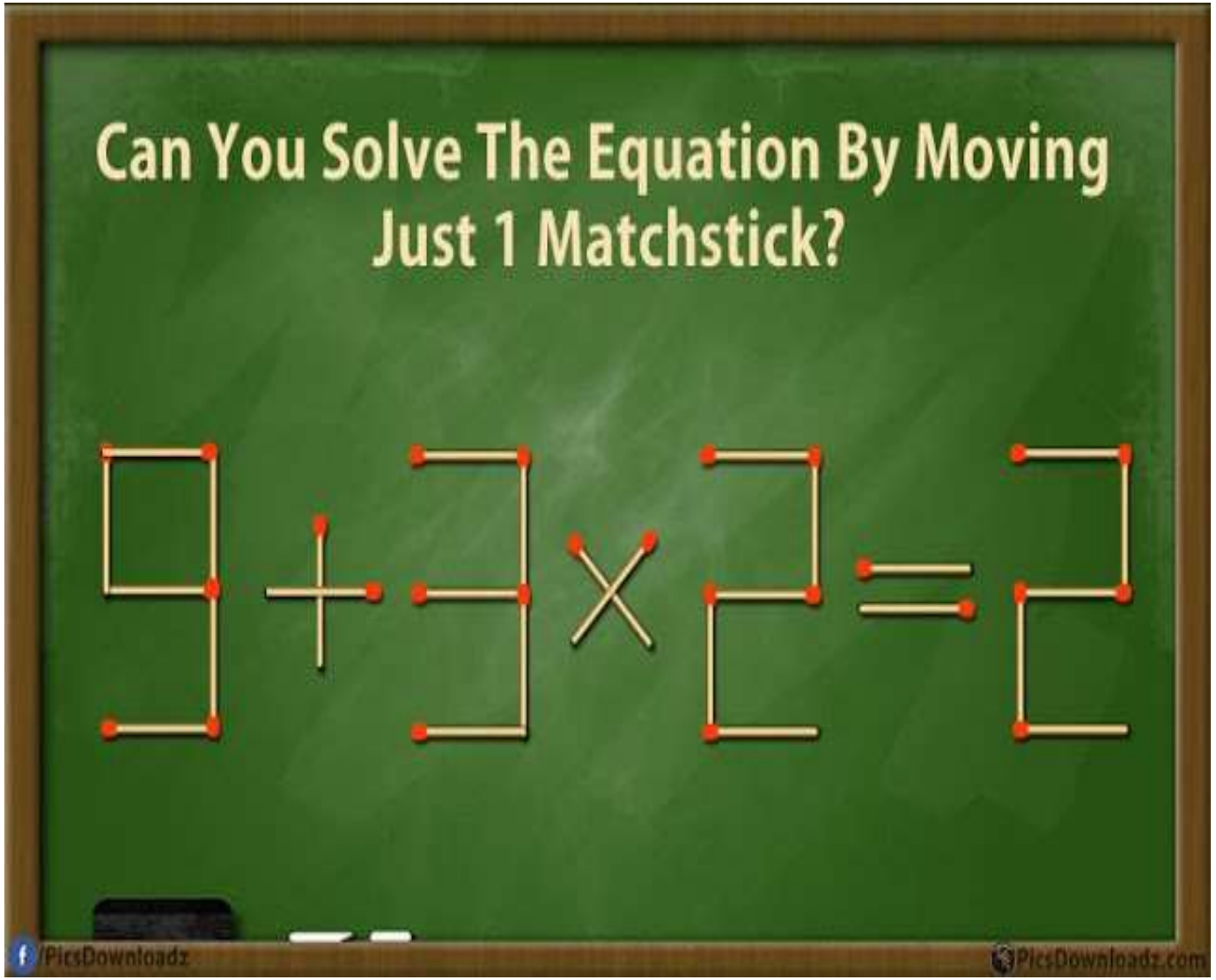

Can You Solve The Equation By Moving Just 1 Matchstick?
9+3×2=2
/PicsDownloadz
PicsDownloadz.com

# Disclosure Belangen Spreker

| | |
|---|---|
| **Geen (potentiële) belangenverstengeling** | |
| **Voor bijeenkomst mogelijk relevante relaties:** | **Bedrijfsnamen:** |
| • Sponsoring of onderzoeksgeld | - |
| • Honorarium of andere (financiële ) vergoedingen | - |
| • Aandeelhouder | - |
| • Andere relatie, namelijk: | - |
| | |

# Statistiek I

Jan Binnekade

# AMC

Together into the future
De sleutel naar gezondheid
V&VN VS

# Wat gaan we doen

- Achtergrond kennis
- Eigenschappen van
- data
- Statistiek
- Tips om het zelf te leren

Statistiek is de wetenschap, de methodiek en de techniek van het verzamelen, bewerken, interpreteren en presenteren van gegevens.

*Onderdeel van de wiskunde*

Theorie – onderzoek – ontwikkeling en testen – toepassen – begrijpen

Twee technieken: beschrijven en verklaren

Wetenschap: verpleegkundige theorieën - kennis verzamelen

waarom vraag

Technologie: patiëntgebonden onderzoek - met wetenschappelijke methoden oplossingen vinden voor praktische problemen

hoe vraag

Techniek: verpleegkundige zorg - de juiste toepassing van vaardigheden

uitvoering (bron voor wetenschap)

# Statistiek is een instrument van de wetenschap

**Wetenschap**

systematisch verkregen en geordende kennis proces van kennisverwerving de gemeenschap

**In je eigen vak**

onderzoek ontwerpen en uitvoeren, lezen van relevante publicaties, begrijpen van studies, toepassen van nieuwe informatie

Publicaties

IMRAD (Introduction, Methods, Results and Discussion)

Introduction: achtergrond van het onderzoek, beschrijving probleem, doelstelling/vraagstelling van het onderzoek

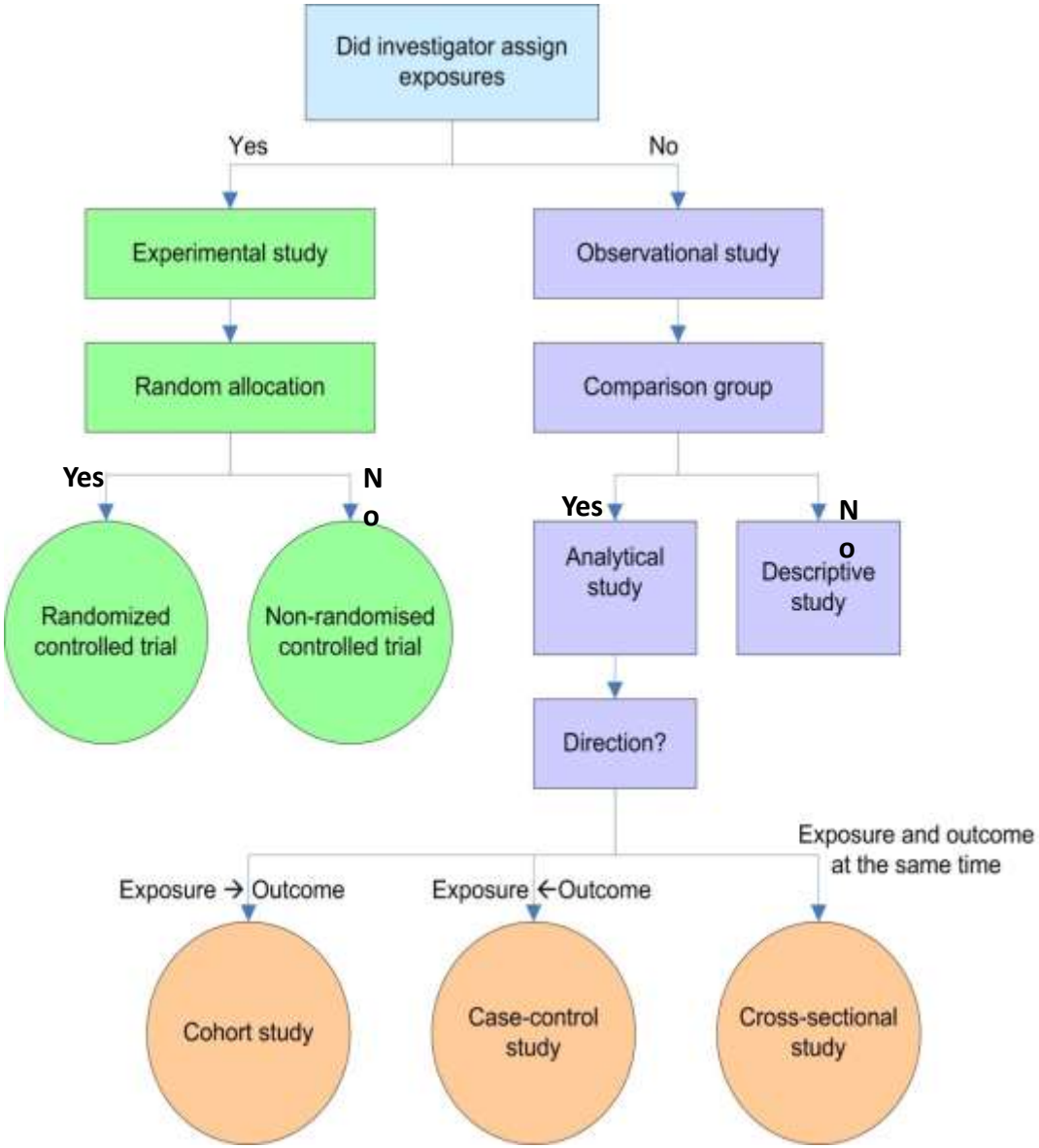
Methods: hoe is het onderzoek uitgevoerd, populatie, technieken,

analyse: statistische paragraaf (study design -> statistiek)

Results: resultaten in getallen, grafieken en afbeeldingen (geen interpretatie hiervan)

Discussion: samenvatting van de resultaten, interpretatie van de resultaten vergeleken met vergelijkbare studies, zwakke punten van het onderzoek, conclusie.

