Qualification Report for the {gsm} R Package

{gsm} v2.2.2

Report Run Date: 2025-02-26

Contents

1 Introduction

Risk-Based Quality Monitoring (RBQM) is a proactive approach to clinical trial monitoring that focuses on identifying and addressing the most critical risks to the integrity of study data and patient safety. This approach aims to ensure that study data are accurate, reliable, and credible while optimizing the use of resources and minimizing the burden on study sites.

The `{gsm}` R package supports RBQM by performing risk assessments primarily focused on detecting differences in quality at the site level. This approach is intended to detect potential issues related to critical data or process(es) across the major risk categories of safety, efficacy, disposition, treatment, and general quality. Each category consists of one or more risk assessment(s). Each risk assessment analyzes the data to flag sites with potential outliers and provides a visualization to help the user understand the issue.

2 Scope

Qualification testing ensures that core functions execute as expected on a system-wide scale. Qualification includes executing various functional, performance, and usability testing. Qualification tests are designed to provide developers with a repeatable process that is easy to update and document. This document summarizes the qualification testing performed on `{gsm}` functions essential to the analysis workflow.

3 Process Overview

Each essential `{gsm}` workflow function is independently qualified using specifications and test cases compiled in this report. Details are provided below.

3.1 Specifications

Specifications capture the most critical use cases for a given function. Each function must have at least one (1) specification, and each specification must have at least one (1) associated test case. Multiple specifications may exist for a function, and multiple test cases may exist for a specification.

Each specification includes the following components:

- **Description:** outlines the use case for the specification
- **Risk Assessment**
  - **Risk Level:** assigned a value of "Low", "Medium", or "High", corresponding to the risk associated with the specification failing
  - **Risk Impact:** assigned a value of "Low", "Medium", or "High", corresponding to the severity of the impact associated with the specification failing
- **Test Cases:** lists measurable test cases associated with the specification

3.2 Test Cases

Test cases translate specifications into testable scripts to confirm that the package functions meet the established requirements. Test cases represent how a user may utilize the function to help identify code gaps and support testing automation.

Function Name	Number of Tests	Number Passed	Number Failed	Number Skipped
Adverse Event Assessment	52	52	0	0
Analysis workflow	22	21	0	0
Analyze_NormalApprox_PredictBounds	19	19	0	0
Data Change Rate Assessment	19	19	0	0
Data Entry Lag Assessment	19	19	0	0
Disposition Assessment	35	35	0	0
Flag_NormalApprox	7	7	0	0
Labs Assessment	19	19	0	0
Mapping workflow	46	46	0	0
Protocol Deviation Assessment	30	30	0	0
Query Age Assessment	35	35	0	0
Query Rate Assessment	35	35	0	0
Summarize	16	16	0	0

Test cases for `{gsm}` are written using the standard `testthat` workflow. A single test script is saved for each test case and is named following the convention `test_qual_{TestID}.R`, where `TestID` is the test case number. Test code within these scripts is written clearly and concisely to facilitate quick execution and interpretability. Note that a single test case may be associated with multiple specifications.

4 Test Results: Overview

5 Test Results: Detailed

5.1 One Row Per Specification

Spec ID	Spec Description	Risk	Impact	Associated Test IDs
S1_1	Given raw participant-level data, all necessary data.frame transformations are made to create input data for all workflows	High	High	T1_1
S2_1	Given raw participant-level data, a properly specified Workflow for a KRI creates summarized and flagged data	High	High	T2_1
S2_2	Given raw participant-level data with missingness, a properly specified Workflow for a KRI creates summarized and flagged data	High	High	T2_2
S3_1	Given pre-processed input data, a properly specified Workflow for a KRI creates summarized and flagged data	High	High	T3_1
S4_1	Given appropriate metadata (i.e. vThresholds), flagged observations are properly marked in summary data	High	High	T4_1
S4_2	Given appropriate metadata (i.e. vThresholds), data.frame of bounds can be created	High	High	T4_2
S5_1	Given appropriate raw participant-level data, flag values can be correctly assigned to records that meet flagging criteria, including custom thresholding.	High	High	T5_1
S5_2	Given appropriate raw participant-level data, flag values are correctly assigned as NA for sites with low enrollment.	High	High	T5_2
S6_1	Given appropriate raw participant-level data, an Adverse Event Assessment can be done using the Normal Approximation method.	High	High	T6_1

