

## Casus behorende bij de vragen 1 t/m 25

Bij een trombosebeen ontstaat er een stolsel in de diepe aderen in meestal het onderbeen. De kans om als Nederlander in een gegeven jaar een trombosebeen te krijgen is ongeveer 1,5‰; het overkomt dus jaarlijks 25.000 Nederlanders. De gangbare behandeling bestaat uit het toedienen van bloedverdunners om ervoor te zorgen dat het stolsel niet groeit, zwachtelen om de eerste zwelling te verminderen, gevolgd door steunkousen om de ophoping van vocht te voorkomen. Het is de bedoeling dat het lichaam het stolsel zelf afbreekt.

Er is echter een kans van tussen de 25 en 50% dat een patiënt een post-trombotisch syndroom (PTS) ontwikkelt. Bij een studie naar het ziekteverloop van trombosebenen werd twee jaar na de start van de behandeling van het trombosebeen bij 125 patiënten nagegaan of ze PTS ontwikkeld hadden. Dit was bij 51 patiënten het geval. Tevens is bij deze patiënten de leeftijd, de Body Mass Index (BMI in kg/m<sup>2</sup>), het geslacht en de trombofiliestatus bepaald (trombofilie is een ziektebeeld waarbij door verschillende veranderingen in de bloedsamenstelling de stolling bevorderd is en dus ook de kans op trombose groter is).

### Aanvulling casus bij vraag 1 en 2

In tabel 1 staan wat karakteristieken van de patiënten op deze variabelen vermeld:

Tabel 1: Karakteristieken van de onderzochte patiënten (N = 125)

Gemiddelde leeftijd (SD)	59,0 (15,2)
Gemiddelde BMI (SD)	30,2 (6,1)
Aantal vrouwen (%)	50 (40)
Aantal Trombofilie patiënten (%)	49 (39)
Post trombotisch syndroom (PTS; %)	51 (41)

### Vraag 1

Hoe heet een tabel zoals tabel 1?

1. Een overzichtstabel.
2. Een kruistabel.
3. Een correlatietabel.
4. Een modustabel.

## Vraag 2

Stel dat bekend is dat de algemene populatie voor 10% bestaat uit personen met trombofilie. Met welke toets kan de onderzoeker dan nagaan of er in verhouding meer personen met trombofilie zijn in de populatie trombosebeënvallenden?

1. De toets voor één proportie.
2. De toets voor twee proporties.
3. De toets voor één gemiddelde.
4. De toets voor twee gemiddeldes.

## Vraag 3

Op basis van de resultaten wordt Tabel 2 bepaald.

Tabel 2

	Geen PTS (N = 74)	Wel PTS (N = 51)
Gemiddelde BMI (SD)	29,4 (6,0)	31,2 (6,0)

Stel dat de onderzoeker op basis van deze resultaten wil nagaan of er een verband is tussen de hoogte van de BMI en het ontwikkelen van PTS, dan moet de onderzoeker gebruik maken van:

1. een gepaarde t-toets.
2. een McNemar test.
3. een t-test voor twee groepen.
4. een chi-kwadraat test.

## Aanvulling casus bij de vragen 4 t/m 6

Op basis van de resultaten van dezelfde studie is ook Tabel 3 bepaald:

Tabel 3

	Geen PTS	Wel PTS	Totaal
Geen trombofilie	38	38	76
Wel trombofilie	36	13	49
Totaal	74	51	125

#### **Vraag 4**

Stel dat de onderzoeker wil nagaan of er een verband is tussen het hebben van trombofilie en het ontwikkelen van PTS, dan moet de onderzoeker gebruik maken van:

1. een gepaarde t-toets.
2. een McNemar test.
3. een t-test voor twee groepen.
4. een chi-kwadraat test.

#### **Vraag 5**

De toets van vraag 4 bleek een significant resultaat te geven. Welke conclusie is dan gerechtvaardigd?

