

Projectnr.: 505.0060
Normalisatie Monsterneming en Analyse
Projectleider: dr W.G. de Ruig

Rapport 90.56 December 1990

Chemometrische evaluatie van
het keuringsonderzoek Hormonen

dr W.G. de Ruig
ir A.A.M. Jansen

Afdeling: Coördinatie Chemometrie

Goedgekeurd door dr F. A. Huf

Rijks-Kwaliteitsinstituut voor land- en tuinbouwprodukten (RIKILT)
Bornsesteeg 45, 6708 PD Wageningen
Postbus 230, 6700 AE Wageningen
Telefoon 08370-75400
Telex 75180 RIKIL
Telefax 08370-17717

Copyright 1990, Rijks-Kwaliteitsinstituut voor land- en
tuinbouwprodukten
Overname van de inhoud is toegestaan, mits met duidelijke
bronvermelding

VERZENDLIJST

INTERN:

directeur
sectorhoofden
coördinatie Chemometrie
programmabeheer en informatieverzorging (2x)
bibliotheek
circulatie

EXTERN:

Dienst Landbouwkundig Onderzoek
Directie Wetenschap en Techniek
Directie Voedings- en Kwaliteitsaangelegenheden
Directie Rijksdienst voor de keuring van Vee en Vlees
Groep Landbouwwiskunde, ir A.A.M. Jansen
Centraal Laboratorium van de Rijksdienst voor de keuring van Vee en
Vlees, L.M.H. Frijns (2x)
Rijksinstituut voor Volksgezondheid en Milieuhygiene, dr R.W. Stephany

INHOUD	1
SAMENVATTING	3
1 INLEIDING	5
2 ONTWIKKELING METHODEN VOOR HORMONENONDERZOEK	5
3 KEURINGSKARAKTERISTIEK VAN EEN KEURINGSONDERZOEK	6
4 KEURING EN ANALYSE	8
5 BEMONSTERINGS-/ ANALYSESTRATEGIE HORMONENONDERZOEK	10
5.1 Inleiding	10
5.2 Eenvoudig model ter verkenning van het probleem	11
5.3 Een realistischer model	13
5.4 Vaststelling van de eigenschappen van de meet- en keuringsprocedure	15
5.5 Multivariate aspecten	16
6 KWALITEITSBORGING	17
7 ONDERZOEKSVORSTEL	17
8 OPTIMALISERING INFORMATIEINHOUD	18
BIJLAGE 1: NORMALE PROCEDURE CLRVV	20
BIJLAGE 2: MATHEMATISCHE BESCHRIJVING VAN HET IN 5.3 BESCHREVEN MODEL	22

()

()

SAMENVATTING

Nadat in 1981 een aantal fout-positieve uitslagen bekend werden, is de uitvoering van de analyses en de daarbij gehanteerde criteria gericht geweest op het voorkómen van fout-positieve resultaten. Dit beleid is succesvol geweest in die zin, dat sindsdien fout-positieve resultaten op theoretische gronden uiterst onwaarschijnlijk zijn en in de praktijk ook niet zijn waargenomen. In de normale routineprocedure wordt een groot aantal controlemonsters meegeanalyseerd voor de borging van de procedure. Op een totaal van 36 monsters per dag worden 33 controlemonsters geanalyseerd.

De keerzijde van het genoemde beleid is, dat er een relatief grote kans op het optreden van fout-negatieve resultaten bestaat. Tot op heden ontbrak elk zicht op de frequentie hiervan.

Bij de keuringsbeslissing worden grote veiligheidsmarges in acht genomen. Dit geschiedt mede omdat de onderbouwing van de keuring t.a.v. de kans op fout-positieve en de kans op fout-negatieve resultaten ontbreekt.

