

- ClinOps Intel: Stakeholder Demo Script
 - The Story
 - Pre-Demo Setup
 - Set the Stage (2 min)
- Part 1: MVR Intelligence Queries (Priority Demo Set)
 - Query 1: MVR Narrative Extraction
 - What You Will See
 - Query 2: Cross-Domain MVR + Data Quality Synthesis
 - What You Will See
 - Query 3: CRA Behavioral Pattern Detection
 - What You Will See
 - Query 4: Full Multi-Agent + MVR Synthesis
 - What You Will See
 - Query 5: Cross-Site MVR Comparison
 - What You Will See
 - Query 6: PI Engagement Temporal Trajectory
 - What You Will See
 - Query 7: Zombie Finding Recurrence
 - What You Will See
 - Browse the Evidence (1 min)
- Part 2: Additional Test Cases
 - Query 8: Broad Triage -- Weekly Attention
 - Query 9: Financial Intelligence
 - Query 10: Phantom Compliance -- Suspiciously Perfect Metrics
 - Query 11: Site-Specific -- CRU Hungary (SITE-074) Deep Dive
 - Query 12: Site-Specific -- Clearview Cancer Institute (SITE-012) Anomaly
 - Query 13: Site-Specific -- SITE-114 Ghost Site
- Quick Reference: All Cached Queries
 - MVR Intelligence Queries (Priority)
 - Additional Test Cases
- Feature Callouts During Demo
- Closing Statement
- Dashboard vs Agent Discovery
- Novel Concepts (Emerged from AI Analysis)

ClinOps Intel: Stakeholder Demo Script

The Story

"Monitoring Visit Reports are the richest qualitative data source in clinical operations -- but nobody reads them systematically. CRAs write them, they get filed, and the insights die in PDFs. We built an agentic AI system that reads every MVR, cross-references it with structured operational data, and surfaces hidden connections that no dashboard or single-person review could find."

Pre-Demo Setup

1. Start backend: `source .venv/bin/activate && python -m uvicorn backend.main:app --host 0.0.0.0 --port 8000 --reload`
 2. Start frontend: `cd frontend && npm run dev`
 3. Navigate to `http://localhost:5000` -> click **STUDY-001** to enter Command Center
 4. Demo queries are pre-cached. Responses appear instantly.
-

Set the Stage (2 min)

Open on the dashboard. Let it breathe for a moment.

Say: "This is a Phase III trial -- 595-patient target across 20+ sites in 5 countries. The dashboard shows our KPIs, the world map, and the sites flagged for attention. Everything here comes from structured data -- database tables, enrollment logs, query counts."

Point to the Needs Attention panel. "These sites are flagged. A traditional approach would have a CODM Lead open a spreadsheet, pull KRI data, and try to figure out what's going on. That takes hours per site."

Point to the "View MVRs" button on one of the attention sites. "Notice these sites also have monitoring visit reports -- PDF documents written by CRAs after every on-site visit. These reports contain qualitative intelligence that no structured KPI captures: Was the PI engaged? Did old findings come back? Was the CRA thorough or rubber-stamping?"

Click "**View MVRs**" on **SITE-033**. **Browse 2-3 reports**. "Here are the actual MVR PDFs -- executive summaries, findings checklists, PI engagement notes. Now imagine reading all of these reports across multiple sites and finding hidden patterns. That's what our AI agents do in seconds."

Part 1: MVR Intelligence Queries (Priority Demo Set)

These queries showcase the core MVR analysis capability -- reading unstructured monitoring visit reports and cross-referencing with structured operational data.

Query 1: MVR Narrative Extraction

Type in the Assistant panel:

Show key findings from the last 2 monitoring visit reports for SITE-033

While it streams, narrate: "The Conductor routes this to the MVR Analysis Agent plus supporting agents. The agent enters its PRPA loop: it *perceives* by reading actual MVR narrative content -- executive summaries, findings checklists, PI engagement fields, SDV results. Then it *reasons* about patterns across visits, *plans* deeper analysis, *acts*, and *reflects* on completeness."

What You Will See

Agents invoked: mvr_analysis, data_quality, enrollment_funnel, clinical_trials_gov, phantom_compliance, site_rescue, vendor_performance (7 agents)

Executive Summary: "SITE-033 has successfully met its enrollment target (21/21) but faces a critical data integrity risk due to ineffective monitoring. While recent visits are actively generating new queries (30-40 per visit), the CRA is failing to verify and close 154 aged 'Answered' queries, creating a 'database lock threat' with zero documented follow-up actions in the last 5 reports."