1. Het hebben van trombofilie is een beschermende factor voor het ontwikkelen van PTS.
2. Het hebben van trombofilie is een risicofactor voor het ontwikkelen van PTS.
3. Het hebben van trombofilie is een beschermende factor voor het ontwikkelen van PTS gegeven dat een persoon een trombosebeen heeft.
4. Het hebben van trombofilie is een risicofactor voor het ontwikkelen van PTS gegeven dat een persoon een trombosebeen heeft.

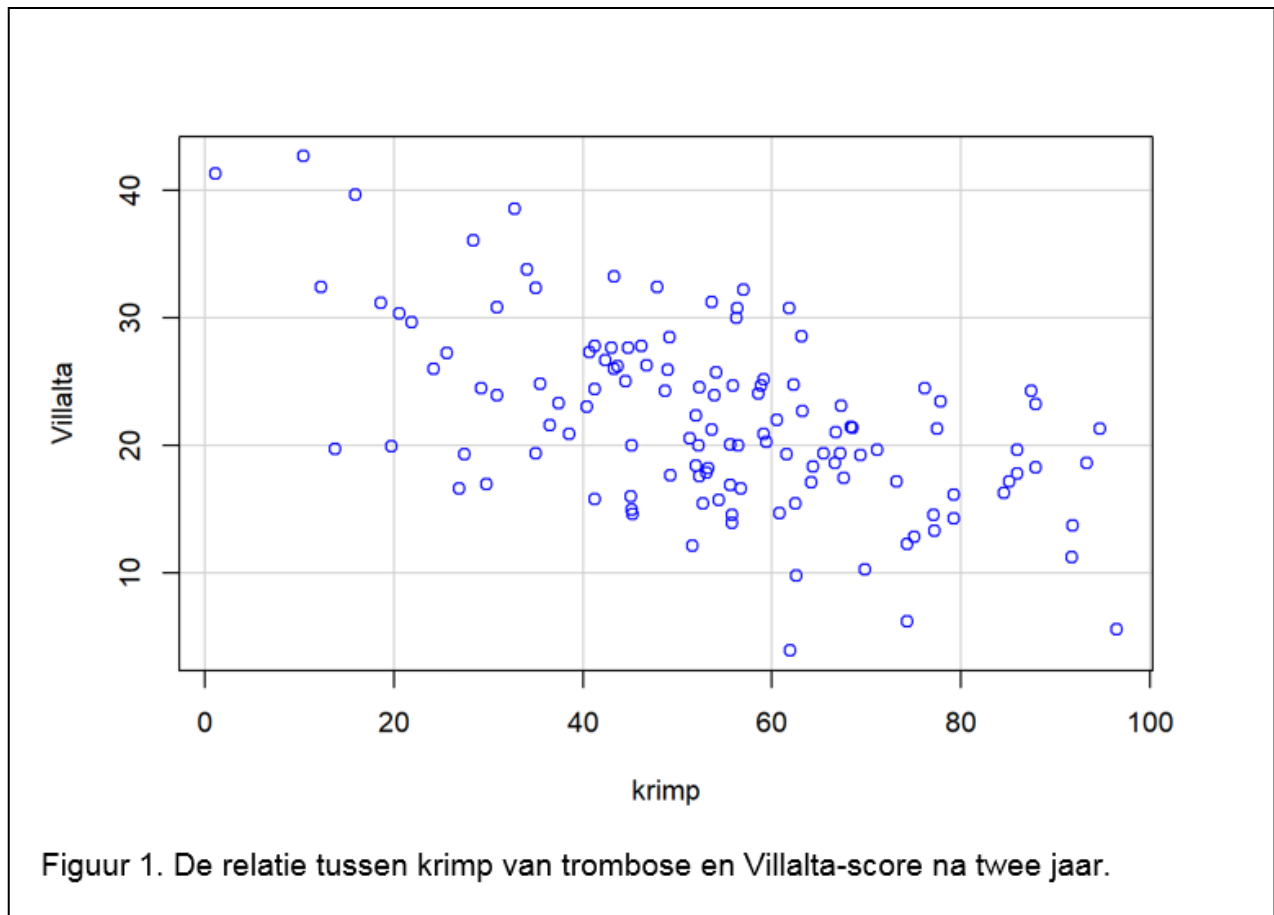
#### **Vraag 6**

Stel dat de onderzoeker op basis van Tabel 3 wil nagaan wat vaker voorkomt bij trombosebeenpatiënten: trombofilie of PTS, dan moet de onderzoeker gebruik maken van:

1. een gepaarde t-toets.
2. een McNemar test.
3. een t-test voor twee groepen.
4. een chi-kwadraat test.

#### **Aanvulling casus bij vraag 7 en 8**

In deze studie werd twee weken na de start van de behandeling middels dopplertechnieken een schatting gemaakt van de grootte van de verstopping. Dit werd herhaalt na zes weken en vervolgens werd de percentuele daling berekend. De ernst van de PTS werd bepaald met behulp van de schaal van Villalta (de definitie van een PTS was gebaseerd mede op deze schaal: als bij een patiënt na twee jaar een waarde boven de afkapwaarde werd gevonden, werd deze geclassificeerd als een PTS patiënt). Om de relatie tussen de daling in trombosegrootte en de ernst van PTS na twee jaar te onderzoeken, maakt de onderzoeker eerst Figuur 1.



### Vraag 7

Hoe heet een figuur zoals figuur 1?