Abstract: samenvatting van de studie




Case reports
*anekdotal*
Case series
*series of anekdotes*

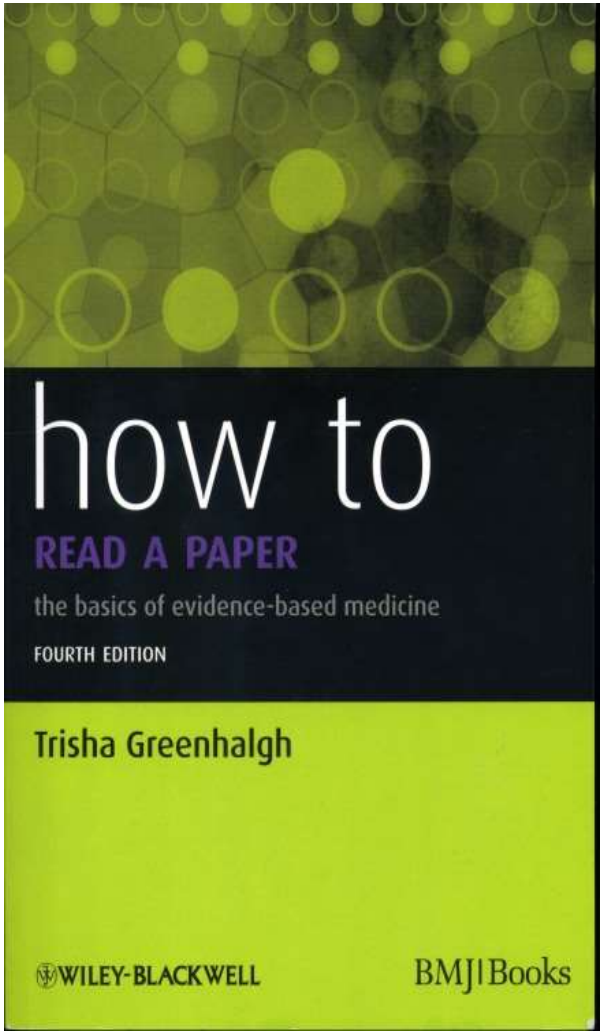
Cross sectional studies
*a snapshot in time*
Case control studies
*thinking backwards*
Cohort studies
*looking forward in time*

Randomised controlled trials *golden standard*

Non-randomised controlled trials *not so gold*

## 3.1 The science of 'trashing' papers

It usually comes as a surprise to students to learn that some (the purists would say up to 99% of) published articles belong in the bin, and should certainly not be used to inform practice. In 1979, the editor of the *BMJ*, Dr Stephen Lock, wrote 'Few things are more dispiriting to a medical editor than having to reject a paper based on a good idea but with irremediable flaws in the methods used'. Fifteen years later Doug Altman was still claiming that only 1% of medical research was free of methodological flaws;[1] and more recently he confirmed that serious and fundamental flaws commonly occur even in papers published in 'quality' journals.[2] Box 3.1



**Basis:  structuur van data**

atomaire eenheden = gebruik de kleinste eenheid van informatie, dus niet verder deelbaar in zinvolle eenheden, per variabele

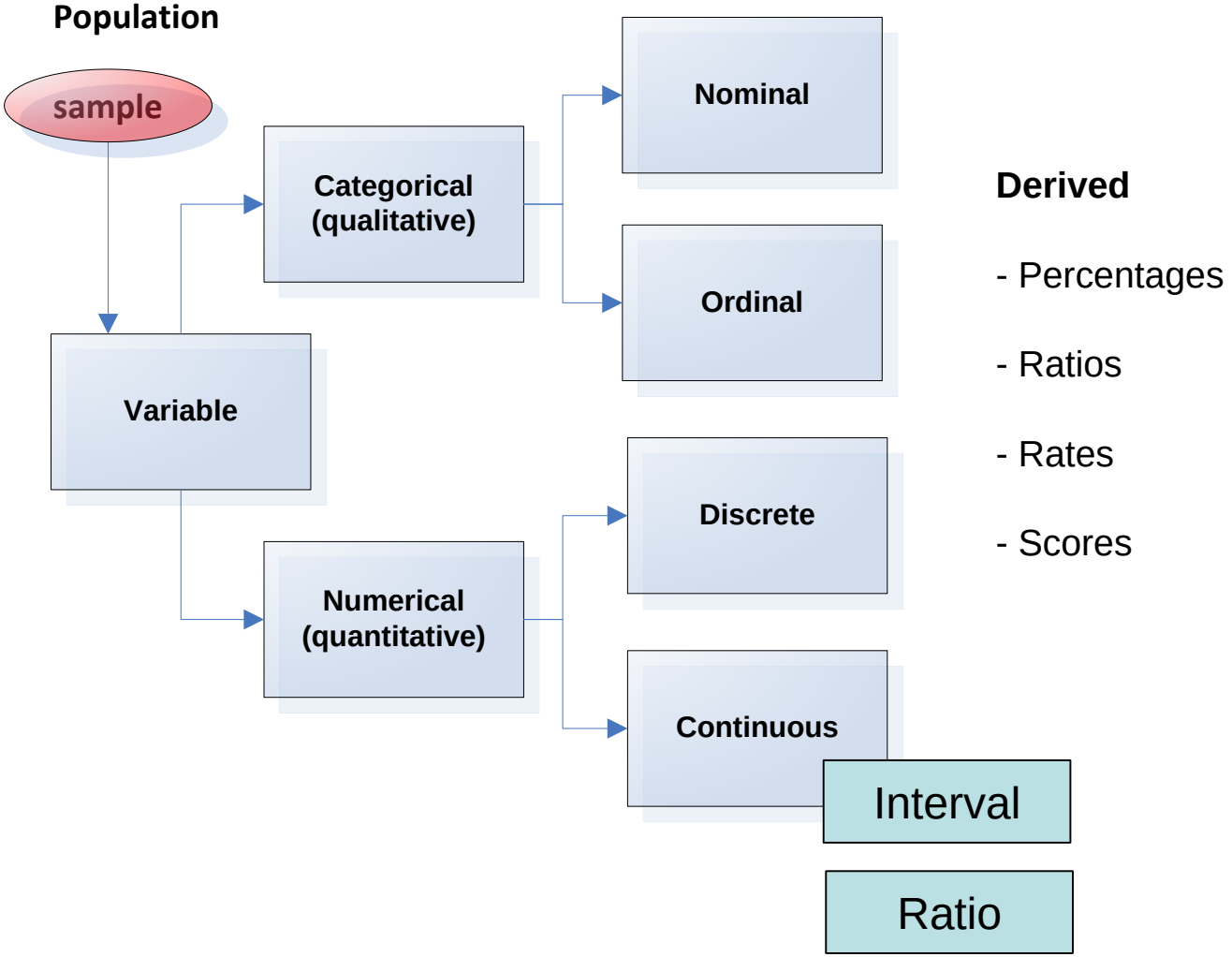
een variabele heeft eigenschappen

deze eigenschappen bepalen het type variabele ofwel het gegevenstype

welke vraag wil je beantwoorden

**Collect data**

**Population**




**Derived**

- Percentages
- Ratios
- Rates
- Scores

Onderzoek de data: meetniveaus

nominal

namen, labels, bijv: bloedgroepen

ordinal

ontevreden, neutraal, tevreden, zeer tevreden

interval

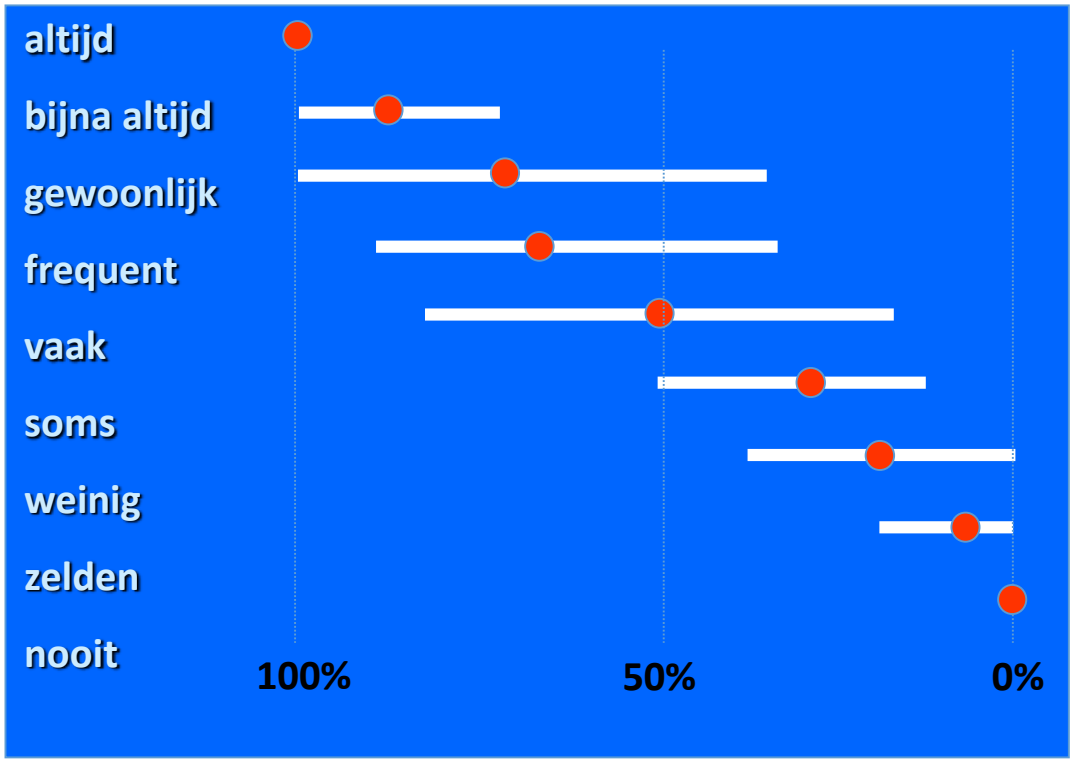
geen gedefinieerd (echt) nulpunt (Celsius,  Fahrenheit)

ratio

gedefinieerd nulpunt, de afwezigheid van iets

Continu  |  discreet

# Meten – operationalisatie - eindpunten




Gemiddelde kwantitatieve scores van veel gebruikte bijwoorden (Lancet 1980)