Deze onderbouwing is wel te leveren. Daarvoor is echter nader onderzoek noodzakelijk aan de hand van monsters met bekende gehalten. Dit onderzoek dient gemengd in de normale routineprocedure te worden uitgevoerd, in casu door het CLRVV. Dit kan zonder de huidige analysecapaciteit in ernstige mate aan te tasten. De onderzoekmonsters zouden namelijk gedeeltelijk in de plaats kunnen komen van de huidige controlemonsters, die geheel gericht zijn op het voorkomen van fout-positieve resultaten. Met de resultaten kan men een *keuringskarakteristiek* opstellen, dat is een figuur die de hoedanigheid van een bepaalde keuringsstrategie weergeeft. Van belang zijn de *onbetrouwbaarheid* van een keuring, d.w.z. de kans op ten onrechte afkeuren en het *onderscheidingsvermogen*, dat is de kans op terecht afkeuren.

Op grond van de onderzoeksresultaten kan voor gekozen kansen op fout-positieve en fout-negatieve resultaten een goed onderbouwde bemonsterings- en keuringsstrategie opgesteld worden.

Ook kan er een optimalisatie plaats vinden in het gebruik van voorhanden informatie. Mogelijkheden hiertoe zijn:

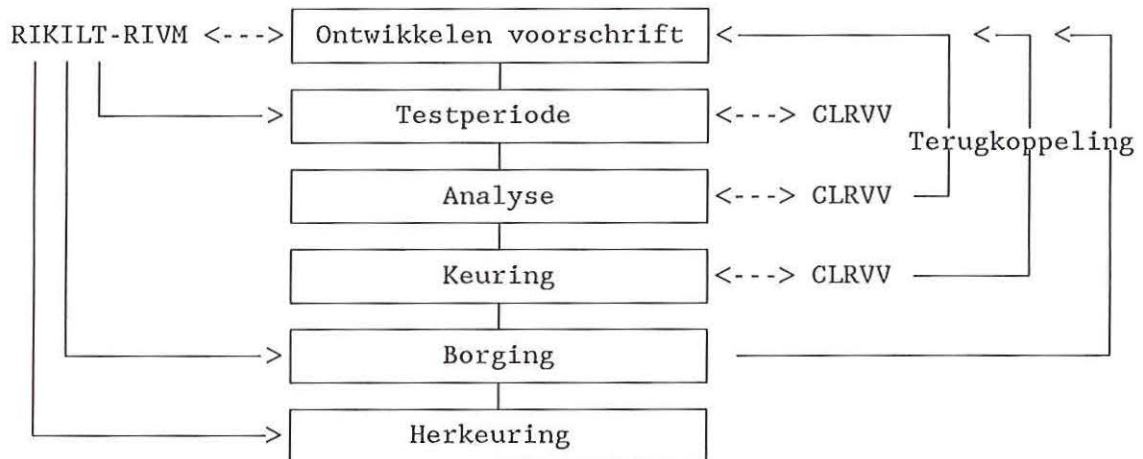
- verbeteren van criteria
- criteria opstellen voor 'verdachte' monsters
- opsporen van hormoonpatronen.

Reeds thans zijn een aantal in de keuring niet-positief bevonden monsters 'verdacht' te noemen, wat aanleiding kan geven tot maatregelen. Aan de hand van meer gegevens kan dit wellicht omgezet worden in 'positief'.

De strategie, die in dit rapport wordt beschreven, is gericht op de keuring van hormonen, maar is daartoe niet beperkt. De aanpak is algemeen toepasbaar voor vergelijkbaar onderzoek, bijvoorbeeld in het kader van KAP en voor EEG regelingen.

1 INLEIDING

De gang van zaken voor het ontwikkelen en toepassen van analysevoorschriften voor keuringsdoeleinden van de RVV voor hormonen kan als volgt worden weergegeven.



Met het oog op de borging en aan de hand van de analyseresultaten en keuringsresultaten dient er een terugkoppeling plaats te vinden. Dit kan leiden tot het aanpassen van bestaande analysevoorschriften of het opstellen van nieuwe analysevoorschriften, zonodig gevolgd door een nieuwe testperiode. Men kan hierbij denken aan aanwijzingen in de analyseresultaten voor nieuwe analyten die nog niet in het pakket van onderzochte analyten opgenomen zijn.

