Package 'medicaldata'

July 22, 2025

Type Package

Title Data Package for Medical Datasets

Version 0.2.0

Date 2021-08-08

Description Provides access to well-documented medical datasets for teaching. Featuring several from the Teaching of Statistics in the Health Sciences website <https://www.causeweb.org/tshs/category/dataset/>, a few reconstructed datasets of historical significance in medical research, some reformatted and extended from existing R packages, and some data donations.

Depends R (>= 3.1)

License MIT + file LICENSE

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

BugReports https://github.com/higgi13425/medicaldata/issues

Suggests knitr, rmarkdown, markdown, learnr

VignetteBuilder knitr

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

NeedsCompilation no

Author Peter Higgins [aut, cre] (ORCID: <https://orcid.org/0000-0002-7111-0077>)

Maintainer Peter Higgins <higgi13425@yahoo.com>

Repository CRAN

Date/Publication 2021-08-16 07:00:06 UTC

Contents

`blood_storage` *Retrospective Cohort Study of the Effects of Blood Storage on Prostate Cancer*

Description

This data set contains data on 316 men who had undergone radical prostatectomy and received transfusion during or within 30 days of the surgical procedure and had available prostate serum antigen (PSA) follow-up data. The main exposure of interest was RBC storage duration group. A number of demographic, baseline and prognostic factors were also collected. The outcome was time to biochemical (PSA) cancer recurrence. The dataset is cleaned and complete. There are no outliers or data problems (more details after variable information).

Usage

```
blood_storage
```

Format

A data frame with 316 observations and 20 variables

`RBC.Age.Group` NA, numeric, range: 1.00- 3

`Median.RBC.Age` NA, numeric, range:10.00- 25

`Age` NA, numeric, range:38.40- 79

`AA` NA, numeric, range: 0.00- 1

`FamHx` NA, numeric, range: 0.00- 1

`PVol` NA, numeric, range:19.40-274

`TVol` NA, numeric, range: 1.00- 3

`T.Stage` NA, numeric, range: 1.00- 2

`bGS` NA, numeric, range: 1.00- 3

`BN+` NA, numeric, range: 0.00- 1

`OrganConfined` NA, numeric, range: 0.00- 1

`PreopPSA` NA, numeric, range: 1.30- 40

`PreopTherapy` NA, numeric, range: 0.00- 1

`Units` NA, numeric, range: 1.00- 19

`sGS` NA, numeric, range: 1.00- 4

`AnyAdjTherapy` NA, numeric, range: 0.00- 1

`AdjRadTherapy` NA, numeric, range: 0.00- 1

`Recurrence` NA, numeric, range: 0.00- 1

`Censor` NA, numeric, range: 0.00- 1

`TimeToRecurrence` NA, numeric, range: 0.27-104

Details

Background:

Prostate cancer is the most common malignant neoplasm in men, and radical prostatectomy is among the primary therapies for localized prostate cancer. The biochemical recurrence-free survival rate 5 years after prostatectomy ranges from 70% to 90%. Improvements in the surgical technique have decreased the amount of intraoperative blood loss occurring during radical prostatectomy; however, substantial numbers of patients still require perioperative blood transfusions.
Blood transfusions are associated with adverse reactions, including postoperative infections and transfusion-related immune perturbations. Allogeneic leukocytes present in the transfused blood are thought to suppress host cellular immune responses. Furthermore, the immunodepressant effect is secondary to an imbalance of accumulated cytokines and proinflammatory mediators in the transfused blood against decreased production of lymphocyte stimulating cell-mediated cytokines, such as interleukin 2 and increased release of immunosuppressive prostaglandins in the patient undergoing transfusion.
In cancer patients, perioperative blood transfusion has long been suspected of reducing long-term survival, but available evidence is inconsistent. It is also unclear which components of transfused blood underlie the cancer-promoting effects reported by some studies. An important factor associated with the deleterious effects of blood transfusion is the storage age of the transfused blood units. It is suspected that cancer recurrence may be worsened after the transfusion of older blood.

Study Objective:

This study evaluated the association between red blood cells (RBC) storage duration and biochemical prostate cancer recurrence after radical prostatectomy. Specifically, tested was the hypothesis that perioperative transfusion of allogeneic RBCs stored for a prolonged period is associated with earlier biochemical recurrence of prostate cancer after prostatectomy.

Subjects and Variables:

Patients were assigned to 1 of 3 RBC age exposure groups on the basis of the terciles (ie, the 33rd and 66th percentiles) of the overall distribution of RBC storage duration if all their transfused units could be loosely characterized as of 'younger,' 'middle,' or 'older' age. Although this approach resulted in the removal of certain patients with wide RBC age distributions, it has the advantage of defining an essentially random and clearly separable exposure.
Prostate-specific antigen (PSA) was used as a biochemical marker of prostate cancer recurrence after prostatectomy. A PSA value of at least 0.4 ng/mL (to convert to microg/L, multiply by 1.0) followed by another increase was considered biochemical cancer recurrence.

The initial population consisted of 865 men who had undergone radical prostatectomy and received transfusion during or within 30 days of the surgical procedure at Cleveland Clinic and had available PSA follow-up data. Of these patients, 110 were excluded from the analysis because they received a combination of allogeneic and autologous blood products. Of the remaining 755 patients, 405 (54%) received solely allogeneic and 350 patients (46%) received solely autologous RBC units. Of the 405 patients who received allogeneic RBC transfusion, 89 were excluded because their transfused RBC age distribution included more than one of the terciles. Thus, this dataset consists of the 316 patients who received solely allogeneic blood products and could be classified into an RBC age exposure group.

Source

Cata et al. 'Blood Storage Duration and Biochemical Recurrence of Cancer after Radical Prostatectomy'. Mayo Clin Proc 2011; 86(2): 120-127.

`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 observations 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 ...

Details

Data on testing for SARS-CoV2 from days 4-107 of the COVID pandemic in 2020. CHOP is a pediatric hospital in Philadelphia, Pennsylvania, USA. These data have been anonymized, time-shifted, and permuted.

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. These data have been anonymized, time-shifted, and permuted.

`cytomegalovirus` *Retrospective Cohort Study of the Effects of Donor KIR genotype on the reactivation of cytomegalovirus (CMV) after myeloablative allogeneic hematopoietic stem cell transplant.*

Description

This data set contains 64 consecutive patients who underwent T-cell replete, matched sibling donor reduced-intensity conditioning allogeneic hematopoietic stem cell transplant. The primary risk factor of interest was the number of activating killer immunoglobulin-like receptors (aKIRs: 1-4 vs. 5-6). (more details after variable information).

Usage

```
cytomegalovirus
```

Format

A data frame with 64 observations and 26 variables

`ID` Patient ID, numeric, range: 1-64

`age` Recipient age at transplant, numeric, range: 29-67

`sex` Recipient sex, numeric, range: 0 (male) - 1(female)

`race` Recipient race, numeric, range: 0 (white) - 1 (african-american)

`diagnosis` type: character, levels: 13

`diagnosis.type` Category of cancer diagnosis, numeric, range: 0 (myeloid) - 1 (lymphoid)

`time.to.transplant` Time from cancer diagnosis to transplant (months), numeric, range: 1.84-173.8

`prior.radiation` Prior radiation therapy, numeric, range: 0 (no) - 1 (yes)

`prior.chemo` Number of prior chemotherapy regimens, numeric, range: 0-8

`prior.transplant` Prior stem cell transplant, numeric, range: 0 (no) - 1 (yes)

`recipient.cmv` Recipient cytomegalovirus seropositive status, numeric, range: 0 (negative) - 1 (positive)

`donor.cmv` Donor cytomegalovirus seropositive status, numeric, range: 0 (negative) - 1 (positive)

`donor.sex` Donor sex, numeric, range: 0 (female) - 1 (male)

`TNC.dose` Total nucleated cell dose (x 10^8/kg), numeric, range: 2.06- 21.0

`CD34.dose` Total CD34+ (stem) cell dose (x 10^8/kg), numeric, range: 2.04- 12.5

`CD3.dose` Total CD3+ (T) cell dose (x 10^8/kg), numeric, range: 1.08- 8.2

`CD8.dose` Total CD8+ cell dose (x 10^8/kg), numeric, range: 0.16- 3.2

`TBI.dose` Total body irradiation dosage (centiGrays), numeric, range:200.00-400.0

`C1/C2` HLA-Cw group, numeric, range: 0 (heterozygous) - 1 (homozygous)

`aKIRs` Number of donor activating killer immunoglobulin-like receptors (hypothesized Predictor), numeric, range: 1.00- 6.0

`cmv` cytomegalovirus reactivation posttransplant (hypothesized Outcome), numeric, range: 0 (No) - 1 (Yes)

`time.to.cmv` Time to cytomegalovirus reactivation (months), numeric, range: 0.43- 84.5

`agvhd` Acute level 2-4 graft versus host disease, numeric, range: 0 (no) - 1 (yes)

`time.to.agvhd` Time to acute level 2-4 graft versus host disease (months), numeric, range: 0.66-85.2

`cgvhd` Chronic graft versus host disease, numeric, range: 0 (no) - 1 (yes)

`time.to.cgvhd` Time to chronic graft versus host disease (months), numeric, range: 0.82- 65.1

Details

A number of demographic, baseline and transplant characteristics were also collected. The primary outcome is presence of and time to cytomegalovirus reactivation. The dataset is cleaned and relatively complete. There are no outliers or data problems.