Meten

*Informatieverlies*
Continue maat met discrete schaal meten

*Error*
X (observatie) = T (score) + E (error)

# Meten – operationalisatie - eindpunten

| *Voorbeeld: Van de 25 patienten kregen 5 een delier* | |
|---|---|
| **Ratio** | De ratio van delier tot geen delier is 5:25 |
| **Proportie** | De proportie van patienten met een delier is 5/25 of 0,2 |
| **Percentage** | Het percentage delier is (5/25)* 100 = 20% |
| **Rate is verbonden aan een tijd** | In 1-maand, 5 van de 25 patienten een delier, is een rate van 20% |
| **Toename of afname percentage** | [eind waarde - begin waarde]/begin waarde maal 100 |

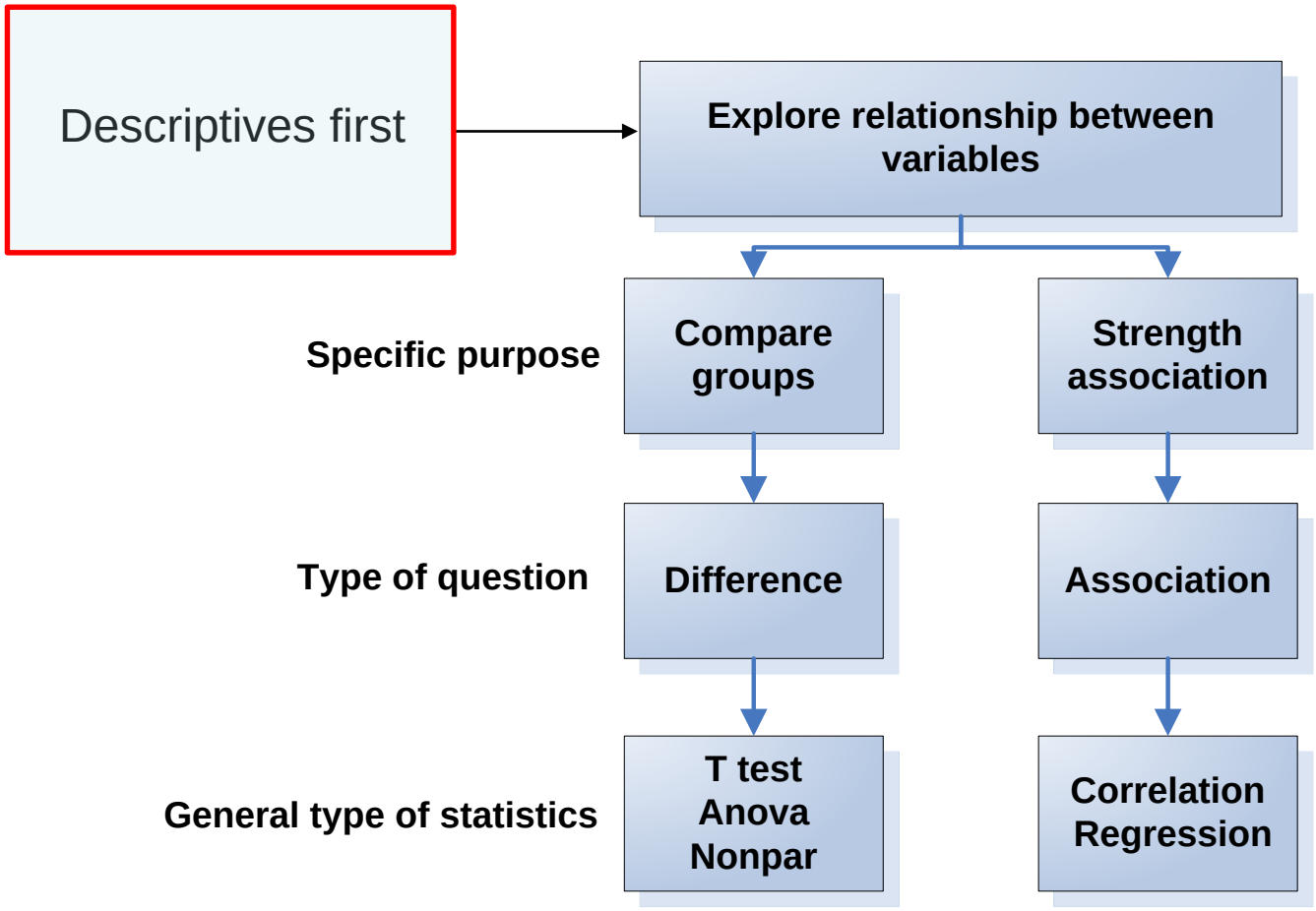
# Studiedesign: vraagstelling – operationalisatie -> uitkomstmaat

Dieet vs gedragstherapie
- afname gewicht = kwantitatief, continu

Supportgroep vs geen supportgroep  chronisch zieke patiënten
- frequentie huisartsbezoek = kwantitatief, discreet, tellen

Chemotherapy vs chirurgie + chemotherapie
- 5 jaar overleving (dood / leven) = logisch, dichotoom

Spreekuur met arts of verpleegkundige
- patiënten tevredenheid = categorie (analyse afhankelijk van aantal categorieën)


Descriptives first
Explore relationship between variables
Specific purpose
Compare groups
Strength association
Type of question
Difference
Association
General type of statistics
T test
Anova
Nonpar
Correlation
Regression

Beschrijven van de dataset

Belangrijke vraag voor iedere continue variabele:

is de data normaal verdeeld?

ja…

**mean (standaard deviatie)**

nee…

**mediaan (interkwartiel range)**

## Wat is een mean (standaard deviatie)

$$\frac{10+12+23+15+20}{5} = 80/5 = 16$$

## Standaard deviatie

$$\text{SQRT}((10-16)^2 + (12-16)^2 + (23-16)^2 + (15-16)^2 + (20-16)^2)$$

$$36 + 16 + 49 + 1 + 36 = 138$$

$$138/5 = 27.6$$

$$\text{SQRT}(27.6) = 5.4$$

# Wat is een mediaan (Interquartiel range)

Sorteer observaties: 10, 12, 23, 15, 20, 8, 30

8, 10, 12, 15, 20, 23, 30

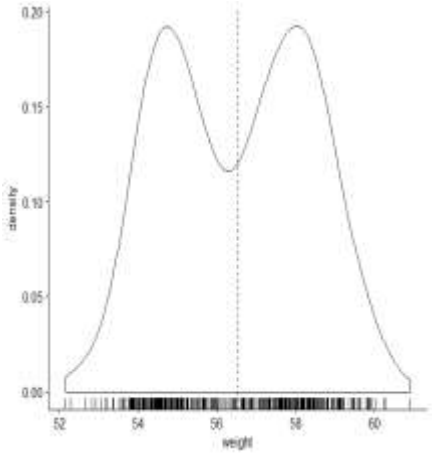
Middelste observatie [8, 10, 12, 15, 20, 23, 30] is de mediaan

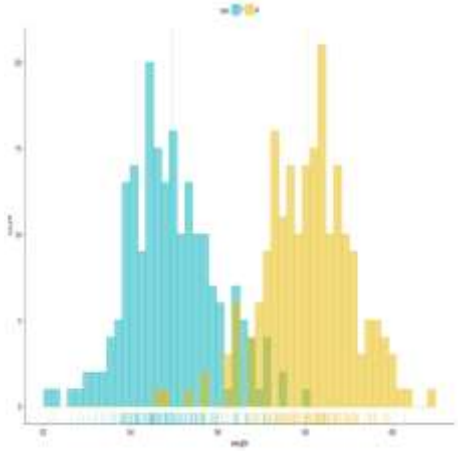
Mediaan (Iqr) = 15 (Iqr 11 – 21,5)

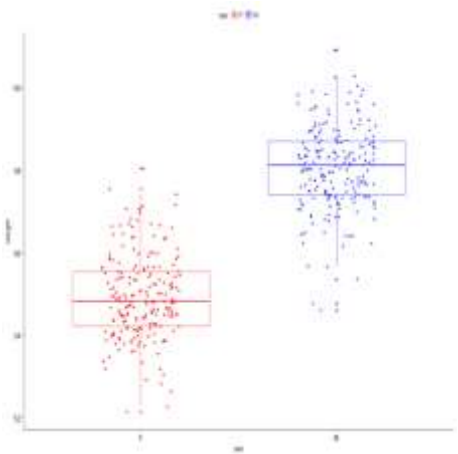
Mean (sd) = 16,9 (7,9)

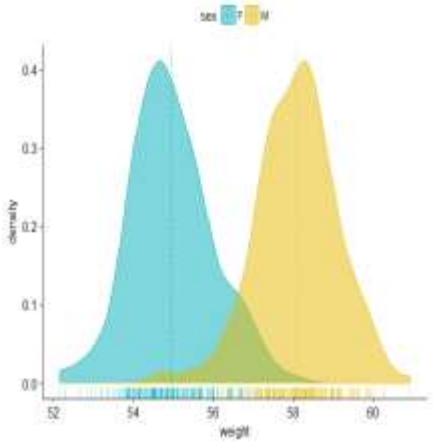
A picture says more than a …….

