

# Package ‘rclinicaltrials’

January 9, 2017

**Type** Package

**Title** Download Aggregate Trial Information and Results from  
ClinicalTrials.gov

**Description** ClinicalTrials.gov is a registry and results database of publicly and privately supported clinical studies of human participants conducted around the world (see <<https://clinicaltrials.gov/>> for more information). Users can search for information about and results from those trials. This provides a set of functions to interact with the search and download features. Results are downloaded to temporary directories and returned as R objects.

**Version** 1.4.7

**Date** 2017-01-09

**Author** Michael C Sachs <sachsmc@gmail.com>

**Maintainer** Michael C Sachs <sachsmc@gmail.com>

**Depends** R (>= 3.0.0)

**License** MIT + file LICENSE

**Imports** httr, XML, plyr

**Suggests** rmarkdown, devtools, knitr, roxygen2, testthat, dplyr,  
ggplot2

**VignetteBuilder** knitr

**LazyData** true

**RoxygenNote** 5.0.1

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2017-01-09 15:17:15

## R topics documented:

rclinicaltrials-package . . . . .	2
advanced_search_terms . . . . .	2

clinicaltrials_count . . . . .	3
clinicaltrials_download . . . . .	4
clinicaltrials_search . . . . .	5
gather_results . . . . .	5
parse_study_xml . . . . .	6

<b>Index</b>	<b>7</b>
--------------	----------

---

rclinicaltrials-package

*An interface to the clinicaltrials.gov API*

---

## Description

ClinicalTrials.gov is a registry and results database of publicly and privately supported clinical studies of human participants conducted around the world. Users can search for information about and results from those trials. This package provides a set of functions to interact with the search and download features. Results are downloaded to temporary directories and returned as R objects.

## Details

`clinicaltrials_search` searches and returns basic study information. `clinicaltrials_count` counts the number of trials that satisfy a specific query. `clinicaltrials_download` downloads the set of trials that satisfy a query, which is passed as a string to the function. See `advanced_search_terms` for information about the search terms.

## Author(s)

Maintainer: Michael Sachs <sachsmc@gmail.com>

## References

<http://clinicaltrials.gov>

---

advanced\_search\_terms *Clinicaltrials.gov advanced search terms*

---

## Description

Clinicaltrials.gov advanced search terms. Their keys, descriptions, and links to help documents.

## Usage

`data(advanced_search_terms)`

**Format**

A data frame with three columns: the key to use in advanced search queries keys, a description of the search term description, and a link to the help document help.

**Source**

<http://clinicaltrials.gov/ct2/search/advanced>

**Examples**

```
advanced_search_terms
## Not run: browseURL(advanced_search_terms["age", "help"])
```

---

clinicaltrials_count	<i>Count number of results for a given search</i>
----------------------	---

---

**Description**

Count the number of results for a given search. Useful to check before attempting to pull down a very large number of records.

**Usage**

```
clinicaltrials_count(query = NULL)
```

**Arguments**

query	Search pattern as a string; a vector of key-value pairs is interpreted as an advanced search and is therefore combined with '&'
-------	---

**Value**

Number of results (integer).

**Examples**

```
# count trials satisfying 'heart disease AND stroke AND California'
## Not run: clinicaltrials_count(query = 'heart disease AND stroke AND California')

# advanced search for open, interventional trials involving melanoma
## Not run: clinicaltrials_count(query = c('recr=Open', 'type=Intr', 'cond=melanoma'))

# can also use a named list
## Not run: clinicaltrials_count(query = list(recr='Open', type='Intr', cond='melanoma'))
```

---

`clinicaltrials_download`

*Downloads detailed information about clinical trials satisfying a query*

---

## Description

Given a query, or a dataframe containing trial identifiers, downloads detailed study information from clinicaltrials.gov. Optionally includes results of completed trials. The URL search is limited to 2000 characters, and count restricted searches are done by NCT ID, which are 11 characters long. Therefore, the effective maximum count allowed is roughly 100. If count is greater than 100, then the first 100 trials will be returned, with a warning. To return all results, use a query string and set count to NULL.