S6_2	Adverse Event Assessments can be done correctly using a grouping variable, such as Site or Country for KRIs, and Study for QTLs, when applicable.	High	High	T6_2
S7_1	Given appropriate raw participant-level data, a Protocol Deviation Assessment can be done using the Normal Approximation method.	High	High	T7_1
S7_2	Protocol Deviation Assessments can be done correctly using a grouping variable, such as Site or Country for KRIs, and Study for QTLs, when applicable.	High	High	T7_2
S8_1	Given appropriate raw participant-level data, a Dispositon Assessment can be done using the Normal Approximation method.	High	High	T8_1
S8_2	Disposition Assessments can be done correctly using a grouping variable, such as Site or Country for KRIs, and Study for QTLs, when applicable.	High	High	T8_2
S9_1	Given appropriate raw participant-level data, a Labs Assessment can be done using the Normal Approximation method.	High	High	T9_1
S9_2	Labs Assessments can be done correctly using a grouping variable, such as Site or Country for KRIs, and Study for QTLs, when applicable.	High	High	T9_2
S10_1	Given appropriate raw participant-level data, a Data Change Rate Assessment can be done using the Normal Approximation method.	High	High	T10_1
S10_2	Data Change Rate Assessments can be done correctly using a grouping variable, such as Site, Country, or Study, when applicable.	High	High	T10_2
S11_1	Given appropriate raw participant-level data, a Data Entry Lag Assessment can be done using the Normal Approximation method.	High	High	T11_1
S11_2	Data Entry Lag Assessments can be done correctly using a grouping variable, such as Site, Country, or Study, when applicable.	High	High	T11_2
S12_1	Given appropriate raw participant-level data, a Query Age Assessment can be done using the Normal Approximation method.	High	High	T12_1
S12_2	Query Age Assessments can be done correctly using a grouping variable, such as Site, Country, or Study, when applicable.	High	High	T12_2
S13_1	Given appropriate raw participant-level data, a Query Rate Assessment can be done using the Normal Approximation method.	High	High	T13_1
S13_2	Query Rate Assessments can be done correctly using a grouping variable, such as Site, Country, or Study, when applicable.	High	High	T13_2

5.2 One Row Per Test

Function	Spec ID	Test ID	Test Description	Test Result
Mapping workflow	S1_1	T1_1	mappings now done by individual domain, test that inputs and outputs of priority 1 mappings are completed as expected	Pass
Mapping workflow	S1_1	T1_1	mappings now done by individual domain, test that inputs and outputs of priority 2 mappings are completed as expected	Pass
Mapping workflow	S1_1	T1_1	mappings now done by individual domain, test that inputs and outputs of priority 3 mappings are completed as expected	Pass
Analysis workflow	S2_1	T2_1	Given raw participant-level data, a properly specified Workflow for a KRI creates summarized and flagged data	Pass
Analysis workflow	S2_2	T2_2	Given raw participant-level data with missingness, a properly specified Workflow for a KRI creates summarized and flagged data	Pass
Analysis workflow	S3_1	T3_1	Given pre-processed input data, a properly specified Workflow for a KRI creates summarized and flagged data	Pass
Flag_NormalApprox	S4_1	T4_1	Given appropriate metadata (i.e. vThresholds), flagged observations are properly marked in summary data	Pass
Analyze_NormalApprox_RelDiffBounds	S4_2	T4_2	Given appropriate metadata (i.e. vThresholds), bounds are properly applied to generate flags	Pass
Summarize	S5_1	T5_1	Given appropriate raw participant-level data, flag values can be correctly assigned to records that meet flagging criteria, including custom thresholding.	Pass
Summarize	S5_2	T5_2	Given appropriate raw participant-level data, flag values are correctly assigned as NA for sites with low enrollment.	Pass
Adverse Event Assessment	S6_1	T6_1	Given appropriate raw participant-level data, an Adverse Event Assessment can be done using the Normal Approximation method.	Pass
Adverse Event Assessment	S6_2	T6_2	Adverse Event Assessments can be done correctly using a grouping variable, such as Site or Country for KRIs, and Study for QTLs, when applicable.	Pass
Protocol Deviation Assessment	S7_1	T7_1	Given appropriate raw participant-level data, a Protocol Deviation Assessment can be done using the Normal Approximation method.	Pass
Protocol Deviation Assessment	S7_2	T7_2	Protocol Deviation Assessments can be done correctly using a grouping variable, such as Site or Country for KRIs, and Study for QTLs, when applicable.	Pass
Disposition Assessment	S8_1	T8_1	Given appropriate raw participant-level data, a Dispositon Assessment can be done using the Normal Approximation method.	Pass
Disposition Assessment	S8_2	T8_2	Disposition Assessments can be done correctly using a grouping variable, such as Site or Country for KRIs, and Study for QTLs, when applicable.	Pass
Labs Assessment	S9_1	T9_1	Given appropriate raw participant-level data, a Labs Assessment can be done using the Normal Approximation method.	Pass