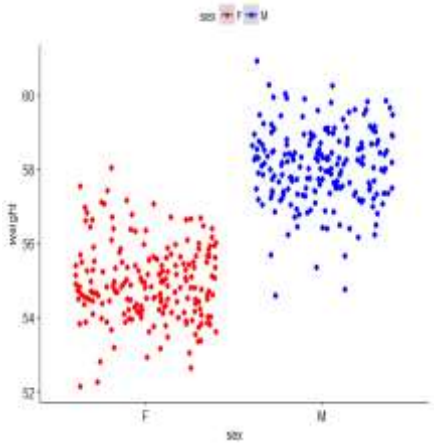
```
sex  weight
1   F
53.79293
2   F
55.27743
3   F
56.08444
4   F
52.65430
5   F
55.42912
6   F
55.50606
7   F
54.42526
8   F
54.45337
9   F
54.43555
10  F
54.10996
11  F
54.52281
12  …. N=400
```



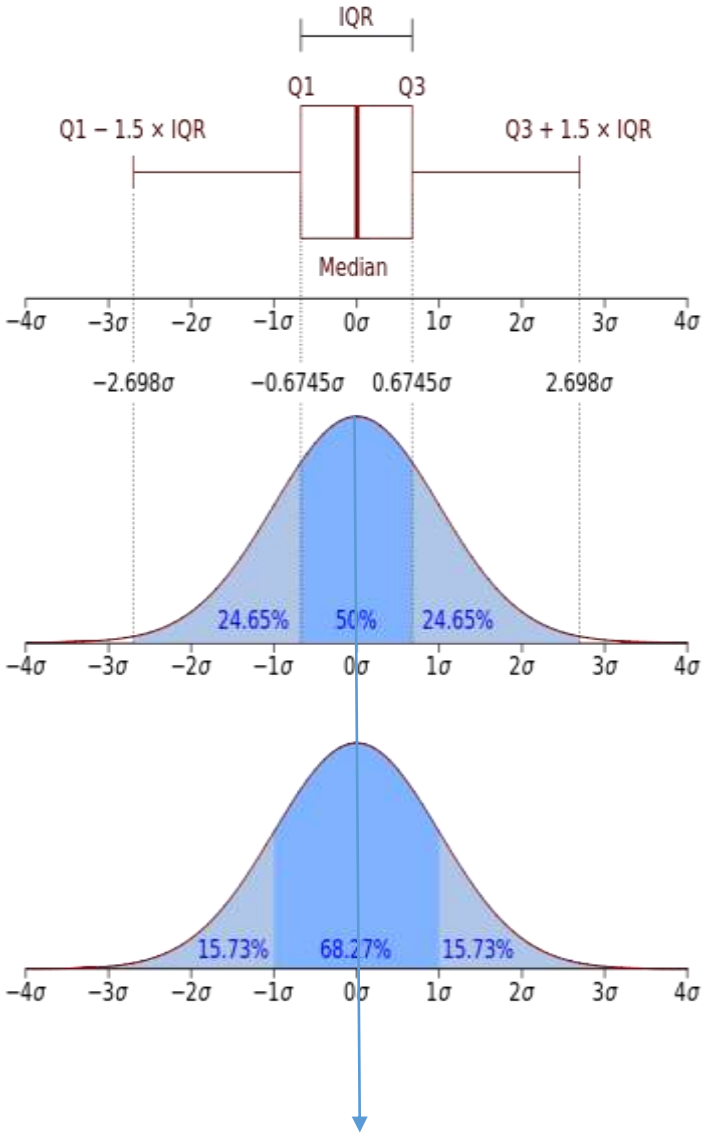







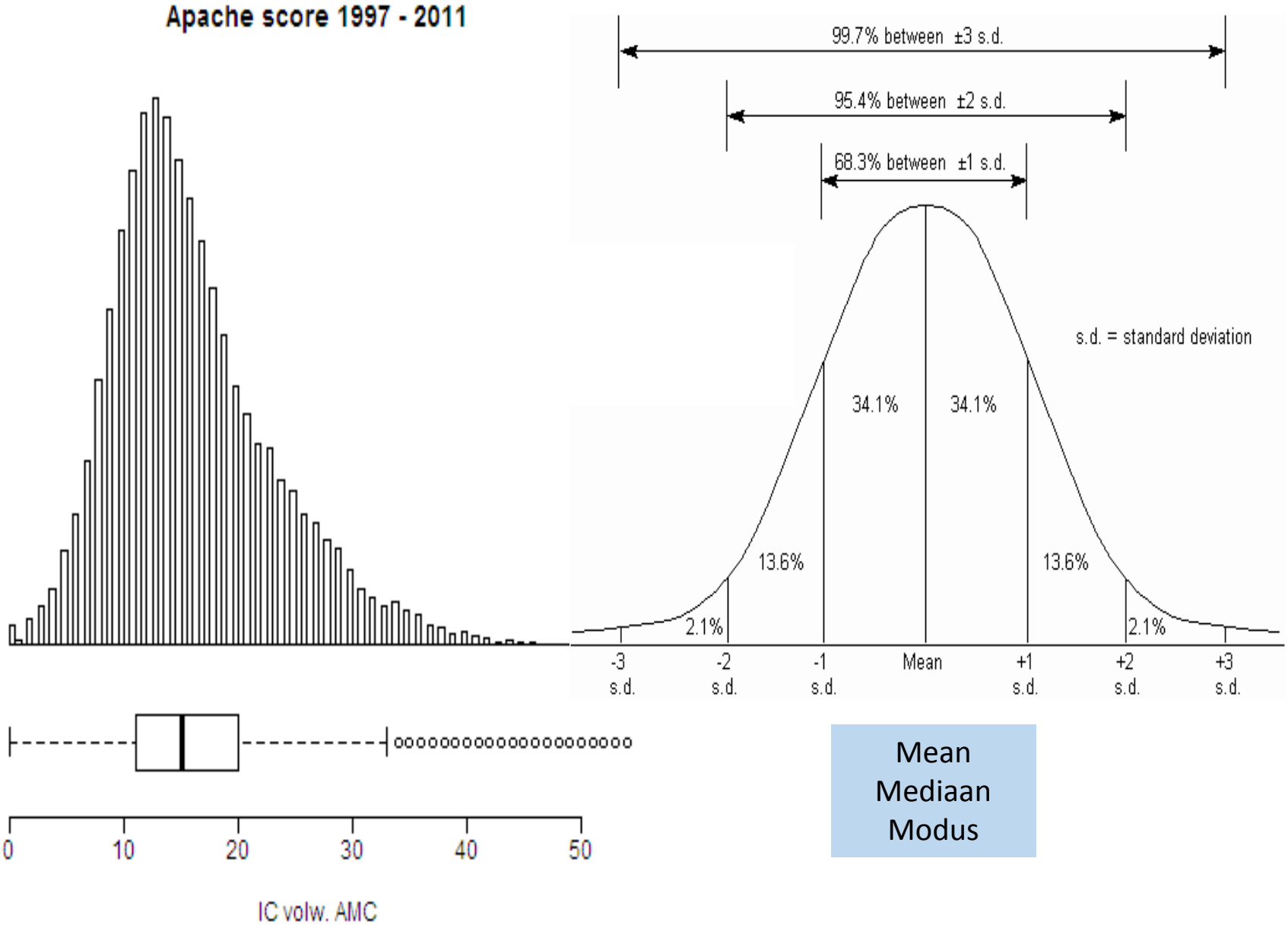
# Normale verdeling




Mean, Mediaan, Modus

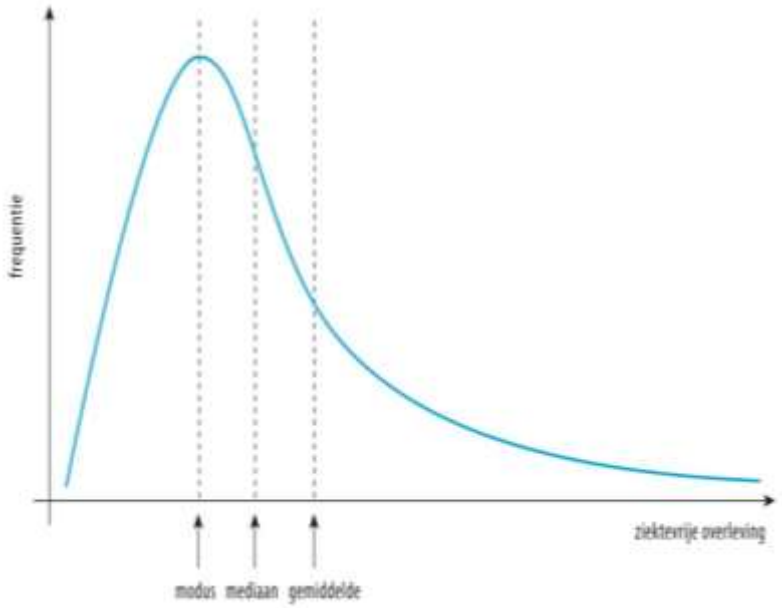


By Jhguch at en.wikipedia, CC BY-SA 2.5, https://commons.wikimedia.org/w/index.php?curid=14524285