## Usage

```
clinicaltrials_download(query = NULL, tframe = NULL, count = 20,  
  include_results = FALSE, include_textblocks = FALSE)
```

## Arguments

<code>query</code>	Search pattern as a string; a vector of key-value pairs is interpreted as an advanced search and is therefore combined with '&'
<code>tframe</code>	Data frame containing trial identifiers, as returned by <a href="#">clinicaltrials_search</a>
<code>count</code>	Limit the results to a specified integer. Set to NULL to include all results.
<code>include_results</code>	Logical. Include results of completed trials
<code>include_textblocks</code>	Logical. Include lengthy text descriptions and eligibility criteria.

## Value

A list of data.frames.

## Examples

```
# trials satisfying 'heart disease AND stroke AND California'  
## Not run: clinicaltrials_download(query = 'heart disease AND stroke AND California', count = 5)  
  
# advanced search for open, interventional trials involving melanoma  
## Not run: clinicaltrials_download(query = c('recr=Open', 'type=Intr', 'cond=melanoma'))  
  
# can also use a named list  
## Not run: clinicaltrials_download(query = list(recr='Open', type='Intr', cond='melanoma'))
```

---

clinicaltrials\_search *Search for clinical trials satisfying a query*

---

**Description**

Searches for results and returns a data.frame with basic study information.

**Usage**

```
clinicaltrials_search(query = NULL, count = 20)
```

**Arguments**

query	Search pattern as a string; a vector of key-value pairs is interpreted as an advanced search and is therefore combined with '&'
count	Limit the results to a specified integer. Set to NULL to include all results.

**Value**

A data.frame with the following columns: NCT study identifier, url, study title, status, condition summary, and date last changed

**Examples**

```
# count trials satisfying 'heart disease AND stroke AND California'
## Not run: clinicaltrials_search(query = 'heart disease AND stroke AND California')

# advanced search for open, interventional trials involving melanoma
## Not run: clinicaltrials_search(query = c('recr=Open', 'type=Intr', 'cond=melanoma'))

# limit to 10 results
## Not run: clinicaltrials_search(query = "colon cancer", count = 10)
```

---

gather\_results *Parses results from an xml object downloaded from clinicaltrials.gov*

---

**Description**

Results of a clinical study are stored in a particular way. This reads and organizes the information and returns it as a list of dataframes. Throws an error if the xml has no clinical\_results node.

**Usage**

```
gather_results(parsed)
```

**Arguments**

parsed                    A parsed XML object, as returned by XML::xmlParse

**Value**

A list of data.frames, participant flow, baseline data, outcome results

---

parse\_study\_xml            *Parses study information for a trial downloaded from clinicaltrials.gov*

---

**Description**

Results of a clinical study are stored in a particular way. This reads and organizes the from a trial xml file (stored in a temporary directory) and returns it as a list of dataframes. Optionally includes long textblock fields and results fields stored as separate data frames.

**Usage**

```
parse_study_xml(file, include_textblocks = FALSE, include_results = FALSE)
```

**Arguments**

file                    Path to xml file  
include\_textblocks            Logical, include long textblock fields in the results  
include\_results            Logical, include results data, if available

**Value**

A list of data.frames, study information, locations, interventions, arms, outcomes, results, and textblocks. Results and textblock frames will be NULL if not requested.

# Index

\*Topic **API**

    rclinicaltrials-package, 2

\*Topic **Internal**

    gather\_results, 5

    parse\_study\_xml, 6

\*Topic **datasets**

    advanced\_search\_terms, 2

advanced\_search\_terms, 2

clinicaltrials\_count, 3

clinicaltrials\_download, 4

clinicaltrials\_search, 4, 5

gather\_results, 5

parse\_study\_xml, 6

rclinicaltrials

    (rclinicaltrials-package), 2

rclinicaltrials-package, 2