Background
:
Hematopoietic stem cell transplantation (HSCT) is the transplantation of multipotent hematopoietic stem cells, from bone marrow, peripheral blood, or umbilical cord blood. It is a medical procedure most often performed for patients with certain cancers of the blood or bone marrow, such as multiple myeloma or leukemia. Allogeneic HSCT involves two people: the (healthy) donor and the (patient) recipient. Allogeneic HSC donors must have a tissue (HLA) type that matches the recipient.
In myeloablative allogeneic HSCT, chemotherapy or irradiation is given immediately prior to a transplant (the conditioning regimen) with the purpose of eradicating the patient's disease prior to the infusion of HSC and to suppress immune reactions. The bone marrow can be ablated (destroyed) with dose- levels that cause minimal injury to other tissues. For many patients who are at high risk for transplant-related mortality with myeloablative allogeneic HSCT, reduced- intensity conditioning allogeneic hematopoietic stem cell transplant has proven effective. Although the reduced-intensity conditioning allogeneic HSCT may avoid many of the organ toxicities associated with myeloablative conditioning, the risk for developing graft-versus-host disease and infection including cytomegalovirus remains significant.
Cytomegalovirus (CMV) is a common virus that can infect almost anyone. Once infected, your body retains the virus for life. Most people don't know they have CMV because it rarely causes problems in healthy people. But if pregnant or having a weakened immune system, CMV is cause for concern. For people with compromised immunity, such as after allogeneic HSCT, CMV infection can be fatal. Natural killer (NK) and T cells provide protection against CMV reactivation. The reactivity of NK cells and some T-cell subsets are regulated by the interaction of killer immunoglobulin-like receptors (KIRs) with target cell HLA class 1 molecules. The donor activating KIR genotype has been implicated as a contributing factor for CMV reactivation after myeloablative allogeneic HSCT.

Study Objective:
This study investigates whether donor KIR genotype influences reactivation of CMV after T-cell replete, matched sibling donor reduced-intensity conditioning allogeneic HSCT.

Subjects and Variables:

The study included 64 consecutive patients who underwent T-cell replete, matched sibling donor reduced-intensity conditioning allogeneic hematopoietic stem cell transplant between January 16, 2000 and April 24, 2007 at the Cleveland Clinic. Human leucocyte antigen (HLA) typing on donors and recipients was performed to allow assessment of killer immunoglobulin-like receptor ligands (KIRs). To allow for comparison with previous studies, donors were categorized as having 1-4 or 5-6 activating killer immunoglobulin-like receptor genes (aKIRs).
CMV reactivation was defined as any detection of cytomegalovirus DNA in the blood; the lower detection limit for this assay was 600 copies/mL.

The initial population consisted of 865 men who had undergone radical prostatectomy and received transfusion during or within 30 days of the surgical procedure at Cleveland Clinic and had available PSA follow-up data. Of these patients, 110 were excluded from the analysis because they received a combination of allogeneic and autologous blood products. Of the remaining 755 patients, 405 (54%) received solely allogeneic and 350 patients (46%) received solely autologous RBC units. Of the 405 patients who received allogeneic RBC transfusion, 89 were excluded because their transfused RBC age distribution included more than one of the terciles. Thus, this dataset consists of the 316 patients who received solely allogeneic blood products and could be classified into an RBC age exposure group.

Source

Sobecks et al. 'Cytomegalovirus Reactivation After Matched Sibling Donor Reduced-Intensity Conditioning Allogeneic Hematopoietic Stem Cell Transplant Correlates With Donor Killer Immunoglobulin-like Receptor Genotype'. Exp Clin Transplant 2011; 1: 7-13.

`esoph_ca` *esoph_ca: Esophageal Cancer dataset*

Description

Data from a case-control study of esophageal cancer in Ille-et-Vilaine, France, evaluating the effects of smoking and alcohol on the incidence of esophageal cancer. Smoking and alcohol are associated risk factors for squamous cell cancer of the esophagus, rather than adenocarcinoma of the esophagus, which is associated with obesity and esophageal reflux (more details available below the variable definitions).

Usage

```
esoph_ca
```

Format

A data frame with 88 rows and 5 variables, with 200 cases and 975 controls.

agegp 6 levels of age: "25-34", "35-44", "45-54", "55-64", "65-74", "75+"; type: ordinal factor

alcgp 4 levels of alcohol consumption: "0-39g/day", "40-79", "80-119", "120+"; type: ordinal factor

tobgp 4 levels of tobacco consumption: "0-9g/day", "10-19", '20-29", "30+"; type: ordinal factor

ncases Number of cases; type: integer

ncontrols Number of controls; type: integer

Details

An original base R dataset, though of somewhat unclear origin. The statistical textbook source is clear, though it is not clear which of the original epidemiological papers on esophageal cancer in Ille-et-Vilaine is referred to by this dataset. The original authors of the medical study were **not** credited in the base R dataset. There are several possible papers in PubMed, none of which quite match up with this dataset. This could be from Tuyns, AJ, et al., Bull Cancer, 1977;64(1):45-60, but this paper reports 778 controls, rather than the 975 found here. A 1975 paper from the same group reported 718 cases (Int J Epidemiol, 1975 Mar;4(1):55-9. doi: 10.1093/ije/4.1.55.). There is also another possible source - a 1975 paper from the same group, *Usefulness of population controls in retrospective studies of alcohol consumption. Experience from a case–control study of esophageal cancer in Ille-et-Vilaine, France*, Journal of Studies on Alcohol, 39(1): 175-182 (1978), which is behind a publisher paywall.

Source

Breslow, N. E. and Day, N. E. (1980) Statistical Methods in Cancer Research. Volume 1: The Analysis of Case-Control Studies. IARC Lyon / Oxford University Press. Originally in base R datasets.

`indometh` *Cohort Study of the Pharmacokinetics of Intravenous Indomethacin*

Description

Results of a Cohort Study of the Pharmacokinetics of Intravenous Indomethacin, with plasma concentrations over time (**more details** available below the variable definitions).

Usage

```
indometh
```

Format

A data frame with 66 observations and 3 variables

Subject subject id number for each participant; type: character

time Time from initial dose in hours; type: double

conc Concentration of indomethacin in the plasma in micrograms per milliliter' type: double

Details

This data set contains data on 6 healthy volunteer subjects who participated in a pharmacokinetic study of intravenous indomethacin. Indomethacin is an anti-inflammatory and pain-relieving non-steroidal medication. It can be administered by the intravenous, oral, or rectal suppository routes. Some of the indomethacin is excreted in the bile and reabsorbed by the intestine. This phenomenon, called enterohepatic circulation, keeps the drug around longer than would be expected otherwise.
Each subject in Study 1 (intravenous route) received a single 50 mg dose of radioactively labeled indomethacin (^14^-carbon-labeled, with each dose containing 25 microCuries of radioactivity). Subjects received a standard meal (one 8-oz can of Metrecal, 8 oz of whole milk, and one medium-size apple) 30 rain prior to medication and 8 oz of water every 2 hr throughout the waking hours to ensure adequate urine output.
Blood samples were taken at frequent intervals over the first 8 hours after dosing, and the quantity of indomethacin in the plasma (as well as stool and urine) at each time point was measured in micrograms per milliliter. This data set only contains the plasma measurements from Table 1 on page 258 of the manuscript. While this paper was published in 1976 (post-Tuskegee reveal), there is no mention of ethics review, IRB review, or consent of the healthy volunteers.
The abstract from the original manuscript:
There are no discernible quantitative differences in the biotransformation and the excretion of indomethacin following oral, rectal, and intravenous administration of indomethacin-2-^14^C. Approximately 50% (range 24-115% for n = 6) of an intravenous dose undergoes enterohepatic circulation. Thus the bioavailability of indomethacin to the systemic circulation may exceed the administered dose. Relative to the intravenous dose, indomethacin is 80 and 100% bioavailable from suppositories and capsules, respectively. Absorption and/or reabsorption appears to be more rapid and uniform by the rectal route. Recognition of the attributes of biliary recycling also helps to explain the observed variability in apparent plasma half-life, while their neglect requires alternative explanations for anomalies between the disappearance rate from plasma and the corresponding appearance rate in urine.

Source

Kwan, Breault, Umbenhauer, McMahon and Duggan (1976) Kinetics of Indomethacin absorption, elimination, and enterohepatic circulation in man. Pharmacokinetics and Biopharmaceutics. 1976 Jun;4(3):255-80. doi: 10.1007/BF01063617.

---

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 (more details available below the variable definitions).

