# Package 'FragiliTidy'

June 25, 2026

**Title** Tidyverse-Compatible Fragility Index Calculations

**Version** 0.1.0

**Description** Provides optimized, Tidyverse-compatible functions for calculating the Fragility Index and Reverse Fragility Index for 2x2 contingency tables from clinical trials. Uses customized hypergeometric and algebraic calculations along with binary search algorithms to achieve substantial speedups over standard implementations, with seamless integration into 'dplyr' pipelines.

**License** GPL-3

**Encoding** UTF-8

**RoxygenNote** 7.3.1

**Imports** dplyr, purrr, rlang, tibble, stats

**Suggests** testthat (>= 3.0.0), knitr, rmarkdown

**VignetteBuilder** knitr

**URL** https://github.com/tomdrake/fragilitidy

**BugReports** https://github.com/tomdrake/fragilitidy/issues

**NeedsCompilation** no

**Author** Tom Drake [aut, cre]

**Maintainer** Tom Drake <t.drake@ed.ac.uk>

**Repository** CRAN

**Date/Publication** 2026-06-25 15:50:21 UTC

## Contents

---

`continuous_fragility` *Continuous Fragility Index*

---

## Description

Implements the Continuous Fragility Index (CFI) of Caldwell, Youssefzadeh and Limpisvasti (J Clin Epidemiol 2021;136:20-25) for two-arm trials with a continuous outcome compared via Welch's t-test.

## Details

The CFI is the minimum number of substitution iterations required to drive a significant Welch t-test result to non-significance, where on each iteration the data point in the higher-mean group that lies closest to but still above that group's mean is moved into the lower-mean group.

When raw data are unavailable, datasets matching the supplied summary statistics are generated by rejection sampling and the procedure is repeated `n_sim` times; the mean CFI across simulations is returned.

---

`continuous_fragility_index`
*Continuous Fragility Index for a Data Frame*

---

## Description

Adds a Continuous Fragility Index column to a data frame of trial summary statistics. Supports tidy evaluation.

## Usage

```
continuous_fragility_index(
  data,
  mean1,
  sd1,
  n1,
  mean2,
  sd2,
```

```
  n2,
  conf.level = 0.95,
  n_sim = 5L,
  tol_mean = 0.01,
  tol_sd = 0.01,
  col_name = "continuous_fragility_index"
)
```

## Arguments

| | |
|---|---|
| `data` | A data frame or tibble. |
| `mean1`, `sd1`, `n1` | Unquoted column names for arm 1 mean, SD, and sample size. |
| `mean2`, `sd2`, `n2` | Unquoted column names for arm 2 mean, SD, and sample size. |
| `conf.level`, `n_sim`, `tol_mean`, `tol_sd` | See `continuous_fragility_index_summary()`. |
| `col_name` | Name of the output column (default `"continuous_fragility_index"`). |

## Value

The input data frame with an additional CFI column.

---

`continuous_fragility_index_raw` *Continuous Fragility Index from raw outcome vectors*

---

## Description

Direct (non-simulated) CFI when raw per-patient outcomes are available.

## Usage

```
continuous_fragility_index_raw(x, y, conf.level = 0.95)
```

## Arguments

| | |
|---|---|
| `x`, `y` | Numeric vectors of outcome values for arms 1 and 2. |
| `conf.level` | Confidence level for the Welch t-test (default `0.95`). |

## Value

A single integer: the number of substitution iterations required to lose significance, or `0` if the baseline test was non-significant.

## Examples

```
set.seed(1)
x <- rnorm(50, 70, 10)
y <- rnorm(50, 50, 10)
continuous_fragility_index_raw(x, y)
```

`continuous_fragility_index_summary` *Continuous Fragility Index from summary statistics*

## Description

Calculates the Continuous Fragility Index (Caldwell et al. 2021) from two-arm summary statistics by simulating compatible datasets and applying the iterative substitution algorithm.

## Usage

```
continuous_fragility_index_summary(
  mean1,
  sd1,
  n1,
  mean2,
  sd2,
  n2,
  conf.level = 0.95,
  n_sim = 5L,
  tol_mean = 0.01,
  tol_sd = 0.01,
  seed = NULL
)
```

## Arguments

| | |
|---|---|
| `mean1`, `sd1`, `n1` | Mean, standard deviation, and sample size of arm 1. |
| `mean2`, `sd2`, `n2` | Mean, standard deviation, and sample size of arm 2. |
| `conf.level` | Confidence level for the Welch t-test (default `0.95`). |
| `n_sim` | Number of simulated datasets to average over (default `5`, matching Caldwell et al.). |
| `tol_mean`, `tol_sd` | Relative tolerances for rejection sampling. |
| `seed` | Optional integer seed for reproducibility. |

## Value

A single numeric value: the mean CFI across `n_sim` simulations. Returns `0` if the baseline Welch test is already non-significant and `NA_real_` if any input is missing.