Key finding to expand (click the finding card):

- **Finding:** Ineffective CRA workflow causes query backlog because the CRA prioritizes generating new queries over verifying 'Answered' ones
- **Naive reading:** "The site is unresponsive to queries"
- **Causal chain:** CRA visits site -> generates new queries -> ignores 'Answered' queue -> backlog accumulates
- **Confirming evidence:** 154 queries in 'Answered' status with mean age >100 days; recent visits generated 37 and 39 new queries; zero follow-up actions in last 5 reports
- **Confidence:** 0.98

Second finding:

- **Finding:** Inadequate prescreening causes avoidable screen failures -- site formally screens patients with clear exclusions (prior chemo)
- **Confidence:** 0.90

How to explain: "Notice the 'naive vs actual' pattern. A dashboard would tell you 'the site isn't answering queries.' But the agent discovered the opposite -- the site IS answering, it's the CRO that's failing to verify. That reversal of blame is something only an agentic system can uncover by reading the actual MVR narratives and cross-referencing with query aging data."

Query 2: Cross-Domain MVR + Data Quality Synthesis

This is the centerpiece. Type:

SITE-033 has had missed monitoring visits. Is there hidden data quality debt that the KPIs aren't showing?

Say: "The Conductor recognizes this needs *multiple* agents. It invokes MVR Analysis, Data Quality, Phantom Compliance, Financial Intelligence, and more -- *in parallel*. Each runs its own autonomous investigation, then the Conductor synthesizes."

What You Will See

Agents invoked: data_quality, enrollment_funnel, clinical_trials_gov, phantom_compliance, site_rescue, vendor_performance, financial_intelligence, mvr_analysis (8 agents)

Executive Summary: "SITE-033 is actively enrolling (100% of target) but carries a critical, hidden data quality debt because the CRO missed 3 consecutive monitoring visits between November 2024 and February 2025. While the site appears responsive with 146 'Answered' queries, these responses remain unverified due to the monitoring gap, creating a backlog of 154-day-old unconfirmed data points. This reactive management style has driven a \$193,618 budget overrun, as CRAs are now forced to clean up aged data rather than preventing errors upfront."

Key finding -- the "aha" moment (click to expand):

- **Finding:** Missed monitoring visits cause hidden data debt because site answers to queries are never verified against source documents
- **Naive reading:** "The site is responsive because they have answered 146 queries and have few 'Open' items"
- **Causal chain:** Monitoring Gaps -> No Source Verification -> Stagnant 'Answered' Queries -> Hidden Quality Debt
- **Confirming evidence:** Visits missed on 2024-11-09, 2024-12-21, 2025-02-08 (325, 283, 234 days overdue); 35 'Missing Data' queries in 'Answered' status with mean age 154.0 days; 0.0% monitoring-triggered query resolution
- **Confidence:** 0.98

Second finding -- budget impact:

- **Finding:** Reactive data management causes budget overruns -- the CRO pays for expensive cleanup instead of cheaper prevention
- **Confirming evidence:** Variance = \$193,618.80 (70.02% over budget); unprompted corrections = 0, query-triggered = 104; post-gap visits generated query spikes of 37, 39, 30
- **Confidence:** 0.90

Third finding -- phantom compliance signal:

- **Finding:** CRA negligence causes 'phantom compliance' at SITE-006 -- 27 findings across 11 visits by CRA-008 but zero EDC queries ever generated
- **Confidence:** 0.99

How to explain: "A dashboard would show green KPIs during the monitoring gap -- query 'Time to Answer' was low, which looks healthy. But the agent discovered a 'false positive' success story. The site answered queries, but nobody verified the answers against source documents. 146 queries are in limbo. And independently, the Financial Intelligence agent traced this to a \$193k budget overrun -- reactive cleanup instead of prevention. No single data source reveals this."

Query 3: CRA Behavioral Pattern Detection

Type:

Are there any CRAs who might be rubber-stamping their monitoring visit reports?

Say: "This is almost impossible to detect with structured data alone. A rubber-stamping CRA files reports on time, the KPIs look fine. But the MVR agent reads the *narrative quality* -- word counts, finding patterns, the specificity of SDV observations."

