Package 'medicaldata'

July 11, 2021

Type Package

Title Data package for Medical Datasets

Version 0.1.0

Maintainer Peter Higgins <higgi13425@yahoo.com>

Description Provides access to medical datasets for teaching.
Some from the tshs website,
and some from R datasets, including Indometh, esoph,
Theoph, lh, infert, and some data donations.

Depends R (¿= 3.1)

License MIT + file LICENSE

Encoding UTF-8

LazyData true

URL https://github.com/higgi13425/medicaldata

RoxygenNote 7.1.1

Roxygen list(markdown = TRUE)

Suggests knitr, rmarkdown, learnr

VignetteBuilder knitr

NeedsCompilation no

Author Peter Higgins [aut, cre]

R topics documented:

`covid_testing` *Deidentified Results of COVID-19 testing at the Children's Hospital of Pennsylvania (CHOP) in 2020*

Description

A dataset containing details of SARS-CoV-2 testing in 2020 at CHOP

Usage

```
covid_testing
```

Format

A data frame with 15524 rows and 17 variables

subject_id id number for each subject; type: numeric

fake_first_name an auto-generated fake first name; type: character

fake_last_name an auto-generated fake last name; character

gender anonymized Gender, levels: female, male; type: character

pan_day day after start of pandemic; type: numeric

test_id test that was performed, levels: covid, xcvd1; type: character

clinic_name Clinic or ward where the specimen was collected, 88 levels; type: character

result result of test, levels: positive, negative, invalid; type: character

demo_group patient group, levels: patient, misc_adult, client, other adult, unidentified; type: character

age Age of subject at time of specimen collection (Anonymized), units = years; type: numeric

drive_thru_ind Whether the specimen was collected via a drive-thru site, levels: 1: Collected at drive-thru site; 0: Not collected at drive-thru site; type: numeric

ct_result Cycle at which threshold reached during PCR, range: 14.05-45; type: numeric

orderset Whether an order set was used for test order, levels: 1: Collected via orderset; 0: Not collected via orderset; numeric

payor_group Payor associated with order, levels: commercial, government, unassigned, medical assistance, self pay, charity care, other; type: character

patient_class Disposition of subject at time of collection, levels: inpatient, emergency, observation, recurring outpatient, outpatient, not applicable, day surgery, admit after surgery-obs, admit after surgery-ip; type: character

col_rec_tat Time elapsed between collect time and receive time, range: 0 - 61370.2, units = hours; type: numeric

rec_ver_tat Time elapsed between receive time and verification time, range: -18.6 - 218.2, units = hours; type: numeric ...

Source

This data set is from Amrom E. Obstfeld, who de-identified data on COVID-19 testing during 2020 at CHOP (Children's Hospital of Pennsylvania). This data set contains data concerning testing for SARS-CoV2 via PCR as well as associated metadata. The data has been anonymized, time-shifted, and permuted.

`indo_rct` *RCT of Indomethacin for Prevention of Post-ERCP Pancreatitis*

Description

Results of a randomized, placebo-controlled, prospective 2-arm trial of rectal indomethacin (100 mg) vs. placebo prevent post-ERCP pancreatitis in 602 participants, as reported by Elmunzer, Higgins, et al. in 2012 in the New England Journal of Medicine.
ERCP, or endoscopic retrograde cholangio-pancreatogram, is a procedure performed by threading an endoscope through the mouth to the opening in the duodenum where bile and pancreatic digestive juices are released into the intestine. ERCP is helpful for treating blockages of flow of bile (gallstones, cancer), or diagnosing cancers of the pancreas, but has a high rate of complications (15-25%).
The occurrence of post-ERCP pancreatitis is a common and feared complication, as pancreatitis can result in multisystem organ failure and death, and can occur in ~ 16% of ERCP procedures.
The inflammatory cytokine storm that can result from this procedural complication can be quite severe. Several small randomized trials suggested that anti-inflammatory NSAID therapies at the time of ERCP could reduce the rate of this complication, but all were rather small single-center studies, and were not sufficiently convincing to change practice. /cr Elmunzer, Higgins, and colleagues performed a meta-analysis of these small trials, which suggested that this was a significant effect, and that indomethacin could result in a 64% reduction in post-ERCP pancreatitis.
The investigators took this as a possible over-estimate of the effect (due to publication bias), and designed a multicenter RCT of a planned 948 patients to see a reduction of 50% from a placebo rate of 10% to an indomethacin rate of 5%. Two interim analyses were performed, after 400 and 600 patients were enrolled, using an alpha spending function. The Data and Safety Monitoring Board stopped the study after 602 participants were enrolled because of the significantly positive effect of indomethacin, which reduced post-ERCP pancreatitis from 16% in the placebo group to 9% in the indomethacin group.
You can find the manuscript at Indomethacin to Prevent Post-ERCP Pancreatitis.

