Academic Station

Intro

* Authors, Journal
* Journal impact factor
* Multi centre authors
* Relevant topic
* Type of study and Oxford hierarchy

Aims

* General
* Primary outcome (hard / soft)
* Secondary outcomes
* Multiple outcomes (data dredging, increases type 1 error)

Methods

* Single / multicentre
* Inclusion and exclusion criteria
* For RCTS – CONSTORT, randomisation, blinding, study size
* For Meta-analysis – PRISMA, fixed or random effects

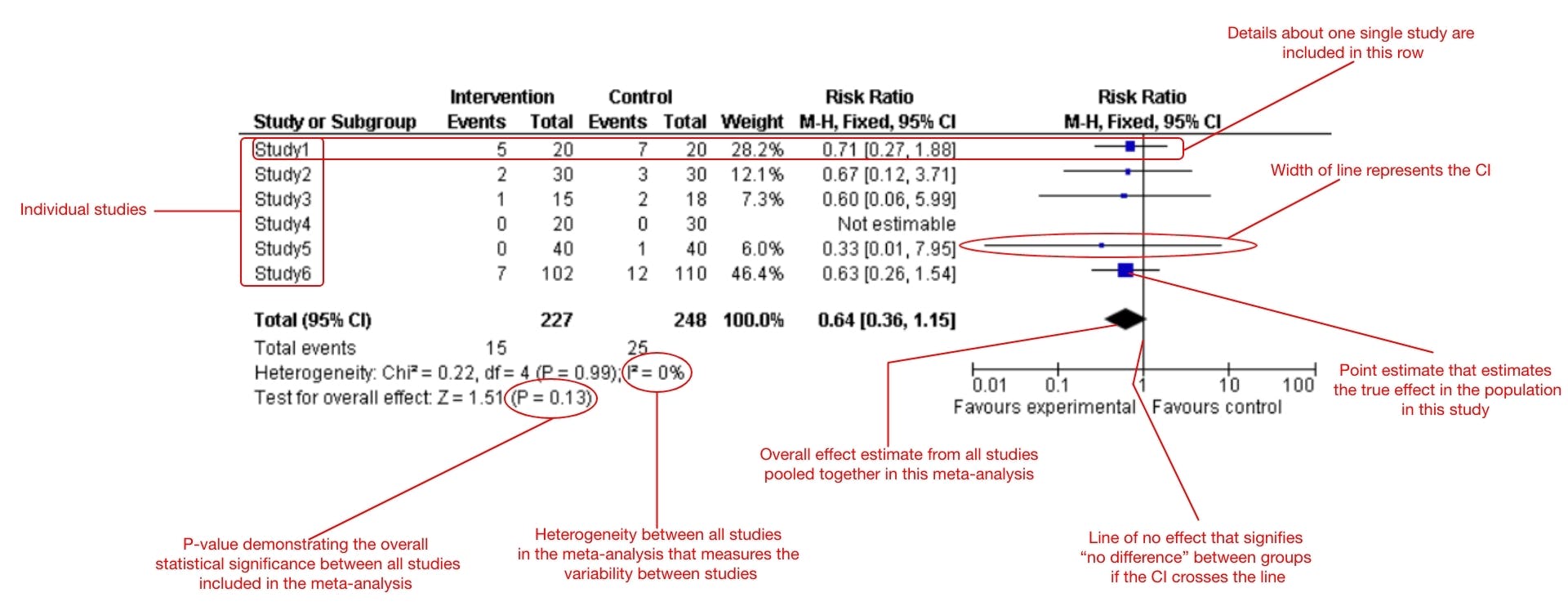
Results

* Number of patients
* Number in each arm (effective randomisation?)
* Table 1 baseline demographics (p value differences, effective randomisation?)
* Length of follow up
* Mean / RR for RCT or cohort
* OR for case control
* P values (significant?)
* Confidence interval (significant?)
* Intention to treat or per protocol
* Hetergeneity
* Publication bias