2 ONTWIKKELING METHODEN VOOR HORMONENONDERZOEK

In 1981 zijn fout-positieve resultaten voor het aantonen van DES in runderurine met een RIA-methode voorgekomen. Statistisch gezien betrof het slechts een zeer gering aantal op de duizenden onderzochte monsters, maar de consequenties waren groot. Er kwamen schadeclaims en de rechter in Almelo verbood 'de Staat der Nederlanden en het RIV' nog langer een dergelijke 'onbetrouwbare methode' te gebruiken, maar sprak het vertrouwen uit dat de Staat der Nederlanden en het RIV in staat zouden zijn een betrouwbare methode te ontwikkelen. Er is toen, binnen de ad hoc werkgroep EVATH, gewerkt aan een methodiek, die fout-positieve resultaten vrijwel uitsloot. Het RIKILT heeft bijgedragen met een methode, die berust op gaschromatografie gecombineerd met laag-oplossend vermogen massaspectrometrie (GC-LRMS); deze methode wordt thans, in aangepaste vorm, door het CLRVV voor het onderzoek op

hormonen gebruikt. De methoden werden nauwkeurig omschreven en de betrouwbaarheid door ringtesten aangetoond. Het ringonderzoek alleen voor de bepaling 'DES in urine' heeft ruim f 500 000 gekost. Voor de talrijke hormonen die gebruikt kunnen worden en vele monstersoorten bleek deze aanpak dan ook onhaalbaar. Om met nieuw-ontwikkelde methoden, die echter niet geringst waren, toch keuringen te kunnen uitvoeren zijn daarop door De Ruig, Stephany en Dijkstra 'criteria' opgesteld voor een reeks analysetechnieken; deze zijn in zeer korte tijd in de EG wetgeving opgenomen (Beschikking 410/87/EEG en 610/89/EEG).

Ingegeven door het DES debacle en het rechterlijke vonnis is vanuit het beleid de eis gesteld, dat fout-positieve resultaten (de stof wordt aangetoond, maar is in werkelijkheid niet aanwezig) niet mochten vóórkomen.

Eerst binnen de EVATH en later bij het opstellen van de criteria voor het aantonen van residuen is derhalve gelet op het voorkómen van fout positieve resultaten.

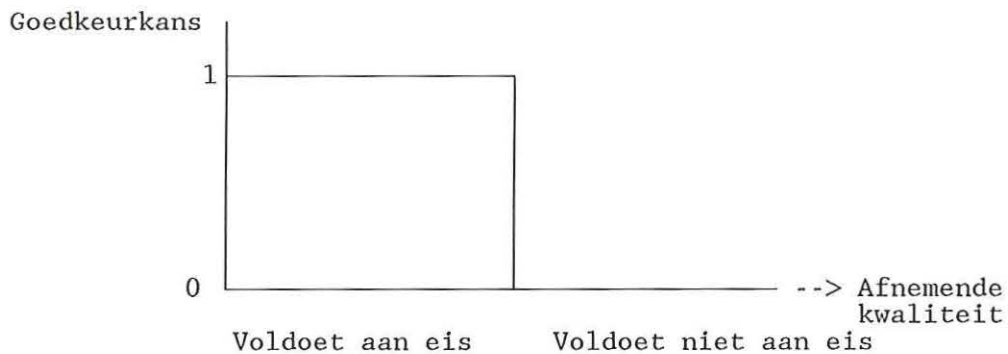
Deze aanpak is in zoverre succesvol geweest, dat het optreden van fout-positieve resultaten op theoretische gronden vrijwel uitgesloten geacht moet worden en in de praktijk tot op heden ook niet is geconstateerd.

De keerzijde is echter, dat de criteria zó scherp gesteld zijn, dat monsters die wel hormonen bevatten in relatief veel gevallen niet als zodanig worden aangemerkt. Tot dusverre ontbrak het zicht op de frequentie van het optreden van deze 'fout-negatieve' resultaten.

3 KEURINGSKARAKTERISTIEK VAN EEN KEURINGSONDERZOEK

Wat men als keurende instantie zou willen, is dat men monsters die aan de gestelde eisen voldoen altijd goedkeurt en monsters die niet aan de eisen voldoen altijd afkeurt, grafisch weergegeven in figuur 1.

