trial and are essential for ensuring data integrity, regulatory compliance, and participant safety and sole responsible for the site administration.

- ### Data Management

  Leverage AI to streamline clinical trial data management by automating data cleaning, validation, and query management orchestration. Enhance data accuracy, reduce manual intervention in query resolution and communication and ensure compliance with regulatory standards and quicker analysis of the data by the downstream teams.

- ### Central Monitoring with Ops, Safety, and Signal Detection

  Use AI-driven central monitoring to track operational metrics, identify safety signals, and detect anomalies in real-time. Improve trial monitoring oversight, ensure participant safety, and mitigate risks proactively.

- ### CDISC Standards Conversions and TLFs automation

  Automate the conversion of clinical trial data into CDISC (Clinical Data Interchange Standards Consortium) formats such as SDTM and ADaM and TLFs generation based on the ADaM data. This ensures seamless regulatory submissions and improves efficiency in data standardization.

- ### CSR (Clinical Study Report) Automation

  Utilize AI to automate the generation of Clinical Study Reports by extracting insights from trial data (TLFs, Protocol, Lab Data…), creating narrative, and

ensuring compliance with regulatory ICH guidelines. Save time and reduce manual effort in report creation.

Note: all of them are inter dependent with an exception of point #2. We can explore if there is a possibility of combining them into two rather than 4.

Detailed explanation of each use case:

1. Functional Area: CRA (Clinical Research Associate)

- Agent CRA
- Persona: Clinical Research Associate

Background and Challenge:

Core Responsibilities of a CRA

A Clinical Research Associate (CRA) plays a pivotal role in the success of clinical trials by managing lot of activities at the clinical site as described below and have lot of manual work such as several documents compliance verification, Source data verification and liaising between various stakeholders in the clinical trial process with a goal of patient safety.

Trial Monitoring

- Conduct on-site and remote visits to ensure trials follow approved protocols and Good Clinical Practice (GCP) guidelines
- Verify that participant rights and safety are protected throughout the study1

Regulatory Compliance

- Ensure all necessary approvals (e.g., ethics committee, IRB) are obtained before trial initiation
- Confirm adherence to local and international regulations like FDA, EMA, or ICH standards1

Data Integrity

- Perform source data verification (SDV) to ensure accuracy between case report forms (CRFs) and original documents
- Identify and resolve discrepancies in clinical data1

Site Management

- Train site staff on protocol requirements and study procedures
- Support recruitment and retention of study participants1

- **Documentation & Reporting**
  - Maintain essential trial documents including consent forms, drug dispensation records, and regulatory forms
  - Prepare monitoring visit reports and final study reports2
- **Study Start-Up Activities**
  - Conduct Site Selection Visits (SSVs) to assess feasibility of potential sites
  - Evaluate staffing, equipment, and facility readiness for trial execution3
- **Communication & Coordination**
  - Act as a liaison between sponsors, CROs, and clinical sites
  - Facilitate clear communication to align all stakeholders
  - Address and follow up on site-based queries

Challenge Goal:

Build an intelligent agent assistant that can take care of the following responsibilities of the CRA:

- Visit schedule monitoring and reminders, capture of the oucome and follow up actions management
- Ability to do a remote monitoring visit
- Source Data verification (SDV) by an agent and create tasks out of it
- Respond to minor and moderate queries received by CRA
- Alert the CRAs on various outliers of the data from specific site
- Auto upload of the documents to eTMF and compliance monitoring
- Automated Protocol compliance monitoring and create an action plan for the same.

Expected outcome of ideal solution in Production:

- Reduce physical site visits by 50%
- Automation of SDV up to 80%
- Automated Email and task monitoring up to 70%
- Minimize the trial execution cost
- Faster therapies to the market.

2. Functional Area: Clinical Data management

- Agent Data Manager
- Persona: Clinical Data Manager

Background & Challenge

Data Manager:

In clinical trials, a data manager ensures the study design, External data specifications, data integrity and accuracy by overseeing data collection, processing, and analysis. They design data management plans, develop entry guidelines, and maintain quality standards, collaborating with stakeholders to ensure regulatory compliance. Data managers also analyse data to identify trends and make recommendations to the trial core team to make amendments.

Here's a more detailed responsibility:

Data Management role:

- Understand the clinical protocol and develop the following docs
    - EDC Study design specifications
    - External data specifications for Labs, Imaging, IRT, ECG, ePRO, eCOA systems
    - Develop Data review plan and report specifications
    - Data blinding/masking management
    - Review the data of individual vendor (Labs, Imaging, IRT,…) as per the spec and raise queries and resolve appropriately
    - Reconcile all the data with the EDC data and raise quires appropriately and close them
    - Annotate the eCRF for the SDTM domain mappings
    - Track the milestones of the data review.