Conclusion

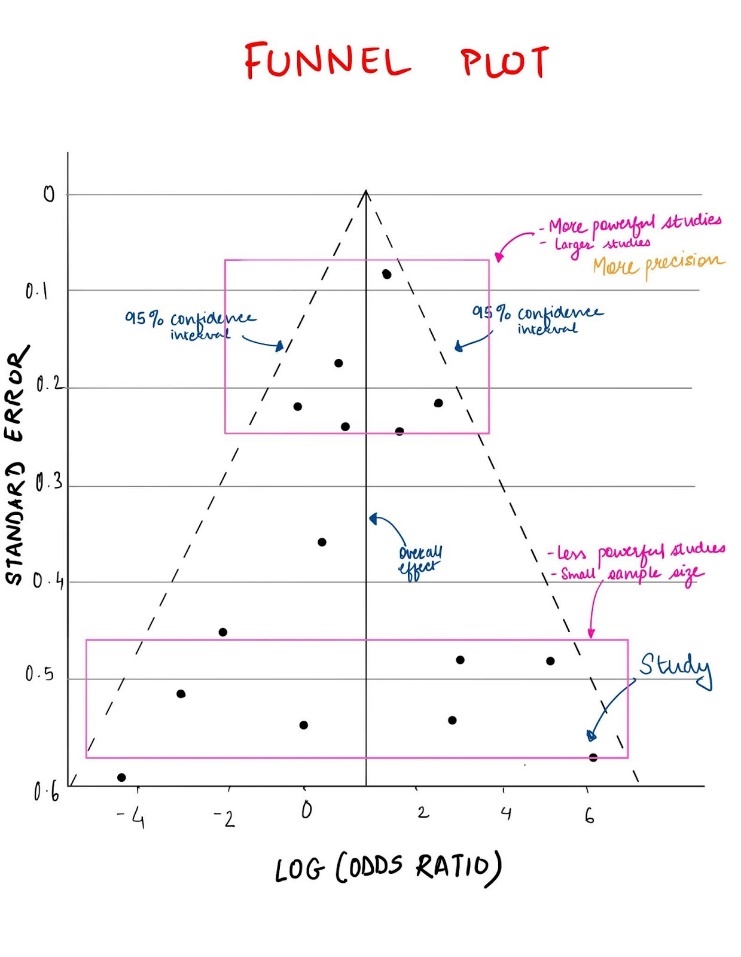
* What did they conclude
* Do you think it is a good paper
* Would it change your practice
  + Does the sample population match your population
  + What does the other evidence say
* How would you improve it
* Further areas for research

|  |  |
| --- | --- |
| Pros | Cons |
| Multiple authors, multiple academic institutes  High IF journal  Meta-analysis (improved accuracy of results)  RCT (randomisation reduces selection bias and confounding factors, blinding reduces observation bias)  Registered trial  Multicentre recruitment (increases external validity)  One relevant hard primary outcome  Large study (narrow confidence interval, reduces type 2 error)  Statistically significant resuls  Intention to treat analysis (reduces attrition bias)  Random effects model (allows for heterogeneity)  Homogeneity | English language only (publication bias)  Multiple outcomes (Data dredging, type 1 errror)  Short follow up  Not statistically significant  Per protocol analysis  Heterogeneity  Publication bias  Competing interest |
| Suggestions to improve paper   * More centres * Randomisation, blinding * Increase sample size * Clinical end point (not surrogate marker) * Longer follow up * Cost analysis | |