## References

Caldwell JE, Youssefzadeh K, Limpisvasti O. A method for calculating the fragility index of continuous outcomes. *J Clin Epidemiol* 2021;136:20-25.

## Examples

```
continuous_fragility_index_summary(
  mean1 = 70, sd1 = 10, n1 = 100,
  mean2 = 50, sd2 = 10, n2 = 100,
  seed  = 1
)
```

---

`continuous_fragility_index_vec` *Vectorised Continuous Fragility Index*

---

## Description

Vectorised wrapper around `continuous_fragility_index_summary()` for use inside `dplyr::mutate()`.

## Usage

```
continuous_fragility_index_vec(
  mean1,
  sd1,
  n1,
  mean2,
  sd2,
  n2,
  conf.level = 0.95,
  n_sim = 5L,
  tol_mean = 0.01,
  tol_sd = 0.01
)
```

## Arguments

`mean1`, `sd1`, `n1`, `mean2`, `sd2`, `n2`
: Numeric vectors of summary statistics.

`conf.level`, `n_sim`, `tol_mean`, `tol_sd`
: See `continuous_fragility_index_summary()`.

## Value

A numeric vector of CFI values.

`fragility_index` *Fragility Index for a Data Frame*

## Description

Computes the fragility index for columns in a data frame. Supports tidy evaluation and integrates with `%>%` or `|>`.

## Usage

```
fragility_index(
  data,
  intervention_event,
  control_event,
  intervention_n,
  control_n,
  conf.level = 0.95,
  verbose = FALSE,
  col_name = "fragility_index"
)
```

## Arguments

| | |
|---|---|
| `data` | A data frame or tibble. |
| `intervention_event` | Column name (unquoted) for the intervention events. |
| `control_event` | Column name (unquoted) for the control events. |
| `intervention_n` | Column name (unquoted) for the intervention group totals. |
| `control_n` | Column name (unquoted) for the control group totals. |
| `conf.level` | Confidence level (default 0.95). Can be a number or a column name. |
| `verbose` | Logical; if TRUE, returns a nested list-column with p-values for each iteration. |
| `col_name` | Name of the output column. Default is `"fragility_index"`. |

## Value

The original data frame with an added column for the fragility index.

## fragility_index_vec *Vectorised Fragility Index Calculation*

### Description

Calculates the fragility index for vector inputs. This is useful for running inside `dplyr::mutate()`.

### Usage

```
fragility_index_vec(
  intervention_event,
  control_event,
  intervention_n,
  control_n,
  conf.level = 0.95,
  verbose = FALSE
)
```

### Arguments

| | |
|---|---|
| `intervention_event` | Vector of events in the intervention group. |
| `control_event` | Vector of events in the control group. |
| `intervention_n` | Vector of total patients in the intervention group. |
| `control_n` | Vector of total patients in the control group. |
| `conf.level` | Significance level / confidence level (default 0.95). |
| `verbose` | Logical indicating if full progression of p-values should be returned. |

### Value

A numeric vector of fragility indices (if `verbose = FALSE`), or a list of tibbles containing step-by-step p-values (if `verbose = TRUE`).

## reverse_continuous_fragility_index *Reverse Continuous Fragility Index for a Data Frame*

### Description

Adds a reverse Continuous Fragility Index column to a data frame of trial summary statistics.

## Usage

```
reverse_continuous_fragility_index(
  data,
  mean1,
  sd1,
  n1,
  mean2,
  sd2,
  n2,
  conf.level = 0.95,
  n_sim = 5L,
  tol_mean = 0.01,
  tol_sd = 0.01,
  max_iter = 10000L,
  col_name = "reverse_continuous_fragility_index"
)
```

## Arguments

| | |
|---|---|
| `data` | A data frame or tibble. |
| `mean1`, `sd1`, `n1` | Unquoted column names for arm 1 summary stats. |
| `mean2`, `sd2`, `n2` | Unquoted column names for arm 2 summary stats. |
| `conf.level`, `n_sim`, `tol_mean`, `tol_sd`, `max_iter` | See `reverse_continuous_fragility_index_summary()`. |
| `col_name` | Output column name (default `"reverse_continuous_fragility_index"`). |

## Value

The input data frame with an additional reverse CFI column.

---

`reverse_continuous_fragility_index_summary` *Reverse Continuous Fragility Index*

---

## Description

Estimates how many additional participants per arm would have been required to render a non-significant Welch t-test significant, given two-arm summary statistics. This is a continuous-outcome analogue of the reverse fragility index for dichotomous outcomes: a measure of how far a non-significant trial was from significance, expressed in participants per arm.

