Retrieval-Grounded Normalization of Clinical Trial Eligibility Criteria

Presented at the Meridian Health Informatics Summit 2026, Track B.

Eligibility criteria in trial registries are written in free prose that resists direct querying. In this work, Hana Kowalski and Robert Adeyemi describe a retrieval-grounded method that maps each criterion to a controlled schema and records the source sentence as provenance. Evaluated on 800 oncology trials, the approach recovered structured inclusion and exclusion rules with 0.89 F1 and flagged 142 ambiguous criteria for human review rather than guessing.

The authors are affiliated with the Riverside Center for Clinical Data Science and the Kadmos Institute. They emphasize that surfacing provenance let reviewers audit disagreements in seconds, which was decisive for clinical adoption.

Terms: clinical trials, eligibility criteria, retrieval grounding, provenance, F1