1. Een box plot.
2. Een histogram.
3. Een scatterplot (strooidiagram).
4. Een staafdiagram.

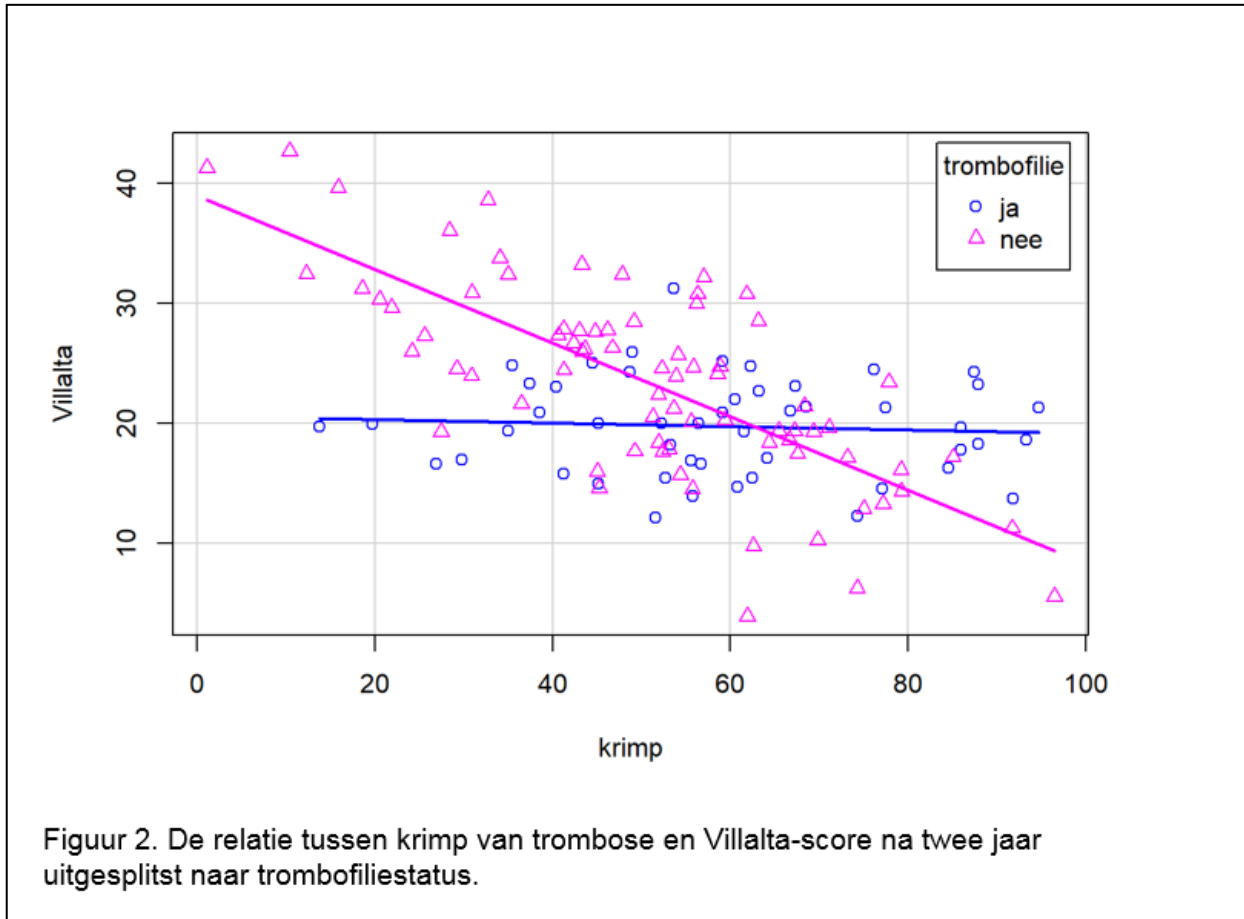
### Vraag 8

Hoe kan het verband tussen de krimp van de trombose en de Villalta-score na twee jaar omschreven worden?

1. Een positief verband.
2. Een negatief verband.
3. Een curvilineair verband.
4. Er is geen verband tussen deze variabelen.

## Aanvulling casus bij vraag 9 en 10

De onderzoeker kijkt ook nog naar Figuur 2. Hierin wordt een onderscheid gemaakt tussen trombofiliepatiënten en niet-trombofiliepatiënten.



### Vraag 9

Hoe heten de lijnen in Figuur 2?

1. Extrapolatielijnen.
2. Interpolatielijnen.
3. Regressielijnen.
4. Exponentiële lijnen.

### Vraag 10

Welke statistische term beschrijft de uitkomst zoals weergegeven in Figuur 2?

1. Confounding.
2. Interactie.
3. Mediatie.
4. Non-inferioriteit.

## Aanvulling casus bij de vragen 11 t/m 14

Het idee ontstaat dat het optreden van PTS niet alleen bepaald wordt door de mate van succes in het opruimen van de trombose zelf. Dit idee leidt uiteindelijk tot een project gericht op ontwikkeling van een medicijn ter voorkoming van PTS. Dit middel grijpt in op Protrombine (stollingsfactor II), een eiwit dat betrokken is bij de bloedstolling. Omdat er geen diermodel is om PTS te