## Usage

```
reverse_continuous_fragility_index_summary(
  mean1,
  sd1,
  n1,
  mean2,
  sd2,
  n2,
  conf.level = 0.95,
  n_sim = 5L,
  tol_mean = 0.01,
  tol_sd = 0.01,
  max_iter = 10000L,
  seed = NULL
)
```

## Arguments

| | |
|---|---|
| mean1, sd1, n1 | Mean, standard deviation, and sample size of arm 1. |
| mean2, sd2, n2 | Mean, standard deviation, and sample size of arm 2. |
| conf.level | Confidence level for the Welch t-test (default 0.95). |
| n_sim | Number of simulated datasets to average over (default 5). |
| tol_mean, tol_sd | Relative tolerances for rejection sampling. |
| max_iter | Maximum additional participants per arm before giving up and returning NA_real_ (default 10000). |
| seed | Optional integer seed for reproducibility. |

## Details

If the original test is already significant the function returns 0. Otherwise additional participants are sampled from each arm's assumed normal distribution (parameterised by the supplied mean and SD) and added one per arm per iteration until significance is reached. The procedure is repeated n_sim times and the mean is returned.

## Value

A single numeric value: mean additional participants per arm required to reach significance across n_sim simulations. Returns 0 if the original test was already significant.

## Examples

```
reverse_continuous_fragility_index_summary(
  mean1 = 55, sd1 = 10, n1 = 30,
  mean2 = 50, sd2 = 10, n2 = 30,
  seed  = 1
)
```

## reverse_continuous_fragility_index_vec *Vectorised Reverse Continuous Fragility Index*

### Description

Vectorised wrapper around `reverse_continuous_fragility_index_summary()` for use inside `dplyr::mutate()`.

### Usage

```
reverse_continuous_fragility_index_vec(
  mean1,
  sd1,
  n1,
  mean2,
  sd2,
  n2,
  conf.level = 0.95,
  n_sim = 5L,
  tol_mean = 0.01,
  tol_sd = 0.01,
  max_iter = 10000L
)
```

### Arguments

`mean1`, `sd1`, `n1`, `mean2`, `sd2`, `n2`
: Numeric vectors of summary statistics.

`conf.level`, `n_sim`, `tol_mean`, `tol_sd`, `max_iter`
: See `reverse_continuous_fragility_index_summary()`.

### Value

A numeric vector of reverse CFI values.

## revfragility_index *Reverse Fragility Index for a Data Frame*

### Description

Computes the reverse fragility index for columns in a data frame. Supports tidy evaluation and integrates with `%>%` or `|>`.

### Usage

```
revfragility_index(
  data,
  intervention_event,
  control_event,
  intervention_n,
  control_n,
  conf.level = 0.95,
  verbose = FALSE,
  col_name = "revfragility_index",
  compatibility_mode = FALSE
)
```

### Arguments

| | |
|---|---|
| `data` | A data frame or tibble. |
| `intervention_event` | Column name (unquoted) for the intervention events. |
| `control_event` | Column name (unquoted) for the control events. |
| `intervention_n` | Column name (unquoted) for the intervention group totals. |
| `control_n` | Column name (unquoted) for the control group totals. |
| `conf.level` | Confidence level (default 0.95). Can be a number or a column name. |
| `verbose` | Logical; if TRUE, returns a nested list-column with p-values for each iteration. |
| `col_name` | Name of the output column. Default is `"revfragility_index"`. |
| `compatibility_mode` | If TRUE, reproduces the original package's bug in verbose mode. |

### Value

The original data frame with an added column for the reverse fragility index.

---

## `revfragility_index_vec` *Vectorised Reverse Fragility Index Calculation*

---

### Description

Calculates the reverse fragility index for vector inputs. This is useful for running inside `dplyr::mutate()`.

## Usage

```
revfragility_index_vec(
  intervention_event,
  control_event,
  intervention_n,
  control_n,
  conf.level = 0.95,
  verbose = FALSE,
  compatibility_mode = FALSE
)
```

## Arguments

| | |
|---|---|
| intervention_event | Vector of events in the intervention group. |
| control_event | Vector of events in the control group. |
| intervention_n | Vector of total patients in the intervention group. |
| control_n | Vector of total patients in the control group. |
| conf.level | Significance level / confidence level (default 0.95). |
| verbose | Logical indicating if full progression of p-values should be returned. |
| compatibility_mode | If TRUE, reproduces the original package's bug in verbose mode. |

## Value

A numeric vector of reverse fragility indices (if `verbose = FALSE`), or a list of tibbles containing step-by-step p-values (if `verbose = TRUE`).

---

# tidyverse_fragility *Tidyverse-Compatible Fragility Index Functions (Binary Search Optimized)*

---

## Description

This file provides optimized, tidyverse-compatible functions for calculating the Fragility Index and the Reverse Fragility Index. It uses customized 2x2 hypergeometric and algebraic calculations to achieve a 25x speedup compared to standard stats package functions, and binary search algorithms to yield an additional 10x-1000x speedup for large trials.

# Index