Forest plot

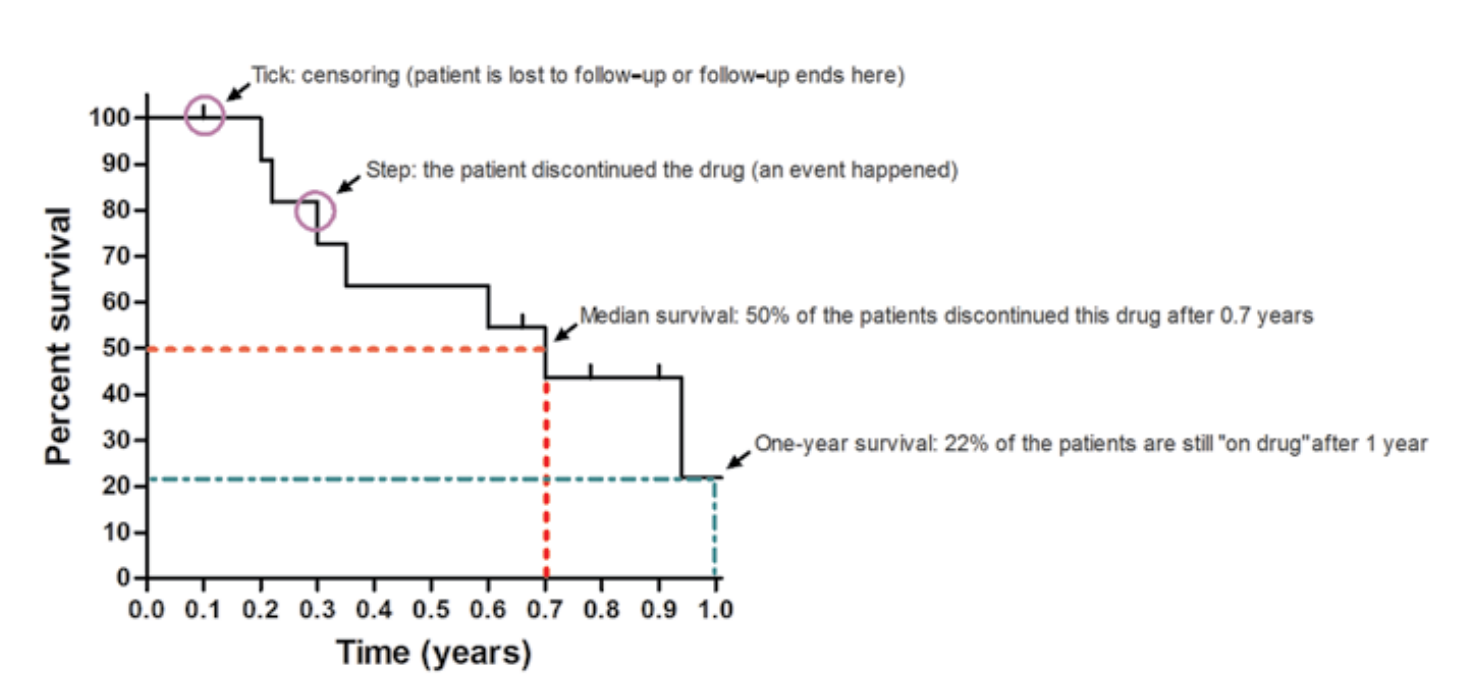
* List of studies arranged by size
* Outcome measure on horizontal axis (OR, RR, HR)
* Line of no effect vertically
* Area of each study box proportional to weight (usually by size)
* 95% confidence intervals show by line
* If 95% CI crosses line of no effect = not statistically significant, sample size too small
* Confidence intervals between studies not overlapping suggests heterogeneity
* Diamond is the pooled outcome
  + Centre is the estimated pooled result
  + Width is the 95% confidence interval
* Heterogeneity
  + I2
  + 0-50 low (fixed effects)
  + >50 medium, 80 high (use random effects – good to use as assumes heterogeneity)
* Meta regression analysis
  + Adjusts for variables at the level of the study rather than at the level of the patient
  + Relates size of a treatment effect to factors within the study rather than for all the studies

Funnel plot



* Used to assess publication bias (tendency to include larger, positive effect studies)
* Study size vertically (smaller studies at the bottom, higher standard error)
* Treatment effect horizontally
* Each dot represents a study
* Hollow dots are imputed studies
* Plot should be symmetrical if there is no publication bias
* Dealing with publication bias
  + Delete studies
  + Impute studies
  + Trim and fill