Labs Assessment	S9_2	T9_2	Labs Assessments can be done correctly using a grouping variable, such as Site or Country for KRIs, and Study for QTLs, when applicable.	Pass
Data Change Rate Assessment	S10_1	T10_1	Given appropriate raw participant-level data, a Data Change Rate Assessment can be done using the Normal Approximation method.	Pass
Data Change Rate Assessment	S10_2	T10_2	Data Change Rate Assessments can be done correctly using a grouping variable, such as Site, Country, or Study, when applicable.	Pass
Data Entry Lag Assessment	S11_1	T11_1	Given appropriate raw participant-level data, a Data Entry Lag Assessment can be done using the Normal Approximation method.	Pass
Data Entry Lag Assessment	S11_2	T11_2	Data Entry Lag Assessments can be done correctly using a grouping variable, such as Site, Country, or Study, when applicable.	Pass
Query Age Assessment	S12_1	T12_1	Given appropriate raw participant-level data, a Query Age Assessment can be done using the Normal Approximation method.	Pass
Query Age Assessment	S12_2	T12_2	Query Age Assessments can be done correctly using a grouping variable, such as Site, Country, or Study, when applicable.	Pass
Query Rate Assessment	S13_1	T13_1	Given appropriate raw participant-level data, a Query Rate Assessment can be done using the Normal Approximation method.	Pass
Query Rate Assessment	S13_2	T13_2	Query Rate Assessments can be done correctly using a grouping variable, such as Site, Country, or Study, when applicable.	Pass

6 Qualification Testing Environment

6.1 Session Information

R version 4.4.1 (2024-06-14 ucrt)

Platform: x86_64-w64-mingw32/x64

locale: *LC_COLLATE=English_Europe.utf8*, *LC_CTYPE=English_Europe.utf8*, *LC_MONETARY=English_Europe.utf8* *LC_NUMERIC=C* and *LC_TIME=English_Europe.utf8*

attached base packages: *tcltk*, *stats*, *graphics*, *grDevices*, *utils*, *datasets*, *methods* and *base*

other attached packages: *gsm.qc(v.0.0.1)*, *riskmetric(v.0.2.4)*, *gh(v.1.4.1)*, *pander(v.0.6.5)*, *knitr(v.1.49)*, *gt(v.0.11.1.9000)*, *stringr(v.1.5.1)*, *yaml(v.2.3.10)*, *glue(v.1.8.0)*, *cli(v.3.6.4)*, *purrr(v.1.0.4)*, *dplyr(v.1.1.4)*, *gsm.reporting(v.0.0.1)*, *gsm.kri(v.0.0.1)*, *gsm.mapping(v.0.0.1)*, *gsm(v.2.2.2)* and *testthat(v.3.2.1.1)*

loaded via a namespace (and not attached): *tidyselect(v.1.2.1)*, *urltools(v.1.7.3)*, *fastmap(v.1.2.0)*, *lazyeval(v.0.2.2)*, *duckdb(v.1.0.0-2)*, *promises(v.1.3.0)*, *rex(v.1.2.1)*, *digest(v.0.6.37)*, *mime(v.0.12)*, *lifecycle(v.1.0.4)*, *waldo(v.0.5.2)*, *ellipsis(v.0.3.2)*, *magrittr(v.2.0.3)*, *compiler(v.4.4.1)*, *rlang(v.1.1.5)*, *tools(v.4.4.1)*, *htmlwidgets(v.1.6.4)*, *curl(v.6.2.0)*, *pkgbuild(v.1.4.4)*, *here(v.1.0.1)*, *xml2(v.1.3.6)*, *cranlogs(v.2.1.1)*, *pkgload(v.1.4.0)*, *miniUI(v.0.1.1.1)*, *covr(v.3.6.4)*, *withr(v.3.0.2)*, *desc(v.1.4.3)*, *triebeard(v.0.4.1)*, *grid(v.4.4.1)*, *urlchecker(v.1.0.1)*, *profvis(v.0.3.8)*, *xtable(v.1.8-4)*, *colorspace(v.2.1-1)*, *log4r(v.0.4.4)*, *ggplot2(v.3.5.1)*, *scales(v.1.3.0)*, *tinytex(v.0.54)*, *rmarkdown(v.2.29)*, *generics(v.0.1.3)*, *remotes(v.2.5.0)*, *rstudioapi(v.0.17.1)*, *httr(v.1.4.7)*, *sessioninfo(v.1.2.2)*, *DBI(v.1.2.3)*, *cachem(v.1.1.0)*, *BiocManager(v.1.30.23)*, *vctrs(v.0.6.5)*, *devtools(v.2.4.5)*, *jsonlite(v.1.8.9)*, *tidyr(v.1.3.1)*, *clindata(v.1.0.5)*, *stringi(v.1.8.4)*, *gtable(v.0.3.6)*, *later(v.1.3.2)*, *munsell(v.0.5.1)*, *tibble(v.3.2.1)*, *pillar(v.1.10.1)*, *htmltools(v.0.5.8.1)*, *brio(v.1.1.5)*, *R6(v.2.6.1)*, *dbplyr(v.2.5.0)*, *rprojroot(v.2.0.4)*, *shiny(v.1.9.1)*, *evaluate(v.1.0.3)*, *backports(v.1.5.0)*, *memoise(v.2.0.1)*, *broom(v.1.0.7)*, *httpuv(v.1.6.15)*, *Rcpp(v.1.0.14)*, *xfun(v.0.50)*, *fs(v.1.6.5)*, *usethis(v.3.1.0)* and *pkgconfig(v.2.0.3)*

7 Pull Request History