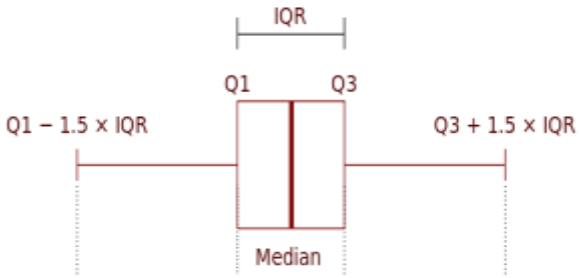


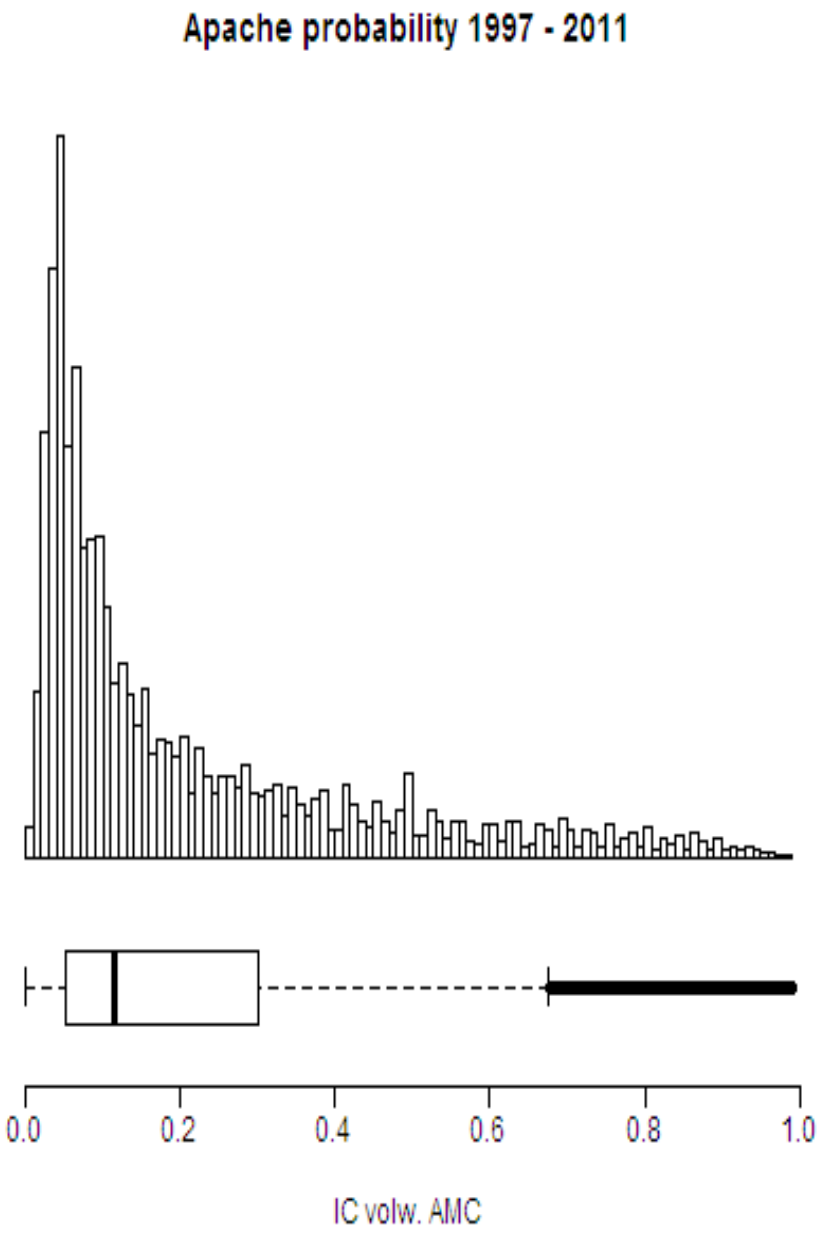

Mean
Mediaan
Modus

En, als de data niet normaal verdeeld is?

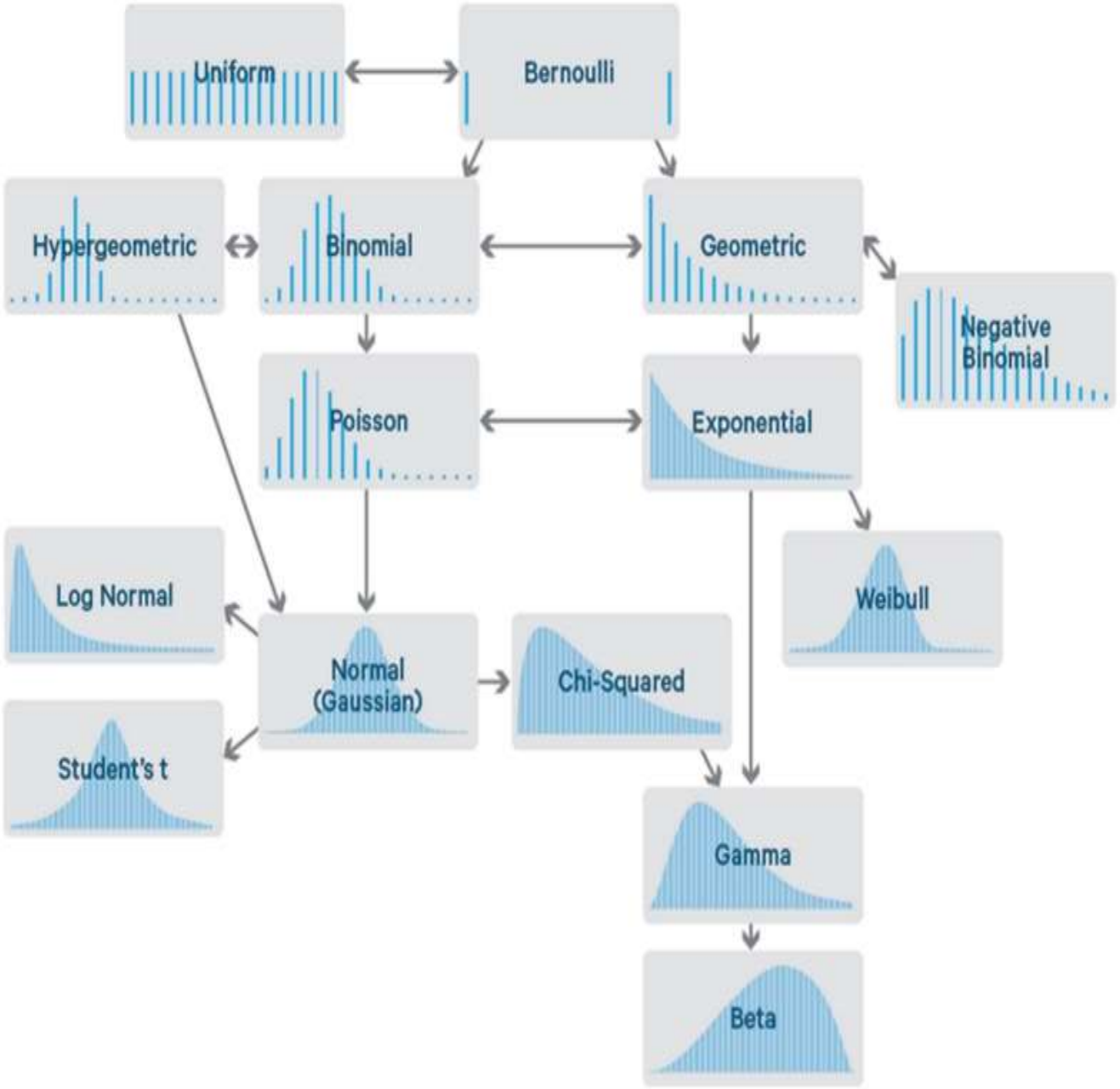



Positief scheef
50% van de data ligt onder mediaan
25% van de data tussen Q1 en mediaan




Apache probability 1997 - 2011
0.0
0.2
0.4
0.6
0.8
1.0
IC volw. AMC

Even tussendoor: distributies



# Verschil mean Apache score geplande opname en acute opname

**Hypothese**

H0: mean apache gepland = mean apache acuut;

H1: mean 1 ≠ mean 2

T.Test voor onafhankelijke groepen

1) Mean 1 = 36.3 (9,6)

2) Mean 2 = 63.6 (30,7)

Verschil: 27.3 (19 tot 35), p <0,0001

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_P^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad df = n_1 + n_2 - 2$$

*Standard error* →

*1 Pooled variance*

$$S_P^2 = \frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

# Generalizing from a sample to a population

De populatie heeft een verdeling met parameters

De sample wordt beschreven door een statistic (mean, median, mode)

Waarnemingen in de sample kun je terug vertalen naar de populatie, maar met een bepaalde onzekerheid

Random selection

Population mean

Inferences

Sample mean

# Standard deviatie ≠ standard error

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad df = n_1 + n_2 - 2$$

*Standard error* →

$$S_p^2 = \frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

*1 Pooled variance*

$$SE_{\bar{x}} = \frac{S}{\sqrt{n}}$$

Populatie -> sample mean* x keer

N means heeft een distributie verdeling

De standaard deviatie daarvan is de standard error

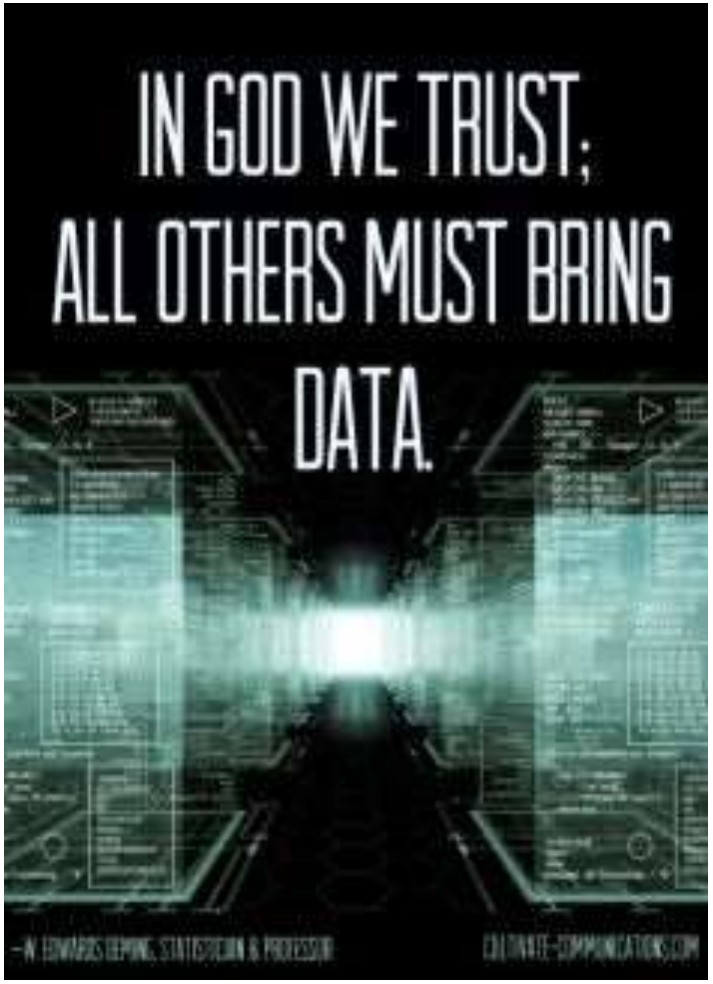
Data analyse

Data set: IC patiënten opgenomen van 2000 tot en met 2014

Vraagstelling

Is er een verschil tussen de opnameduur van geplande – en ongeplande (acute opname) IC-patiënten

Welke vragen ga je stellen?



