Why AI matters in early phase clinical trials

Early phase clinical trials generate intensive, high-resolution datasets, including laboratory results, pharmacokinetic samples, continuous physiological measurements, and task-based or exploratory biomarker data. These datasets are heterogeneous, time-sensitive, and subject to strict regulatory standards.

Ensuring data integrity requires consistent quality control, cross-source reconciliation, structured workflows, and timely identification of meaningful patterns. Artificial intelligence (AI), including machine learning and advanced automation techniques, can support these activities by improving efficiency, consistency, and scalability while maintaining human oversight and regulatory accountability. Within a validated and compliant framework AI is positioned as a support tool. It is not a replacement for expert judgment.

How AI supports data recording and quality

Risk-based data quality review and validation

Early phase trials often involve dense sampling schedules and rapid data turnaround. AI can assist by:

- Prioritizing participants, visits, or time windows for review based on anomaly detection or uncertainty scoring
- Identifying unusual patterns across laboratory values, ECGs, vital signs, or pharmacokinetic profiles
- Detecting inconsistencies across data sources (e.g., timing mismatches or protocol deviations)

These systems support a risk-based review strategy. They do not replace data management or medical oversight. Final decisions, however, remain with qualified personnel (consistent with Good Clinical Practice (GCP) and data integrity principles).

Supporting query generation

When data checks trigger discrepancies, AI systems, including large language models, can draft structured query text based on predefined templates and contextual data. Potential benefits include:

- Faster query turnaround
- Standardized language across studies
- Reduced administrative burden

All queries remain subject to human review and approval before transmission. AI supports drafting, not autonomous decision-making or query closure.

How AI supports data analysis

Pharmacokinetics and pharmacodynamics (PK/PD)

Model-informed drug development in early phase trials relies heavily on structured, repetitive workflows such as:

- Compartmental and non-compartmental analyses
- Nonlinear mixed-effects (NLME) modeling
- Bayesian adaptive dose-escalation support

AI and machine learning approaches can assist in:

- Automating data preprocessing and dataset assembly
- Supporting model selection or covariate screening
- Accelerating parameter estimation workflows
- Identifying atypical model behavior

Established pharmacometric methods remain the foundation of regulatory submissions. Experts interpret results, assess model adequacy, approve structural assumptions, and make dosing recommendations. AI may accelerate routine steps but does not replace pharmacometric expertise or regulatory accountability.

Exploratory biomarker discovery

Early phase programs increasingly incorporate exploratory biomarkers, including:

- EEG or imaging-derived signals
- Genomics, proteomics, metabolomics
- Digital or wearable-derived physiological data

Machine learning techniques such as clustering, dimensionality reduction, and classification models can analyze high-dimensional datasets to identify potential signals associated with:

- Drug response
- Toxicity
- Disease progression

These applications remain exploratory. Findings require independent validation, biological plausibility assessment, and careful control of overfitting risk. AI can generate hypotheses, but clinical interpretation and validation remain essential before any biomarker is used for decision-making.

What AI is not ready to do (yet)

Fully automated decision-making

AI does not replace human responsibility on critical decisions such as dose selection, safety assessments, stopping rules, or regulatory strategy. Sponsors, investigators, and regulators require clear (human) accountability. AI can inform and support these decisions, but final responsibility remains with qualified experts.

Autonomous data modification or query closure

Automatically approving data corrections or closing queries without human review is not appropriate in regulated clinical research. Such actions raise compliance, liability, and auditability concerns. AI-generated outputs must remain traceable, reviewable, and embedded within validated systems that preserve audit trails and data integrity.

Conclusion

Within an early phase CRO environment, AI can support:

- Risk-based data quality review
- Structured query drafting

- PK/PD workflow acceleration
- Exploratory biomarker analysis

When implemented within validated, compliant frameworks and under expert supervision, AI improves efficiency and consistency while preserving scientific rigor, data integrity, and human accountability. AI is best positioned as a structured augmentation tool, enhancing expert capability rather than replacing it.