

STATISTISCH VERSLAG BETREFFENDE EEN MICROBIOLOGISCHE VERGELIJKING TUSSEN HET VITAMINE B₁-GEHALTE IN BLOED VAN OUDE EN VAN JONGE MANNEN *)

door

Jkvr. H.SANDBERGH EN A.BENARD

1. Teneinde te onderzoeken of het vitamine B₁-gehalte van bloed daalt met de leeftijd werd in de loop van enige maanden van 45 gezonde jonge mannen en 45 gezonde oude mannen (verdeeld in 9 groepjes, corresponderend met 9 verschillende data van bloedafname) eenmaal bloed afgenomen. In elk bloedmonster werden op de dag van afname en op enige volgende dagen bepalingen van het B₁-gehalte verricht. Bij beide groepen, die we met „oud” en „jong” zullen aangeven, werd op dezelfde data bloed afgenomen en bij telkens 5 personen van iedere groep en de verkregen bloedmonsters werden zoveel mogelijk gelijk behandeld wat betreft de tijdstippen van de B₁-bepalingen per bloedmonster. De leeftijden van beide groepen lagen ver uiteen (17 - 42 jaar en 50 - 91 jaar).

Het statistisch onderzoek werd gesplitst in

1. een vooronderzoek naar mogelijke systematische invloeden van de wijze van waarneming op de uitkomsten, waardoor de resultaten beïnvloed zouden kunnen worden en
2. de eigenlijke vergelijking tussen jong en oud.

Bij elke toegepaste statistische toetsingsmethode wordt vermeld:

H₀: d.i. de hypothese die getoetst wordt,

H₁: d.i. de alternatieve hypothese ten gunste waarvan H₀ eventueel verworpen wordt en

k : d.i. de overschrijdingskans, die bij toepassing van de toets gevonden wordt.

De overschrijdingskans van een waarnemingsresultaat is een maat voor de aannemelijkheid van de getoetste hypothese. Het is de kans op een resultaat, dat even weinig als of nog minder dan het gevonden resultaat met de getoetste hypothese te rijmen valt. Is deze kans klein, dan gaat men tot verwerping van H₀ over ten gunste van de beschouwde alternatieve hypothese of hypothesen. Als grens voor verwerping van H₀ wordt vaak $k = 0,05$ genomen. Zie verder bv. [1] of [2] **).

*) Rapport S 146 van de statistische afdeling van het Mathematisch Centrum te Amsterdam. De afdeling staat onder leiding van Prof. Dr D.van Dantzig.

**) Cijfers tussen vierkante haken verwijzen naar de literatuurlijst aan het einde van dit verslag.

2. Het *statistisch vooronderzoek*, dat voor jong en oud apart werd uitgevoerd, betrof de volgende punten:

1) H_0 : het bewaren van een bloedmonster gedurende enkele dagen voor het verrichten van een waarneming heeft geen invloed op de waarnemingsuitkomst.

H^1 : het bewaren is wel van invloed.

Om dit te onderzoeken werd in de eerste plaats de toetsingsmethode der m rangschikkingen [3] toegepast op de bloedmonsters waarin 3 B_1 -bepalingen werden verricht. Het resultaat van de toets was bij oud: $k \approx 0,90$ en bij jong: $k \approx 0,80$. Er is dus geen aanleiding H_0 te verwerpen.

Aangezien niet van alle bloedmonsters 3 waarnemingen beschikbaar waren werd ook nog de symmetrietoets van J. HEMELRIJK (zie [4] en [5]) toegepast op het verschil tussen de eerste twee waarnemingen en op dat der tweede en derde apart. Hiermee wordt dezelfde hypothese getoetst. De uitkomsten waren:

oud; vergelijking tussen 1e en 2e bepaling; $k = 0,90$

jong; vergelijking tussen 1e en 2e bepaling; $k = 0,19$

oud; vergelijking tussen 1e en 2e bepaling; $k = 0,73$

jong; vergelijking tussen 1e en 2e bepaling; $k = 0,41$

Ook hier dus grote overschrijdingskansen en dus geen aanleiding H_0 te verwerpen.

2) H_0 : er bestaat geen verschil tussen de bepaling van de 9 verschillende groepen bloedmonsters.

H^1 : er zijn groepen waarvan de bepalingen systematisch hoger of lager dan het gemiddelde liggen.

