



PARZIVALPRIME

Team: Riyaz, Dhruv, Preetam

TrialPlus NEXUS

AI-Powered Clinical Trial Intelligence Platform

58,085

Patients

3,401

Sites

23

Studies

4

AI Agents

5

ML Models



THE PROBLEM

The multi-source, heterogeneous data from clinical trials remains inconsistent and siloed leading to a delay in identification of operational bottlenecks, inconsistent data quality and limited visibility for scientific decision making. Currently this is being done manually by different teams leading to fragmented communication and increase in cycle times.

SOLUTION: Layer Architecture and Data Flow

LAYER 6 Presentation

Role Based Dashboards | AI Assistant | Patient/Study/Site Level Views | Visualization (Forms, Patient Status, DQI, Lab Issues, Open Queries, etc) | React + Vite + NodeJS

LAYER 5 Collaboration

Investigation Rooms | @Tagging & Mentions | SLA-Monitored Escalation Pipeline | Team Presence Tracking | Evidence Voting | Resolution Workflow

LAYER 4 Agentic & Generative AI

4 Agents (ReAct + 20+ Tools) | LangGraph Orchestration | ChromaDB for RAG | 12 Types Context-Aware Report Generation | Groq & Ollama Models | Unslot Fine-tuning

LAYER 3 ML Models

Patient Risk Classifier | Issue Type Detector | Resolution Time Predictor | Site Risk Ranker | Anomaly Detector | 5 Models | SHAP Explainability | XGBoost | LightGBM | Governance

LAYER 2 Metrics/Analytics

8-Component DQI | Clean Patient | DB Lock Ready | Cascade Intelligence | Resolution Genome | Causal Hypothesis Engine | Monte-Carlo Simulation | What-If Simulation

LAYER 1 Data Foundation

9 Sources → 264 Features → Unified Patient Record (57,974 patients) | Neo4j



Cascade Intelligence

Creates a NetworkX/Neo4j Dependency Graph of issues. And understands how fixing 1 issue → unblocks 10+ downstream issues.

E.g. "Fix 12 queries at JP-101

→ Unlocks 8 blocked PI signatures

→ Clears 3 SAE reviews waiting on data

→ NET IMPACT: +14 DQI points for the site"



Resolution Genome

Every issue resolution becomes knowledge for similar future issues. This module captures and matches patterns.

E.g. "This issue type resolved 847 times. Top solution: 94% success rate"



Causal Hypothesis Engine & What-If Scenario Simulator

E.g.

> Hypothesis: PI Absence Cascade at Site JP-101: Confidence: 82% | Evidence: Overdue signatures +14, No PI login 10 days

> What-If: Add 2 CRAs → 73% chance of DB Lock by March 15

ML MODELS: TRAINING & EVALUATION

Methodology: 58,095 patients \times 264 features | 70/15/15 Split (Stratified) | 5-Fold CV | Optuna Tuning | SHAP Explainability

Site Risk Ranker

Type: Learning-to-Rank (Pairwise)

NDCG@10: 0.84 | MAP: 0.85

Model: XGBoost Ranker

Top Features: 141 aggregated from 27 raw features

Patient Risk Classifier

Type: 4-Class Classification

ROC-AUC: 0.91 | F1: 0.78 | Recall: 84%

Model: XGBoost + LightGBM Ensemble

Top Features: Query density, SDV rate, Signature delays

Resolution Time Predictor

Type: Quantile Regression

MAE: 3.2 days | Coverage: 82%

Model: XGBoost Quantile [0.1-0.9]

Top Features: Issue type, Site history, Complexity

Multi-Label Issue Detector

Type: 14 Binary Classifiers

Avg AUC: 0.90 | Avg F1: 0.77

Model: XGBoost (Binary Relevance)

Top Features: Per-issue optimized thresholds

Anomaly Detector

Type: 4-Class Unsupervised Classification

Anomalies: 5800 (10%) | Critical : 580 (1%)

Model: Isolation Forest, DBScan, Autoencoder

Top Features: Per-issue optimized thresholds

Model Comparison (Risk Classifier)

Model	ROC-AUC	F1
Random Baseline	0.50	0.21
Logistic Regression	0.72	0.48
Decision Tree	0.68	0.44
Random Forest	0.81	0.62
XGBoost	0.89	0.74
Ensemble ✓	0.91	0.78



EXPLAINABILITY (SHAP)

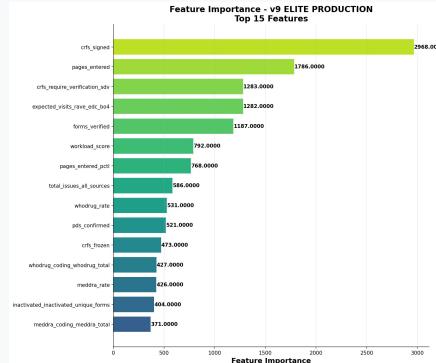
Every prediction includes feature contributions:

"High Risk (78%) because:

- 5 open queries (+18% risk)
- 2 missing visits (+12% risk)
- PI signature overdue 45+ days (+8%)



SHAP Feature Importance



AGENTIC AI ARCHITECTURE

SUPERVISOR

Routes, Plans, Orchestrates

DIAGNOSTIC

NL to SQL, Fetch Data points

RESOLVER

Action Recommendations

COMMUNICATOR

Final Response

EXAMPLE: "Why is site JP-101 underperforming?"