onderzoeken, kan de werkzaamheid niet via dierproeven worden vastgesteld. Er wordt echter wel gekeken naar de veiligheid. Een mogelijk risico is dat het nieuwe middel zal leiden tot een hogere kans op een nieuwe trombose.

Via de methode van Pierangeli wordt bij muizen in één van de aderen langs de femur een trombose opgewekt. Hiertoe wordt de ader vrijgemaakt en wordt er met een tang een standaard-kneep van  $1500 \text{ gr/cm}^2$  gegeven. De muis wordt vervolgens dicht gemaakt en krijgt een standaardmedicatie van bloedverdunners in de ene conditie (TAU) en een behandeling met bloedverdunners en het nieuwe medicijn in de andere conditie (EXP). Na drie weken wordt de andere femur-ader vrijgemaakt en wordt hier een kneep van  $750 \text{ gr/cm}^2$  gegeven en wordt hier gekeken middels een videomicroscoop bijgelicht door een fiberoptisch apparaat hoe de vorming van een mogelijke trombose verloopt. Men meet de grootte (variabele Grootte in  $\mu\text{m}^3$ ) van de nieuwe trombose en de tijd (variabele Tijd in secondes) dat het geduurd heeft totdat de trombose de helft van deze grootte behaald heeft. Er geldt voor Grootte: hoe groter, hoe minder gunstig, terwijl voor Tijd geldt: hoe langer hoe gunstiger. In Tabel 4 staan wat karakteristieken van de uitkomsten vermeld.