Usage

```
indo_rct
```

Format

A data frame with 602 observations and 33 variables

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

site study site (center), factor, 1 = University of Michigan, 2= Indiana University, 3 = University of Kentucky, 4 = Case Western

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

risk risk score, numeric, range: 1-5.5

gender male or female, factor, levels: 1_female, 2_male

sod sphincter of oddi dysfunction was present, a risk factor favoring post-ERCP pancreatitis, factor, levels: 0_no, 1_yes

pep previous post-ERCP pancreatitis (PEP), a risk factor for future PEP, factor, levels: 0_no, 1_yes

recpanc Recurrent Pancreatitis, a risk factor for future PEP, factor, levels: 0_no, 1_yes

psphinc a Pancreatic Sphincterotomy was performed, a risk factor for PEP, factor, levels: 0_no, 1_yes

precut a sphincter pre-cut was needed to enter the papilla, a risk factor for PEP, factor, levels: 0_no, 1_yes

difcan Cannulation of the papilla was difficult, a risk factor for PEP, factor, levels: 0_no, 1_yes

pneudil Pneumatic dilation of the papilla was performed, a risk factor for PEP, factor, levels: 0_no, 1_yes

amp An Ampullectomy was performed for dysplasia or cancer, which could be a risk factor for PEP, factor, levels: 0_no, 1_yes

paninj Contrast was injected into the pancreas during the procedure, a risk factor for PEP, factor, levels: 0_no, 1_yes

acinar The pancreas appeared to have acinarization on imaging, which could be a risk factor for PEP, factor, levels: 0_no, 1_yes

brush Brushings were taken from the pancreatic duct, a possible risk factor favoring post-ERCP pancreatitis. factor, levels: 0_no, 1_yes

asa81 Aspirin was used at a dose of 81 mg per day, which may increase the risk of bleeding. factor, levels: 0_no, 1_yes

asa325 Aspirin was used at a dose of 325 mg per day, which may increase the risk of bleeding. factor, levels: 0_no, 1_yes

asa Aspirin was used (at a dose of 325 mg per day(at any dose), which may increase the risk of bleeding. factor, levels: 0_no, 1_yes

prophystent 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, levels: 0_no, 1_yes

therastent A pancreatic duct stent was placed in order to treat a clinically significant narrowing of the pancreatic duct, a potential protective effect against PEP, factor, levels: 0_no, 1_yes

pdstent A pancreatic duct stent was placed at the end of the procedure for any reason, a potential protective effect against PEP, factor, levels: 0_no, 1_yes

sodsom Sphincter of oddi manometry was performed during the procedure for SOD, a risk factor for PEP, factor, levels: 0_no, 1_yes

bsphinc A biliary sphincterotomy was performed, which could be a risk factor for PEP, factor, levels: 0_no, 1_yes

bstent A biliary stent was placed to relieve significant biliary obstruction, factor, levels: 0_no, 1_yes

chole Choledocholithiasis (gallstones blocking the biliary duct) was present, factor, levels: 0_no, 1_yes

pbmal Malignancy of the biliary duct or pancreas was found, factor, levels: 0_no, 1_yes

train A trainee participated in the ERCP, which could be a risk factor for PEP, factor, levels: 0_no, 1_yes

outcome outcome of post-ercp pancreatitis, factor, levels: 0_no, 1_yes

status outpatient status, factor, levels: 0_inpatient, 1_outpatient

type Sphincter of Oddi dysfunction type/level - higher numbers are more severe with greater association with PEP, factor, levels: 0_no SOC, 1_type 1, 2_type 2, 3_type 3

rx treatment arm, factor, levels: 0_placebo, 1_indomethacin

bleed A gastrointestinal bleed occurred (which could be a complication of indomethacin therapy), factor, levels: 1. no, 2. yes

Details

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.

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.

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.

`laryngoscope` *Randomized, Comparison Trial of Video vs. Standard Laryngoscope*

Description

This data set contains 99 adult patients with a body mass index between 30 and 50 kg/m2 who required orotracheal intubation for elective surgery. Patient demographics, airway assessment data, intubation success rate, time to intubation, ease of intubation, and occurrence of complications were recorded. The dataset is cleaned and complete. There are no outliers or data problems (**more details** available below the variable definitions).

Usage

```
laryngoscope
```

Format

A data frame with 99 observations and 22 variables

`age` Age (years), numeric, range: 20-77

`gender` Gender, numeric, 0 = female; 1 = male

`asa` American Society of Anesthesiologists physical status(1-4), range: 2-4

`BMI` Body Mass Index (kg/m^2), numeric, range: 31-61

`Mallampati` Mallampati score predicting ease of intubation 1 = Full visibility of tonsils, uvula and soft palate (easy intubation); 2 = Visibility of hard and soft palate, upper portion of tonsils and uvula; 3 = Soft and hard palate and base of the uvula are visible; 4 = Only Hard Palate visible (difficult intubation), numeric, range: 1-4

`Randomization` Laryngoscope randomized, numeric, range: 0 = Standard Macintosh #4, 1 = AWS Pentaz Video

`attempt1_time` First intubation attempt time (seconds), numeric, range: 9-113

`attempt1_S_F` Successful intubation first attempt, numeric, 0 = no, 1 = yes

`attempt2_time` Second intubation attempt time (seconds), numeric, range: 11- 60

`attempt2_assigned_method` Second intubation attempt made with assigned laryngoscope, numeric, 0 = no, 1 = yes

`attempt2_S_F` Successful intubation second attempt, numeric, 0 = no, 1 = yes, numeric, range: 0 = no, 1 = yes

`attempt3_time` Third intubation attempt time (seconds), numeric, range: 15- 30

`attempt3_assigned_method` Third intubation attempt made with assigned laryngoscope, numeric, 0 = no, 1 = yes

`attempt3_S_F` Successful intubation third attempt, numeric, 0 = no, 1 = yes, numeric, range: 1-1

`attempts` Number of intubation attempts, numeric, range: 1-3

`failures` Number of intubation failures, numeric, range: 0-2

`total_intubation_time` Total Intubation time (second), numeric, range: 9-100

`intubation_overall_S_F` Overall successful intubation, numeric, 0 = no, 1 = yes

`bleeding` Bleeding (trace), numeric, 0 = no, 1 = yes

`ease` Ease of tracheal intubation, 0 = extremely easy to 100 = extremely difficult, numeric, range: 0-100

`sore_throat` Severity of postoperative sore throat, 0 = none; 1 = mild; 2 = moderate; 3 = severe, numeric, range: 0- 3

`view` Cormack-Lehane grade of glottic view 0 = "not good" Cormack- Lehane grade 1 or 2; 1 = "good" Cormack-Lehane grade 3 or 4, numeric, range: 0- 1

Details

The Laryngoscope dataset was contributed by Dr. Amy Nowacki, Associate Professor, Cleveland Clinic. Please refer to this resource as: Amy S. Nowacki, 'Laryngoscope Dataset', TSHS Resources Portal (2017). Available at https://www.causeweb.org/tshs/laryngoscope/.
Difficult and failed tracheal intubations are among the principal causes of anesthetic-related mortality and morbidity. Because a good laryngeal view facilitates successful tracheal intubation, new technologies have been introduced to improve visualization. Video laryngoscopes, for example, often use miniature cameras to facilitate visualization of the laryngeal inlet with no need to align the oral, pharyngeal, and tracheal axes.
The Pentax AWS is a novel video laryngoscope, available in Japan since 2006, which is designed to facilitate intubation by providing a video image of the glottis. It incorporates a miniature video camera and a battery-powered, built-in LCD monitor. A disposable blade is attached to the base system. Incorporation of an LCD display makes it possible to view the glottis simultaneously with insertion of the endotracheal tube (ETT). In this regard, it differs from some other video laryngoscope designs that use external monitors. The Pentax AWS also differs in having a side channel that positions and guides the ETT. Reports suggest that the Pentax AWS can help intubate, but randomized data remain sparse.cr This study tested the hypothesis that intubation with the Pentax AWS would be easier and faster than with a standard Macintosh laryngoscope with a #4 blade.

Source

These are data from a study by Abdallah et al. A Randomized Comparison between the Pentax AWS Video Laryngoscope and the Macintosh Laryngoscope in Morbidly Obese Patients. Anesthesia Analgesia 2011; 113: 1082-7.

`licorice_gargle` *Randomized, Controlled Trial of Licorice Gargle before Intubation for Elective Thoracic Surgery*

Description

This study enrolled 236 adult patients undergoing elective thoracic surgery requiring a double-lumen endotracheal tube. Gender, physical status, BMI, age, Mallampati score, smoking status, preoperative pain, surgery size, intervention and the outcomes (cough, sore throat and pain swallowing at various time points) are provided. The dataset is cleaned and complete (missing outcomes for 2 patients). There are no outliers or data problems (**more details** available below the variable definitions).