Usage

```
indo_rct
```

Format

A data frame with 602 rows and 25 variables

`subject` subject id, first integer indicates center, integer, range:1001-4003

`age` age in years, numeric, range: 19-90

`risk` risk score, numeric, range: 1-5.5

`gender` male or female, factor, ordered: FALSE, levels: 1. 1_female, 2. 2_male

`outcome` outcome of post-ercp pancreatitis, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`group` treatment arm, factor, ordered: FALSE, levels: 1. 1_placebo, 2. 2_indomethacin

`sod` sphincter of oddi dysfunction was present, a risk factor favoring post-ERCP pancreatitis, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`pep` previous post-ERCP pancreatitis (PEP), a risk factor for future PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`recpancreatitis` Recurrent Pancreatitis, a risk factor for future PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`psphinct` a Pancreatic Sphincterotomy was performed, a risk factor for PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`precut` a sphincter pre-cut was needed to enter the papilla, a risk factor for PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`diffcann` Cannulation of the papilla was difficult, a risk factor for PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`pancinj` Contrast was injected into the pancreas dring the procedure, a risk factor for PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`somsod` Sphincter of oddi manometry was performed during the procedure for SOD, a risk factor for PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`pdstent` A pancreatic duct stent was placed at the end of the procedure per the judgement of the endoscopist (more often in high-risk cases), a potential protective effect against PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`trainee` A trainee participated in the ERCP, which could be a risk factor for PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`bilsphinct` A billiary sphincterotomy was performed, which could be a risk factor for PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`ampullectomy` An Ampullectomy was performed for dysplasia or cancer, which could be a risk factor for PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`acinarization` The pancreas appeared to have acinarization on imaging, which could be a risk factor for PEP, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`bilstent` A biliary stent was placed, which could be protective. factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`choledocho` Choledocholithiasis (gallstones blocking the biliary duct) was present, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`cancer` Cancer of the biliary duct or pancreas was found, factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`pbrush` Brushings were taken from the pancreatic duct, a possible risk factor favoring post-ERCP pancreatitis factor, ordered: FALSE, levels: 1. 1_yes, 2. 2_no

`gibleed` A gastrointestinal bleed occurred (which could be a complication of indomethacin therapy), factor, ordered: FALSE, levels: 1. 0_no, 2. 1_yes

`mspep` NA, factor, ordered: FALSE, levels: 1. 0_no, 2. 1_yes

Source

This data set is sourced from the authors of the 2012 manuscript in the New England Journal of Medicine, entitled, A Randomized Trial of Rectal Indomethacin to Prevent Post-ERCP Pancreatitis, pages 1414-1422 volume 366, in the April 12, 2012 edition, authored by the Elmunzer, BJ, Higgins PDR, et al.

You can find the manuscript at Indomethacin to Prevent Post-ERCP Pancreatitis.

`polyps`	*RCT of Sulindac for Polyp Prevention in Familial Adenomatous Polyposis*

Description

Results of a randomized, placebo-controlled trial of sulindac in the reduction of colonic polyps in Familial Adenomatous Polyposis (FAP).
FAP is an inherited condition caused by mutations in the APC (Adenomatous Polyposis Coli) gene that leads to early and frequent formation of precancerous polyps of the colon at a young age, and invariably leads to the development of colon cancer at a young age.
Early, frequent surveillance colonoscopy and polyp removal is helpful, but this study examined whether there is a beneficial effect of preventive medical therapy with the nonsteroidal pain reliever, sulindac, versus placebo in a RCT vs placebo in 22 participants, with polyp number measured (via colonoscopy) at baseline, 3 months, and 12 months after starting the study drug.