Hierbij werd de „ k -steekproeven toets” van KRUSKAL (zie [6]) gebruikt met als resultaten:

oud; 1e bepaling; $k = 0,10$ } gecombineerd $k = 0,01$
jong; 1e bepaling; $k = 0,014$ }

oud; 2e bepaling; $k = 0,57$ } gecombineerd $k = 0,025$
jong; 2e bepaling; $k = 0,005$ }

oud; 3e bepaling; $k = 0,10$ } gecombineerd $k = 0,045$
jong; 3e bepaling; $k = 0,09$ }

(De combinatie geschiedt door optelling van de gevonden waarden van de toetsingsgrootte, niet door samenvoegen van de waarnemingen). Op grond van deze resultaten zullen we dus overgaan tot verwerping van H_0 ten gunste van H^1 . Het blijkt dat bij alle 6 groepen de bepalingen van de 8e bloedmonstergroep zeer laag zijn. Op grond van mededelingen van de onderzoeker kan aangenomen worden dat er bij deze groep iets bijzonders bij de bepalingen is opgetreden. Daarom is bij het verdere onderzoek deze bloedmonstergroep niet met de andere tot één groep samengevoegd. Na weglating van deze 8e groep werd nog eens bovenstaande toets toegepast met als resultaten:

II

Tabel IV

Nummers van de proefpersonen	Oud					Jong				
	Bloedgehaltenes in my thiamine/ml									
	1	2	3	4	5	6	7	8	9	0
4171-4170	76	61	68	62	82	-	88	85	99	93
	64	87	86	46	62	92	69	78	65	64
4221-4220	72	100	81	67	83	82	64	58	58	73
5061-5060	82	90	70	73	-	90	81	65	91	-
	59	75	54	75	-	77	71	66	75	-
5131-5130	90	63	-	95	75	108	89	-	109	62
	89	73	60	75	105	95	79	-	100	96
6101-6100	50	68	85	65	89	117	89	89	62	60
	72	71	89	-	-	-	104	80	-	-
7221-7220	-	69	78	76	63	71	70	60	66	77
	60	56	92	54	50	55	60	72	71	-
7291-7290	57	108	50	84	42	48	91	102	97	81
	94	67	61	86	57	90	62	78	103	99
	66	67	49	58	57	-	53	61	87	62
8051-8050	44	51	34	-	44	44	40	34	-	53
	42	85	40	78	51	27	43	29	45	61
	54	54	-	168	43	35	38	25	140	39
8121-8120	77	-	146	74	75	-	65	87	77	77
	76	50	83	78	70	109	65	67	56	77
	82	78	87	92	-	77	73	98	107	107
	Gem. 73,5 my thiamine/ml					Gem. 79,4 my thiamine/ml				

$$\begin{array}{l} \text{oud; 1e bepaling; } k = 0,69 \\ \text{jong; 1e bepaling; } k = 0,14 \end{array} \left. \vphantom{\begin{array}{l} \text{oud; 1e bepaling; } k = 0,69 \\ \text{jong; 1e bepaling; } k = 0,14 \end{array}} \right\} \text{gecombineerd } k = 0,34$$

$$\begin{array}{l} \text{oud; 2e bepaling; } k = 0,54 \\ \text{jong; 2e bepaling; } k = 0,095 \end{array} \left. \vphantom{\begin{array}{l} \text{oud; 2e bepaling; } k = 0,54 \\ \text{jong; 2e bepaling; } k = 0,095 \end{array}} \right\} \text{gecombineerd } k = 0,17$$

(Van de 3e bepalingen waren nu te weinig waarnemingen over om toetsing nog mogelijk te maken). Er is nu geen reden meer om de ge-
toetste hypothese te verwerpen.

3. *Vergelijking tussen jong en oud*

Op grond van het vooronderzoek, waarbij gebleken is, dat groep 8 duidelijk van de overige groepen verschilt - zodat ook de aanwezigheid van (kleinere) verschillen tussen de andere groepen niet uitgesloten geacht kan worden - is de vergelijking tussen jong en oud in de eerste plaats uitgevoerd met eliminatie van de invloed van deze (eventuele) verschillen. Daartoe is de toets van WILCOXON [7, 8], toegepast voor iedere groep apart en de voor de toetsingsgrootheid gevonden waarden zijn vervolgens door optelling gecombineerd tot een gezamenlijke toetsingsgrootheid voor alle groepen tesamen. Daarbij is gebruik gemaakt van de gemiddelden der bepalingen per bloedmonster. Dit laatste wordt gerechtvaardigd doordat bij het vooronderzoek geen reden aanwezig bleek te zijn, om aan de aequivalentie van de op elkaar volgende bepalingen bij een bloedmonster te twijfelen, terwijl verder aangenomen kan worden, dat het mislukken der bepalingen, dat nu en dan optrad, slechts toevalligerwijze voorkwam en niet in een van beide categorieën (jong en oud) vaker dan in de andere. Er werd gevonden een overschrijdingskans $k = 0,07$; in de steekproef ligt het B_1 -gehalte bij de oudere mannen lager dan bij de jongere.

Conclusie. Indien men de gebruikelijke onbetrouwbaarheidsdrempel $\alpha = 0,05$ aanhoudt, kan op grond van dit resultaat de hypothese, dat er geen verschil bestaat tussen de B_1 -gehalten van jong en oud, *niet* verworpen worden. Anderzijds kan opgemerkt worden dat de overschrijdingskans betrekkelijk laag is en dat dit materiaal dus wel een zwakke aanwijzing voor een verschil bevat. In de tweede plaats werd verworpen, dat er, afgezien van groep 8, geen duidelijke aanwijzing is voor verschil tussen de groepen. Daarom zijn, ter nader onderzoek, ook nog eens zoveel mogelijk de groepen tesamen gevoegd tot nieuwe groepen.