What You Will See

Agents invoked: mvr_analysis, phantom_compliance, data_quality, vendor_performance, financial_intelligence (5 agents)

Executive Summary: "We have confirmed 'rubber-stamping' behavior by three specific CRAs exhibiting two distinct patterns of negligence. CRA-008 (SITE-006) and CRA-009 (SITE-007) are documenting critical findings in reports to satisfy administrative requirements but generating zero EDC queries to actually fix the data. Conversely, CRA-117 (SITE-074) is filing 'perfect' zero-finding reports with copied text, indicating they are likely not reviewing source data at all."

Three rubber-stamping patterns discovered:

1. **CRA-008 + CRA-009: "Document but Don't Act"** (Confidence: 0.95)

- CRA-008 at SITE-006: 11 visits, 27 documented findings (3 critical), but **zero EDC queries** ever generated

- CRA-009 at SITE-007: 6 visits, 16 findings (1 critical), zero queries -- then a forced spike of 9 queries on a single day (2025-04-18)
- **Naive:** "Zero queries indicate these sites are performing exceptionally well with clean data"
- **Actual:** KPI pressure drives CRAs to document findings in reports (to pass review) but skip the EDC work that would actually fix the data

2. **CRA-117: "Ghost Monitoring"** (Confidence: 0.95)

- SITE-074: 100% zero-finding visits across 3 consecutive trips (Mar-Sep 2025)
- Executive summaries contain identical phrases: "No concerns noted" and "All source data verified... no discrepancies noted"
- Average word count: 149 (extremely low); 0 action items
- **Naive:** "SITE-074 is a 'star site' with perfect compliance"
- **Actual:** CRA copies previous report templates instead of verifying current data

3. **SITE-085: Phantom Compliance signal** -- robotic batch data entry (50% on weekends, 0.84 day variance) suggesting CRA is missing massive back-entry issues

How to explain: "The agent cross-referenced every CRA's *entire portfolio* -- not just one site, but all their assignments. It found two completely different failure modes: document-but-don't-act and template-copying ghost monitoring. A structured dashboard would show all three CRAs as 'on track' because they're filing their reports on time. Only by reading the *content* and cross-referencing with EDC activity can you see what's happening."

Query 4: Full Multi-Agent + MVR Synthesis

Type:

Why is SITE-022 underperforming? Include insights from the monitoring visit reports.

Say: "Now we're asking for the full multi-agent synthesis with MVR integration."

What You Will See

Agents invoked: data_quality, enrollment_funnel, clinical_trials_gov, phantom_compliance, site_rescue, vendor_performance, financial_intelligence, mvr_analysis (8 agents)

Executive Summary: "SITE-022 is not a failure but a stalled high-performer (20/24 enrolled) masking a 'shadow screening' process. The site maintains an artificial 91% conversion rate by pre-screening patients offline, hiding true recruitment effort and failures. Operational health has degraded following a CRA transition (CRA-004 to CRA-005), resulting in a collapse of PI engagement and a spike of 46 findings in the most recent visit."

Four convergent findings from independent agents:

1. "Shadow Screening" Detection (Enrollment Funnel + Phantom Compliance, Confidence: 0.95)

- 91% conversion rate (20/22) vs regional peer SITE-012 (77%); only 9.1% screen failures
- 100% of failures are for a single reason ('SF_ORGAN' -- Renal); zero subjective failures (ECOG, Smoking)
- 2 active Phase III NSCLC competitors in Bakersfield saturating the pool
- **Naive:** "The site has excellent patient selection and high efficiency"
- **Actual:** "The site is 'shadow screening' offline to inflate conversion rates, likely due to competitive pressure"

2. CRA Transition -> PI Disengagement (MVR Analysis, Confidence: 0.90)

- PI engagement dropped from 'High' to 'Low/Medium' after CRA-004 replaced by CRA-005
- MVR word count dropped from 452 (Visit 4) to 162 (Visit 5) immediately post-transition
- Visit 8 (CRA-005): findings spiked to 46; PI rated 'Medium' with deferred questions
- **Naive:** "The site's performance naturally degraded over time"
- **Actual:** "The CRA transition broke PI rapport, causing operational decline"

3. Competitive Intelligence (ClinicalTrials.gov Agent)

- Competitor trial NCT06345729 (Merck) activated 8 days after SITE-022

4. **Recovery Signal** -- Velocity stabilized at 0.5 patients/week with positive slope (+0.083)

How to explain: "Eight separate agents investigated independently, and the Conductor found convergent root causes. The enrollment agent detected statistical anomalies in conversion rates. The competitive intelligence agent found rival trials in the same geography. And the MVR agent traced PI engagement decline to a specific CRA transition date. No single data source reveals this picture."