Kaplan Meier curve



* Looks at cumulative event probability (usually survival)
* Median survival time (time taken until 50% of the population survive)
* Survival time (time taken for a certain proportion of the population to survive)
* Survival probability (at a given time point what is the probability that an individual will have survived)
* Tick censoring (patient drops out of lost to follow up), allows their data to be included up to the time they survived
* Veritical step is each time an event occurs e.g. patient dies
* Probability of an event is recalculated each time an event occurs

|  |  |
| --- | --- |
| Common journals IF | Lancet 200  NEJM 176  JAMA 157  BMJ 96  Nature 69  Annals of surgery 13  BJS 6.9  International Journal Surgery 13  World journal surgery 3  American journal surgery 3  **Colorectal disease 3.9** |
| Impact Factor | 2022 = no citations in 2022 to articles published in 21/20 / total number of articles published in 21/20 |
| H index | Cumulative impact of an author's scholarly output and performance; measures quantity with quality by comparing publications to citations |
| Oxford hierarchy | 1a SR / MA of RCTs  1b single RCT  2a SR / MA of cohort  2b Cohort  3 Case control (or SR/MA of case control)  4 Case series  5 Expert opinion |
| Guidelines for reporting | PRISMA SR / MA  CONSORT RCT  STROBE Observational (cohort, case control)  MOOSE Observational meta-analysis  STARD Diagnostic accuracy |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-analyses |
| Internal validity | To what extent did the study show what it set out to |
| External validity | Generalizability of the results to wider population |
| Confounding factors | A factor that distorts the relationship between exposure and outcome and leads to **wrong conclusions** |
| Bias | A mistake in the study process that leads to **wrong results** |
| List the types of bias, examples and how to prevent them | Selection bias (randomisation and concealed allocation)  Observation bias (blinding)  Publication (funnel plot)  Attrition (intention to treat)  Recall bias (seen in case controls) |
| Observational studies | Cohort (look prospectively, use RR)  + reduces recall bias  + reduces confusion about causality  - time consuming  - expensive  - long follow up  Case control (look retrospectively, use odds ratio)  + quick and cheap  - recall bias present  - uncertain causality |
| Experimental trial | RCT gold standard  Uses relative risk |
| Sampling methods | Simple random  Systematic e.g. every nth person selected for trial |
| Randomisation | Reduces selection bias  Fixed – simple (random number generation), block (to ensure equal numbers into each arm)  Adaptive – randomisation method changes throughout trial  Concealed allocation – researchers cannot predict which group patient will get allocated to due to the randomisation process |
| Blinding | Reduces observation bias  Open label no blinding  Single – either patient or researcher  Double – patient and researcher  Triple – patient, researcher, analyst/statistician  Hawthorne effect – patients alter their behaviour because they are aware hey are being observed |
| Surrogate endpoint | Not a clinical outcome  e.g. tumour shrinkage for survival |
| Accuracy and precision | Accuracy – how close to true value  Precision – how close repeat measurements are to each other |
| Incidence | Number of new cases per population size over a period of time |
| Prevalence | Number of current caseses per population size over a period of time  Affected by disease duration |
| Data types | Qualitative  Quantitative (Discrete e.g. number. days, continuous e.g. height)  Data distribution  Normal = Gaussian = parametric  Symmetrical distribution, mean=median=mode  99% within 3 SD’s of mean  **95% within 2 SD’s of mean**  68% within 1 SD of mean  Non-normal = non parametric  e.g. length of stay, QOL   |  |  |  | | --- | --- | --- | | **Qualitative** | **Quantitative** | | |  | **Parametric** | **Non-parametric** | | ꭓ2 | T test (paired and unpaired) | Sign | | Fishers exact | ANOVA | Mann-Whitney U | | McNemars (paired) |  | Wilcoxon (paired) | |  |  | ANOVA |   Paired data – same patients, two different measurements e.g. pre and post chemo |
