

Een nieuw scoresysteem voor de Stichting Subcommissie Stolling van de CCKL?

G. L. A. REIJNIERSE¹ en A.M.H.P. van den BESSELAAR²

Scoresysteem

Het scoresysteem van de Federatie van Nederlandse Trombosediensten en dat van de Subcommissie Stolling gaat uit van een normale verdeling van de uitslagen. Bij betrekkelijk grote aantallen uitslagen per rondzending (fibrinogeen, ca. 110; antitrombine-III, ca. 60; factor VIII, ca. 35) gaat deze veronderstelling redelijk op. Bij kleinere aantallen uitslagen is deze aanname problematischer (protrombinetijd, APTT en INR, die per preparaat worden beoordeeld).

Het toekennen van scores verloopt in twee stappen: eerst wordt de tolerantie-interval-methode toegepast om uitbijters aan te wijzen en om het consensusgemiddelde en de consensusstandaardafwijking te berekenen. Daarna wordt van elke deelnemer de Z-score of excentriciteit berekend en de bijbehorende score afgelezen in een tabel.

Bij de tolerantie-interval-methode (1, 2), die ook wordt gebruikt bij het Combi-programma van de SKZL (3), worden de volgende stappen doorlopen:

1. Bereken het gemiddelde van de waarden van alle deelnemers.
2. Verwijder alle meetuitkomsten die groter zijn dan tweemaal het onder stap 1 berekende gemiddelde (evidente "grove fouten").
3. Bereken het gemiddelde en de standaardafwijking van de resterende meetuitkomsten.
4. Stel vast welke meetuitkomst het verst van het (geschoonde) gemiddelde verwijderd ligt. Dit kan zowel een lage als een hoge waarde betreffen. Plaats deze meetuitkomst apart.
5. Bereken het gemiddelde \bar{x} en de standaardafwijking s van de dan resterende meetuitkomsten.
6. Bereken een interval van de vorm $\bar{x} \pm k.s$ op basis van de onder stap 5 berekende \bar{x} en s . De waarde van k als functie van n (het aantal meetuitkomsten minus de kandidaat-uitbijter) wordt afgelezen uit een tabel (1, 2). De factor k is zo gekozen dat gemiddeld 99% van de populatie te verwachten is tussen $\bar{x} \pm k.s$. Hierbij wordt geaccepteerd dat gemiddeld 1 % van de tot de onderzochte populatie behorende meetuitkomsten ten onrechte als uitbijter zal worden aangemerkt. Het genoemde interval

$\bar{x} \pm k.s$ heet een tolerantie-interval. Indien de extreme meetuitkomst hierbinnen valt, dan geldt deze niet als uitbijter, en rapporteert men gemiddelde en standaardafwijking als berekend onder stap 3. Indien de extreme meetuitkomst wel buiten het berekende interval valt, dan geldt deze als uitbijter en wordt verder buiten beschouwing gelaten.

7. De stappen 4 tot en met 6 worden herhaald.
8. Van de uiteindelijk overgebleven meetuitkomsten rapporteert men het consensusgemiddelde en de consensusstandaardafwijking.

Vervolgens wordt van elke deelnemer de Z-score of excentriciteit berekend volgens de formule: Z-score = verschil van de waarde van de deelnemer minus het consensusgemiddelde, gedeeld door de consensusstandaardafwijking. Scores worden toegekend volgens tabel 1.

Gevolgen van het systeem

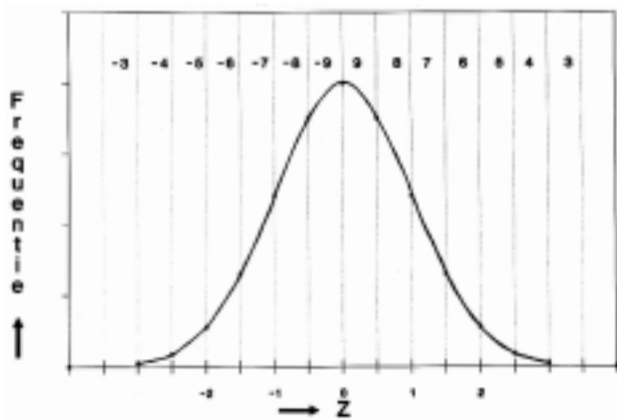
Eén van de eigenschappen van de normale verdeling is dat het oppervlak onder de curve kleiner wordt indien de curve vanuit het midden naar buiten gaand wordt afgesneden per halve standaardafwijking (figuur 1). We verwachten dat 38,3% van de waarnemingen een 9 zal ontvangen; 30,0% een 8; 18,4% een 7; 8,8% een 6; 3,3% een 5; 1,0 % een 4; 0,2% een 3; 0,04% een 2, en 0,006 % een 1. De te verwachten gemiddelde score zal 7,9 bedragen indien er geen uitbijters zijn. Karakteristieken van dit scoresysteem zijn, dat 95,5% voldoende scoort; 3,3% twijfelachtig en 1,2% onvoldoende (figuur 2). Overigens zullen genoemde percentages bij de rondzendingen iets anders liggen: a) bij kleine aantallen monsters hebben we niet te maken met een normale verdeling, maar met een Student-verdeling, die afhankelijk van het

