*Category: Clinical Trial Operations (Trial Start-up, Regulatory, Data Management, IITs) – Completed Project*

**Taking Data Validation to the Next Level: Automating Data Validation Using CDASH-Standardized Global eCRFs - Phase II**

S. Rachuri, K. Douglas, L. Barry, J. Tewell, A. Patwardhan, R. Johnson, T. Carroll, R. Gattas, K. Kennedy, S. Balu, K. Morrison, E. Crecelius

*UNC Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill*

**1. Background**

Ensuring the accuracy and reliability of clinical trial data is crucial for the integrity of research outcomes. Manual data validation processes are time-consuming and susceptible to human error, which can compromise data quality. The Lineberger Comprehensive Cancer Center (LCCC) recognized the need for an automated data validation system to address these challenges. This system aims to enhance efficiency, reduce the risk of errors, and improve the overall quality of clinical trial data. By automating the validation process, LCCC can detect and resolve data issues more quickly and accurately, leading to cleaner and more reliable data. This initiative is essential for streamlining the validation workflow and ensuring robust data management practices in clinical research.

**2. Goals**

- Streamline the data validation process by reducing manual efforts, improving accuracy, and increasing efficiency.
- Decrease validation time by quickly identifying and resolving more errors.
- Enhance real-time data monitoring and traceability with automated reports and detailed logs.

**3. Solutions and Methods**

The data management team submits a Data Validation Plan (DVP) in Excel for programming. This spreadsheet contains a data dictionary with variables and metadata for each protocol. The data manager then provides programming logic for automated edit checks. The system generates reports highlighting discrepancies, enabling faster issue resolution. To ensure consistency across datasets, cross-form data validation detects and addresses discrepancies automatically. The validation report is integrated into Statistical Analysis Software (SAS) reports, which can be run through a web-based reporting interface for real-time monitoring. The Clinical Data Acquisition Standards Harmonization (CDASH) eCRF Global Library standardizes SAS programs, making them reusable across multiple studies. To measure success, we compared manual and automated data review times for similar trial subjects. A survey of data managers was conducted to gather user feedback on accuracy and efficiency improvements.

**4. Outcomes**

Metrics collected from Clinical Data Management Associates (CDMAs) highlight a positive impact of automation on the validation process. The data indicates that automation has reduced validation time by 50 percent on average, representing a substantial efficiency gain. Additionally, feedback from CDMAs reflects a considerable improvement in accuracy, with the majority reporting a "significant" improvement, while others acknowledge a "moderate" increase. Issue detection has also seen notable advancements, ranging from moderate to significant improvements. This suggests that automation accelerates the process and enhances the overall validation quality. Furthermore, the reconciliation process has been well-received, with most respondents rating it as "much" more helpful than "enough," reinforcing its effectiveness in streamlining workflows. These findings underscore the substantial

benefits of automation, which lead to increased efficiency, improved accuracy, and a more effective reconciliation process—all of which contribute to a more seamless and reliable validation workflow.

**5. Lessons Learned and Future Directions**

Successful adoption required training to ensure a smooth transition to automated workflows. Continuous performance monitoring was essential for optimization. While automation significantly improved efficiency and accuracy, certain checks still require manual validation due to specific protocol requirements. Overall, automation has strengthened clinical data validation at LCCC by reducing errors, streamlining workflows, and improving data quality. Future efforts will focus on refining automated processes, expanding functionality, and integrating additional validation checks to further enhance efficiency.

**Figure**