| Mean  Mode  Median  Inter quartile range  Standard deviation  Standard error | Mean and median are measures of central tendency  Mean – sum of value / no. of values, used for NORMAL  Mode – most common value  Median – middle value (50th centile), NON-NORMAL  IQR – measure of spread of middle 50% of values  Difference between 1st and 3rd quartiles  Robust to outliers  Standard deviation – degree of spread of data about the mean. A measure of PRECISION  Standard error – measure of how variable the mean is. A measure of PRECISION  SE = SD / square root sample size |
| Confidence interval | A measure of precision  The range in which we can be certain that the true result lies  The range in which we are 95% sure the true result lies  Absolute value = 0 or OR/RR = 1 means NO statistical significance |
| Intention to treat | Analysis method of choice  All subjects who were randomised are included in the results regardless if they dropped out |
| Per protocol analysis | Only data from subjects who completed the study are analysed  Gives a true treatment effect BUT attrition bias |
| Relative risk | For cohort  Risk in experimental / risk in control  >1 then in increased risk in experimental  e..g. RR 2 twice the risk  RR 0.5 half the risk  ARR risk control – risk experimental  RRR risk control-risk experimental / risk control  NNT 1 / ARR |
| Odds ratio | For case control  Odds in experimental / odds in control  >1 then in increased risk in experimental |
| Null hypothesis | States that any difference between the results is due to chance  Alpha level – threshold below which results are unlikely due to chance. Often set to 0.05 (5% probability the results occurred by chance) |
| P value | The probability that differences occurred by chance  With alpha set to 0.05  P<0.05 indicates statistical significance and the null hypotheseis is rejected |
| Type 1 error  Type 2 error  Power | Type 1 error  (False positive) A difference does not exist but the study shows one, usually due to data dredging  Incorrectly rejects null hypothesis  Larger alpha = type 1 error more likely  Type 2 error  (False negative) A difference exists but the study fails to show it, usually due to small sample size  Fails to reject null hypothesis  Power  The probability that a type 2 error will NOT be made (1-beta)  Ranges from 0-1, 0.8 accetable for most studies (80% chance of no type 2 error)  Measure of the studies ability to detect small differences, if a difference exists  Power depends on alpha, sample size, standard deviation of outcome |
| Correlation | The association between two variables |
| Correlation coefficient | The strength of association between two variables  R (-1 to 1)  R = 0 no correlation  R = -1/1 perfect correlation |
| Regression | Expressing the relationship between >2 variables |
| Linear regression | Relationship between a single independent variable and a continuous dependent variable  e.g. y = mx + b  where m is the slope |
| Multivariate regression | Relationship between multiple independent variables and a continuous dependent variable  Allows you to deal with multiple confounding factors |
| Logistic regression | Relationship between multiple independent variables and a BINARY dependent variable |
| Diagnostic studies  Sensitivity  Specificity  PPV  NPV | Sensitivity – if you have the disease, the chance you have a positive test  e.g. FIT 97% (positive in 97% of colorectal cancers)  Specificity – if you don’t have the disease, the chance you have a negative test  PPV – if you have a positive test, the chance you have the disease  e.g. FIT PPV 10% (there is a 10% chance of cancer if you have a positive test)  NPV – if you have a negative test, the chance you don’t have the disease (rule out)  e.g. FIT 99% (negative test means 99% sure you don’t have cancer) |
| Draw 2 x 2 table | |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  |  | DISEASE | |  |  | | TEST |  | + | - |  |  | | + | TP | FP | PPV |  | |  | - | FN | TN | NPV |  | |  |  | Sensitivity | Specificity |  |  | |
| Receiver operator curve | Sensitivity y axis, 1-specificity x axis (true positives against false positives  To find the optimum cut off point for a test  Receiver operating characteristic - WikipediaAUC 1 = perfect test  Line of unity – test is no better than chance |
| Hazard rate  Hazard ratio  Log rank test  Cox proportional hazards regression | Probability of an event occuring / duration of the time interval  e.g. death per time period  Cumulative over the entire study, therefore varies throughout study  Hazard rate experimental / hazard ratio control  >1 experimental arm has increased hazard ratio  Significance test for survival difference between two groups  Multivariate version of log rank used to assess survival difference |

