Package 'CareDensity'

January 20, 2025

Title Calculate the Care Density or Fragmented Care Density Given a
Patient-Sharing Network

Version 0.1.0

Maintainer Robin Denz <robin.denz@rub.de>

Description Given a patient-sharing network, calculate either the classic care density as
proposed by Pollack et al. (2013) <doi:10.1007/s11606-012-2104-7> or the
fragmented care density as proposed by Engels et al. (2024) <doi:10.1186/s12874-023-02106-0>.
By utilizing the 'igraph' and 'data.table' packages, the provided functions scale well for
very large graphs.

License GPL (>= 3)

URL https://github.com/RobinDenz1/CareDensity,

https://robindenz1.github.io/CareDensity/

BugReports https://github.com/RobinDenz1/CareDensity/issues

Depends MatrixExtra

Imports data.table, igraph (>= 2.0.0), utils

Suggests knitr, rmarkdown, testthat (>= 3.2.1), covr

VignetteBuilder knitr

Config/testthat/edition 3

Contact <robin.denz@rub.de>

Encoding UTF-8

RoxygenNote 7.3.1

NeedsCompilation no

Author Robin Denz [aut, cre]

Repository CRAN

Date/Publication 2024-11-18 13:50:11 UTC

Contents

---

care_density	*Calculate the Care Density for all Patients*

Description

This function calculates the classic Care Density Index as defined by Pollack et al. (2013) for each patient in the supplied dataset. Works well with large patient-sharing networks.

Usage

```
care_density(data, pat_col=1, data_frame=TRUE)
```

Arguments

data	A `data.frame` like object containing exactly two columns. One should include only patient IDs and the other one only provider IDs. Each row should denote one patient-provider contact. Multiple contacts (same rows) are allowed but have no effect on the outcome. Both patient and provider IDs should be unique, which means that one ID may not be in both rows.
pat_col	Specifies which column of `data` includes the patient IDs. If the first column contains the patient IDs this should be kept at 1, if the second column contains the patient IDs it should be set to 2.
data_frame	Set this argument to `TRUE` to return a `data.frame` instead of the `data.table` format that is used under the hood.

Details

The Care Density (C_p) is "a patient-level measure that quantifies the amount of patient-sharing among his or her providers" (DuGoff et al. 2018). Higher care densities have been posited to reflect greater connections among a patients "care team". Formally, it is defined as:

$$C_p = \frac{\sum_{i=1}^{m} w_{p,j}}{n_p(n_p - 1)/2}$$

with n_p being the number of providers a patient has visited, m defined as the number of all possible combinations of length two and $w_{p,j}$ being the number of patients that a pair of provider is sharing. An example is given below and explained more thoroughly in the vignette of this package.

Under the hood, this function uses the `igraph` package to construct a patient-sharing network from the provided `data` to calculate the weights. It then uses the `data.table` package to efficiently calculate the care densities from a resulting edge list with weights.

Value

Returns a single `data.frame` (or `data.table`) containing the sum of all weights (`"sum_weights"`), the number of providers seen by each patient (`"n"`) and the calculated Care Density (`"care_density"`).

Author(s)

Robin Denz

References

Pollack, Craig Evan, Gary E. Weissman, Klaus W. Lemke, Peter S. Hussey, and Jonathan P. Weiner. (2013). "Patient Sharing Among Physicians and Costs of Care: A Network Analytic Approach to Care Coordination Using Claims Data". Journal of General Internal Medicine 28 (3), pp. 459-465.

DuGoff, Eva H., Sara Fernandes-Taylor, Gary E. Weissman, Joseph H. Huntley, and Craig Evan Pollack. (2018). "A Scoping Review of Patient-Sharing Network Studies Using Administrative Data". Translational Behavioral Medicine 8 (4), pp. 598-625.

See Also

fragmented_care_density

Examples

```
library(CareDensity)
library(data.table)
library(igraph)

# some arbitrary patient-provider contact data
data <- data.frame(PatID=c("1", "1", "1", "2", "2", "3", "3", "4", "5"),
                   ArztID=c("A", "C", "D", "A", "D", "A", "D", "D", "C"))

# calculate the care densities
care_density(data)
```

---

fragmented_care_density

Calculate the Fragmented Care Density for all Patients

Description

This function calculates the Fragmented Care Density Index as defined by Engels et al. (2024) for each patient in the supplied dataset. Works well with large patient-sharing networks.

Usage

```
fragmented_care_density(data, pat_col=1, weights, type,
                        by_connection=FALSE, data_frame=TRUE)
```

Arguments

data	A `data.frame` like object containing exactly two columns. One should include only patient IDs and the other one only provider IDs. Each row should denote one patient-provider contact. Multiple contacts (same rows) are allowed but have no effect on the outcome. Both patient and provider IDs should be unique, which means that one ID may not be in both rows.
pat_col	Specifies which column of `data` includes the patient IDs. If the first column contains the patient IDs this should be kept at 1, if the second column contains the patient IDs it should be set to 2.
weights	A `data.frame` containing three columns called `"from"` (character), `"to"` (character) and `"weight"` (numeric). The first two columns should contain types of providers, thus defining different provider connections. All possible non-redundant connections need to be specified this way. The `"weight"` column should include the weight associated with that connection. When using `by_connection=TRUE` this argument can be set to `NULL`, because it won't be needed then. See examples for more information.
type	A `data.frame` containing two columns called `"ID"` (containing all provider IDs) and `"Type"` (containing the type of the provider). Both columns should be character vectors.
by_connection	Either `TRUE` or `FALSE` (default). If `TRUE` this function returns the person and connection-specific sums of weights and simple care densities instead of returning the fragmented care density directly. This may be useful to estimate weights for the `weights` argument.
data_frame	Set this argument to `TRUE` to return a `data.frame` instead of the `data.table` format that is used under the hood.