Een dergelijke ideale 'keuringskarakteristiek' is alleen te verwezenlijken bij een volledig betrouwbaar onderzoek van alle elementen uit een populatie. In de praktijk, wanneer men een steekproef neemt, en bij het vaststellenn van de resultaten ook meetonnauwkeurigheden een rol spelen, zal de keuringsbeslissing niet altijd juist zijn.



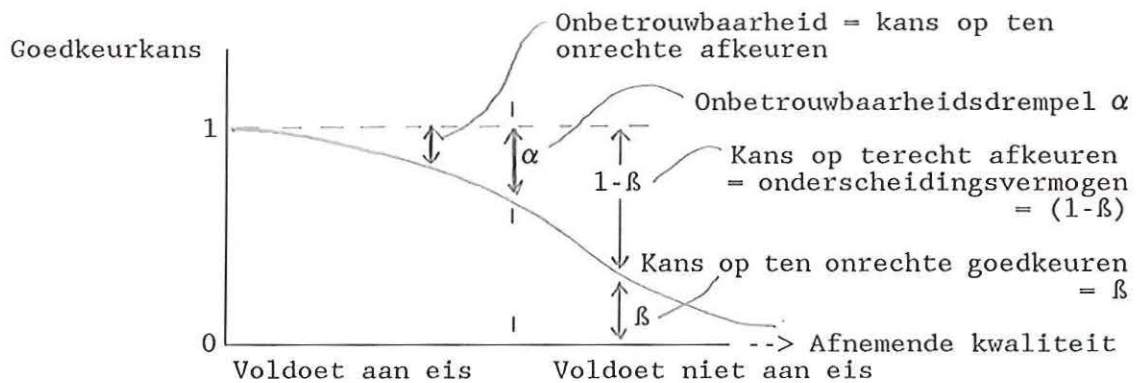
Figuur 1. Ideale keuringskarakteristiek

In het geval van hormonen betekent dit (zie schema 1), dat bij werkelijke aanwezigheid van hormonen er toch altijd een kans op goedkeuren aanwezig is ('fout-negatief') en bij werkelijke afwezigheid een kans op afkeuren ('fout-positief').

Schema 1. Kans op goed- en afkeuren in relatie tot werkelijke situatie

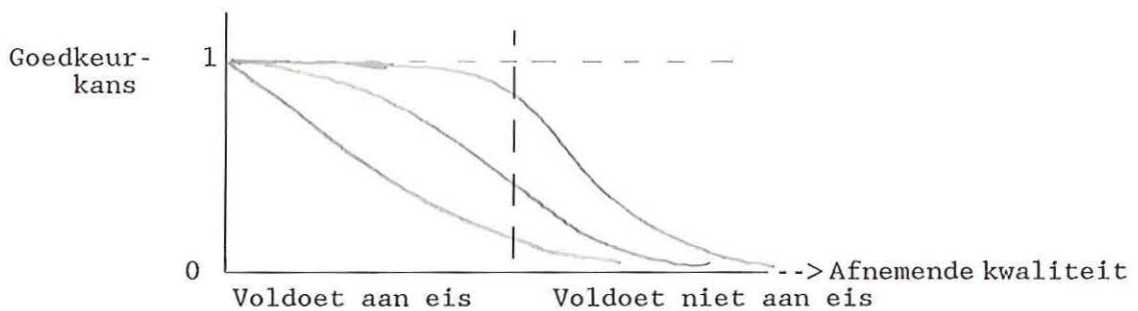
		Werkelijkheid	
		Afwezig	Aanwezig
Keuringsbeslissing	Goedkeuren	'Negatief' Juist kans: $(1-\alpha)$	Fout-negatief Fout vd 2e soort kans: β
	Afkeuren	Fout-positief Fout vd 1e soort kans: α	'Positief' Juist kans: $(1-\beta)$

Hierdoor krijgt de keuringskarakteristiek een vorm als in figuur 2. De kans op terechte afkeuring wordt het *onderscheidingsvermogen* van de keuringsprocedure genoemd. De maximale kans op ten onrechte afkeuren is de *onbetrouwbaarheidsdrempel*, α .