FRCS statistics definitions

* Internal validity
  + To what extent does the study measure what it set out to
* External validity
  + To what extent are the results generalisable to other populations
* Impact factor
  + No citations in 2022 to papers published in 2020+2021
  + Total number articles published in 2020+2021
* Confounders
  + A factor that has a relationship between exposure and outcome but is not the actual cause
  + Eliminate / randomise to match confounders between groups / use multivariate regression analysis
* Bias
  + Mistakes that lead to wrong results
  + Selection
    - Randomisation
  + Observation
    - Blinding
  + Publication
    - Funnel plot
  + Attrition
    - Intention to treat (all subjects regardless of completion included)
* Forest plot
  + Heterogeneity (I2 index)
  + Forest plot CI don’t overlap
  + 0-50 low used fixed effects
  + 50-75 high used random effects
* Funnel plot
  + Eggers test (funnel plot asymmetry)
  + Study size (precision) on y axis, treatment effect on x
  + Asymmetry at the wide part = publication bias
* Observational studies
  + Cohort (prospective) RR
  + Case control (retrospective) OR
* Blinding
  + Single (either researcher or patient)
  + Double (researcher and patient)
  + Triple (researcher patient statistician)
* Incidence
  + New cases (risk of disease)
  + Number of new cases over a period of time per population size
* Prevalence
  + Current cases (relates to duration of disease)
  + Number of patients with a disease at a given time per population size
* Data types
  + Non numerical / categorical / qualitative
  + Numerical / quantitative (either discrete or continuous)
* Data distribution
  + Parametric = normal = Gaussian
  + Non parametric
* Parametric
  + Symmetrical about mean
  + Mean = median = mode
  + 95% data within 2 SD’s
  + 68% within 1 SD
  + 99% within 3 SDs
* Standard deviation
  + Measure of spread about the mean
* Standard error
  + Measure of how variable the mean is
* Non-parametric = non normal
  + Median and interquartile range
  + IQR is the spread of the middle 50% of values
* Confidence interval
  + Range in which 95% of the data lies
  + RR or OR = 1 no difference
* Relative risk
  + RCT or Cohort prospective
  + Outcome in experimental / outcome in control
  + >1 risk more in experimental
  + 2 twice the risk
  + 0.5 half the risk
* Odds ratio
  + Case control retrospective
  + Odds in experimental / odds in control
  + >1 risk more in experimental
* Number needed to treat
  + Number of patients needed to treat to have ONE beneficial outcome
  + 1/ARR
* Null hypothesis
  + No difference between the two groups
  + Any difference between the two groups is due to chance
* Alpha
  + Threshold below which results unlikely due to chance
  + 0.05 is convention (results would occur by chance 5% of the time)
* P value
  + Probability that the events occurred by chance
  + P<0.05 statistical significance and null hypothesis rejected
* Type 1 error
  + False positive
  + Shows there is a difference when actually there is not
  + Wrongful rejection of null hypothesis
  + BIAS + DATA DREDGING + CONFOUNDERS
* Type 2 error
  + False negative
  + There is a difference but the study fails to show it
  + Wrongful acceptance of null hypothesis
  + SMALL SAMPLE SIZE
* Power
  + Probability that a type 2 error will not be made (1-beta)
  + Determines sample size
  + 0.8 acceptable
  + 80% probability of no type 2 error
* Correlation
  + Strength of association
* Sensitivity
  + If you have the **disease** how likely is it the test will be positive
  + TP/TP+FN
* Sensitivity
  + If you don’t have the **disease** how likely is it the test will be negative
  + TN/TN+FP
* Positive predictive value
  + If the **test** is positive how likely is it you have the disease
  + TP/TP+FP
* Negative predictive value
  + If the **test** is negative how likely is it you don’t have the disease
  + TN/TN+FN
* Kaplan Meier curve
  + Displays cumulative survival probability
  + Steps when an event occurs
  + Ticks are censors (drop out, loss to follow up or die)
  + Median survival time
    - Time at which 50% of the population are alive
* Log rank test
  + Tests whether differences in survival times between two groups are statistically significant
* Cox proportional hazards regression
  + Used to test effect of other variables on survival times of the groups
  + Multivariate analysis
* Hazard
  + Probability of the endpoint in a time interval / duration of time interval
* Hazards ratio
  + Hazard in experimental / hazard in control
  + >1 hazard greater in experimental