Usage

```
licorice_gargle
```

Format

A data frame with 235 observations and 19 variables

`preOp_gender` Gender, numeric, 0 = Male; 1 = Female

`preOp_asa` American Society of Anesthesiologists physical status, numeric, 1 = a normal healthy patient; 2 = a patient with mild systemic disease; 3 = a patient with severe systemic disease

`preOp_calcBMI` Body mass index (kg/m^2), numeric, range:16-36

`preOp_age` Age (years), numeric, range:18-86

`preOp_mallampati` Mallampati score, with 1 = easy to intubate, 4= difficult intubation, numeric, 1 = soft palate, fauces, uvula, pillars visible; 2 = soft palate, fauces, uvula visible; 3 = soft palate, base of uvula visible; 4 = soft palate not visible at all

`preOp_smoking` Smoking status, numeric, 1 = Current; 2 = Past; 3 = Never

`preOp_pain` Preoperative pain, numeric, 0 = No; 1 = Yes

`treat` Intervention, 0 = Sugar 5g; 1 = Licorice 0.5g

`intraOp_surgerySize` Surgery size, numeric, 1 = Small (thoracoscopy); 2 = Medium (thoracotomy < 3 h); 3 = Large (thoracotomy > 3 h or blood loss > 1000 mL)

`extubation_cough` Amount of coughing immediately after extubation, numeric, 0 = No cough; 1 = Mild; 2 = Moderate; 3 = Severe

`pacu30min_cough` Amount of coughing at 30 minutes after arrival in PACU, numeric, 0 = No cough; 1 = Mild; 2 = Moderate; 3 = Severe

`pacu30min_throatPain` Sore throat pain score at rest at 30 minutes after arrival in PACU (11 point Likert scale, 0=no pain, 10 = worst pain)

`pacu30min_swallowPain` Sore throat pain score during swallowing at 30 minutes after arrival in PACU (11 point Likert scale, 0=no pain, 10 = worst pain), numeric, range: 0-10

`pacu90min_cough` Amount of coughing at 90 minutes after arrival in PACU, numeric, 0 = No cough; 1 = Mild; 2 = Moderate; 3 = Severe

`pacu90min_throatPain` Sore throat pain score at rest at 90 minutes after arrival in PACU (11 point Likert scale, 0=no pain, 10 = worst pain), numeric, range: 0-6)

`postOp4hour_cough` Amount of coughing at 4 hours after surgery, numeric, 0 = No cough; 1 = Mild; 2 = Moderate; 3 = Severe, range: 0-2

`postOp4hour_throatPain` Sore throat pain score at rest at 4 hours after surgery (11 point Likert scale, 0=no pain, 10 = worst pain), numeric, range: 0-6), numeric, range: 0- 7

`pod1am_cough` Amount of coughing on the first postoperative morning, 0 = No cough; 1 = Mild; 2 = Moderate; 3 = Severe, numeric, range: 0- 3

`pod1am_throatPain` Sore throat pain score at rest on the first postoperative morning (11 point Likert scale, 0=no pain, 10 = worst pain), numeric, range: 0-6), numeric, range: 0- 6

Details

The Licorice Gargle dataset was contributed by Dr. Amy Nowacki, Associate Professor, Cleveland Clinic. Please refer to this resource as: Amy S. Nowacki, 'Licorice Gargle Dataset', TSHS Resources Portal (2017). Available at https://www.causeweb.org/tshs/licorice-gargle/.

Postoperative sore throat is a common and annoying complication of endotracheal intubation. Intubation with double-lumen tubes, which are much larger than conventional single-lumen tubes, are especially likely to provoke sore throats, with a reported incidence up to 90%. Presumably, postoperative sore throats are a consequence of local tissue trauma, due to laryngoscopy and/or endotracheal intubation, leading to inflammation of pharyngeal mucosa.

Nonpharmacological methods for preventing an intubation-related sore throat include using smaller-sized endotracheal tubes, lubricating the endotracheal tube with water-soluble jelly, and careful airway instrumentation as examples. Pharmacological measures for attenuating postoperative sore throats include inhalation of beclomethasone or fluticasone propionate; gargling with azulene sulfonate, aspirin, or ketamine; and gargling or spraying benzydamine hydrochloride on the endotracheal cuff for example. Each of these approaches and others not listed, however, has limitations and variable success rates; thus none has become established or is in routine clinical use.

Recently, a study reported that gargling with licorice halves the risk of sore throat after intubation with conventional endotracheal tubes, based on a study of just 40 patients. A number of active ingredients have been isolated from licorice, including glycyrrhizin, liquilitin, liquiritigenin, and glabridin. The glycyrrhizin component reportedly has anti-inflammatory and antiallergic properties. Liquilitin and liquiritigenin have peripheral and central antitussive properties. Glabridin has significant antioxidant and ulcer-healing properties, which might help heal pharyngeal and tracheal mucosa after minor injuries that often complicate laryngoscopy, intubation, and endotracheal tube cuff inflation.

This study tested the hypothesis that gargling with licorice solution immediately before induction of anesthesia prevents sore throat and postextubation coughing in patients intubated with double-lumen tubes.

Source

These are data from a study by Ruetzler et al. 'A Randomized, Double-Blind Comparison of Licorice Versus Sugar-Water Gargle for Prevention of Postoperative Sore Throat and Postextubation Coughing'. Anesth Analg 2013; 117: 614 – 21.

opt *Obstetrics and Periodontal Therapy Dataset*

Description

The objective of this randomized controlled trial was to determine whether treatment of maternal periodontal disease can reduce risk of preterm birth and low birth weight (**more details** available below the variable definitions).

Usage

```
opt
```

Format

A data frame with 823 observations and 171 variables

`PID` Participant ID, First digit indicates enrollment center (1 = NY, 2 = MN, 3 = KY, 4 = MS); Next 4 digits are sequential; Sixth digit is a check digit; There are no missing data, numeric, range: 100034-402477

`Clinic` Enrollment Center, factor, NY = Harlem Hospital, MN = Hennepin County Center; KY = University of Kentucky; MS = University of Mississippi Medical Center; There are no missing data

`Group` Randomized treatment assignment, factor, T = Intervention; C = Control; There are no missing data

`Age` Age of participant at baseline (years), numeric, range: 16-44

`Black` Black participant (self-identified), factor; Yes, No

`White` White participant (self-identified), factor; Yes, No

`Nat.Am` Native American participant, incl. Latin Americans with aboriginal origin(self-identified), factor; Yes, No

`Asian` Asian participant (self-identified), factor; Yes, No

`Hisp` Hispanic participant (self-identified), factor; Yes, No

`Education` Education level of participant, factor; LT 8 yrs = Less than 8 years; 8-12 yrs = 8 to 12 years; MT 12 yrs = More than 12 yrs; blank = Missing

`Public.Asstce` Public Assistance: Whether a government agency paid for the delivery, factor; Yes, No;

`Hypertension` Whether participant had chronic hypertension at baseline, factor; Yes, No

`Diabetes` Whether participant had diabetes at baseline (self-reported), factor; Yes, No

`BL.Diab.Type` Baseline Diabetes Type: Type of diabetes, for participants having diabetes at baseline (self-reported), factor; Type I; Type II; Blank = No diabetes at baseline (variable 13 = No)

BMI NA, numeric, range: 15.000-68.0

Use.Tob Self-reported participant history of tobacco use, factor; Yes, No; Blank = Missing

BL.Cig.Day Self-reported number of cigarettes per day for those with tobacco use history, numeric, range: 1-30; Blank = Missing (variable 16= Yes or blank) or non-smoker (variable 16 = No)

Use.Alc Self-reported participant history of alcohol use, factor; Yes, No; Blank = Missing

BL.Drks.Day , Blank = Missing (variable 18 = Yes or blank) or non-drinker (variable 18 = No)

Drug.Add Self-reported participant history of drug addiction, factor; Yes, No; Blank = Missing

Prev.preg Any previous pregnancy, factor; Yes, No; No missing data

N.prev.preg Number of previous pregnancies for those with any previous pregnancy, numeric, range: 1-11; Blank = Missing (variable 21 = Yes) or no previous pregnancies (variable 21 = No)

Live.PTB Previous live preterm birth for those with any previous pregnancy, factor; Yes; No = No previous live preterm birth (variable 21 = Yes) or no previous pregnancies (variable 21 = No)

Any.stillbirth Previous stillbirth, factor; Yes; No = No previous stillbirth (variable 21 = Yes) or no previous pregnancies (variable 21 = No)

Spont.ab Previous spontaneous abortion, factor; Yes; No; Blank = Missing (variable 21 = Yes) or no previous pregnancies (variable 21 = No)

Induced.ab Previous induced abortion, factor; Yes; No; Blank = Missing (variable 21 = Yes) or no previous pregnancies (variable 21 = No)

Any.live.ptb.sb.sp.ab.in.ab Any previous live pre-term birth, stillbirth, spontaneous abortion, or induced abortion, factor; Yes; No = No live pre-term birth/stillbirth/abortion (variable 21 = Yes) or no previous pregnancies (variable 21 = No)

N.living.kids Number of living children the subject had at baseline, numeric, range: 0-9; Blank = Missing (variable 21 = Yes) or no previous pregnancies (variable 21 = No)

Tx.comp. Whether treatment plans were completed by participants in treatment group, factor, Yes = Completed; No = Not completed; Und = Some therapy (unknown whether completed); Blank = Withdrew from treatment (variable 3 = T) or no periodontal therapy (variable 3 = C)

Local.anes Whether any local anesthetic used during periodontal therapy for participants in treatment group, factor, Yes; No = No local anesthetic used or withdrew from treatment (variable 3 = T); Blank = No periodontal therapy (variable 3 = C)

Topical.Anest Whether any topical anesthetic used during periodontal therapy for participants in treatment group, factor, Yes; No = No topical anesthetic used or withdrew from treatment (variable 3 = T); Blank = No periodontal therapy (variable 3 = C)