Query 5: Cross-Site MVR Comparison

Type:

Compare the quality of monitoring across SITE-012, SITE-033, and SITE-074 based on their visit reports

What You Will See

Agents invoked: mvr_analysis, data_quality, phantom_compliance, vendor_performance, enrollment_funnel, clinical_trials_gov (6 agents)

Executive Summary: "Monitoring quality varies critically across the three sites, ranging from negligence to abandonment. SITE-074 presents the highest risk with 'rubber-stamp' monitoring (CRA-117) masking egregious safety violations like enrolling a 'Never Smoker' in a lung cancer trial. SITE-033 suffers from operational abandonment with a 174-day monitoring gap following CRA turnover, while SITE-012 exhibits high-effort but inefficient oversight characterized by recurring 'zombie findings' that are fixed individually but never resolved at the root cause."

Three failure mode archetypes:

Site	Failure Mode	Key Evidence	Confidence
SITE-074	Rubber-Stamping	CRA-117: 0 findings across 3 consecutive visits; 'Never Smoker' enrolled despite protocol; mean entry lag 1.48 days (too fast to be real)	0.95

Site	Failure Mode	Key Evidence	Confidence
SITE-033	Abandonment	CRA-006 replaced by CRA-007; 174-day max monitoring gap; 154 queries in 'Answered' status waiting for CRA closure	0.98
SITE-012	Reactive Churn	Same lab unit conversion error recurs across Visits 4, 6, and 9; CRA fixes symptoms without retraining staff; 5 confirmed 'zombie patterns'	0.90

How to explain: "The agent didn't just compare numbers -- it identified three completely different *archetypes* of monitoring failure. SITE-074 looks perfect in a dashboard because it has zero findings. But the CRA approved a patient who was ineligible ('Never Smoker' in a lung cancer trial). SITE-033 was abandoned by its CRA for 174 days. And SITE-012 has an overactive CRA who fixes errors one by one without ever retraining the staff. Three sites, three different diseases, all invisible to a KRI dashboard."

Query 6: PI Engagement Temporal Trajectory

Type:

Has PI engagement declined at any of our monitored sites?

What You Will See

Agents invoked: mvr_analysis, enrollment_funnel, data_quality, site_rescue, vendor_performance, phantom_compliance, clinical_trials_gov (7 agents)

Executive Summary: "We have confirmed tangible declines in PI engagement at SITE-012 and SITE-055, directly impacting study performance. At SITE-012, the PI's shift to 'passive signing' has allowed critical safety data errors to recur ('zombie findings') despite prior retraining, while at SITE-055, the PI's absence for three consecutive visits has caused enrollment to flatline. Conversely, the apparent disengagement at SITE-033 is largely an operational artifact of a CRO monitoring backlog leaving 154-day-old queries unverified."

Key findings:

1. **SITE-012: PI Delegates -> Errors Recur** (Confidence: 0.95)

- PI 'delegated immediate resolution to the SC' (Visit 6); errors recurred by Visit 9 for Subjects 101-019 to 101-023
- Action items escalated from 4-5 to 12 as PI engagement dropped to 'medium'
- **Naive:** "The site staff are incompetent and need more training"
- **Actual:** "PI delegation breaks accountability -- coordinator reverts workflow after retraining fails"

2. **SITE-055: PI Absent -> Enrollment Stalls** (Confidence: 0.90)

- "PI has not been available for the last 3 monitoring visits" (Visit 9)
- "Sub-I reports difficulty in patient recruitment without PI involvement" (Visit 7)
- **Naive:** "The site has exhausted its patient pool"
- **Actual:** "PI unavailability disconnects the referral network; enrollment flatlines"

3. **SITE-033: False Signal** (Confidence: 0.90)

- 35 'Missing Data' queries in 'Answered' status with 154-day mean age
- But this is a CRO verification backlog, not site disengagement

4. **CRA Transitions Cause Stalls** (Confidence: 0.80)

- SITE-017: transition to CRA-037 -> velocity dropped to 0.0 with -0.06 slope
- SITE-009: transition to CRA-013 -> similar pattern

How to explain: "The agent traced PI engagement trajectories across time using MVR narrative signals -- not just a snapshot, but a progression. It found two sites where PI decline directly caused operational failure, one site where disengagement was actually a false signal (it was the CRO's fault), and a pattern where CRA transitions consistently break PI rapport."