# Data IC opnames van 2000 tot en met 2014 (27923 records)

| pin | sexe | age | plan_opn | | opn.icu | los.icu | apache4 | died.icu | jaar |
|---|---|---|---|---|---|---|---|---|---|
| 1 | M | 18 | 0 | 2000-01-01 | 04:00:00 | 1.3 | 47 | 0 | 2000 |
| 2 | M | 74 | 0 | 2000-01-01 | 04:20:00 | 1.6 | 90 | 0 | 2000 |
| 3 | M | 25 | 0 | 2000-01-01 | 07:00:00 | 0.7 | 39 | 0 | 2000 |
| 4 | F | 45 | 0 | 2000-01-01 | 22:34:00 | 5.5 | 29 | 0 | 2000 |
| 5 | F | 38 | 0 | 2000-01-01 | 23:30:00 | 3.5 | 101 | 0 | 2000 |
| 6 | M | 56 | 0 | 2000-01-02 | 10:45:00 | 0.4 | 49 | 0 | 2000 |
| 7 | F | 80 | 0 | 2000-01-02 | 15:00:00 | 10.8 | 56 | 0 | 2000 |
| 8 | M | 64 | 1 | 2000-01-03 | 12:30:00 | 2.9 | 47 | 0 | 2000 |
| 9 | M | 63 | 1 | 2000-01-03 | 14:00:00 | 0.8 | 31 | 0 | 2000 |
| 10 | F | 54 | 1 | 2000-01-03 | 16:30:00 | 7.9 | 55 | 0 | 2000 |
| 11 | M | 60 | 1 | 2000-01-03 | 17:30:00 | 1.9 | 46 | 0 | 2000 |
| 12 | F | 46 | 0 | 2000-01-03 | 18:30:00 | 14.6 | 30 | 0 | 2000 |
| 13 | F | 68 | 1 | 2000-01-03 | 20:45:00 | 0.6 | 49 | 0 | 2000 |
| 14 | M | 73 | 0 | 2000-01-03 | 23:45:00 | 68.4 | 77 | 1 | 2000 |
| 15 | F | 56 | 0 | 2000-01-04 | 11:00:00 | 4.1 | 45 | 0 | 2000 |
| 16 | F | 53 | 1 | 2000-01-04 | 13:00:00 | 16.9 | 82 | 0 | 2000 |
| 17 | F | 44 | 1 | 2000-01-04 | 13:30:00 | 0.9 | 39 | 0 | 2000 |
| 18 | F | 77 | 1 | 2000-01-04 | 14:00:00 | 29.9 | 96 | 0 | 2000 |
| 19 | M | 69 | 0 | 2000-01-04 | 15:15:00 | 6.8 | 83 | 0 | 2000 |
| 20 | M | 69 | 1 | 2000-01-04 | 15:30:00 | 0.8 | 87 | 0 | 2000 |
| 21 | M | 67 | 1 | 2000-01-04 | 19:00:00 | 0.7 | 41 | 0 | 2000 |
| 22 | M | 34 | 0 | 2000-01-05 | 06:30:00 | 0.3 | 106 | 1 | 2000 |
| 23 | F | 54 | 1 | 2000-01-05 | 14:00:00 | 4.0 | 21 | 0 | 2000 |
| 24 | M | 72 | 1 | 2000-01-05 | 15:15:00 | 1.0 | 64 | 0 | 2000 |
| 25 | M | 85 | 1 | 2000-01-05 | 16:30:00 | 3.2 | 76 | 0 | 2000 |
| 26 | M | 63 | 1 | 2000-01-05 | 17:10:00 | 0.8 | 28 | 0 | 2000 |
| 27 | F | 79 | 1 | 2000-01-05 | 19:00:00 | 1.7 | 39 | 0 | 2000 |
| 28 | M | 82 | 0 | 2000-01-06 | 05:00:00 | 2.2 | 61 | 0 | 2000 |
| 29 | M | 76 | 1 | 2000-01-06 | 12:45:00 | 1.2 | 49 | 0 | 2000 |
| 30 | F | 68 | 1 | 2000-01-06 | 15:15:00 | 4.0 | 48 | 0 | 2000 |

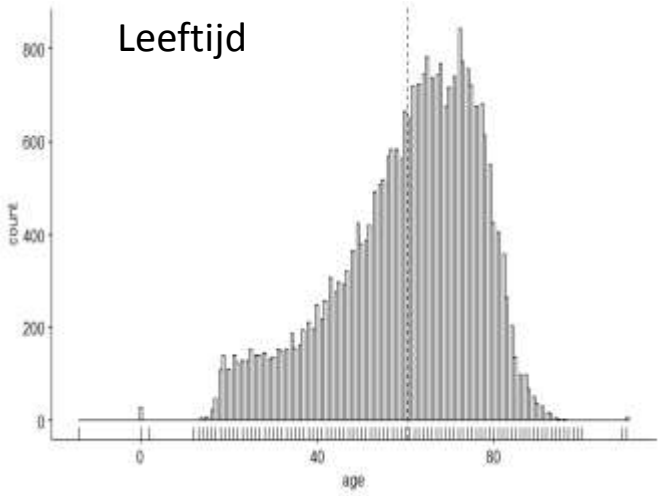
# Beschrijven van de dataset

| | | Gepland | Acuut | Verschil |
|---|---|---|---|---|
| Patiënten | N = 27923 | 11996 | 15927 | |
| Vrouwen | % | 33% (3891/11969) | 40% (6398/15896) | p < 0,0001 |
| Leeftijd | Mean (sd) | 63 (14) | 58 (18) | p <0,0001 |
| Ernst van ziekte, apache 4 | Mean (sd) | 50 (20) | 71 (32) | p <0,0001 |
| Opnameduur | Mean (sd) | 2,5 (4,9) | 5,5 (8,3) | p <0,0001 |

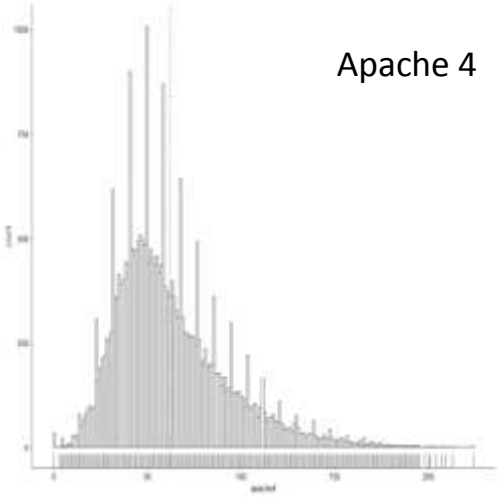
Is er een verschil tussen de opnameduur van geplande – en ongeplande (acute opname) IC-patiënten

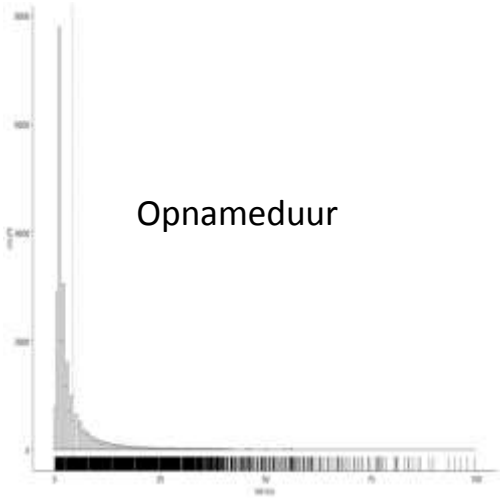
Wordt de data correct beschreven?

Leeftijd




Apache 4





Opnameduur

Welke data is normaal verdeeld?

Hoe streng moet je zijn?

