

Embedding AI-Assisted Review of Unstructured Clinical Data Into Oncology Trial Operations

N. Bouvier, L. Yuravlivker, R. Singh, B. Young Kim, B. Liebe, M. Buckley, R. Panchal

Memorial Sloan Kettering Cancer Center

1. Background

At Memorial Sloan Kettering Cancer Center (MSK), performance status (PS) is a required data element in oncology clinical trials and is commonly used for eligibility, treatment assessment, and safety reporting. In practice, PS is often documented only in unstructured clinical notes, requiring data managers (DMs) to manually search and review clinical documentation before entering data into sponsor electronic data capture (EDC) systems. This process is time-consuming, error-prone, and contributes to operational burden. Although artificial intelligence (AI) has potential to assist with unstructured data, many solutions fail to achieve sustained use due to poor workflow fit, limited transparency, and concerns about automation replacing DM judgment.

2. Goals

The goal of this work was to develop and operationalize an AI-assisted feature that augments the review of unstructured clinical data within existing workflows, while preserving source transparency, regulatory accountability, and DM decision ownership.

3. Solutions and Methods

We implemented an AI-assisted review feature as a native capability within MSK's clinical trials data hub (CTDataHub), a web-based application for DMs that surfaces ready-to-enter clinical trial data from source systems. PS data was selected as the initial use case; prior to this capability, CTDataHub primarily supported structured data and retrieving PS required DMs to manually search unstructured notes in the EHR.

The feature extends CTDataHub by surfacing unstructured clinical information that requires review, specifically, AI-identified ECOG, KPS and Lansky PS scores documented within clinical notes. These values are presented to DMs for review and confirmation, rather than generated as free-text output. It was designed using a human-centered, human-in-the-loop approach to support review and confirmation while preserving DM decision ownership. The AI-assisted review feature was deployed into production in October 2025 within a secure, on-premise environment.

4. Outcomes

In the first four months following launch (October 2025-February 2026), the feature was used across 23 oncology service groups by 225 DMs, with 18 percent adoption (41 active DM users) and 339 confirmed PS scores. Review-without-confirmation interactions were common during early adoption, which we

hypothesize reflects DMs' exploratory use of the feature. During this initial evaluation period, no AI-identified performance status values presented for confirmation were corrected.

5. Lessons Learned and Future Directions

Early experience suggests that the AI-assisted approach is most effective when implemented as a review aid embedded within existing data management workflows rather than as an automated solution.

Review-without-confirmation behavior highlighted the importance of change management, including clear workflow guidance and onboarding, to support user confidence while preserving data manager decision ownership. Future work will expand this feature to additional unstructured data elements and continue evaluation of user interaction to assess its impact on trial operations.