Query 7: Zombie Finding Recurrence

Type:

Are there zombie findings at SITE-012 -- issues that keep coming back despite being marked resolved?

What You Will See

Agents invoked: mvr_analysis, data_quality, enrollment_funnel, clinical_trials_gov, phantom_compliance, site_rescue, vendor_performance, financial_intelligence (8 agents)

Executive Summary: "SITE-012 exhibits confirmed 'zombie findings' in Laboratory Results and Pre-screening, where identical errors recur across visits despite being marked 'resolved' by the CRA. While the site successfully reached its enrollment cap (28/28), this operational success masks a degrading quality process: the Site Coordinator is fixing individual data errors without correcting the underlying workflow, and PI oversight has visibly declined from 'High' to 'Medium'."

Key findings:

1. Lab Unit Zombie Pattern (Confidence: 0.95)

- Visit 4 (2024-10-02): "Correct Lab Results CRF entries" marked Resolved
- Visit 6 (2025-01-08): Same error recurred for Subjects 101-008 to 101-011
- Visit 9 (2025-06-11): Error persists for Subjects 101-019 through 101-023
- **Naive:** "The site is responsive because they resolve findings immediately"
- **Actual:** "SC fixes the data point, but the process remains broken -- error recurs for every new patient"

2. Pre-screening Failures (Confidence: 0.90)

- 7/10 screen failures are for hard criteria (2x Histology, 3x ECOG, 2x Prior Chemo) -- all verifiable from a chart review before consent
- PI engagement declined from 'High' (V1-V4) to 'Medium' (V6-V9)
- **Naive:** "High screen failure rates are a byproduct of high enrollment volume"
- **Actual:** "Site relies on formal screening to catch exclusions instead of chart-based pre-screening"

3. False Stall Alert (Confidence: 0.95)

- "SITE-012 has stalled" is a false positive -- cumulative randomized is 28/28 (target met)
- Velocity dropped to zero because enrollment is complete, not because of failure

How to explain: "The 'stalled enrollment' alert is a great example of why you need agentic investigation. The dashboard sees zero velocity and flags it. But the agent discovered the site already hit its 28-patient target -- it's *supposed* to have zero velocity. The real problem is the zombie findings: the same lab unit error keeps coming back because the CRA fixes individual data points without ever retraining the coordinator."

Browse the Evidence (1 min)

Click "View MVRs" on SITE-033 in the Needs Attention panel.

Say: "For full transparency, here are the actual source documents. You can verify everything the agent found by reading the original reports."

Click through 2 reports -- point out executive summary differences, action counts, word counts. **Resize the Assistant panel wider** to show investigation results alongside the MVR browser.

Part 2: Additional Test Cases

These queries demonstrate broader agentic capabilities beyond MVR analysis -- financial intelligence, triage, phantom compliance, and site-specific deep dives.

Query 8: Broad Triage -- Weekly Attention

Which sites need attention this week?

Agents invoked: data_quality, enrollment_funnel, clinical_trials_gov, phantom_compliance, site_rescue, vendor_performance, financial_intelligence, mvr_analysis (8 agents)

Executive Summary: "Critical operational failures are driving a \$3.2M projected shortfall and data integrity risks. SITE-033 faces a database lock threat not due to site performance, but because the CRO has failed to verify site responses for over 5 months."

Meanwhile, SITE-031 and SITE-009 are 'zombie sites' burning resources with negligible enrollment, requiring immediate closure."

Key findings (7 total):

- SITE-033: 154 'Answered' queries waiting for CRA verification, mean age 154 days, max monitoring gap 174 days
- SITE-074: 100% of 4 visits resulted in zero findings; average word count 182 vs active site average >350
- SITE-119: 22 patients stuck post-screening (26 passed, only 4 randomized) -- process failure, not drug supply
- SITE-012: 51% over budget (\$181k overrun); recurring screen failures for hard exclusion criteria
- SITE-031: \$3.1M projected financial liability from enrollment stagnation
- SITE-085: 35.7% of data entries on Sundays -- phantom compliance signal
- Smoking history criterion systematically misunderstood, causing 100% screen failure for new candidates

How to explain: "Eight agents ran simultaneously and the Conductor ranked findings by urgency. Notice the mix: financial liability (\$3.1M), data integrity risk (154 stale queries), phantom compliance (Sunday batch entry), and process failures (22 stuck patients). A triage dashboard can flag sites; this system explains *why* they need attention and *what specific action* to take."