Tabel 4 Karakteristieken van Grootte en Tijd

	TAU (N = 32)	Exp (N = 32)
Gemiddelde Grootte (SD)	1000 (190)	1050 (210)
Gemiddelde Tijd (SD)	1000 (900)	950 (750)

### Vraag 11

Welke variabele is scheef verdeeld, gegeven Tabel 4?

1. Grootte.
2. Tijd.
3. Dat kan op basis van deze gegevens niet bepaald worden.

### **Vraag 12**

Na inspectie van de histogrammen van de uitkomsten besluit een onderzoekster uit dit team voor één van de variabelen een 95%-betrouwbaarheidsinterval op te stellen voor het verschil van de gemiddeldes. Gehoorzamen aan een aloude psychologische wet wordt het interval zo opgesteld dat positieve waarden betekenen dat de nieuwe behandeling (conditie EXP) beter is dan de controlebehandeling (conditie TAU). Het betrouwbaarheidsinterval is: (-150; 50).

Kan op basis van dit interval geconcludeerd worden dat één van beide behandelingen significant beter is dan de andere?

1. Ja, het nieuwe middel (conditie EXP) is significant beter.
2. Ja, het nieuwe middel (conditie EXP) is significant slechter.
3. Nee, want er is geen toets uitgevoerd.
4. Nee, er is onvoldoende bewijs om een conclusie te trekken over welke van de twee beter is.

### **Vraag 13**

Een collega van de onderzoekster stelt dat men eigenlijk pas tevreden zou moeten zijn wanneer men significant bewijs kan vinden dat het nieuwe middel minstens even goed/veilig is als de controle conditie. Deze collega is pas tevreden wanneer dit middels een toets/betrouwbaarheidsinterval aangetoond kan worden.

De onderzoekster vermoedt dat het nieuwe middel niet zal leiden tot extra veiligheidsproblemen en dus precies even goed zal scoren als het oude middel, omdat het nieuwe middel wordt toegevoegd aan de behandeling met het oude middel.

Stel dat de onderzoekster gelijk heeft en toevoegen van het nieuwe middel precies even goed/veilig is als de oude behandeling.

Hoe groot is dan de kans dat een onderzoek ook haar collega tevreden zou stellen?

1. 5% of minder.
2. Ongeveer 20%.
3. Ongeveer 50%.
4. Ongeveer 80%.
5. 95% of meer.

### **Vraag 14**

Op de variabele die scheef verdeeld is, wordt een logtransformatie (natuurlijk logaritme) toegepast. De zo berekende variabele heet LogUitkomst en er wordt weer een 95%-betrouwbaarheidsinterval voor het verschil in gemiddelde LogUitkomst opgesteld, zodanig dat een positieve uitkomst staat voor een beter effect van het nieuwe middel. De grenzen van dit betrouwbaarheidsinterval worden vervolgens teruggetransformeerd (via e-tot-de-macht). Dit teruggetransformeerde 95%-betrouwbaarheidsinterval is: (0,85; 1,05).

Welke conclusie wordt door de data ondersteund?

1. Omdat het hele betrouwbaarheidsinterval positief is, is aangetoond dat het nieuwe middel beter scoort dan de bestaande behandeling.
2. Omdat het hele betrouwbaarheidsinterval positief is, kan geen conclusie getrokken worden over welke behandeling beter is.
3. Omdat het betrouwbaarheidsinterval de waarde 1 bevat, is aangetoond dat het nieuwe middel beter scoort dan de bestaande behandeling.
4. Omdat het betrouwbaarheidsinterval de waarde 1 bevat, kan geen conclusie getrokken worden over welke behandeling beter is.