Tx.time Total treatment time for participants in treatment group (hours), numeric, range: 0.117-5.8; Blank = Withdrew from treatment (variable 3 = T and variable 29 = blank) or no periodontal therapy (variable 3 = C)

EDC.necessary. Whether patient required1 essential dental care (EDC), factor, Yes; No; Blank = Missing

Completed.EDC Did patient complete EDC before 20 weeks gestational age?, factor, Yes; No; Blank = Missing

N.extractions Number of teeth extracted during EDC, numeric, range: 0-20; Blank = Missing

N.perm.restorations Number of permanent restorations carried out as a part of EDC, numeric, range: 0-18; Blank = Missing

N.qualifying.teeth Number of teeth meeting OPT (Obstetrics and Periodontal Therapy Study) criteria for having periodontal disease at baseline, numeric, range: 3.000-28.0

BL.GE Whole-mouth average gingival index at baseline, numeric, range: 0.429-3.0, Silness-Lowe Gingival Index: Higher value indicates more severe inflammation; 0 = Normal gingiva; There are no missing data

BL..BOP Percentage of sites bleeding on probing at baseline, numeric, range:33.951-100.0

BL.PD.avg Whole-mouth average pocket depth at baseline (mm), numeric, range: 1.851-7.0

BL..PD.4 Percentage of sites with pocket depth greater than or equal to 4mm at baseline, numeric, range: 3.571-99.2

BL..PD.5 Percentage of sites with pocket depth greater than or equal to 5mm at baseline, numeric, range: 0-91.7

BL.CAL.avg Whole-mouth average clinical attachment level at baseline (mm), numeric, range: 0.185-5.1

BL..CAL.2 Percentage of sites with clinical attachment level greater than or equal to 2 mm at baseline, numeric, range: 2.381-100.0

BL..CAL.3 Percentage of site with clinical attachment level greater than or equal to 3 mm at baseline, numeric, range: 0-94.9

BL.Calc.I Whole-mouth average calculus index at baseline, Simplified Oral Hygiene Index (OHI-S): Higher value indicates more calculus; 0 = No calculus present; numeric, range: 0-3.0

BL.Pl.I Whole-mouth average plaque index at baseline, Silness-Lowe Gingival Index:Higher value indicates more severe inflammation, 0= normal gingiva, numeric, range: 0.056-3.0

V3.GE Whole-mouth average gingival index at Visit 3, numeric, range: 0.030-3.0

V3..BOP Percentage of sites bleeding on probing at Visit 3, numeric, range: 0.725-100.0, Blank = Missing

V3.PD.avg Whole-mouth average pocket depth at Visit 3 (mm), numeric, range: 1.601-5.5, Blank = Missing

V3..PD.4 Percentage of sites with pocket depth greater than or equal to 4mm at Visit 3, numeric, range: 0-83.9, Blank = Missing

V3..PD.5 Percentage of sites with pocket depth greater than or equal to 5mm at Visit 3, numeric, range: 0-77.4, Blank = Missing

V3.CAL.avg Whole-mouth average clinical attachment level at Visit 3 (mm), numeric, range: 0.036-3.9, Blank = Missing

V3..CAL.2 Percentage of sites with clinical attachment level greater than or equal to 2 mm at visit 3, numeric, range: 0-97.8, Blank = Missing

V3..CAL.3 Percentage of sites with clinical attachment level greater than or equal to 3 mm at visit 3, numeric, range: 0-85.7, Blank = Missing

V3.Calc.I Whole-mouth average calculus index at visit 3, numeric, range: 0-2.6, Simplified Oral Hygiene Index (OHI-S): Higher value indicates more calculus; 0 = No calculus present; Blank = Missing

V3.Pl.I Whole-mouth average plaque index at visit 3, numeric, range: 0-2.6, Silness-Lowe Plaque Index: Higher value indicates more abundant plaque; 0 = No plaque in gingival area; Blank = Missing

V5.GE Whole-mouth average gingival index at visit 5, numeric, range: 0.190-2.7, Silness-Lowe Gingival Index: Higher value indicates more severe inflammation; 0 = Normal gingiva; Blank = Missing

V5..BOP Percentage of sites bleeding on probing at visit 5, numeric, range: 3.571-100.0, Blank = Missing

V5.PD.avg Whole-mouth average pocket depth at visit 5, numeric, range: 1.536-5.4, Blank = Missing

V5..PD.4 Percentage of sites with pocket depth greater than or equal to 4mm at Visit 5, numeric, range: 0-83, Blank = Missing

V5..PD.5 Percentage of sites with pocket depth greater than or equal to 5mm at Visit 3, numeric, range: 0-75.6, Blank = Missing

V5.CAL.avg Whole-mouth average clinical attachment level at visit 5 (mm), numeric, range: 0.018-4.3, Blank = Missing

V5..CAL.2 Percentage of sites with clinical attachment level greater than or equal to 2 mm at visit 5, numeric, range: 0.000-99.2, Blank = Missing

V5..CAL.3 Percentage of sites with clinical attachment level greater than or equal to 3 mm at visit 5, numeric, range: 0.000-85.0, Blank = Missing

V5.Calc.I Whole-mouth average calculus index at visit 5, numeric, range: 0.0-2.6, Simplified Oral Hygiene Index (OHI-S): Higher value indicates more calculus; 0 = No calculus present; Blank = Missing

V5.Pl.I Whole-mouth average plaque index at visit 5, numeric, range: 0.0-2.5, Silness-Lowe Plaque Index: Higher value indicates more abundant plaque; 0 = No plaque in gingival area; Blank = Missing

N.PAL.sites Number of sites for which attachment loss increased from baseline by greater than or equal to 3 mm, numeric, range: 0-33, 0 = No sites; Blank = Missing

Birth.outcome Birth outcome, factor, Elective abortion; Live birth; Lost to FU = Lost to Follow-Up; Non-live birth = Stillbirth or spontaneous abortion; There are no missing data

Preg.ended...37.wk Whether the pregnancy ended before gestational age 37 weeks (259 days), factor, Yes; No; Blank = Lost to Follow-Up

GA.at.outcome Gestational age at end of pregnancy, or at mother's last follow-up visit if lost to follow-up, numeric, range: 103-302

Birthweight Infant birth weight at time of birth, abstracted from obstetrical records (grams), numeric, range: 101-5160, Blank = Missing

Fetal.congenital.anomaly Fetal/congenital anomaly identified at birth or during pregnancy?, factor, Yes; No; There are no missing data

Apgar1 Apgar score, a summary of a newborn infant's 'Appearance, Pulse, Grimace, Activity, Respiration' at 1 minute Score interpretation: less than or equal to 3: Critically low 4-6: Fairly low greater than or equal to 7: Normal, numeric, range: 0-10, Blank = Missing

Apgar5 Apgar score at 5 minutes, numeric, range: 0-10, Blank = Missing

Any.SAE. Whether participant experienced any serious adverse events (e.g. lost pregnancies) factor, Yes; No; There are no missing data

GA...1st.SAE Gestational age of first SAE (serious adverse event), integer, range: 96-467, 259 = No SAE (variable 76 must = No); There are no missing data

`Bact.vag` Whether mother had bacterial vaginosis during pregnancy, factor, Yes; No; Blank = Missing

`Gest.diab` Whether mother had gestational diabetes during pregnancy, factor, Yes; No; Blank = Missing

`Oligo` Whether mother had oligohydramnios during pregnancy, factor, Yes; No; Blank = Missing

`Polyhyd` Whether mother had polyhydramnios during pregnancy, factor, Yes; No; Blank = Missing

`Gonorrhea` Whether mother had gonorrhea during pregnancy, factor, Yes; No; Blank = Missing

`Chlamydia` Whether mother had chlamydia during pregnancy, factor, Yes; No; Blank = Missing

`Strep.B` Whether mother had strep B colonization during pregnancy, factor, Yes; No; Blank = Missing

`Traumatic.Inj` Whether mother had a traumatic injury during pregnancy, factor, Yes; No; Blank = Missing

`UTI` Whether mother had a urinary tract infection during pregnancy, factor, Yes; No; Blank = Missing

`Pre.eclamp` Whether mother had pre-eclampsia, a pregnancy condition characterized by high blood pressure and associated with fetal growth restriction during pregnancy, factor, Yes; No; Blank = Missing

`Mom.HIV.status` HIV status of mother during pregnancy, factor, Yes = HIV-positive; No = HIV-negative or unknown (question answered but HIV status at delivery not recorded); Blank = Missing (question not answered)

`BL.Anti.inf` Did participant report use of antiinflammatory medication at or less than 6 months before baseline?, integer, 0 = No; 1 = Yes; There are no missing data

`BL.Cortico` Did participant report use of corticosteroids at or less than 6 months before baseline?, integer, 0 = No; 1 = Yes; There are no missing data

`BL.Antibio` Did participant report use of antibiotics at or less than 6 months before baseline?, integer, 0 = No; 1 = Yes; There are no missing data

`BL.Bac.vag` Did participant report use of bacterial vaginitis treatments at or less than 6 months before baseline?, integer, 0 = No; 1 = Yes; There are no missing data

`V3.Anti.inf` Did participant report use of antiinflammatory medication between baseline and visit 3?,integer, 0 = No; 1 = Yes; There are no missing data