Query 9: Financial Intelligence

What is driving budget variance?

Agents invoked: financial_intelligence, vendor_performance, enrollment_funnel, data_quality, clinical_trials_gov, mvr_analysis, phantom_compliance (7 agents)

Executive Summary: "Budget variance is primarily driven by *3.1M in projected delay costs from enrollment stagnation at SITE – 031 and 2.5M in reactive change orders due to initial strategy failures. Additionally, operational inefficiencies are bleeding cash: VEND-001 is 405k over budget on monitoring due to 'zombie finding' rework loops (SITE – 012) and missed visits (SITE – 033), while SITE – 033 alone is 193k over budget due to a 'double-touch' data correction workflow.*"

Key findings (5 total):

1. Poor Site Selection -> Delay Costs (Confidence: 0.95)

- SITE-031 projected delay cost = \$3,116,742; 83% screen failure rate on basic criteria
- SITE-006 'Zombie Site' -- 0 enrolled in 12 weeks
- **Naive:** "Budget variance is due to unexpected high patient volume"
- **Actual:** "Non-performing sites burn fixed costs without delivering enrollments"

2. Zombie Finding Rework Loops (Confidence: 0.90)

- VEND-001 CRO monitoring overage = \$405,830
- SITE-012: lab unit error recurred across multiple visits -- CRA repeats work each time
- SITE-033: CRA-007 missed 3 visits, causing 156 queries to stall

3. Reactive Data Correction (Confidence: 0.90)

- SITE-033 budget overrun = 193,618(7022,386
- SITE-041, SITE-108, SITE-053: 100% query-triggered corrections (zero self-correction)

How to explain: "The financial agent traced every dollar of variance to a root cause.

3.1M isn't higher patient care costs' -- it's zombie sites burning cash. 405k isn't 'complex protocol' -- it's CRAs repeating the same rework loops. \$193k isn't 'high enrollment costs' -- it's reactive cleanup from missed monitoring visits. Every number has a story."

Query 10: Phantom Compliance -- Suspiciously Perfect Metrics

Give me some examples of suspiciously perfect data entry metrics?

Agents invoked: phantom_compliance, data_quality, mvr_analysis (3 agents)

Executive Summary: "We identified three clear instances of 'suspicious perfection' where flawless metrics are masking operational failures. SITE-125 maintains a 0% correction rate in the EDC despite monitors finding ~3 issues per visit, suggesting data is being manipulated or fixed 'off the books.' SITE-006 shows zero open queries despite 27 documented monitoring findings, indicating the CRA is identifying issues but failing to log them in the system."

Key findings (4 total):

1. **SITE-125: Offline Corrections** (Confidence: 0.90)

- 0.0% correction rate across 12 entries, yet monitoring visits recorded 2-3 findings each
- **Actual:** Staff fix errors offline to evade audit trails, creating a 'clean' EDC facade

2. **SITE-006: CRA Never Queries** (Confidence: 0.95)

- 27 total findings across 11 visits (including 2 critical), yet 0 queries in EDC
- **Actual:** CRA-008 identifies issues in trip reports but never enters them into the system

3. **SITE-007: Speed Over Quality** (Confidence: 0.90)

- Entry lag: 2.9 days mean; CRA-009 had 0 queries for first 6 visits despite 16 findings, then sudden spike of 9 queries on a single day
- **Actual:** CRA rubber-stamped rapid entry for months, then was forced to act by an external trigger

4. **IQVIA Reporting Bug:** Protocol Deviation Rate of -0.6% -- a mathematically impossible value indicating a broken KPI algorithm

Query 11: Site-Specific -- CRU Hungary (SITE-074) Deep Dive

About CRU Hungary Egeszsegugyi Kft.: What's causing the enrollment issues at this site?

Agents invoked: enrollment_funnel, clinical_trials_gov, data_quality, site_rescue, mvr_analysis (5 agents)

Executive Summary: "SITE-074 is not failing to enroll; it has actually exceeded its target with 22 subjects (110% of goal), but this activity is invisible in central dashboards due to severe data entry lags. The critical issue is 'rubber-stamping' by the CRA, who has reported zero findings across 7 consecutive visits despite this high volume, creating a massive unverified data integrity risk."

