{gsm} Operations: Implementing an Open Central Monitoring Framework

PHUSE 2026 Draft

Abstract

The Good Statistical Monitoring (`{gsm}`) framework, introduced in 2023, has evolved into a scalable, open-source solution for Central Monitoring (CM), supporting reproducible, GxP-compliant risk signal management with over 50 releases across 10 packages and is now deployed across 20+ clinical studies.

This paper introduces the `{grail}` R package (Gsm Risk And Issue Listing), which provides business logic for actioning risk signals and enables systematic and transparent oversight. `{grail}` demonstrates how analytics and business processes can be aligned to strengthen clinical trial quality monitoring.

1. Introduction

Risk-Based Quality Management (RBQM) is now a foundational expectation for modern clinical trial oversight. Regulatory guidance from ICH E6(R2)/(R3), FDA, and EMA emphasizes proactive detection of quality and safety risks, with centralized analytics serving as a key mechanism for identifying potential issues earlier than traditional on-site approaches alone. In practice, however, many organizations still face a gap between statistical detection of anomalies and operational actioning of those findings.

The open-source `{gsm}` ecosystem was developed to address this gap. At its core, `{gsm.core}` provides an analytics engine with a broad library of Key Risk Indicators (KRIs). Supporting packages, including `{gsm.mapping}`, `{gsm.kri}`, and `{gsm.reporting}`, provide data preparation, configuration, and output tooling. While this ecosystem can reliably identify potential risks, CM teams still require consistent business rules and workflows to convert raw metric flags, or "Risk Signals", into actionable, traceable quality management tasks.

The `{grail}` package was created to operationalize this business process. It transforms analytic outputs into structured risk signals, enriches those signals with relevant context and metadata, and supports triage workflows that can be integrated with practical action logs (for example, spreadsheets, JIRA, or Azure DevOps). This paper describes the `{grail}` approach, its YAML-based workflow architecture, and its role in supporting scalable, inspection-ready CM operations.

2. Background

The initial implementation of `{gsm}` relied heavily on analyst-driven spreadsheet processes to track and triage signals. Although effective for pilot studies, this approach became difficult to sustain as portfolio size and cadence increased. As deployment expanded to 20+ studies, the need for reproducibility, standardization, and auditability became central.

Three operational requirements emerged:

1. **Consistency**: Apply the same risk signal logic across studies, snapshots, and teams.
2. **Transparency**: Preserve clear provenance from source metric to action decision.

3. **Scalability**: Support frequent refreshes and growing study volume without proportional manual effort.

`{grail}` addresses these requirements by formalizing post-analytics business logic into reproducible, configurable workflows.

3. Methods

3.1 End-to-end operating model

The CM operating model uses a staged approach:

1. Generate KRI and related outputs using core `{gsm}` workflows.
2. Use `{grail}` workflows to augment and standardize flagged results into risk signals.
3. Export those signals to an action management layer for triage, assignment, and follow-up.
4. Evaluate and resolve signals using standardized triage and review procedures.
5. Feed resolved and ongoing signal data into a **Risk Signal Reporting Layer** for operational oversight, trend analysis, and governance reporting.

This model separates analytics from business process while keeping both components tightly connected through structured metadata.

The **Risk Signal Reporting Layer** (Step 5) builds on the structured, standardized outputs from `{grail}` to enable downstream reporting and process monitoring. Because `{grail}` produces consistent, machine-readable risk signal records with rich metadata, it becomes straightforward to build supporting tools that automate and scale CM operations. Examples include:

- **Templated Risk Review Meeting Slide Decks**: Automatically generate standardized slide decks for recurring risk review meetings, pre-populated with current signal summaries, trend charts, and action status updates.
- **Signal Actionability Dashboards**: Visualize the rate of "actionability" of risk signals by metric type, study, or region—tracking how often flagged signals result in meaningful follow-up versus resolution without action.
- **Triage Timeliness Metrics**: Monitor operational KPIs such as the percentage of risk signals awaiting triage for more than 90 days, enabling teams to identify bottlenecks and ensure timely review cycles.