Tabel 1. Huidige systeem van score-toekenning, waarbij $Z_i = (x_i - \bar{x}) / s$

	$ Z_i $		Score
0,0	tot	0,5	9
0,5	tot	1,0	8
1,0	tot	1,5	7
1,5	tot	2,0	6
2,0	tot	2,5	5
2,5	tot	3,0	4
3,0	tot	3,5	3
3,5	tot	4,0	2
>4,0			1

Rijnland Ziekenhuis, Leiderdorp¹ en Stichting RELAC, Leiden²

Correspondentie: Dr. G.L.A. Reijnierse, Klinisch Chemisch en Medisch Microbiologisch Laboratorium, Rijnland Ziekenhuis, Postbus 4220, 2350 CC Leiderdorp.
Ingekomen: 03.09.96



Figuur 1. Normale verdeling van de Z-score met daarin de toekenning van de scores van 3 tot en met 9

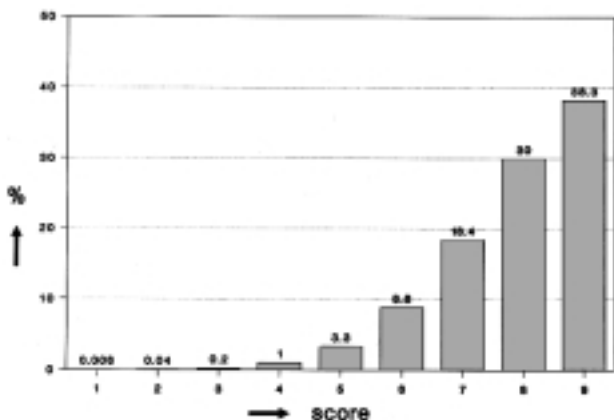
aantal waarnemingen breder en platter is dan een normale verdeling; b) uitbijters met lage scores maken de verdeling nog platter, zodat er meer lage scores dan verwacht, zullen worden toegekend. Dit verlaagt de te verwachten gemiddelde score.

Voorgesteld scoresysteem

Onze ervaring is dat sommige deelnemers aan ons programma de toegekende scores interpreteren als vroeger de schoolcijfers werden beleefd. In dat geval verschaffen wij hun een te rooskleurige blik op hun resultaten. Op school waren er meer zessen dan negens; bij ons zijn er meer negens dan zessen.

De Subcommissie Stolling meent dat de scoretoekenning niet tot te hoge scores moet leiden. Beter is het om aansluiting te zoeken bij de schoolcijfers. De score 6 moet het meeste voorkomen, er moeten minder zevens dan zessen zijn, minder achten dan zevens, enzovoort. Hierbij dient een arbitraire keuze gemaakt te worden, die vermeld staat in tabel 2. De te verwachten gemiddelde score zal (zonder uitbijters) 6,9 bedragen. Karakteristieken van dit scoresysteem zijn, dat 90% voldoende scoort; 5% twijfelachtig en 5% onvoldoende (figuur 3). De relatie tussen Z-score en toegekende score staat vermeld in tabel 3.

De resultaten van toepassing van het huidige en het voorgestelde systeem van scoretoekenning op de INR van de rondzending van maart 1996 staan vermeld in tabel 4.



Figuur 2. Karakteristieken van het huidige scoresysteem

Tabel 2. Voorgestelde scoretoekenning

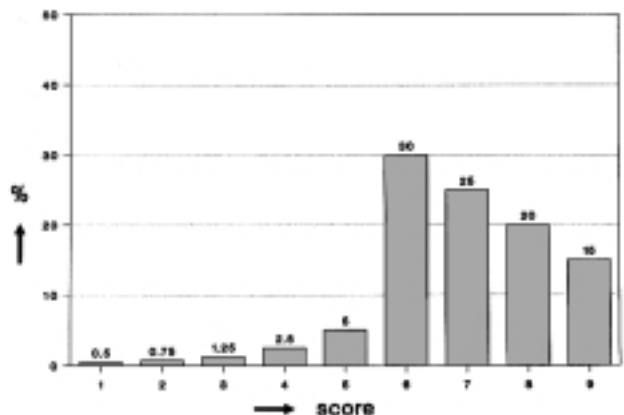
Percentage deelnemers	Absolute score
15	9
20	8
25	7
30	6
5	5
2,5	4
1,25	3
0,75	2
0,50	1