Key findings:

- CRA negligence causes false enrollment alerts -- dashboard shows empty, but 22 patients (110% target) are enrolled
- CRA-117: 100% zero-finding visit rate; word count dropped 282 -> 133 with repetitive 'Remarkable recruitment' phrasing
- No competing NSCLC trials in the Miskolc region -- performance issues are internal, not environmental

How to explain: "This is a perfect example of why dashboards lie. The system flagged SITE-074 as 'enrollment issues' -- but the agent discovered it's actually *exceeding* its target. The real issue is the CRA rubber-stamping visits, making the data invisible upstream."

Query 12: Site-Specific -- Clearview Cancer Institute (SITE-012) Anomaly

Is there some data anomaly with Clearview Cancer Institute?

Agents invoked: data_quality, enrollment_funnel, clinical_trials_gov, phantom_compliance, site_rescue, vendor_performance, financial_intelligence, mvr_analysis (8 agents)

Executive Summary: "The 'anomaly' at Clearview Cancer Institute (SITE-012) is a reporting artifact masking a high-performing but operationally expensive site. The site has fully met its enrollment target (28/28) and generated 2,725 data queries that were not correctly reflecting in summary reports due to a vendor data integration failure. The site is

now in critical status due to a 51% budget overrun (\$181k) and a heavy data cleaning backlog."

Key findings (5 total):

- Vendor integration failure causes missing query metrics -> false 'perfect data' signal (Confidence: 0.95)
- Enrollment velocity zero because protocol cap reached, not performance failure (Confidence: 0.95)
- Weak pre-screening: 2 failures for squamous histology, 2 for prior chemo -- all hard exclusions catchable earlier (Confidence: 0.90)
- 51% over budget (\$181,412 overrun) from unbudgeted screen failure pass-through costs
- IQVIA reported mathematically impossible -0.6% Protocol Deviation Rate -- systemic data corruption

Query 13: Site-Specific -- SITE-114 Ghost Site

Show key findings from the last 2 monitoring visit reports for SITE-114

Agents invoked: mvr_analysis, data_quality (2 agents)

Executive Summary: "There are no monitoring visit reports for SITE-114 because the site has enrolled zero patients despite being active for over 12 months. Unlike peer site SITE-110, which has received 7 monitoring visits and generated ~45 queries, SITE-114 is operationally dormant ('Ghost Site'). The lack of reports is not a documentation failure but a direct result of total recruitment stagnation."

How to explain: "The agent was asked about MVRs, but instead of saying 'no data found,' it investigated *why* there are no reports and compared against a peer site. It reframed the absence of data as itself a finding -- this is a ghost site that should be closed."

Quick Reference: All Cached Queries

MVR Intelligence Queries (Priority)

#	Query	Agents	Key Insight	Naive vs Actual
1	Show key findings from the last 2 monitoring visit reports for SITE-033	7	CRA generates new queries but never closes 154 'Answered' ones	Naive: "Site unresponsive" / Actual: "CRO failing to verify"
2	SITE-033 has had missed monitoring visits. Is there hidden data quality debt...?	8	100% enrolled but 146 unverified queries, \$193k overrun	Naive: "KPIs green" / Actual: "Flying blind on 21 patients"
3	Are there any CRAs who might be rubber-stamping their monitoring visit reports?	5	Two patterns: Document-but-Don't-Act (CRA-008/009) and Ghost Monitoring (CRA-117)	Naive: "CRAs file on time" / Actual: "Three CRAs gaming the system differently"
4	Why is SITE-022 underperforming? Include insights from the monitoring visit reports.	8	Shadow screening (91% conversion), CRA transition broke PI rapport	Naive: "83% enrolled" / Actual: "Pre-screening offline, PI disengaged"
5	Compare the quality of monitoring across SITE-012, SITE-033, and SITE-074	6	Three failure archetypes: Rubber-Stamping, Abandonment, Reactive Churn	Naive: "SITE-074 is cleanest" / Actual: "SITE-074 is highest liability"
6	Has PI engagement declined at any of our monitored sites?	7	SITE-012 PI delegates -> errors recur; SITE-055 PI absent -> enrollment stalls	Naive: "Sites naturally slow down" / Actual: "CRA changes and PI absence have causal impact"

#	Query	Agents	Key Insight	Naive vs Actual
7	Are there zombie findings at SITE-012...?	8	Same lab error recurs V4->V6->V9; enrollment 'stall' is false positive (28/28 met)	Naive: "Site has stalled" / Actual: "Enrollment complete; real issue is broken workflow"