`V3.Cortico` Did participant report use of corticosteroids between baseline and visit 3?, integer, 0 = No; 1 = Yes; There are no missing data

`V3.Antibio` Did participant report use of antibiotics between baseline and visit 3?, integer, 0 = No; 1 = Yes; There are no missing data

`V3.Bac.vag` Did participant report use of bacterial vaginitis treatments between baseline and visit 3?, integer, 0 = No; 1 = Yes; There are no missing data

`V5.Anti.inf` Did participant report use of antiinflammatory medication between visit 3 and visit 5?, integer, 0 = No; 1 = Yes; There are no missing data

`V5.Cortico` Did participant report use of corticosteroids between visit 3 and visit 5?, integer, 0 = No; 1 = Yes; There are no missing data

`V5.Antibio` Did participant report use of antibiotics between visit 3 and visit 5?, integer, 0 = No; 1 = Yes; There are no missing data

V5.Bac.vag Did participant report use of bacterial vaginitis treatments between visit 3 and visit 5?, integer, 0 = No; 1 = Yes; There are no missing data

X..Vis.Att Visit attendance: Number of study visits attended AFTER baseline, integer, Range: 0-5

X..Vis.Elig Number of visits for which participant was eligible (could become ineligible after miscarriage or early delivery), integer, Range: 0-5

X1st.Miss.Vis First missed visit. No one missed the baseline visit, so this variable takes values 2, 3, 4, 5, 6, and 100 (no eligible visits missed), integer, Range: 2-6, 100

OAA1 Serum IgG (immunoglobulin) antibodies to A. actinomycetemcomitans at baseline, factor (actually numeric or missing), dot(.) = Missing

OCR1 Serum IgG (immunoglobulin) antibodies to C. rectus at baseline, factor (actually numeric or missing), dot(.) = Missing

OFN1 Serum IgG (immunoglobulin) antibodies to F. nucleatum at baseline, factor (actually numeric or missing), dot(.) = Missing

OPG1 Serum IgG (immunoglobulin) antibodies to P. gingivalis at baseline, factor (actually numeric or missing), dot(.) = Missing

OPI1 Serum IgG (immunoglobulin) antibodies to P. intermedia at baseline, factor (actually numeric or missing), dot(.) = Missing

OTD1 Serum IgG (immunoglobulin) antibodies to T. denticola at baseline, factor (actually numeric or missing), dot(.) = Missing

OTF1 Serum IgG (immunoglobulin) antibodies to T. forsythus at baseline, factor (actually numeric or missing), dot(.) = Missing

OCRP1 Serum measure for C-reactive protein (CRP) at baseline, factor (actually numeric or missing), dot(.) = Missing

O1B1 Serum measure for Interleukin(IL)-1b at baseline, factor (actually numeric or missing), dot(.) = Missing

O61 Serum measure for Interleukin(IL)-6 at baseline, factor (actually numeric or missing), dot(.) = Missing

O81 Serum measure forInterleukin(IL)-8 at baseline, factor (actually numeric or missing), dot(.) = Missing

OPGE21 Serum measure for Prostaglandin E2 at baseline, factor (actually numeric or missing), dot(.) = Missing

OTNF1 Serum measure for tumor necrosis factor (TNF)-alpha at baseline, factor (actually numeric or missing), dot(.) = Missing

OMMP91 Serum measure for gelatinase (MMP9) at baseline, factor (actually numeric or missing), dot(.) = Missing

ETXU_CAT1 Serum endotoxin level at baseline, factor (actually numeric or missing), dot(.) = Missing

OFIBRIN1 Serum measure for fibrinogen at baseline, factor (actually numeric or missing), dot(.) = Missing

OAA5 Serum IgG (immunoglobulin) antibodies to A. actinomycetemcomitans at visit 5, factor (actually numeric or missing), dot(.) = Missing

OCR5 Serum IgG (immunoglobulin) antibodies to C. rectus at visit 5, factor (actually numeric or missing), dot(.) = Missing

OFN5 Serum IgG (immunoglobulin) antibodies to F. nucleatum at visit 5, factor (actually numeric or missing), dot(.) = Missing

OPG5 Serum IgG (immunoglobulin) antibodies to P. gingivalis at visit 5, factor (actually numeric or missing), dot(.) = Missing

OPI5 Serum IgG (immunoglobulin) antibodies to P. intermedia at visit 5, factor (actually numeric or missing), dot(.) = Missing

OTD5 Serum IgG (immunoglobulin) antibodies to T. denticola at visit 5, factor (actually numeric or missing), dot(.) = Missing

OTF5 Serum IgG (immunoglobulin) antibodies to T. forsythus at visit 5, factor (actually numeric or missing), dot(.) = Missing

OCRP5 Serum measure for C-reactive protein (CRP) at visit 5, factor (actually numeric or missing), dot(.) = Missing

O1B5 Serum measure for Interleukin(IL)-1b at visit 5, factor (actually numeric or missing), dot(.) = Missing

O65 Serum measure forInterleukin(IL)-6 at visit 5, factor (actually numeric or missing), dot(.) = Missing

O85 Serum measure forInterleukin(IL)-8 at visit 5, factor (actually numeric or missing), dot(.) = Missing

OPGE25 Serum measure for Prostaglandin E2 at visit 5, factor (actually numeric or missing), dot(.) = Missing

OTNF5 Serum measure for tumor necrosis factor (TNF)-alpha at visit 5, factor (actually numeric or missing), dot(.) = Missing

OMMP95 Serum measure for gelatinase (MMP9) at visit 5, factor (actually numeric or missing), dot(.) = Missing

ETXU_CAT5 Serum endotoxin level at visit 5, factor (actually numeric or missing), dot(.) = Missing

OFIBRIN5 Serum measure for fibrinogen at visit 5, factor (actually numeric or missing), dot(.) = Missing

BL.DNA Total amount of bacterial DNA extracted from plaque as a measure of total bacterial concentration at baseline (ng/mL), numeric, range: 0-5750.0

BL.Univ Count of all bacteria detected by universal primer at baseline, numeric, range: 1,890,000-1,070,000,000, Blank = Missing

BL.AA Count of A. actinomycetemcomitans bacteria at baseline, numeric, range: 0-7,970,000, Blank = Missing

BL.PG Count of P. gingivalis bacteria at baseline, numeric, range: 0-167,000,000, Blank = Missing

BL.TD Count of T. denticola bacteria at baseline, numeric, range: 0-50,500,000, Blank = Missing

BL.TF Count of T. forsythus bacteria at baseline, numeric, range: 0-40,200,000, Blank = Missing

BL.PI Count of P. intermedia bacteria at baseline, numeric, range: 0-87,500,000, Blank = Missing

BL.CR Count of C. rectus bacteria at baseline, numeric, range: 0-32,600,000, Blank = Missing

BL.FN Count of F. nucleatum bacteria at baseline, numeric, range: 67,300- 152,000,000, Blank = Missing

BL.S7 Sum of the 7 species-specific bacterial counts (variables 138-144) at baseline, rounded to 3 significant figures, numeric, range: 87,000-391,000,000, Blank = Missing

V5.DNA Total amount of bacterial DNA extracted from plaque as a measure of total bacterial concentration at visit 5 (ng/mL), numeric, range: 0-5750.0

V5.Univ Count of all bacteria detected by universal primer at visit 5, numeric, range: 1,890,000-1,070,000,000, Blank = Missing

V5.AA Count of A. actinomycetemcomitans bacteria at visit 5, numeric, range: 0-40,200,000, Blank = Missing

V5.PG Count of P. gingivalis bacteria at visit 5, numeric, range: 0-40,200,000, Blank = Missing

V5.TD Count of T. forsythus bacteria at visit 5, numeric, range: 0-40,200,000, Blank = Missing

V5.TF Count of T. forsythus bacteria at visit 5, numeric, range: 0-40,200,000, Blank = Missing

V5.PI Count of P. intermedia bacteria at visit 5, numeric, range: 0-87,500,000, Blank = Missing

V5.CR Count of C. rectus bacteria at visit 5, numeric, range: 0-32,600,000, Blank = Missing

V5.FN Count of F. nucleatum bacteria at visit 5, numeric, range: 67,300- 152,000,000, Blank = Missing

V5.S7 Sum of the 7 species-specific bacterial counts (variables 138-144) at visit 5, rounded to 3 significant figures, numeric, range: 87,000-391,000,000, Blank = Missing

BL..AA Percent of A. actinomycetemcomitans out of total DNA (variable 146) at baseline, numeric, range: 0-8.9, Blank = Missing

BL..PG Percent of P. gingivalis out of total DNA at baseline, numeric, range: 0-37.3, Blank = Missing

BL..TD Percent of T. denticola out of total DNA at baseline, numeric, range: 0-13.2, Blank = Missing

BL..TF Percent of T. forsythus out of total DNA at baseline, numeric, range: 0-17.7, Blank = Missing

BL..PI Percent of P. intermedia out of total DNA at baseline, numeric, range: 0-46.3, Blank = Missing

BL..CR Percent of C. rectus out of total DNA at baseline, numeric, range: 0-10.5, Blank = Missing

BL..FN Percent of F. nucleatum out of total DNA at baseline, numeric, range: 0.330-63.2, Blank = Missing

BL..S7 Sum of the percents for the 7 species (AA, PG, TD, TF, PI, CR, and FN) at baseline, numeric, range: 0.420-86.3, Blank = Missing

V5..AA Percent of A. actinomycetemcomitans out of total DNA at visit 5, numeric, range: 0-16.1, Blank = Missing

V5..PG Percent of P. gingivalis out of total DNA at visit 5, numeric, range: 0-59.7, Blank = Missing

V5..TD Percent of T. denticola out of total DNA at visit 5, numeric, range: 0-20.5, Blank = Missing

V5..TF Percent of T. forsythus out of total DNA at visit 5, numeric, range: 0-19.3, Blank = Missing

V5..PI Percent of P. intermedia out of total DNA at visit 5, numeric, range: 0-40.7, Blank = Missing

V5..CR Percent of C. rectus out of total DNA at visit 5, numeric, range: 0-14.6, Blank = Missing

V5..FN Percent of F. nucleatum out of total DNA at visit 5, numeric, range: 0-49.9, Blank = Missing

V5..S7 Sum of the percents for the 7 species (AA, PG, TD, TF, PI, CR, and FN) at visit 5, numeric, range: 2.560-80.8, Blank = Missing

Details

Background::

Randomized Clinical Trial on the Effect of Treatment of Maternal Periodontal Disease Can Reduce Preterm Birth Risk.