### **Aanvulling casus bij vraag 15 en 16**

Bij dit experiment werd telkens eerst bij de linkerfemur de oorspronkelijke trombose gemaakt waarna de metingen gedaan werden bij de rechterfemur. Men had er ook voor kunnen kiezen om bij de helft van de muizen de oorspronkelijke trombose bij de rechterfemur te maken en dan bij de linkerfemur de metingen te doen. Men zorgt ervoor dat er even veel muizen uit de twee condities (TAU, EXP) in beide groepen (Meting Links of Meting Rechts) zitten.

### **Vraag 15**

Het design dat men dan krijgt heet:

1. een parallelle groepen design.
2. een cross-over design.
3. een factorieel design.



### Vraag 16

Stel dat men wél besluit om zowel rechts als links de metingen te verrichten, maar besluit om dezelfde analyses uit te voeren als in vraag 12 en 14 besproken zijn: al dan niet teruggetransformeerde 95%-betrouwbaarheidsintervallen voor het verschil in gemiddelde. Wat zal het gevolg zijn?

1. Omdat de standaarderror kleiner wordt zal het betrouwbaarheidsinterval smaller worden.
2. Omdat de standaarderror kleiner wordt zal het betrouwbaarheidsinterval breder worden.
3. Omdat de standaarderror groter wordt zal het betrouwbaarheidsinterval smaller worden.
4. Omdat de standaarderror groter wordt zal het betrouwbaarheidsinterval breder worden.

## **Aanvulling casus bij de vragen 17 t/m 19**

Na nog meer veiligheidsonderzoek bij dieren en gezonde vrijwilligers, volgen de eerste studies bij patiënten. Deze zijn gericht op het vinden van een geschikte dosering, het onderzoeken van de veiligheid en het eerste bewijs van een mogelijk effect. De uitkomsten van deze onderzoeken zijn uitermate voorspoedig. Vandaar dat gestart wordt met een grotere studie. In deze studie zullen 120 patiënten met een diagnose acuut trombosebeen random verdeeld worden over drie groepen. NB: patiënten met trombofilie worden uitgesloten van deelname aan de studie.

De eerste groep (Normaal) krijgt bloedverdunners, een korte periode zwachtels, gevolgd door 3 maanden steunkousen en een placebomedicijn dat dagelijks geslikt moet worden. Daarna wordt geëvalueerd of er doorgedaan moet worden met de steunkousen of zelfs een invasieve ingreep nodig is en volgt een routinematige controle.

De tweede groep (NormaalPlus) krijgt bloedverdunners, een korte periode zwachtels, gevolgd door 3 maanden steunkousen en het nieuwe middel dat dagelijks geslikt moet worden. Ook in deze groep wordt na drie maanden geëvalueerd of er doorgedaan moet worden met de steunkousen of zelfs een invasieve ingreep nodig is en volgt er daarna een routinematige controle.

De derde groep (Plus) krijgt bloedverdunners, een korte periode zwachtels, gevolgd door 3 maanden waarin het nieuwe middel dagelijks geslikt moet worden (geen steunkousen dus). Ook deze groep wordt na drie maanden geëvalueerd en er kan dan besloten worden om steunkousen te gaan gebruiken of zelfs een invasieve ingreep te plannen. Daarna volgt ook hier een routinematige controle.