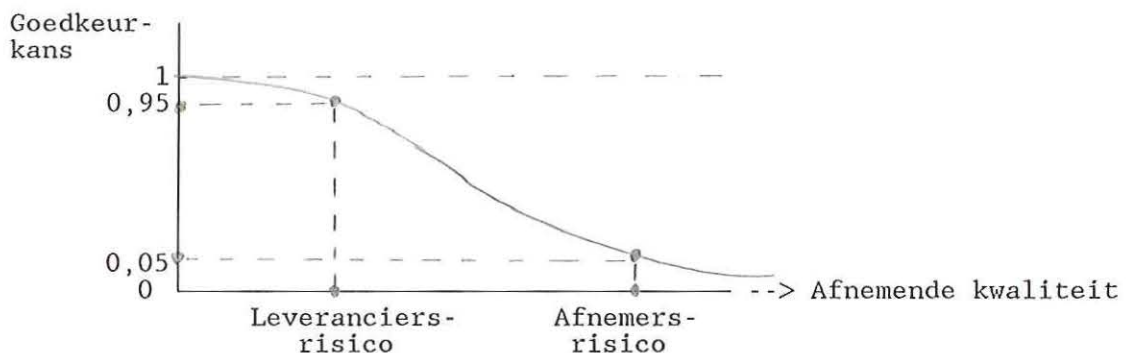
Figuur 2. Keuringskarakteristiek in de praktijk

Men kan door de wijze van keuren de vorm van de keuringskarakteristiek beïnvloeden: meer of minder steil, α groter of kleiner, β groter of kleiner, zie figuur 3.



Figuur 3. Verschillende keuringskarakteristieken

In het handelsverkeer tussen producenten en afnemers legt men vaak twee punten van de keuringskarakteristiek vast: een punt bij hoge kwaliteit waarvoor een kleine afkeurkans (meestal 5%) moet gelden (het leveranciersrisico) en een punt bij lage kwaliteit waarvoor een kleine goedkeurkans (ook 5%) moet gelden (het afnemersrisico), zie figuur 4.



Figuur 4. Keuringskarakteristiek met leveranciers- en afnemersrisico

De redeneringen uit het handelsverkeer zijn ook toepasbaar op de overheidscontrole, zoals in het geval van de hormonen. Ook hier kan men eisen stellen aan α en β en de consequenties daarvan voor de keuringsprocedure nagaan. In §6 wordt dit voor het hormonenonderzoek nader uitgewerkt.

4 KEURING EN ANALYSE

Ten behoeve van de keuring wordt analytisch onderzoek verricht. We gaan er hierbij stilzwijgend van uit, dat er tussen beide een één-één correspondentie bestaat: als het analyseresultaat positief is, is het keuringsresultaat dat ook en hetzelfde geldt voor het negatieve geval. In het geval van de hormonen is dit helaas niet het geval.

De keuringsvraag luidt: is er een illegale toepassing van hormonen?

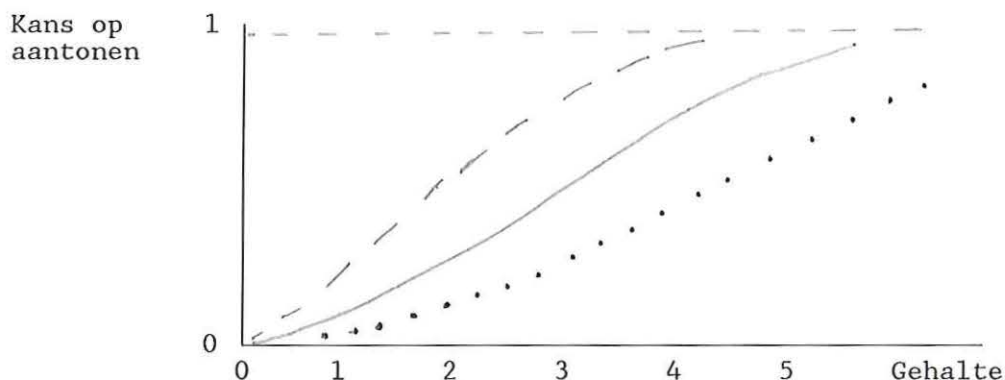
Het analytische antwoord is: in het monster zijn hormonen (niet) aangetoond. De aanwezigheid van hormonen is echter nog geen bewijs van illegale toepassing: ze kunnen ook van nature aanwezig zijn. Dit betreft dus de natuurlijke hormonen.