- Supervisor
 - Thought: Complex diagnostic query. Need metrics + root cause
 - Action: route_to_agent("DIAGNOSTIC", query)
- Diagnostic
 - Thought: Need current metrics for JP-101
 - Action: query_site_metrics(site_id="JP-101")
 - Observation: DQI=71, Queries=23, Missing Sigs=12
 - Thought: High signatures missing, Check PI activity
 - Action: get_pi_activity(site_id="JP-101", days=14)
 - Observation: PI signed 0 does since Nov 1 (conference)
 - Hypothesis: PI Absence Pattern, 82% confidence
- Resolver
 - Action: predict_recovery(site="JP-101", issue="pi_absence")
 - Prediction: 12 days to recovery, CI: [8, 18] days
- Communicator
 - Action: search_resolution_genome(pattern="pi_absence")
 - Recommendation: [Batch signature session, Delegate to sub-I]

Time taken ~ 6 seconds

24 AGENT TOOLS

-  Data: get_patient, get_site_summary, get_high_priority_patients
-  Search: search_knowledge (RAG), search_resolutions, search_patterns
-  Analytics: get_cascade_path, get_cascade_impact, get_dblock_projection
-  ML: predict_risk, detect_issues, predict_resolution_time
-  Genome: get_resolution_recommendation, get_patient_resolution_plan

AUTONOMY MATRIX (5×4)

- $\geq 95\%$ + Low Risk → AUTO-EXECUTE (Immediate)
- 80-94% + Medium → AUTO-DRAFT (72h Human Review)
- 60-79% + High → ESCALATE (24h Manager)
- <60% + Critical → BLOCK (Sponsor Required)
-  NEVER-AUTO (12 Types):
 - SAE Causality, Protocol Deviation, Unblinding,
 - Regulatory, Site Closure, DB Lock, Consent,
 - Medical Judgment, Safety Signal, Locked Data...

TECHNOLOGY STACK

- LLM: Groq (Llama 3.3 70B) + Ollama (fallback)
- Agents: Custom ReAct Loop + NL-to-SQL Grounding
- ML: XGBoost + LightGBM + Sentence Transformers
- Frontend: React 18 + Vite + TailwindCSS + Recharts
- Database: PostgreSQL + ChromaDB (RAG) + Neo4j (Cascade)
- Deploy: Docker + FastAPI

RESULTS AND NEW FEATURES

66.7%

Clinical Clean Rate
38,684 patients

78.0%

Operational Clean
41,583 patients

73.2%

DB Lock Ready
10,401 patients

32,287+

Issues Tracked
Real-time monitoring

⚠ CHALLENGES OVERCOME

- 9 data sources → Unified Patient Record (PostgreSQL)
- Class imbalance (2-45%) → SMOTE + Class Weights (20:15:8:1)
- Multi-class risk → Cascade Prediction (Critical→High→Medium→Low)
- Noisy labels → Percentile thresholds (P50/P80/P95)
- Feature leakage → Outcome features excluded
- Model drift → Low-class recall as stability detector

📌 ROADMAP

-  NEXT: Real-time Streaming (Kafka) | AWS Deployment (SageMaker)
- Expanded Agent Autonomy | Multi-tenant Architecture
- Production SLA Monitoring | Notification Service | Electronic Signatures

97.5

8-point Data Query Index

🐳 DB & Deployment

- PostgreSQL DB
- Dockerized service for DB
- FastAPI Backend



Monte-Carlo Simulation

- 10,000 runs with uncertainty quantification. What-if scenarios to predict DB lock timeline, site closure impact, transfer success rate, etc.

E.g. Timeline: P50 = 14.2 days | 95% CI: [8.1, 28.7 days] Deadline Probability: 73.2% on-time | Risk Score: 42.5 (MODERATE)

🌊 PSI Drift Detection

- PSI-Based Monitoring with 21 CFR Part 11 Audit. Queries Postgresql for live DQI metrics and logs to audit_logs.

E.g. "drift_detection": {"monitor_class": "Low", "reason": "Low class has 98%+ recall stability", "action": "Alert on >3% week-over-week drop"}



Frontend Changes

- React18 + Vite + Typescript
- Tailwind CSS + Radix UI
- Tanstack query + table
- Recharts
- Framer Motion
- Zustand