This reporting layer closes the feedback loop in the CM operating model, ensuring that risk signal management is not only consistent at the individual signal level, but also observable and improvable at the portfolio level.

3.2 YAML workflow architecture

`{grail}` uses YAML-defined workflows to provide a reproducible, configurable pipeline. Each workflow contains:

1. Workflow metadata and versioning context.
2. Input object definitions and required data structures.
3. Ordered processing steps with explicit dependencies.
4. Output schemas and quality checks.

This architecture aligns with patterns from the {workr} package used across the broader `{gsm}` ecosystem, improving transparency, change control, and cross-study portability.

3.3 Creating Risk Signals from KRI outputs

Raw KRI outputs are valuable for statistical detection, but they are not sufficient on their own for consistent operational decision-making. Monitoring teams need additional context to understand whether a signal is clinically meaningful, data-related, persistent across snapshots, and actionable within existing governance processes. `{grail}` provides workflows to standardize this translation from analytic flags to operationally usable risk signals so that triage decisions are more reproducible, auditable, and scalable across studies.

The core KRI Risk Signal workflow (at `inst/workflows/3_reporting/RiskSignals_KRI.yaml`) converts raw analysis outputs into triage-ready risk signals through nine major steps:

1. **Capture Context**: Extract study ID and snapshot date from input data.
2. **Filter Results**: Restrict to valid site/country-level KRIs with relevant flags (± 1, ± 2).
3. **Augment Results**: Add calculated operational fields and required metadata.
4. **Add Metric Metadata**: Join metric definitions, labels, model information, and report links.
5. **Add Group Metadata**: Add site/country descriptors (for example, names, regions, enrollment status).
6. **Add Action Metadata**: Attach recommended actions based on metric and flag direction.
7. **Add Signal Context**: Generate a human-readable signal narrative for reviewers.
8. **Identify Repeat Signals**: Detect recurring signals across snapshots and classify persistence.
9. **Enforce KRI Requirements**: Apply final quality rules and standardized output structure.

The output is a comprehensive risk signal table, where each row represents one flagged metric for one site or country at one snapshot date. Table 1 shows a representative example, along with the original KRI result for comparison. This structured output can then be shared with users, exported to Excel, or integrated with work-management platforms for triage and follow-up.

Table 1: Example KRI Result and Corresponding Augmented Risk Signal Record

Field	KRI Result (`gsm.core`)	Risk Signal Record (`grail`)
`StudyID`	`STUDY-001`	`STUDY-001`
`SnapshotDate`	`2024-03-15`	`2024-03-15`
`MetricID`	`Analysis_kri0001`	`Analysis_kri0001`
`GroupLevel`	`Site`	`Site`
`GroupID`	`123`	`123`
`Numerator`	`45`	`45`
`Denominator`	`180`	`180`
`Metric`	`0.25`	`0.25`
`Score`	`2.8`	`2.8`
`Flag`	`2`	`2`
`Metric_Abbreviation`	*not present*	`AE`
`Metric_Label`	*not present*	`Adverse Event Rate`
`Numerator_Label`	*not present*	`Adverse Events`
`Denominator_Label`	*not present*	`Days on Study`
`Model`	*not present*	`Normal Approximation`
`Site_Name`	*not present*	`Regional Medical Center`
`Country`	*not present*	`United States`
`Region`	*not present*	`North America`
`Site_Status`	*not present*	`Active`
`Enrolled`	*not present*	`25`
`Action`	*not present*	`Review for data quality issues or protocol compliance`

Field	KRI Result (gsm.core)	Risk Signal Record (grail)
Priority	*not present*	High
Assigned_To	*not present*	Clinical Monitor
Signal_Description	*not present*	Contextual narrative for triage and escalation
Signal_Context	*not present*	Site-specific investigation context
Recommended_Action	*not present*	CRA follow-up and targeted process review
State	*not present*	Awaiting Triage
IsRepeatSignal	*not present*	FALSE