In verband met het verschillende aantal bepalingen per bloedmonster waarvan steeds het gemiddelde genomen is, zijn de groepen 1, ..., 7 en 9 tesamen in 3 deelgroepen van bloedmonsters verdeeld: die met een, met twee en met 3 bepalingen. Op ieder daarvan is apart

de toets van WILCOXON toegepast en daarna zijn de uitkomsten van deze 3 groepen en die van groep 8 door optelling van de toetsingsgrootheden gecombineerd. Hierbij werd een overschrijdingskans $k = 0,15$ gevonden, hetgeen dus bevestigt, dat het waarnemingsmateriaal geen voldoende aanwijzingen bevat om tot een verschil tussen jong en oud te besluiten.

4. *Betrouwbaarheidsinterval* voor het verschil in B_1 -gehalte. Om een indruk te verkrijgen van de grootte van een eventueel toch aanwezig verschil in B_1 -gehalte van het bloed tussen jong en oud werd voor dit verschil, d , een betrouwbaarheidsinterval bepaald, gebaseerd op boven gebruikte toets van WILCOXON. Onderstellen we nl., dat het (theoretisch) gemiddelde van het B_1 -gehalte bij de jonge mannen d hoger ligt dan bij de oude mannen, dan kunnen we deze hypothese toetsen door van alle bepalingen bij de jonge mannen d af te trekken en daarna oud en jong met de toets van WILCOXON te vergelijken. Het betrouwbaarheidsinterval voor het verschil, met onbetrouwbaarheid α wordt nu gevormd door alle waarden van d , welke bij toepassing van de toets een overschrijdingskans $k \geq \alpha$ geven.

Daar het bepalen van een dergelijke betrouwbaarheidsinterval het eenvoudigst is als men van twee groepen waarnemingen uitgaat zonder deze in deelgroepen te splitsen, is hierbij niet van gemiddelden van waarnemingen, maar alleen van de eerste bepaling bij ieder bloedmonster gebruik gemaakt, met weglating van de 8e bloedmonstergroep. Een meer verfijnde techniek, analoog aan de in de vorige paragraaf beschrevene, zou mogelijk zijn geweest maar kon door tijdgebrek niet toegepast worden. Als interval werd op deze manier gevonden: $(-4, +13)$ met onbetrouwbaarheid 0,04. In overeenstemming met het resultaat uit paragraaf 3 bevat het interval dus de waarde 0. In woorden uitgedrukt ligt dus het verschil in vitamine B_1 -gehalte in bloed bij oude en bij jonge mannen volgens dit waarnemingsmateriaal en behoudens een onbetrouwbaarheid 0,04 tussen: jong 4 lager dan oud en jong 13 hoger dan oud.

5. *Samenvatting der conclusies*

- a. Er is geen systematisch verschil gevonden tussen de verschillende bepalingen per bloedmonster.
- b. Er is wel een systematisch verschil gevonden tussen de groepen bloedmonsters, die op verschillende dagen zijn afgenomen; in het bijzonder bleek groep 8 sterk van de overige af te wijken.
- c. Het waarnemingsmateriaal geeft geen aanleiding om tot een verschil tussen jong en oud te concluderen. Een zwakke aanwijzing hiervoor is aanwezig. Mogelijkerwijze zou een uitgebreider waar-

nemingsmateriaal leiden tot de conclusie, dat er een klein systematisch verschil is. Volgens het huidige onderzoek zou dit verschil tussen jong en oud een waarde tussen -4 en +13 bezitten.

L iteratuur

- [1] W.J. DIXON and F.J. MASSEY Jr., Introduction to Statistical Analysis, New York, Toronto, London, 1951, Mc Graw-Hill Book Company Inc.
- [2] J. NEYMAN, First course in probability and statistics, New York, 1950, Henry Bolt and Company.
- [3] M.G. KENDALL, Rank correlation methods, London, 1948, hoofdstuk 6.
- [4] J. HE MEL RIJK, A family of parameterfree tests for symmetry with respect to a given point I, II, Proceedings van de Kon. Ned. Ak. v. Wet., 53 (1950), pp. 945 - 955
Indagationes Mathematicae, 12 (1950), pp. 340 - 350.
- [5] J. HE MEL RIJK, Symmetrietoetsen, Diss., Den Haag 1950, Excelsior.
- [6] W.H. KRUSKAL and W.A. WALLIS, Use of ranks in one criterion Analysis of variance, Journ. Am. Stat. Ass., 47 (1952), pp. 538 - 621.
- [7] F. WIL COXON, Individual comparisons by ranking methods, Biometrics 1, (1945), pp. 30 - 83.
- [8] MANN en WHITNEY, On a test whether one of two random variables is statistically larger than the other, Ann. Math. Sta. 18(1947), pp. 50 - 60.