Maternal periodontal disease has been linked in observational studies to preterm birth (< 37 weeks) and low birth weight (< 2500 g) outcomes. The Obstetrics and Periodontal Therapy study was a multi-center randomized trial evaluating the effect of nonsurgical periodontal treatment intervention on preterm birth, comparing outcomes of women treated before 21 weeks gestation (treatment) to those treated after delivery (control).

Preterm birth, defined as delivery before 37 weeks of gestation, is a growing problem. In some cases, preterm birth can lead to infant death; in others, its consequences may include neurodevelopmental disabilities, cognitive impairment, and/or respiratory disorders in the child. Many risk factors for preterm birth have already been identified, including maternal age, drug use, and diabetes. However, such factors are exhibited in only about half of preterm birth mothers, highlighting a need to expand our understanding of what contributes to preterm birth risk.

Several observational studies have suggested an association between maternal periodontal disease and preterm birth. Periodontal disease is an inflammatory condition characterized by the destruction of tissue and/or bone around the teeth. A major component of periodontal disease is oral colonization by gram-negative bacteria; systemic release of cytokines and/or lipopolysaccharides from these bacteria may impact fetal condition.

Inoculation of the periodontal pathogen P. gingivalis into pregnant animals does have a dose-dependent effect on birth weight and preterm birth signaling, but no such causal link has been shown in humans, only some associations. Though not definitive, the possibility of a significant relationship raises the question of whether treatment of maternal periodontal disease can decrease preterm birth risk.

Participants::

823 participants enrolled at 4 centers underwent stratified randomization, resulting in 413 women assigned to the treatment group and 410 to control. All participants were 13-16 weeks pregnant at time of randomization (baseline/visit 1) and went on to attend monthly follow-up visits defined as visits 2, 3, 4, and 5 corresponding to gestational age ranges of 17-20, 21-24, 25-28, and 29-32 weeks.

Treatment::

The treatment group received periodontal treatment, oral hygiene instruction, and tooth polishing at their follow-ups, while those assigned to control underwent only brief oral exams. Data collection occurred at visits 1 (baseline), 3, and 5. The primary outcome of interest is gestational age at end of pregnancy. Additional outcomes include birthweight, clinical measures of periodontal disease, and various microbiological and immunological outcomes.

Analysis::

Statistical analyses were carried out on an intent-to-treat basis. Gestational age can be thought of as 'time until end of pregnancy,' for which certain survival analysis methods would be appropriate. The study used a log-rank test stratified by center to compare time until end of pregnancy for treatment and control groups.

A semiparametric proportional hazards model was also used for this purpose and incorporated maternal risk factors as predictors. For the study's main analyses, gestational age was censored at 37 weeks (259 days) because the interest was in extending pregnancies that would otherwise end pre-term, not extending pregnancies generally.
Though not used in the study itself, logistic regression is another method that could be applied: for example, to gestational age, dichotomized as 'preterm' or 'not preterm' according to a gestational age cutoff, or to birthweight dichotomized as 'low' or 'high' at the 2500 g or other cutoff (2500 g would be in keeping with the World Health Organization's definition for low birth weight). Changes in clinical measures of periodontal disease from baseline to visits 3 or 5 could be analyzed using mixed effects linear models. The dataset also features a number of baseline characteristics, which could be compared in treatment and control groups via Student t-tests, Wilcoxon rank sum tests, Fisher's exact tests or Pearson's chi-square tests, as appropriate.

Publishing::

The nonsurgical periodontal treatment involving scaling and root planing induced significant improvements in periodontal health. The study did not however find a significant relation between periodontal treatment and preterm birth risk. The results of this study were published in 2006 by Michalowicz et al., 'Treatment of periodontal disease and the risk of preterm birth', in The New England Journal of Medicine. The Obstetrics and Periodontal Therapy Dataset contains the data used in this study.
The obstetrics and periodontal therapy dataset was contributed by Dr. Ann Brearley, Assistant Professor, Division of Biostatistics, School of Public Health, University of Minnesota and her colleagues. Please refer to this resource as: Meredith Hyun, James S. Hodges and Ann M. Brearley, 'Obstetrics and Periodontal Therapy Dataset', TSHS Resources Portal (2019). Available at https://www.causeweb.org/tshs/obstetrics-and-periodontal-therapy/.

Source

Michalowicz et al., 'Treatment of periodontal disease and the risk of preterm birth', N Engl J Med 2006; 355:1885-1894. DOI: 10.1056/NEJMoa062249

`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) (**more details** available below the variable definitions).

Usage

```
polyps
```

Format

A data frame with 22 observations 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

Details

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. Note that one subject did not return for the 12 month colonoscopy.

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 (**more details** available below the variable definitions).

Usage

```
scurvy
```

Format

A data frame with 12 observations 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

Details

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.

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.

`smartpill` *Prospective Cohort Study of Intestinal Transit using a SmartPill to Compare Trauma Patients to Healthy Volunteers*

Description

This study evaluated gastric emptying, small bowel transit time, and total intestinal transit time in 8 critically ill trauma patients. These data were compared with those obtained in 87 healthy volunteers from a separate trial. Data were obtained with a motility capsule that wirelessly transmitted pH, pressure, and temperature to a recorder attached to each subject's abdomen. Transit times were available for almost all patients, however, pH, pressure and temperature data is missing for all critically ill patients and sparsely missing for the healthy volunteers (**more details** available below the variable definitions)

Usage

```
smartpill
```

Format

A data frame with 95 obsrvations and 22 variables

`Group` Study group, numeric, 0 = Critically Ill Trama Patient, 1 = Healthy Volunteer

`Gender` Gender, numeric, range: 0 = Female, 1 = Male

`Race` Race, numeric, 1 = White, 2 = Black, 3 = Asian/Pacific Islander, 4 = Hispanic, 5 = Other

`Height` Height (centimeters), numeric, range: 132.1-193.0

`Weight` Weight (kilograms), numeric, range: 44.9-127.0

`Age` Age (years), numeric, range: 18.0-72.0

`GE.Time` Gastric Emptying Time is time from ingestion to gastric emptying (hours), numeric, range: 1.7-74.3

`SB.Time` Small Bowel Transit Time is time from gastric emptying to ileocecal junction (hours), numeric, range: 1.8-13.8

`C.Time` Colonic Transit Time is time from ileocecal junction to body exit (hours), numeric, range: 0.7-118.9

`WG.Time` Whole Gut Time is time from ingestion to body exit (hours), numeric, range: 6.0-816.0

`S.Contractions` Stomach contractions are counted if the peak amplitude of the contraction is over 10 mmHg and under 300 mmHg, numeric, range: 47.0-1665.0

`S.Sum.of.Amplitudes` Stomach sum of amplitudes (mm Hg), numeric, range: 655.6-33800.3

`S.Mean.Peak.Amplitude` Stomach mean peak amplitude is the sum of amplitudes divided by number of contractions (mm Hg), numeric, range: 4.6-43.4

`S.Mean.pH` Stomach mean pH is the average pH over the whole recording time in the stomach with normal ~ 1.5-3.5, numeric, range: 1.5-5.9

`SB.Contractions` Small Bowel contractions are counted if the peak amplitude of the contraction is over 10 mmHg and under 300 mmHg, numeric, range: 223.0-2375.0

`SB.Sum.of.Amplitudes` Small Bowel sum of amplitudes (mm Hg), numeric, range:3899.4-41122.5

`SB.Mean.Peak.Amplitude` Small Bowell mean peak amplitude is the sum of amplitudes divided by number of contractions (mm Hg), numeric, range: 15.0-27.9

`SB.Mean.pH` Small Bowel mean pH is the average pH over the whole recording time in the small bowel, normal ~ 6-7.4, numeric, range: 4.7-8.6

`Colon.Contractions` Colon contractions are counted if the peak amplitude of the contraction is over 10 mmHg and under 300 mmHg, numeric, range: 41.0-2672.0

`Colon.Sum.of.Amplitudes` Colon sum of amplitudes (mm Hg), numeric, range:1872.6-117707.5

`C.Mean.Peak.Amplitude` Colon mean peak amplitude is the sum of amplitudes divided by number of contractions (mm Hg), numeric, range: 32.8- 64.2

`C.Mean.pH` Colon mean pH is the average pH over the whole recording time in the colon, normal ~ 5-7-6.7, numeric, range: 3.9-8.1

Details

The Smart Pill dataset was contributed by Dr. Amy Nowacki, Associate Professor, Cleveland Clinic. Please refer to this resource as: Amy S. Nowacki, 'Smart Pill Dataset', TSHS Resources Portal (2017). Available at https://www.causeweb.org/tshs/smart-pill/.

