Package 'hybridEHR'

May 8, 2026

Type Package

Title Synthetic Hybrid Electronic Health Record Generation for SARS-Related Research and CT Views

Version 0.2.0

Description Generates synthetic electronic health record data, including patients, encounters, vitals, laboratory results, medications, procedures, and allergies. The package supports optional SARS-focused and computed tomography (CT) research views and export to CSV, SQLite, and Excel formats for research and development workflows.

License MIT + file LICENSE

Encoding UTF-8

Depends R (>= 4.1.0)

Imports dplyr, tidyr, tibble, lubridate, magrittr, jsonlite, openxlsx, DBI, RSQLite, rlang

Suggests knitr, rmarkdown

RoxygenNote 7.3.3

VignetteBuilder knitr

NeedsCompilation no

Author Dennis Opoku Boadu [aut, cre], Isaac Osei [aut], Justice Kwame Appati [aut]

Maintainer Dennis Opoku Boadu <doboadu@st.ug.edu.gh>

Repository CRAN

Date/Publication 2026-04-06 11:00:02 UTC

Contents

`generate_hybrid_ehr` *High-level wrapper to generate and export a hybrid EHR dataset*

Description

High-level wrapper to generate and export a hybrid EHR dataset

Usage

```
generate_hybrid_ehr(
  n_patients = 50000,
  n_sites = 10,
  SARS_focused = TRUE,
  include_ct_links = FALSE,
  output_dir = "hybrid_ehr_dataset",
  seed = NULL,
  verbose = TRUE
)
```

Arguments

`n_patients`	Number of unique patients.
`n_sites`	Number of sites/hospitals to simulate.
`SARS_focused`	Logical; if TRUE, use SARS-era encounter and lab patterns.
`include_ct_links`	Logical; if TRUE, add CT timing variables and a CT severity score in the CT research view.
`output_dir`	Directory to write export files into.
`seed`	Optional integer used to set the random seed for reproducibility.
`verbose`	Logical; if TRUE, print progress messages to the console.

Value

A list with:

dataset The in-memory dataset list (as from `generate_hybrid_ehr_dataset`).

output_dir The output directory path where files were written.

Examples

```
## Not run:
  # Quick SARS CT research dataset
  res <- generate_hybrid_ehr(
    n_patients = 10000,
    n_sites = 5,
    SARS_focused = TRUE,
```

```
  include_ct_links = TRUE,
  output_dir = "SARS_ct_research_dataset",
  seed = 42
)

# General EHR dataset
res2 <- generate_hybrid_ehr(
  n_patients = 50000,
  n_sites = 10,
  SARS_focused = FALSE,
  include_ct_links = FALSE,
  output_dir = "general_ehr_dataset",
  seed = 123
)

## End(Not run)
```

`generate_hybrid_ehr_dataset` *Generate synthetic hybrid EHR tables*

Description

Generate synthetic hybrid EHR tables

Usage

```
generate_hybrid_ehr_dataset(
  n_patients = 1e+05,
  n_sites = 6,
  SARS_focused = TRUE,
  include_ct_links = FALSE,
  seed = NULL,
  verbose = TRUE
)
```

Arguments

`n_patients`	Number of unique patients.
`n_sites`	Number of sites/hospitals to simulate.
`SARS_focused`	Logical; if TRUE, use SARS-era encounter and lab patterns.
`include_ct_links`	Logical; if TRUE, add CT timing variables and a CT severity score in the CT research view.
`seed`	Optional integer used to set the random seed for reproducibility.
`verbose`	Logical; if TRUE, print progress messages to the console.

Value

A list with elements:

tables Named list of core EHR tables (patients, encounters, vitals, labs, medications, procedures, allergies).

research Named list with `ct_research_view` (if SARS_focused) and `ml_flat_view` (aggregated ML-ready table).

metadata List of high-level generation settings and table metadata.

Index