Usage

```
polyps
```

Format

A data frame with 22 rows and 7 variables

participant_id id number for each participant; type: character

sex participant sex, levels: female, male; type: factor

age age in years; type: numeric

baseline number of colonic polyps at baseline; type: numeric

treatment treatment assignment, levels: sulindac, placebo; type: factor

number3m number of colonic polyps at 3 months; type: numeric

number12m number of colonic polyps at 12 months; type: numeric

Source

This data set is from a study published in 1993 in the New England Journal of Medicine,

F. M. Giardiello, S. R. Hamilton, A. J. Krush, S. Piantadosi, L. M. Hylind, P. Celano, S. V. Booker, C. R. Robinson and G. J. A. Offerhaus (1993), Treatment of colonic and rectal adenomas with sulindac in familial adenomatous polyposis. New England Journal of Medicine, 328(18), 1313–1316.

This dataset is derived from and improved upon from the HSAUR package.

scurvy *Randomized Trial of Six Therapies for Scurvy*

Description

Results of a randomized, 6-arm comparator-controlled trial of 6 interventions to treat scurvy in 12 disabled seamen, as reported by James Lind in 1757.

Scurvy was a common affliction of seamen on long voyages, leading to mouth sores, skin lesions, weakness of the knees, and lassitude. Scurvy could be fatal on long voyages. James Lind reported the treatment of 12 seamen with scurvy in 1757, in *A Treatise on the Scurvy in Three Parts*. This 476 page bloviation can be found scanned to the Google Books website A Treatise on the Scurvy. Pages 149-153 are a rare gem among what can be generously described as 400+ pages of evidence-free blathering, and these 4 pages may represent the first report of a controlled clinical trial.

Lind was the ship's surgeon on board the HMS Salisbury, and had a number of scurvy-affected seamen at his disposal. Many remedies had been described and advocated for, with no more than anecdotal evidence. On May 20, 1747, Lind decided to try the 6 available therapies at his disposal in a comparative study in 12 affected seamen. He selected 12 with roughly similar severity, with notable skin and mouth sores, weakness of the knees, and significant lassitude, making them unfit for duty. They each received the standard shipboard diet of gruel and mutton broth, supplemented with occasional biscuits and puddings. Each treatment was a dietary supplement (including citrus fruits) or a medicinal.

This data frame was reconstructed from Lind's account as recorded on these 4 pages, with his estimates of severity translated to a 4 point Likert scale (0-3) for each of the symptoms he described at his chosen endpoint on day 6. A fanciful study_id variable was added, along with detailed descriptions of the dosing schedule of each treatment.

Of note, there is some dispute about whether this was truly the first clinical trial, or whether it actually happened. See link about the historical debate. Lind reported that the seamen treated with 2 lemons and an orange daily did best, followed by those treated with cider. Those treated with elixir of vitriol only had improvement in mouth sores. One imagines that acidic substances (like dilute sulfuric acid, vinegar, cider, and citrus fruits) might have been rather painful on these mouth sores. Unfortunately, the burial of 4 valuable pages of data in 476 pages of noise, a publication delay of 10 years, and Lind's half-hearted conclusions, meant that it took until 1795 before the British Navy mandated daily limes for seamen.

Usage

```
scurvy
```

Format

A data frame with 12 rows and 8 variables

study_id invented id number for each participant; type: character

treatment assigned treatment, levels: cider, dilute_sulfuric_acid, vinegar, sea_water, citrus, purgative_mixture; type: factor

dosing_regimen_for_scurvy details on daily dosing and schedule; type: character

gum_rot_d6 rating of symptom of rotting of gums; type: factor, with levels: 0=none, 1=mild, 2=moderate, 3=severe

skin_sores_d6 rating of symptom of skin sores; type: factor, with levels: 0=none, 1=mild, 2=moderate, 3=severe

weakness_of_the_knees_d6 rating of symptom of weakness of the knees (ability to stand); type: factor, with levels: 0=none, 1=mild, 2=moderate, 3=severe

lassitude_d6 rating of symptom of lassitude (generalized weakness); type: factor, with levels: 0=none, 1=mild, 2=moderate, 3=severe

fit_for_duty_d6 dichotomous fitness for duty as a seaman; type: factor: 0_no, 1_yes

Source

This data set is faithfully reconstructed from a report published in 1757 as *A Treatise on the Scurvy in 3 Parts*, by James Lind, pp. 149-153, and you can find a scan of the source document that you can read yourself on Google Books here.