Check op normale verdeling

# Beschrijven van de dataset

| | | Gepland | Acuut | Verschil |
|---|---|---|---|---|
| Patiënten | N = 27923 | 11996 | 15927 | |
| Vrouwen | % | 33% (3891/11969) | 40% (6398/15896) | p < 0,0001 |
| Leeftijd | Mean (sd) | 63 (14) | 58 (18) | p <0,0001 |
| Ernst van ziekte, apache 4 | Mean (sd) | 50 (20) | 71 (32) | p <0,0001 |
| Opnameduur | Median(IQr) | 1,1 (1 -2) | 2,6 (1 - 6) | p <0,0001 |

# Invloed van grote aantallen observaties op statistieken

| | | Gepland | Acuut | Verschil |
|---|---|---|---|---|
| Patiënten | N = 27923 | 11996 | 15927 | |
| Vrouwen | % | 33% (3891/11969) | 40% (6398/15896) | p < 0,0001 |
| Leeftijd | Mean (sd) | 63 (14) | 58 (18) | p <0,0001 |
| Ernst van ziekte, apache 4 | Mean (sd) | 50 (20) | 71 (32) | p <0,0001 |
| Opnameduur | Median(IQr) | 1,1 (1 -2) | 2,6 (1 - 6) | p <0,0001 |

| | | Gepland | Acuut | Verschil |
|---|---|---|---|---|
| Patiënten | N = 50 | 24 | 24 | |
| Vrouwen | % | 38% (9/24) | 50% (12/24) | p 0,38 |
| Leeftijd | Mean (sd) | 65 (12) | 52 (20) | p 0,01 |
| Ernst van ziekte, apache 4 | Mean (sd) | 54 (20) | 66 (37) | p <0,178 |
| Opnameduur | Median(IQr) | 1,1 (1 -2) | 2 (1 - 7) | p <0,27 |

Tot nu toe: verschil means met t-test

Nu verschil twee medianen (continue data, niet normaal verdeeld)

## 2 groepen observaties

Groep 1 = 0.6, 1.1, 3.1, 4.3, 8.7, 10, 12.2, 15.9, 20.1, 23.6, 30

Groep 2 = 11, 11.8, 13.3, 16.2, 19, 24, 32.7, 80, 90.3, 200, 250, 270, 310

Combineer de data, sorteer en rank, split de data weer naar de originele groepen, en toets op de ranks

Wilcox.test of Mann-Whitney U test

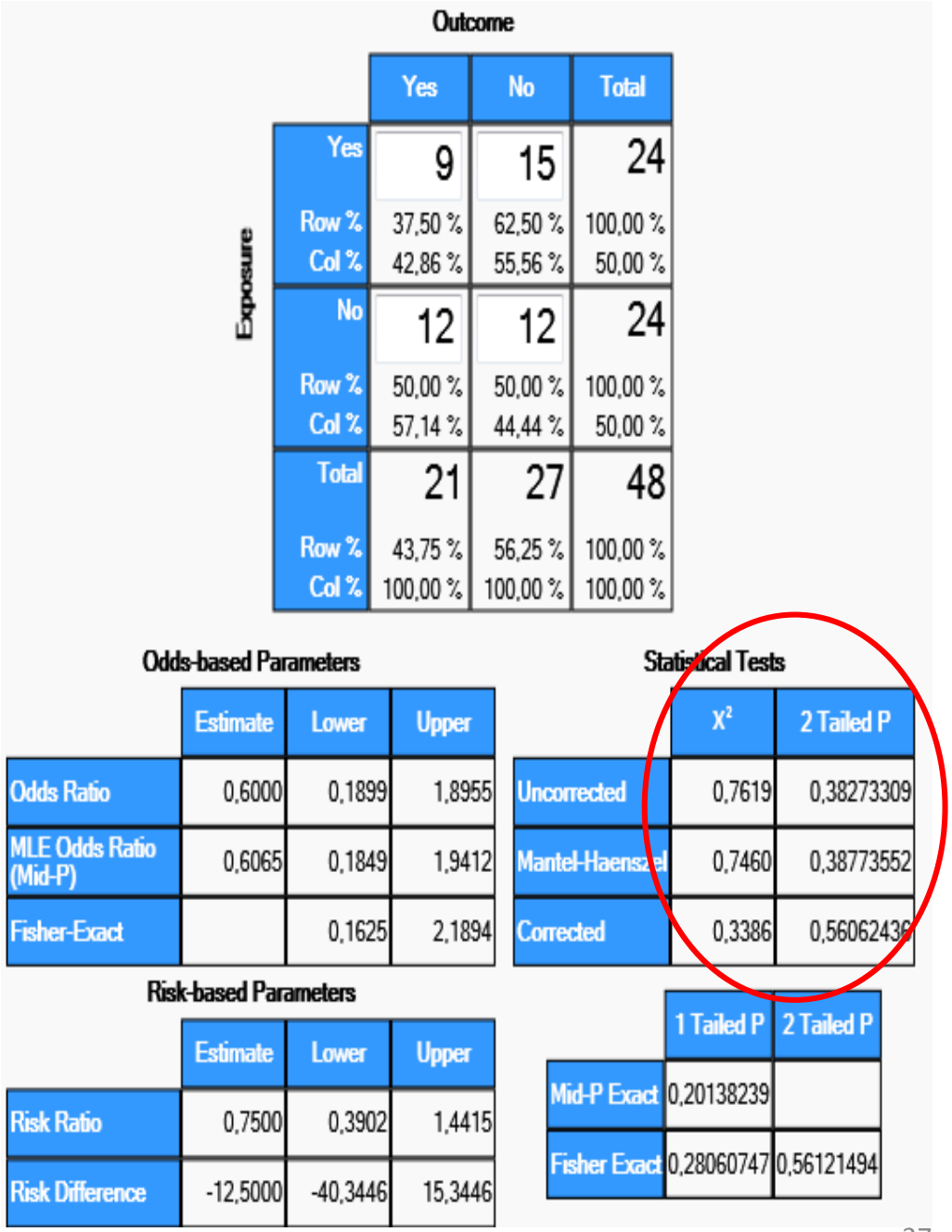
*p*-value = 0.002485

| | val | tel | vn | rank | |
|---|---|---|---|---|---|
| 1 | | 0.6 | 1 | gr1 | 1 |
| 2 | | 1.1 | 2 | gr1 | 2 |
| 3 | | 3.1 | 3 | gr1 | 3 |
| 4 | | 4.3 | 4 | gr1 | 4 |
| 5 | | 8.7 | 5 | gr1 | 5 |
| 6 | | 10.0 | 6 | gr1 | 6 |
| 12 | | 11.0 | 1 | gr2 | 7 |
| 13 | | 11.8 | 2 | gr2 | 8 |
| 7 | | 12.2 | 7 | gr1 | 9 |
| 14 | | 13.3 | 3 | gr2 | 10 |
| 8 | | 15.9 | 8 | gr1 | 11 |
| 15 | | 16.2 | 4 | gr2 | 12 |
| 16 | | 19.0 | 5 | gr2 | 13 |
| 9 | | 20.1 | 9 | gr1 | 14 |
| 10 | | 23.6 | 10 | gr1 | 15 |
| 17 | | 24.0 | 6 | gr2 | 16 |
| 11 | | 30.0 | 11 | gr1 | 17 |
| 18 | | 32.7 | 7 | gr2 | 18 |
| 19 | | 80.0 | 8 | gr2 | 19 |
| 20 | | 90.3 | 9 | gr2 | 20 |
| 21 | 200.0 | | 10 | gr2 | 21 |
| 22 | 250.0 | | 11 | gr2 | 22 |
| 23 | 270.0 | | 12 | gr2 | 23 |
| 24 | 310.0 | | 13 | gr2 | 24 |

Percentages?

Niet meten maar tellen

Output van Epi-Info ver-7



Outcome

| Exposure | | Yes | No | Total |
|---|---|---|---|---|
| Yes | | 9 | 15 | 24 |
| | Row % | 37,50 % | 62,50 % | 100,00 % |
| | Col % | 42,86 % | 55,56 % | 50,00 % |
| No | | 12 | 12 | 24 |
| | Row % | 50,00 % | 50,00 % | 100,00 % |
| | Col % | 57,14 % | 44,44 % | 50,00 % |
| Total | | 21 | 27 | 48 |
| | Row % | 43,75 % | 56,25 % | 100,00 % |
| | Col % | 100,00 % | 100,00 % | 100,00 % |

Odds-based Parameters

| | Estimate | Lower | Upper |
|---|---|---|---|
| Odds Ratio | 0,6000 | 0,1899 | 1,8955 |
| MLE Odds Ratio (Mid-P) | 0,6065 | 0,1849 | 1,9412 |
| Fisher-Exact | | 0,1625 | 2,1894 |

Statistical Tests

| | $X^2$ | 2 Tailed P |
|---|---|---|
| Uncorrected | 0,7619 | 0,38273309 |
| Mantel-Haenszel | 0,7460 | 0,38773552 |
| Corrected | 0,3386 | 0,56062436 |

Risk-based Parameters

| | Estimate | Lower | Upper |
|---|---|---|---|
| Risk Ratio | 0,7500 | 0,3902 | 1,4415 |
| Risk Difference | -12,5000 | -40,3446 | 15,3446 |

| | 1 Tailed P | 2 Tailed P |
|---|---|---|
| Mid-P Exact | 0,20138239 | |
| Fisher Exact | 0,28060747 | 0,56121494 |

Wat hebben we gedaan tot nu toe

Percentages berekend per groep

Gemiddelde en standaard deviatie

Mediaan en interkwartiel range

Verschil tussen

2 percentages: Chi square test

twee gemiddelden: t-test

twee medianen: Wilcoxon-test of Mann-Whitney test

Verschil in resultaten bij een grote N en bij een kleine N

P waarden?

## H0: geen verschil.

## H1: wel verschil

P waarde is de kans op een onderzoeksresultaat onder de aanname dat de H0 waar is.

Dus als je een kleine kans vindt, kleiner dan 0.05, dan is het onwaarschijnlijk dat de H0 waar is.