Delayed gastric emptying is a well-known problem in critically ill patients and is associated with feeding disturbances and inadequate nutrition. However, evaluating gastrointestinal function remains challenging in critically ill patients who are mechanically ventilated. Many tests that are practical and accurate under standardized, controlled conditions often fail in the critical care setting. For example, the consensus recommendations for gastric emptying scintigraphy are impractical in intubated patients because they recommend low-fat, egg white meal with imaging at 0, 1, 2, and 4 hours after meal ingestion. Another test, the lactulose hydrogen breath test, relies on prompt bacterial breakdown of lactulose in the colon; however, changes in bacterial flora - which are presumably common in critical care patients - can produce false transit times.

The 13C-octanoic acid breath test was reported as successful when used bedside to measure gastric emptying. However, manometry only assesses the upper gastrointestinal function, mainly esophagus, stomach, and proximal small bowel. Finally, video capsule technology has been used to determine small bowel transit time and pathomorphology in critically ill patients, although inadequate battery lifespan of the capsule (approximately 8-10 hours) could prevent complete examination in some cases.

An alternative technique, wireless capsule technology, may be useful for evaluating gastrointestinal motility in critical care patients. A newly developed motility capsule for assessing gastric emptying in patients with suspected gastroparesis has been available since 2006. It is a wireless capsule that transmits pH, pressure, and temperature.

This study describes the first use of a novel motility capsule to compare gastric emptying and small bowel transit times in critically ill trauma patients with intracranial hemorrhage with times recorded previously in healthy volunteers. Secondly, this study compares critically ill patients and volunteers

on whole-gut transit time.

Source

Rauch et al. 'Use of Wireless Utility Capsule to Determine Gastric Emptying and Small Intestinal Transit Times in Critically Ill Trauma Patients'. Journal of Critical Care 2012; 27(5): 534.e7-534.e12.

`strep_tb` *RCT of 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 (**more details** available below the variable definitions).

Usage

```
strep_tb
```

Format

A data frame with 107 observations 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.

Details

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.

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.

---

supraclavicular *Study of Supraclavicular Anesthesia*

Description

This data set contains 103 patients who were scheduled to undergo an upper extremity procedure suitable for supraclavicular anesthesia. Patients were randomly assigned to either (1) combined

group-ropivacaine and mepivacaine mixture; or (2) sequential group-mepivacaine followed by ropivacaine. A number of demographic and post-op pain medication variables (fentanyl, alfentanil, midazolam) were collected. The primary outcome is time to 4-nerve sensory block onset. The dataset is cleaned and relatively complete. There are no outliers or data problems (**more details** available below the variable definitions).

Usage

```
supraclavicular
```

Format

A data frame with 103 observations and 17 variables

`subject` Subject ID, numeric, range: 1-103

`group` Anesthetic group, numeric, 1 = Mixture; 2 = Sequential

`gender` Gender, numeric, 1 = Male; 0= Female

`bmi` Body mass index (kg/m^2), numeric, range:19-43.5

`age` Age (years), numeric, range:18-74

`fentanyl` Fentanyl pain medication (micrograms), numeric, range: 0-250.0

`alfentanil` Alfentanil pain medication (milligrams), numeric, range: 0-4.3

`midazolam` Midazolam hypnotic-sedative medication, numeric, range: 0-9.0

`onset_sensory` Time to 4 nerve sensory block onset or, if onset_sensory block failed the observed worst outcome of minutes for any patient (50 minutes), numeric, range: 0-50.0

`onset_first_sensory` Time to first sensory block in minutes, or if block failed, a value of 15 minutes, numeric, range: 6-15.0

`onset_motor` Time to complete motor block or, if motor block failed, the observed worst outcome of minutes for any patient (50 minutes), numeric, range: 1-50.0

`nerve_block_censor` block failed, numeric, 0 = nerve block succeeded, 1 = block failed (censored)

`med_duration` Time from the onset of 4 nerve sensory block until the first request for an analgesic medication (hours), numeric, range: 0-48.0

`med_censor` Patients who did not take an analgesic were censored at 48 hours, numeric, 0 = nerve succeeded, 1 = block failed (censored)

`vps_rest` Maximum postop verbal pain score (at rest), on 11 point Likert scale (0-10), numeric, range: 0-10

`vps_movement` Maximum postop verbal pain score (with movement), on 11 point Likert scale (0-10), numeric, range: 0-10

`opioid_total` Total opioid consumption in milligrams, numeric, range: 0-225.0

Details

The choice of anesthetic technique combined with a suitable plan for postoperative analgesia can facilitate early discharge, improve patient comfort, and increase overall satisfaction. Patients having painful procedures who undergo general anesthesia have a 2- to 5-fold greater risk of unplanned overnight admissions compared with those having regional anesthesia. Regional anesthetic techniques and peripheral nerve blocks are especially favored for surgeries on the extremities. Both rapid onset of the block and prolonged postoperative analgesia are desired characteristics of regional anesthesia.
The choice of local anesthetics or combinations thereof can greatly influence the effectiveness of the block, onset time, duration of postoperative analgesia, need for opioid use, and patient satisfaction. Mepivacaine and ropivacaine are commonly used in peripheral nerve blocks, their drawbacks being a short duration with 1.5% mepivacaine and a delayed onset with 0.5% ropivacaine. An ideal local anesthetic with high potency, low toxicity, rapid onset, and prolonged duration does not exist yet. Investigators have therefore tried mixtures of local anesthetics in an attempt to combine their advantages with conflicting results. A potential problem is that mixing drugs dilutes the effects of each. Thus, a mixture of a rapid-onset drug such as mepivacaine with a long-acting one such as ropivacaine may well result in slower onset than mepivacaine alone and shorter duration of action than ropivacaine alone. In contrast, sequential administration of the same amounts of the same drugs may preserve the desirable features of each.
Objective: This study investigates whether sequential supraclavicular injection of 1.5% mepivacaine followed 90 seconds later by 0.5% ropivacaine provides a quicker onset and a longer duration of analgesia than an equidose combination of the 2 local anesthetics.

Source

These are data from a study by Roberman et al. 'Combined Versus Sequential Injection of Mepivacaine and Ropivacaine for Supraclavicular Nerve Blocks'. Reg Anesth Pain Med 2011; 36:145-50.

`theoph` *Cohort Study of the Pharmacokinetics of Oral Theophylline*

Description

Results of a Cohort Study of the Pharmacokinetics of Oral Theophylline, with plasma concentrations over time (**more details** available below the variable definitions).

Usage

```
theoph
```

Format

A data frame with 132 observations and 5 variables

Subject subject id number for each participant; type: ordinal factor

Wt Weight in kilograms; type: double

Dose Dose in milligrams per kilogram; type: double

Time Time from initial dose in hours; type: double

conc Concentration of theophylline in the plasma in micrograms per milliliter' type: double

Details

This data set is from a pharmacokinetic study of oral dosing of the anti-asthma medication, theophylline, in 12 subjects over 25 hours, published By Dr. Robert A. Upton around 1980. The original publication, if any, is unclear and not cited. These data were used in a package named `nlme`, and reported in Boeckmann, A.J., et al.Dr. Upton did publish several papers on theophylline pharmacokinetics around 1980-1984, and these data could have been from one of these.

Theophylline is an methylxanthine anti-asthma medication, which acts as a bronchodilator, with secondary effects to strengthen diaphragm contraction, reduce pulmonary artery pressures, and reduce mast cell release. It can be administered by the intravenous, oral, or rectal suppository routes. Each subject in this Study (oral route) received a single oral dose of theophylline. Blood samples were taken at frequent intervals over the first 25 hours after dosing, and the quantity of theophylline in the plasma at each time point was measured in micrograms per milliliter.

Unfortunately, the theophylline plasma level in blood varies considerably between patients, because of differences in drug clearance, which is affected by body mass, age, smoking, liver and heart function, and viral infections. To complicate this drug further, it has important interactions with a number of other common medicines which can increase or decrease the drug level. Each subject in this study received a single oral dose of 300 mg of theophylline, which has been converted to a milligrams per kilogram dose. Blood samples were taken at frequent intervals over the next 25 hours after dosing, and the quantity of theophylline in the plasma at each time point was measured in micrograms per milliliter of plasma.

Source

Boeckmann, A. J., Sheiner, L. B. and Beal, S. L. (1994), NONMEM Users Guide: Part V, NONMEM Project Group, University of California, San Francisco. Note that the original data collector, Robert A. Upton, is not credited, nor is the original work cited.

Index