`strep_tb` *Streptomycin Therapy for Tuberculosis*

Description

Results of a randomized, placebo-controlled, prospective 2-arm trial of streptomycin 2 grams daily (arm A2) vs. placebo (arm A1) to treat tuberculosis in 107 young patients, as reported by the Streptomycin in Tuberculosis Trials Committee in 1948 in the British Medical Journal.

The Streptomycin for Tuberculosis trial in 1948 was considered the first modern randomized, placebo-controlled clinical trial, which could be done in part because there were very limited supplies of streptomycin in the UK after World War II.

This publication seems a bit primitive today, without standard features like a proper Table 1, and some creative use of graphs to display baseline characteristics of the study sample

More strikingly, there is no ethics committee approval, or consent.

You can read the pdf of the original journal article at Streptomycin in TB Study.

This was the first of a series of 3 trials, in which the initial effectiveness of Streptomycin was established, but rapid resistance developed, and significant side effects occurred at a dose of 2 grams of streptomycin. This type of resistance also occurred with another new anti-tubercular therapy at the time, PAS (Para-Amino-Salicylate). Subsequent trials B and C evaluated different doses and combinations of Streptomycin and PAS, and were published together in 1952 in the BMJ, with the pdf available here 1952 Three Streptomycin in TB Studies Summarized.

Commentary on the conduct of these trials from one of the MD investigators can be found at MD Clinical Trialist Commentary.

Commentary on the design and analysis of these trials from statistician A. Bradford Hill can be found at Statistican Commentary.

Usage

```
strep_tb
```

Format

A data frame with 107 rows and 13 variables

patient_id invented id number for each participant; type: character

arm assigned treatment arm, Streptomycin or Control; type: factor

dose_strep_g grams, dose of Streptomycin: numeric, 0, 1, or 2 grams

dose_PAS_g grams, dose of PAS (Para-Amino-Salicylate): numeric, 5, 10, or 20 grams. Note that no one in this intial study (study A) received PAS. This was added for combination therapy in studies B and C, as reported in 1952.

gender gender, dichotomous (this was in 1948); type: factor, with levels: M = Male, F= Female

baseline_condition Condition of the Patient at Baseline, 3 levels, 1_Good, 2_Fair, 3_Poor; type: factor

baseline_temp temperature at baseline in degrees fahrenheit or celsius, but categorized into 4 levels (afebrile level apparently were cases not measured with a thermometer): factor, with levels: 1_afebrile, 2_¡99F/¡37.2C, 3_99-99.9F/37.2-37.75C, 4_100F+/37.7C+

baseline_esr Erythrocyte Sedimentation Rate in mm per hour, categorized into 4 levels, from 0-51+ mm per hour; type: factor, with levels: 1_0-10, 2_11-20, 3_21-50, 4_51+

baseline_cavitation dichotomous presence of cavitation on the baseline chest x-ray; type: factor: 0_no, 1_yes

strep_resistance streptomycin resistance after 6 months of therapy, measured on a 0-100+ scale, categorized into 3 levels - sensitive, moderate, and resistant; type: factor: 1_sens_0-8, 2_mod_8-99, 3_resist_100+

radiologic_6m Likert score rating of radiologic response on chest x-ray at 6 months; type: factor: 1_Death, 2_Considerable_deterioration, 3_Moderate_deterioration, 4_No_change, 5_Moderate_improvement, 6_Considerable_improvement

rad_num Likert score numeric rating of radiologic response on chest x-ray at 6 months; type: numeric: 1-6, from Death to Considerable Improvement

improved Dichotomous outcome of improvement (equal to rad_num of 5-6); type: logical, TRUE or FALSE. 55 of the 107 participants were improved.

Source

This data set is reconstructed to the best of my ability from the paper in the British Medical Journal from 1948, entitled, Streptomycin Treatment of Pulmonary Tuberculosis, pages 769-782 in the October 30, 1948 edition, authored by the Streptomycin in Tuberculosis Trials Committee. You can find the pdf at Streptomycin in TB.

Index