Maar, je kan er nog steeds naast zitten:

| | **H0 = TRUE** | **H0 = FALSE** |
|---|---|---|
| Test verwerpt HO niet | 1 - α | β |
| Test verwerpt H0 wel | α | 1-β |

*α = fout positief, β = fout negatief*

The probability of rejecting the null hypothesis when it is actually false is called the **POWER** of the study (Power = $1-\beta$)

| | | The null hypothesis is actually: | |
|---|---|---|---|
| | | **False** | **True** |
| Based on the test results you decide to: | **Reject the null hypothesis** (i.e. the p-value is less than or equal to $\alpha$, usually set to 0.05) | Conclusion is correct. Have rejected the null when it is in fact false ($1-\beta$) | False positive / type I error / $\alpha$ |
| | **Not reject the null hypothesis** (i.e. the p-value is greater than $\alpha$, usually set to 0.05) | False negative / type II error / $\beta$ | Conclusion is correct. Have not rejected the null hypothesis when it is in fact true |

The p-value is the probability of obtaining your results or results more extreme **if the null hypothesis is true**. It is the probability of committing a false positive error i.e. of rejecting the null when it is in fact true.

| Welke toets bij welke vergelijking | | | type vergelijking | | | | |
|---|---|---|---|---|---|---|---|
| | | | 1 groep | 2 groepen | | >2 groepen | |
| | | | vs. referentie | gepaard | ongepaard | gepaard | ongepaard |
| type data | numeriek (continu) | normaal verdeeld | 1 sample t-toets | gepaarde t-toets | ongepaarde t-toets | linear mixed models | One-way ANOVA |
| | | niet normaal verdeeld | tekentoets | Wilcoxon signed rank toets | Mann-Whitney U toets | Friedman toets | Kruskal Wallis |
| | categorisch (discreet) | binair | z-test voor proporties | McNemar toets | Chi-kwadraat toets/ Fisher's exact toets | Cochran's Q toets | Chi-kwadraat toets/ Fisher-Freeman-Halton exact toets |
| | | nominaal / ordinaal | x | McNemar toets / Wilcoxon signed rank toets | Chi-kwadraat toets (trend) | GLMM / GEE | Chi-kwadraat toets (trend) |

https://wikistatistiek.amc.nl/index.php/KEUZE_TOETS

Welke analyse bij welke associatie

| | | | type associatie | | | |
|---|---|---|---|---|---|---|
| | | | crosssectioneel | | longitudinaal | |
| | | | een onafhankelijke variabele | meer onafhankelijke variabelen | een onafhankelijke variabele | meer onafhankelijke variabelen |
| type data | numeriek (continu) | normaal verdeeld | Pearson's rho/enkelvoudige lineaire regressie | meervoudige lineaire regressie | herhaalde metingen | |
| | | niet normaal verdeeld | Spearman's rho | x | | |
| | gecensureerd numeriek | | | | Kaplan Meier analyse | Cox regressie |
| | categorisch (discreet) | binair | OR/ RR/ RRR / ARR / NNT | logistische regressie | GLMM / GEE | |
| | | nominaal /ordinaal | x | multinomiale / ordinale logistische regressie | | |

Klinische epidemiologie

Twee bij twee tabellen

- Relatieve risico
- Absolute risico verschil
- Risico reductie
- Number Needed To Treat

en meer…

| | Outcome + | Outcome - | |
|---|---|---|---|
| Exposure | A | B | A+B |
| Non exposure | C | D | C+D |
| | A+C | B+D | A+B+C+D |

# Example: Smoking & Mortality

| | Dead | Alive | | Risk |
|---|---|---|---|---|
| Smoker | 139 | 443 | 582 | 0.24 |
| Non-Smoker | 230 | 502 | 732 | 0.31 |
| | 369 | 945 | 1314 | |

20-year follow-up Whickham UK

**Relative Risk 0,76 (0,63 / 0.91)**

# Example: Smoking & Mortality

| *18-24* | Dead | Alive | Risk |
|---|---|---|---|
| Smoker | 2 | 53 | 0.04 |
| Non-Smoker | 1 | 61 | 0.02 |

**Relative risk 2,25 (0,21 / 24,2)**

| *45-54* | Dead | Alive | Risk |
|---|---|---|---|
| Smoker | 27 | 103 | 0.21 |
| Non-Smoker | 12 | 66 | 0.15 |

**Relative risk 1,35 (0,73 / 2,5)**

**20-year follow-up Whickham UK**

# Number Needed To Treat

Hoeveel patiënten moet je behandelen om 1 negatieve uitkomst te voorkomen

NNT = 1/ARR  (ARR = events controls – events experiment)

ARR = hoeveel risico reductie

Infectie rate van 50% naar 30% = ARR 20%

NNT = 1/0.2 = 5

5 patienten behandelen om 1 infectie te voorkomen

## Samenvattend

- Wat is statistiek

- Wetenschap – communicatie – studiedesign – statistische keuzes

- Eigenschappen van data

- Beschrijvende statistieke – normaal verdeling – hypotheses

- Een paar categorische associaties – ARR – RR - NNT

Bronnen om te leren

Google, Coursera, Edx, EMGO VUMC, Khan academy

WIKI statistiek van de CRU (Clinical Research Unit) AMC

https://wikistatistiek.amc.nl/index.php/KEUZE_TOETS

BMJ Education and Debate: Statistics Notes

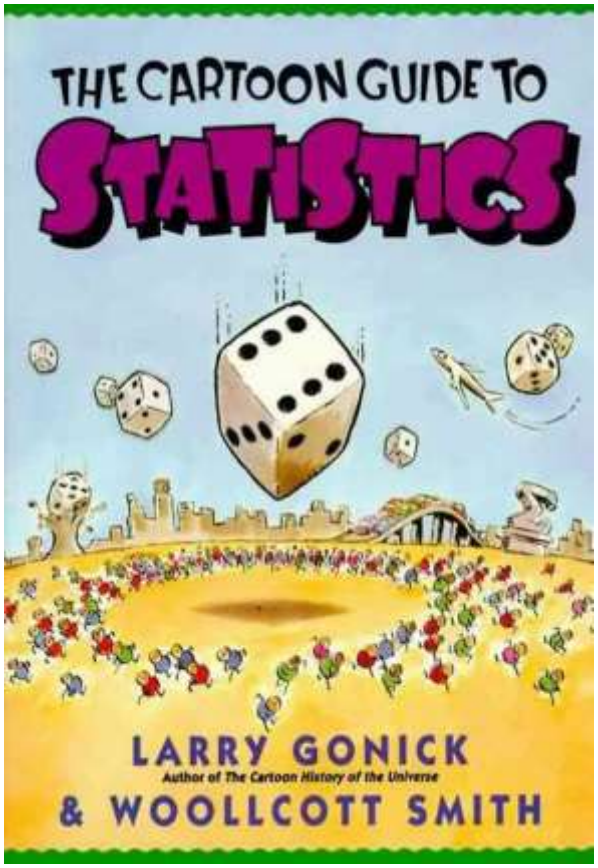
http://www.bmj.com/specialties/statistics-notes

Khan academy

https://nl.khanacademy.org/

Boeken

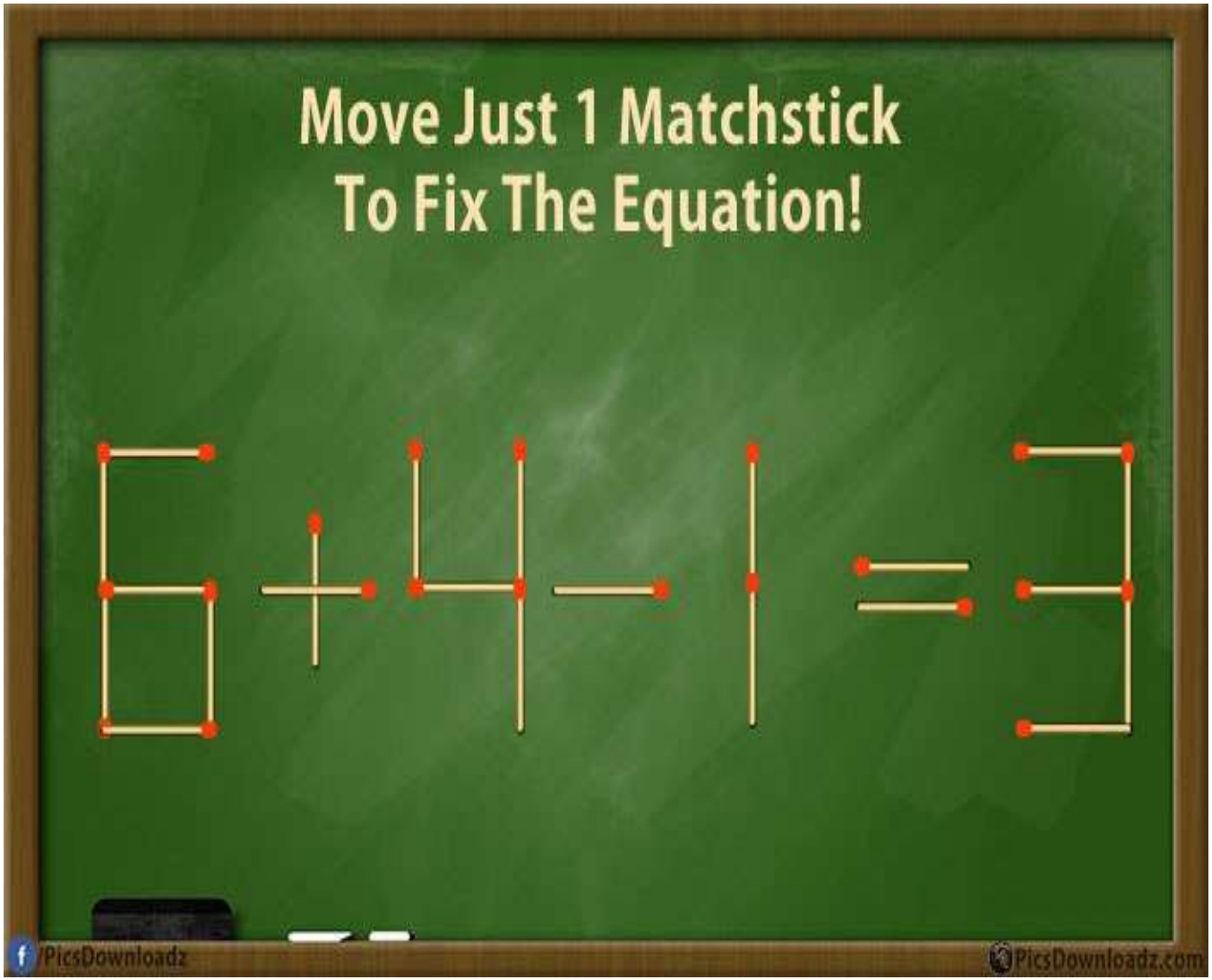
Andy Field. Statistics SPSS , R en SAS



# Learn from the mistakes of others

You can’t live long enough to make
Them all yourself

*Eleanor Roosevelt*


Move Just 1 Matchstick
To Fix The Equation!
6+4-1=9
PicsDownloadz
PicsDownloadz.com