Category: Clinical Trial Operations (Trial Start-up, Regulatory, Finance, Data Management, IITs) – Work in progress

Optimizing Data Collection in Investigator-Initiated Trials: Assessing Alignment Between Collected Data and Study Endpoints at UF Health Cancer Institute

N. McGrew, L. Napier, T. George, P. Crispen, A. Ivey, E. Monari, A. Anderson, K. Barker, C. Cline, A. Kukulka, J. Walsh, S. Alford

UF Health Cancer Institute

1. Background

Inconsistencies between collected data and study endpoints drive inefficiencies in Investigator Initiated Trials (IITs), increasing workload, extending analysis timelines and reducing data utility. A 2024 Tufts Center for the Study of Drug Development (Tufts CSDD[1]) analysis with 15 TransCelerate BioPharma companies, reported that up to 40 percent of data collected in academic trials are not linked to primary or secondary endpoints. Additional evidence shows that 25 percent of phase III and 18 percent of phase II trial data support supplementary or exploratory endpoints, rather than endpoints for safety. To support data-driven, endpoint-focused trial design, we evaluated endpoint alignment across selected UF Health Cancer Institute (UFHCC) IITs and developed a standardized assessment tool.

2. Goals

- Assess proportion of data elements aligned to study endpoints
- Develop a standardized Data Alignment Assessment Tool (DAAT) for use during protocol development
- Pilot DAAT to identify opportunities to streamline data collection and enhance endpoint relevance

3. Solutions and Methods

Informed by a UFHCI Compliance Office presentation highlighting endpoint misalignment as a contributor to trial inefficiencies, a retrospective review of four UFHCI IITs approved between 2021 and 2025 was conducted. Two protocols were evaluated prior to (Trials 1 and 2), and two after (Trials 3 and 4), pilot implementation of the DAAT tool. For each protocol, all planned data elements were extracted from the EDC and protocol and categorized using a standardized framework as Directly Tied, Supporting, Exploratory, or Non-Essential, based on relationship to primary and secondary endpoints. Dual independent review ensured classification consistency with discrepancies resolved through consensus. Aggregate analyses quantified the proportion of endpoint-aligned verses non-aligned data elements. Findings informed development of a data mapping process, subsequently refined into a standardized worksheet designed to guide endpoint-focused data planning during protocol development.

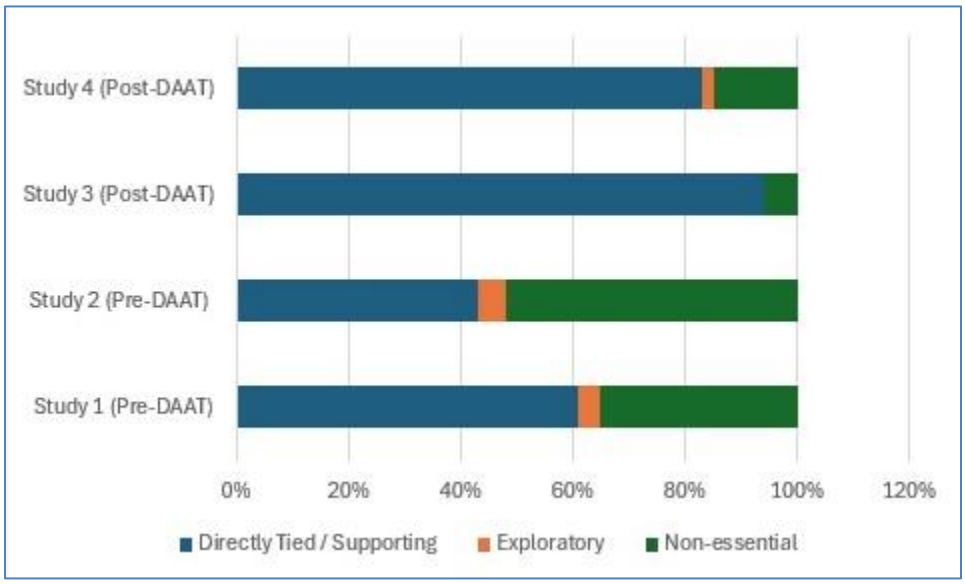
4. Outcomes

A 33-percentage point reduction in non-essential data points was observed in post-DAAT implemented studies (average of 43.5 percent non-essential data points in pre-DAAT trials, compared to 10.5 percent in post-DAAT trials). Commensurately, directly tied or supporting data point percentage improved by 36.5 percentage points (average 52 percent directly tied or supporting data points for pre-DAAT trials, and average 88.5 percent in post-DAAT trials.)

---

[1] *Getz K, Botto E, Arques AC, Galuchie L, Sanmiguel NC, Sheetz N, Smith Z. Insights Informing Strategies for Optimizing the Collection of Clinical Trial Data. Ther Innov Regul Sci. 2026 Mar;60(2):563-574. doi: 10.1007/s43441-025-00899-4. Epub 2025 Dec 29. PMID: 41462003.*

Category: Clinical Trial Operations (Trial Start-up, Regulatory, Finance, Data Management, IITs) – Work in progress



Pilot DAAT implementation demonstrated that it is feasible to identify non-essential data elements early in protocol development, resulting in measurable reductions in non-essential data collection while maintaining endpoint integrity and regulatory compliance.

5. Lessons Learned and Future Directions

Many IITs collect exploratory and non-essential data that increase workload without clear benefit. Findings highlight importance of early, endpoint-based data mapping to promote clarity, efficiency, and regulatory alignment. The DAAT is a practical, scalable approach for improving data relevance and supporting endpoint-focused trial design, with future efforts focusing on institutional adoption to further assess feasibility, impact on data volume and effectiveness, with and overarching goals of improving data quality, reduction of staff workload, and enhanced feasibility and efficiency of IITs.