Challenge Goal:

Build an intelligent agent/DM assistant that can take care of the following responsibilities of the DM:

- Author the study spec based on the protocol document for EDC and vendors
- Monitors incoming EDC and external data streams
- Automatically flags inconsistent/missing data in individual EDC and vendor data sets
- Perform reconciliation across the data
- Suggests or drafts queries in natural language in the context of the identified issues
- Review the queries and provide an update or close the query based on the response.
- Learns from query history to evolve smarter edit check rules

Expected outcome and impact:

- Reduces manual query volume by up to 75%, accelerates data cleaning, and improves data quality, consistency leading to faster interim analysis and database lock.
- Minimize the trial execution cost
- Faster therapies to the market.

Important links for the reference and learning:

https://clinicaltrials.gov/

https://pubmed.ncbi.nlm.nih.gov/

Data Management for Clinical Research

https://greatonlinetraining.com/course/clinical-data-management-training/

3. Functional Area: Clinical Trial Monitoring (Operational, Safety.)

- Agent Monitor
- Persona: CRA, Medical Monitoring, Primary Investigator

Background & Challenge:

Clinical Trail monitoring is very important in during the trial execution across geographies in terms of patient enrolment, Screening, Randomization, Patient visit compliance, Site performance, Protocol deviations, Patient safety monitoring to ensure the clinical trial successfully runs. Today's situation there is lot of manual activity to

monitor the trail, and data is not real time most of the times. This contributes to delays in identifying risks like underperforming sites, protocol deviations, or missed milestones can derail timelines.

Challenge Goal:

- Develop an Agentic AI system that:
- Ingests real-time operational data (site logs, enrolment, deviation reports)
- Identifies risk patterns based on the data such as Site is low in-patient recruitment, data fraudulent ..
- Identify safety issue based on the data and assign the tasks to Monitoring team
- Flags potential dropouts, adverse event risks, or protocol deviations
- Predict early assessment of the drug safety and efficacy
- Sends daily updates and alerts to study team members

Expected outcome and impact:

Enables proactive trial management, reduces protocol violations, and improves team decision-making with intelligent risk recommendations, early identification safety issue and mitigation plan.

Useful links:

https://pmc.ncbi.nlm.nih.gov/articles/PMC9526458/

https://careers.iconplc.com/blogs/2023-11/clinical-monitoring-understanding-the-key-responsibilities

https://www.fda.gov/media/116754/download

4. Functional Area: Conversion of raw data into CDISC Data standards and data Analysis

- Agent CDISC Automation
- Persona: Standards Lead / Stats Programmer/Statistician

Background & Challenge:

In clinical trials, converting raw data to CDISC-compliant datasets (SDTM and ADaM) and generating Tables, Listings, and Figures (TLFs) for submission is a resource-intensive and extensive double programming effort in SAS. It requires deep domain knowledge, CDISC standards, specifications Authoring, and collaboration across

multiple functions. This manual workflow often leads to delays, inconsistencies, and quality issues.

Challenge Goal:

Build an end-to-end CDISC Automation Agent that:

- Learns from historical SDTM and ADaM specifications, annotations, and mapping logic
- Automatically maps raw EDC and external data to SDTM domains
- Derives ADaM datasets using metadata, SAP, and TLF mock shells
- Generates draft TLFs based on ADaM outputs and visualizes results based on SAP
- Validates datasets against CDISC and regulatory standards using P21
- Supports human-in-the-loop refinement of specs and outputs
- Generate the documentation such as Define.XML, SDRG, ADRG.

Expected outcome and impact:

Accelerates the data pipeline, reduces programming burden, improves submission readiness, and enhances consistency across SDTM → ADaM → TLF workflows.

Useful Links:

https://www.transceleratebiopharmainc.com/initiatives/clinical-data-standards/

https://www.cdisc.org/standards

https://www.allucent.com/resources/blog/what-cdisc-and-what-are-cdisc-data-standards

https://www.fda.gov/industry/fda-data-standards-advisory-board/study-data-standards-resources

5. CSR Automation Agent

- Agent CDISC Automation
- Persona: Medical Writer / Clinical Scientist

Background & Challenge:

Writing Clinical Study Reports (CSRs) involves collating data from TLFs, interpretation of the protocol, ICH guidelines and writing multiple narrative sections and

summaries—a process that can take several weeks per study yet some challenges in the interpretation and things may go wrong because of human oversight.

Challenge Goal: Create a generative AI agent that does below activities,

- Ingests statistical outputs, key efficacy and safety findings
- Summarizes results into structured CSR sections
- Suggests draft text with inline references to tables
- Supports dynamic editing and version tracking
- Maintain the documents digitally and enable workflow
- Enable audit trial and e Signatures and compliant with the part 11

Expected outcome and impact:

Enable the complete digital document of the CSR and workflow process, minimizes the CSR drafting time by 40–60%, ensures consistency in interpretation, and allows writers to focus on scientific insights rather than repetitive reporting. Ease of regulatory sharing the documents. Effective communication among the stakeholders.

Useful links:

https://www.udemy.com/course/an-introduction-to-medical-writing/

https://www.youtube.com/watch?v=I2HZs01N8F0

https://cdn.clinicaltrials.gov/large-docs/60/NCT03150160/Prot_000.pdf

https://www.sanofi.com/assets/dotcom/content-app/clinical-studies/pharma/aflibc06561_summary.pdf