Omgekeerd hoeft illegale toepassing van hormonen niet altijd te leiden tot hogere gehalten. Bij toepassing van natuurlijke hormonen is in monsters ook verlaging van de gehalten geconstateerd.

De analytische uitspraak 'aangetoond' zal dus met voorzichtigheid moeten worden vertaald in de keuringsbeslissing 'illegale toepassing'.

Wel is het zo, dat bij illegale toepassing het gehele hormonenpatroon verstoord raakt. Het gehalte van één hormoon mag dan niet zoveel zeggen, uit het patroon van gehalten van een aantal hormonen kunnen wellicht toch conclusies voor de keuring worden getrokken.

Ook de analyse zelf is niet volmaakt. Zie figuur 5. Hoge gehalten zullen doorgaans wel positief bevonden worden. Zeer lage gehalten, beneden de aantoonbaarheidsgrens, worden nooit aangetoond. Daartussen is een gebied, waarvoor bij een bepaald gehalte een fractie positief bevonden wordt. Deze fractie is afhankelijk van het gehalte en neemt bij toenemende gehalten toe. De kans op aantonen wordt niet alleen bepaald door het gehalte, maar ook van de aard van de analyt en is voorts sterk afhankelijk van de matrix. Bij een schone matrix zal de curve meer naar links liggen en steiler lopen: lagere gehalten worden in veel gevallen ook nog gedetecteerd (curve ----); voor een vuile matrix ligt de curve meer naar rechts (curve).



Figuur 5. Kans op aantonen in afhankelijkheid van het gehalte en de matrix.
 - - - - = schone matrix
 = vuile matrix
 ————— = vuilheid matrix tussen beide in

5 BEMONSTERINGS-/ ANALYSESTRATEGIE HORMONENONDERZOEK

5.1 Inleiding

De detectie van de aanwezigheid van hormonen in urinemonsters gebeurt m.b.v. een combinatie van gaschromatografie en massaspectrometrie. Vastgesteld worden de retentietijd (R_t) en de intensiteiten van de 4 belangrijkste massafragmenten (gemeten als piekhoogte cq. piekoppervlak).

Elementen van de keuringsprocedure zijn:

- 1) R_t moet specifiek zijn voor hormoon (R_t van het analyt mag niet meer dan een afgesproken waarde afwijken van de R_t van het standaardanalyt),
- 2) de voor een hormoon specifieke brokstukken moeten in de specifieke verhoudingen aanwezig zijn (afwijkingen tot 10% zijn toelaatbaar),
- 3) het gehalte wordt vastgesteld als het produkt van het gehalte in een standaardmonster en de verhouding van de piekoppervlakken van te meten analyt en het standaardanalyt,
- 4) afkeuring volgt als is voldaan aan 1) en 2) en als het gehalte hoger is dan 5 ppb voor niet-natuurlijke en hoger dan ..ppb voor natuurlijke hormonen.

Van de keuringsprocedure is niet exact bekend wat de relatie is tussen de werkelijke hormoongehalten in de urinemonsters en de bij de keuring vastgestelde hormoongehalten. Het is daarom niet mogelijk uit de bij de keuring verkregen gegevens de werkelijke situatie in de praktijk te reconstrueren.

Het doel van deze paragraaf is na te gaan en aan te geven:

- hoe men, als men de eigenschappen van de keuringsprocedure kent, de werkelijke situatie in de populatie kan vaststellen;
- hoe men de eigenschappen van de keuringsprocedure onder praktijkomstandigheden in handen krijgt;
- hoe beleidskeuzen en keuringsprocedure op elkaar dienen te worden afgestemd.