Figure 1: Sample Risk Signal: High Adverse Event Rate at Site Level *Initial Analysis Result (from gsm.core):*

```
StudyID: STUDY-001
SnapshotDate: 2024-03-15
MetricID: Analysis_kri0001
GroupLevel: Site
GroupID: 123
Numerator: 45        # AE count
Denominator: 180     # Days on study
Metric: 0.25         # Rate
Score: 2.8           # Statistical score
Flag: 2              # High flag
```

Enhanced Risk Signal (from grail):

```
[All original fields, plus:]

Metric Metadata:
  - Abbreviation: "AE"
  - Metric_Label: "Adverse Event Rate"
  - Numerator_Label: "Adverse Events"
  - Denominator_Label: "Days on Study"
  - Model: "Normal Approximation"

Group Metadata:
  - Site_Name: "Regional Medical Center"
  - Country: "United States"
  - Region: "North America"
  - Status: "Active"
  - Enrolled: 25

Action Metadata:
  - Action: "Review for data quality issues or protocol compliance"
  - Priority: "High"
  - Assigned_To: "Clinical Monitor"

Signal Context:
  - Signal Description: "AE reporting rate is high relative to other sites which may be a safety risk o
  - Signal Context: "This may be a safety risk or could indicate data quality issues e.g., reporting sy
  - Recommended Action: "<Central Monitor: Update recommended based on initial investigation into the da
```

```
- State: "Awaiting Triage"
- IsRepeatSignal: FALSE
```

3.5 Reporting modules

`{grail}` provides standardized Excel reporting modules for different operational and governance needs. Reports are generated from the same underlying risk signal data, but with different filtering and formatting logic to support specific review contexts:

- **AllRisksAndActionsReport**: Complete listing for documentation and eTMF filing.
- **OpenActionReport**: Focused view of currently active risk signals.
- **CROOpenActionReport**: Active items assigned to CRO partners.
- **ImportantDeviationsReport**: QTL deviation capture and end-of-study assessment support.
- **OngoingQTLReport**: Ongoing QTL status monitoring against thresholds.

These modules support recurring review cycles while preserving traceability and consistency.

4. Results and Discussion

Across deployment to more than 20 studies, the integrated `{gsm}` and `{grail}` model has improved consistency and governance of central monitoring operations by standardizing how statistical findings are translated into triage-ready risk signals. The combination of workflow-driven augmentation, role-relevant metadata, and structured output formats has reduced variation in reviewer interpretation and improved continuity from detection through action. In practice, this has enabled a shift from ad hoc spreadsheet-centric processes toward a more reproducible operating model with clearer lifecycle tracking.

The addition of a **Risk Signal Reporting Layer** has further strengthened the operating model by making signal management observable at scale. Templated slide decks for risk review meetings reduce preparation time and ensure consistent communication across studies. Dashboards tracking signal actionability by metric type provide CM leadership with a portfolio-level view of where signals are driving meaningful action versus noise. Triage timeliness metrics—such as the percentage of signals awaiting triage for more than 90 days—surface process bottlenecks before they compromise oversight quality. Together, these reporting capabilities transform `{grail}` from a signal-processing engine into a foundation for continuous improvement in CM operations.

Future development is expected to focus on deeper integration with enterprise work-management platforms, additional automation of action routing, expanded support for trend-aware prioritization across recurring signals, and richer reporting layer tooling—while preserving flexibility for study-specific risk profiles and governance requirements.

In conclusion, `{grail}` operationalizes statistical monitoring outputs within an end-to-end RBQM workflow. By converting raw analytic flags into contextualized, action-ready risk signals, it enables CM teams to execute risk-based oversight more consistently and transparently. As part of the broader open-source `{gsm}` framework, `{grail}` demonstrates that robust analytics and practical business processes can be integrated to support scalable, documented quality monitoring in clinical development.

Acknowledgments

The authors acknowledge the broader `{gsm}` contributor community and study teams who operationalized and refined these workflows across real-world clinical programs.

References (Draft placeholders)

1. ICH E6(R2) and E6(R3) Good Clinical Practice guidance.
2. FDA guidance related to risk-based monitoring and quality management in clinical trials.
3. EMA reflection papers and guidance relevant to risk-based trial oversight.
4. Internal and open-source documentation for `{gsm}` and `{grail}` workflows.