Details

The Fragmented Care Density is an extension of the classic Care Density (see [`care_density`](#)) and was proposed by Engels et al. (2024). It is also a measure of care coordination, but it allows a lot more flexibility by using different weights for different provider-type connections. For example, it may make sense to weight the amount of patients shared by two general providers differently than the amount of patients shared by a general provider and a specialist. Formally, the fragmented care density is defined as:

$$FC_p = \sum_{j=1}^{k} w_j \frac{s_j}{n_p(n_p - 1)/2},$$

where n_p is the number of different providers patient p visited and w_j are some connection specific weights. k is defined as:

$$k = \binom{l}{2} + l,$$

with l being the number of different provider types. Finally, s_j is the sum of the number of patients shared by all doctors of a specific connection type. See Engels et al. (2024) for more information.

Under the hood, this function uses the igraph package to construct a patient-sharing network from the provided data to calculate the weights. It then uses the data.table package to efficiently calculate the care densities from a resulting edge list with weights.

Value

Returns a single data.frame (or data.table) containing output depending on the specification of the by_connection argument.

When by_connection=FALSE was used the output only includes the patient id ("PatID") and the calculated fragmented care densities ("fragmented_care_density").

When by_connection=TRUE was used instead, the output includes the patient id ("PatID"), the connection-type ("connection") the sum of all weights ("sum_weights"), the number of providers seen by each patient ("n") and the calculated simple care density ("care_density").

Author(s)

Robin Denz

References

Pollack, Craig Evan, Gary E. Weissman, Klaus W. Lemke, Peter S. Hussey, and Jonathan P. Weiner. (2013). "Patient Sharing Among Physicians and Costs of Care: A Network Analytic Approach to Care Coordination Using Claims Data". Journal of General Internal Medicine 28 (3), pp. 459-465.

Engels, Alexander, Claudia Konnopka, Espen Henken, Martin Härter, and Hans-Helmut König. (2024). "A Flexible Approach to Measure Care Coordination Based on Patient-Sharing Networks". BMC Medical Research Methodology 24 (1), pp. 1-12.

See Also

care_density

Examples

```
library(CareDensity)
library(data.table)
library(igraph)

# some arbitrary patient-provider contact data
data <- data.frame(PatID=c("1", "1", "1", "2", "2", "3", "3", "4", "5"),
                   ArztID=c("A", "C", "D", "A", "D", "A", "D", "D", "C"))

# defining the provider types
d_type <- data.frame(ID=c("A", "C", "D"),
                     Type=c("GP", "GP", "Psychiatrist"))

# defining the connection-specific weights
d_weights <- data.frame(from=c("GP", "GP", "Psychiatrist"),
                        to=c("GP", "Psychiatrist", "Psychiatrist"),
                        weight=c(1.1, 0.8, 1.3))
```

```
# calculate the fragmented care densities
fragmented_care_density(data, type=d_type, weights=d_weights)

# calculate only the connection-specific sums and care-densities per patient
# NOTE: "weights" can be set to NULL here because they won't be used
fragmented_care_density(data, type=d_type, weights=NULL, by_connection=TRUE)
```

---

project_to_one_mode   *Project a Bipartite-Network to a Single Mode*

Description

This function takes a bipartite network created using the igraph package and returns an adjacency matrix of one of its underlying modes. By directly allowing sparse matrices it is faster and more RAM efficient than some other available versions.

Usage

```
project_to_one_mode(g, mode, sparse=TRUE)
```

Arguments

g	An igraph object with two different node types, defining a bipartite graph.
mode	Either "rows" or "cols", defining which mode should be calculated.
sparse	Whether to use sparse matrix representations or not. Must be either TRUE or FALSE.

Details

A bipartite graph only has connections between two types of nodes. For example, one type of node may be the providers and the other type may be patients. In health-care databases we would usually see which patient visited which providers, but we would not see any direct links between providers and patients. This type of graph can be projected into two subgraphs. One which only includes the providers and one which only includes the patients. This function efficiently creates adjacency matrices of these projections.

The resulting adjacency matrix is a symmetric square matrix that contains the number of shared patients (in provider level graphs) or the number of shared providers (in the patient level graph). See the examples below.

Value

Returns a sparse adjacency matrix of the specified mode.

Author(s)

Robin Denz

References

Landon, Bruce E., Nancy L. Keating, Michael L. Barnett, Jukka-Pekka Onnela, Sudeshna Paul, A. James O'Malley, Thomas Keegan, and Nicholas A. Christakis. (2012). "Variation in Patient-Sharing Networks of Physicians Across the United States". JAMA 308 (3): 265–73.

See Also

project_to_one_mode

Examples

```
library(CareDensity)
library(igraph)

# some arbitrary patient-provider contact data
data <- data.frame(PatID=c("1", "1", "1", "2", "2", "3", "3", "4", "5"),
                   ArztID=c("A", "C", "D", "A", "D", "A", "D", "D", "C"))

# create graph
g <- graph_from_data_frame(data, directed=FALSE)

# add type
V(g)$type <- bipartite_mapping(g)$type

## NOTE: we use sparse=FALSE here to show the resulting matrix directly,
#        but it would be more efficient to keep it at sparse=TRUE
# project to patient-level
project_to_one_mode(g, mode="rows", sparse=FALSE)

# project to provider-level
project_to_one_mode(g, mode="cols", sparse=FALSE)
```

Index