Twee jaar na de diagnose wordt bij alle patiënten de Villalta-schaal afgenomen om de ernst van PTS te bepalen. NB: hoe hoger de Villalta-score, hoe ernstiger de PTS. Delen van de output van de analyses staan hieronder weergegeven.

```

          Df Sum Sq Mean Sq F value    Pr(>F)
Conditie    2  535.6   267.78    10.4 0.0000699 ***
Residuals 117 3013.7    25.76
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residuals:
    Min       1Q   Median       3Q      Max
-13.9716  -3.3999  -0.1003   3.9646  11.7556

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    22.4559    0.8025  27.983 < 2e-16 ***
Conditie[T.Normaal]  0.6362    1.1349   0.561 0.576136
Conditie[T.NormaalPlus] -4.1293    1.1349  -3.639 0.000409 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.075 on 117 degrees of freedom
Multiple R-squared:  0.1509, Adjusted R-squared:  0.1364
F-statistic: 10.4 on 2 and 117 DF, p-value: 0.00006988

```

### Vraag 17

Op basis waarvan kan geconcludeerd worden dat niet alle drie de populatiegemiddeldes gelijk zijn aan elkaar? Op basis van de p-waarde...

1. van Conditie in de ANOVA-tabel.
2. van het Intercept in de coëfficiëntentabel.
3. van het regressiegewicht horend bij conditie Normaal in de coëfficiëntentabel.
4. van het regressiegewicht horend bij conditie NormaalPlus in de coëfficiëntentabel.

### Vraag 18

Welke conditie heeft het hoogste steekproefgemiddelde?

1. De groep Normaal.
2. De groep NormaalPlus.
3. De groep Plus.
4. Dat is op basis van deze output niet na te gaan.

### Vraag 19

Is er een significant verschil tussen de NormaalPlus en de Plus conditie?

1. Ja.
2. Nee.

## Aanvulling casus bij vraag 20

Zoals reeds is aangegeven wordt op basis van de Villalta-schaal beoordeeld of een patiënt geclassificeerd moet worden als een PTS-patiënt. In Tabel 5 staat het resultaat van deze classificatie gegeven.

Tabel 5: PTS-patiënten per conditie

	Geen PTS	Wel PTS (%)	Totaal
Normaal	26	14 (35%)	40
NormaalPlus	28	12 (30%)	40
Plus	26	14 (35%)	40
Totaal	80	40 (33.3%)	120

## Vraag 20

Met welke test kan nagegaan worden of er verschil zit tussen de percentages PTS-patiënten in de drie condities?

1. Een z-test voor één proportie.
2. Een z-test voor twee proporties.
3. Een chi-kwadraat test.
4. Een McNemar test.

## Aanvulling casus bij vraag 21 en 22

De resultaten vallen de onderzoeker een beetje tegen; hij had gehoopt minstens 10% verschil te vinden. Hij kijkt daarom ook naar de uitkomsten alleen voor de subgroep die een trombose in het bovenbeen had; dit is een duidelijke risicofactor voor het ontwikkelen van PTS. De resultaten staan in Tabel 6.

Tabel 6: PTS patiënten per conditie binnen de groep bovenbeen patiënten

	Geen PTS	Wel PTS (%)	Totaal
Normaal	6	12 (66,7%)	18
NormaalPlus	8	10 (55,6%)	18
Plus	6	12 (66,7%)	18
Totaal	20	34 (63,0%)	54

Nieuwsgierig geworden door deze uitkomst, wil de onderzoeker nagaan of het schijnbaar grotere effect van het nieuwe middel binnen de groep bovenbeen patiënten ook aangetoond kan worden op de continue Villalta-score. Daarom neemt hij naast de factor conditie ook de factor bovenbeen (variabele Bovenbeen met als waarden 'ja' en 'nee') en de interactie tussen conditie en bovenbeen op. Delen van de output van de analyses staan hieronder vermeld:

Anova Table (Type III tests)