Additional Test Cases

#	Query	Agents	Key Insight
8	Which sites need attention this week?	8	\$3.2M shortfall, SITE-033 database lock risk, zombie sites SITE-031/SITE-009
9	What is driving budget variance?	7	3.1M <i>delaycosts</i> + 405k rework loops + \$193k reactive cleanup
10	Give me some examples of suspiciously perfect data entry metrics?	3	SITE-125 offline corrections, SITE-006 zero queries despite 27 findings
11	About CRU Hungary: What's causing the enrollment issues?	5	Not failing -- 110% enrolled but invisible due to CRA rubber-stamping
12	Is there some data anomaly with Clearview Cancer Institute?	8	Reporting artifact masking 28/28 enrollment; vendor data integration bug
13	Show key findings from last 2 MVRs for SITE-114	2	Ghost site -- no MVRs because zero enrollment for 12 months

Feature Callouts During Demo

When	Do This	Say This
Any time	Click a finding card to expand	"Every finding is explainable -- here's the evidence trail"
After Query 2	Click "View MVRs" button	"Here are the actual source documents for verification"
Any time	Drag the panel edge	"Resize the investigation panel to fit your workflow"
During streaming	Point to phase indicator	"The agent cycles through perceive-reason-plan-act-reflect in real time"
After results	Show naive vs actual	"A dashboard would tell you X, but the agent discovered Y"
After any finding	Point to confidence score	"0.98 means near-certain; 0.70 means suggestive but not conclusive"
After Query 3	Scroll to Next Best Actions	"These name the specific CRA, site, and intervention -- not generic advice"

Closing Statement

"What you just saw is an AI system that:

1. **Reads unstructured documents** -- monitoring visit reports across multiple sites -- and extracts patterns no human would catch reading them one at a time
2. **Cross-references qualitative signals with structured data** -- connecting what CRAs *wrote* with what the database *shows*
3. **Surfaces hidden connections** -- monitoring gaps causing invisible debt, CRA transitions degrading report quality, rubber-stamping hiding risk
4. **Provides full explainability** -- every finding shows what was investigated, what evidence confirms or refutes it, and what the naive vs correct interpretation is
5. **Runs autonomously** -- each agent decides what to investigate, what tools to use, and whether its investigation is complete, through the PRPA loop

This is not a chatbot answering questions. This is an agentic intelligence system that *investigates* -- like having 8 specialist analysts working in parallel, reading every document, querying every table, and reporting their cross-referenced findings in seconds."

Dashboard vs Agent Discovery

Dashboard View	Agent Discovery
SITE-033: "Low query rate"	154 answered queries the CRO never verified -- \$193k hidden liability
SITE-074: "Zero findings"	CRA rubber-stamping -- enrolled an ineligible 'Never Smoker'
SITE-022: "83% enrolled"	Shadow screening hides recruitment difficulty; PI disengaged after CRA transition
SITE-006: "Compliant, reports on time"	27 findings documented in reports, zero entered into EDC
CRA-117: "Reports filed on schedule"	Identical copy-paste narratives, 149 words average, zero action items
SITE-031: "Underperforming"	Zombie site -- 1 patient in 19 months, \$3.1M projected waste
SITE-012: "Data anomaly"	Reporting artifact -- actually 28/28 enrolled; vendor integration bug
SITE-114: "No MVR data"	Ghost site -- zero patients for 12 months, not a documentation gap

Novel Concepts (Emerged from AI Analysis)

- **"Silent Monitoring"** -- CRA visits frequently but generates zero queries (CRA-008)
- **"Shadow Screening"** -- Site pre-filters patients offline to inflate conversion rates (SITE-022)

- **"Phantom Compliance"** -- Data looks too perfect due to batch entry or suppression (SITE-085, SITE-125)
- **"Zombie Site"** -- Site burns budget for months with near-zero enrollment (SITE-031)
- **"Zombie Findings"** -- Same errors recur across visits despite resolution (SITE-012)
- **"False Positive Success"** -- KPIs show green but material risks are hidden (SITE-033)
- **"Reactive Churn"** -- CRA catches errors after entry instead of preventing them (SITE-012)
- **"Ghost Site"** -- Zero patients, zero activity, but site remains active (SITE-006, SITE-114)
- **"Ghost Monitoring"** -- CRA copies templates instead of reviewing source data (CRA-117)
- **"Document but Don't Act"** -- CRA records findings in reports but never enters EDC queries (CRA-008, CRA-009)