Tabel 3. Relatie tussen Z-score en toegekende score

$ Z_i $		Score	
0,0	tot	0,189	9
0,189	tot	0,454	8
0,454	tot	0,842	7
0,842	tot	1,645	6
1,645	tot	1,960	5
1,960	tot	2,240	4
2,240	tot	2,500	3
2,500	tot	2,818	2
>2,818			1

Tabel 4. Scores met het huidige en voorgestelde systeem (INR maart 1996)

Score	Aantal huidige systeem	Aantal voorgesteld systeem
9	19	7
8	21	12
7	10	15
6	6	20
5	2	2
4	0	2
	58	58



Figuur 3. Karakteristieken van het voorgestelde scoresysteem

We merken op, dat de scores in benedenwaartse richting zijn bijgesteld. De gemiddelde score was 7,84 (verwacht 7,9) en wordt bijgesteld naar 6,93 (verwacht 6,9), waarbij de verwachte karakteristieken van het voorgestelde systeem redelijk uitkomen.

Scoresystemen elders

In andere landen worden andere criteria ter beoordeling van individuele laboratoria gebruikt. In het Verenigd Koninkrijk wordt in het "U.K. National External Quality Assessment Scheme for Blood Coagulation" de INR als voldoende ("within consensus") of onvoldoende ("outwith consensus") beoordeeld. Onvoldoende is gedefinieerd als een afwijking groter dan 15% van de mediaan van de INR-uitslagen met een bepaald preparaat. Bij de andere screening tests (PT, APTT en trombinetijd) wordt een soortgelijk criterium voor de score gebruikt. Per preparaat wordt de mediaan van de stollingstijdratio berekend mits het aantal deelnemers met dit preparaat minimaal 10 is. De afwijking (in %) ten opzichte van de mediaan van de ratio bepaalt de score (voldoende of onvoldoende). Bij de andere bepalingen, zoals fibrinogeen en factor VIII, wordt uitgegaan van de mediaan van alle resultaten als referentiepunt. De uitslagen worden ingedeeld in 5 fracties. De dichtstbijzijnde 50% van de resultaten rond de mediaan krijgt een A; de volgende 20%, naar buiten gaand, krijgt een B; de volgende 10%, naar buiten gaand, krijgt een C; de volgende 10% een D; en de laatste 10% een E. Op basis van de combinatie van scores die een deelnemer in twee opeenvolgende rondzendingen ontvangen heeft, wordt een resultaat toegekend. Als onvoldoende resultaat geldt de combinatie DD, CE, EC, DE, ED en EE.

Vergelijking van het Nederlandse en Britse systeem levert de volgende verschillen op:

1. Nederland: rekenkundig gemiddelde als consensus.
Verenigd Koninkrijk: mediaan als consensus.
2. Nederland: uitbijters worden verwijderd voor berekening van gemiddelde, standaardafwijking en variatiecoëfficiënt.

Verenigd Koninkrijk: uitbijters worden niet verwijderd.

3. Nederland: score wordt toegekend op basis van excentriciteit en kwantilenindeling.
Verenigd Koninkrijk: score wordt toegekend op percentage afwijking (screeningstesten), of op basis van rangnummer en kwantilenindeling (overige bepalingen).
4. Nederland: de score van één rondzending bepaalt het resultaat.
Verenigd Koninkrijk: niet alleen de score van één rondzending, maar ook de combinatie van scores in opeenvolgende rondzendingen bepaalt het resultaat van een deelnemer.

Ter volledigheid wordt nog vermeld dat in Zwitserland scores voor de INR worden toegekend zoals in het Verenigd Koninkrijk, en dat in Frankrijk alleen Z-scores worden toegekend.

Beschouwing

Dit brengt ons terug bij het begin. Het toekennen van scores is arbitrair. De Subcommissie Stolling meent echter dat het huidige systeem enige bijstelling behoeft. Mede tengevolge van opmerkingen dienaangaande van deelnemers aan het Symposium "Bloedstolling: Kwaliteitsbewaking en apparatuur-evaluatie, behandelmethoden en ziektebeelden" (Apeldoorn, 7 juni 1996), zal de Subcommissie hierover op korte termijn een aantal voorstellen uitbrengen.

Literatuur

1. Reijnierse GLA, Besselaar AMHP van den, Hermans J. Een nieuw verwerkingsprogramma van ingezonden uitslagen in het kader van externe kwaliteitsbewaking. Ervaringen in 1988. Tijdschrift NVKC 1989; 14: 122-127.
2. Reijnierse GLA, Velde EA van der. Statistiek voor Laboratoriumonderzoek Deel 2. Groningen, Wolters-Noordhoff 1992; 178-181.
3. Steigstra H, Jansen RTP, Baadenhuijsen H. Combi Scheme: New Combined Internal/External Quality-Assessment Scheme in The Netherlands. Clin Chem 1991; 37: 1196-1204.