Response: Villalta

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	55316	1	3162.9921	< 2.2e-16 ***
Bovenbeen	976	1	55.8270	1.738e-11 ***
Conditie	555	2	15.8621	8.361e-07 ***
Bovenbeen:Conditie	44	2	1.2507	0.2902
Residuals	1994	114		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	19.2041	0.8916	21.539	< 2e-16 ***
Bovenbeen[T.Ja]	7.2263	1.3291	5.437	0.000000312 ***
Conditie[T.Normaal]	1.3138	1.2609	1.042	0.2996
Conditie[T.NormaalPlus]	-2.7916	1.2609	-2.214	0.0288 *
Bovenbeen[T.Ja]:Conditie[T.Normaal]	-1.5057	1.8796	-0.801	0.4248
Bovenbeen[T.Ja]:Conditie[T.NormaalPlus]	-2.9727	1.8796	-1.582	0.1165

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.182 on 114 degrees of freedom

Multiple R-squared: 0.4383, Adjusted R-squared: 0.4137

F-statistic: 17.79 on 5 and 114 DF, p-value: 5.289e-13

## Vraag 21

Zijn er indicaties dat het effect van conditie op de continue Villalta-schaal groter zijn voor de bovenbeen-patiënten dan voor de onderbeen-patiënten?

1. Ja, er is aangetoond dat dit het geval is; dat zie je aan de p-waarde die vermeld staat bij Conditie in de ANOVA tabel.
2. Ja, er is aangetoond dat dit het geval is; dat zie je aan de p-waarde die vermeld staat bij Bovenbeen in de ANOVA tabel.
3. Nee, er is aangetoond dat dit niet het geval is; dat zie je aan de p-waarde die hoort bij de interactie tussen Conditie en Bovenbeen.
4. Dat kan op basis van deze analyse niet beslist worden.

### Vraag 22

Hoe groot is de gemiddelde score op de Villalta schaal voor de bovenbeen-patiënten in de NormaalPlus-conditie?

1.  $19,2 + 7,2$
2.  $19,2 + 7,2 + (-2,8)$
3.  $19,2 + 7,2 + (-2,8) + (-3,0)$
4.  $19,2 + 7,2 + (-2,8) + (-3,0) + 1,3 + (-1,5)$

### Vraag 23

Wanneer het artikel wordt ingediend, wordt het bekeken door twee onafhankelijke reviewers. De eerste reviewer merkt op dat in zijn ervaring de leeftijd van patiënten ook een risicofactor voor PTS kan zijn.

Hij vermoedt dat de verschillen tussen de condities die gerapporteerd worden het gevolg zijn van leeftijdsverschillen. Dit is...

1. een mogelijke verklaring aangezien de gemiddelde leeftijden niet precies gelijk zullen zijn in de drie condities.
2. waarschijnlijk gezien het feit dat leeftijd vrijwel altijd een invloed op resultaten heeft.
3. niet erg waarschijnlijk aangezien de groepen door toeval tot stand zijn gekomen.
4. niet erg waarschijnlijk aangezien de leeftijd een constante voor een patiënt gedurende de studie is.

### Vraag 24

De tweede reviewer merkt op dat zij gevonden heeft dat er een tamelijk sterk verband tussen BMI en de continue Villalta-score bestaat. Zij raadt aan om deze samenhang te gebruiken om een krachtigere analyse te doen.

Hoe zou deze analyse eruit moeten zien?

1. Er worden vershilscores berekend en deze worden geanalyseerd.
2. De variabele BMI wordt als covariaat aan de analyses toegevoegd.
3. De variabele BMI wordt als factor aan de analyses toegevoegd.
4. Eerst wordt gekeken of er een significant verschil tussen de condities in BMI is; alleen als zo'n verschil er is wordt de variabele BMI gebruikt in de analyse.

### Vraag 25

Als bij een toets de nulhypothese waar is, bestaat er toch een kans dat deze verworpen wordt. Deze kans...

1. heet het significantieniveau en de grootte ervan hangt af van de studiegrootte.
2. heet het significantieniveau en de grootte ervan hangt niet af van de studiegrootte.
3. heet de power en de grootte ervan hangt af van de studiegrootte.
4. heet de power en de grootte ervan hangt niet af van de studiegrootte.