5.2 Een eenvoudig model ter verkenning van het probleem

Eenvoudigheidshalve wordt in eerste instantie aangenomen dat het gaat om slechts 1 hormoon en dat de gehalten van dit hormoon in urinemonsters behoren tot 1 van de volgende 3 klassen:

- 0 het hormoon is niet aanwezig,
- 1 het hormoon is aanwezig maar de hoeveelheid is minder dan een aangenomen drempelwaarde,
- 2 het hormoon is in een hoeveelheid $>$ de drempelwaarde aanwezig.

Het zou fraai zijn als de keuringsprocedure deze 3 klassen feilloos zou kunnen onderscheiden. De werkelijkheid is echter dat meet- en beslissingsprocedures meestal een element van onzekerheid bevatten. De keuringsprocedure zal daarom een urinemonsters uit klasse 0 met bepaalde kansen h_{00} , h_{01} en h_{02} toewijzen aan de klassen 0, 1 of 2. Evenzo worden urinemonsters uit klasse 1 met kansen h_{10} , h_{11} en h_{12} en uit klasse 2 met kansen h_{20} , h_{21} en h_{22} toegewezen aan de klassen 0, 1 of 2. De kansen h_{ij} beschrijven de eigenschappen van de procedure. Als men deze kent is het mogelijk te voorspellen hoe de verdeling van de keuringsuitslagen (y_0 , y_1 , y_2) zal worden bij een gegeven verdeling (fracties p_0 , p_1 , p_2) van de hormoongehalten in de populatie. Omgekeerd is het ook mogelijk uit de verdeling van de keuringsuitslagen de werkelijke verdeling te berekenen. Dit laatste komt neer op het oplossen van een eenvoudig stelsel van 3 vergelijkingen met 3 onbekenden:

$$y_0 = h_{00}p_0 + h_{10}p_1 + h_{20}p_2$$

$$y_1 = h_{01}p_0 + h_{11}p_1 + h_{21}p_2$$

$$y_2 = h_{02}p_0 + h_{12}p_1 + h_{22}p_2$$

of in matrixnotatie:

$$y = Hp, \text{ waaruit volgt } p = H^{-1}y \text{ als } H \text{ niet singulier is.}$$

Ideale keuringskarakteristiek (matrix H):

W e r k e l i j k h e i d

Klasse	0	1	2
	Afwezig	< drempelwaarde	> drempelwaarde
Gevonden	0	Kans: 1	0
	1	0	1
	2	0	0

Werkelijke keuringskarakteristiek:

W e r k e l i j k h e i d

Klasse	0		1		2		
	Afwezig		< drempelwaarde	> drempelwaarde	> drempelwaarde		
Gevonden	0	h00	0,90	h10	0,2	h20	0,1
	1	h01	0,09	h11	0,6	h21	0,3
	2	h02	0,01	h12	0,2	h22	0,6

Getallenvoorbeeld: laat in werkelijkheid de frequentieverdeling zijn: klasse 0: $p_0 = 0,50$, klasse 1: $p_1 = 0,20$ en klasse 2: $p_2 = 0,30$. Als h_{00} enz. de waarden aannemen, vermeld in het schema, dan wordt de gevonden frequentieverdeling:

	<u>Klasse 0</u>	<u>Klasse 1</u>	<u>Klasse 2</u>
Klasse 0: werkelijke fractie 0,50			
90 % blijft in klasse 0 =	0,45		
9 % gaat naar klasse 1 =		0,045	
1 % gaat naar klasse 2 +			0,005
Klasse 1: werkelijke fractie 0,20:			
20 % gaat naar klasse 0 =	0,04		
60 % blijft in klasse 1 =		0,12	
20 % gaat naar klasse 2 =			0,04
Klasse 2: werkelijke fractie 0,30:			
10 % gaat naar klasse 0 =	0,03		
30 % gaat naar klasse 1 =		0,09	
60 % blijft in klasse 2 =			0,18
Dus gevonden in klasse	0	1	2
een fractie	0,52	0,255	0,225

Deze berekeningen kan men ook omgekeerd uitvoeren om de verdeling in de populatie af te leiden uit de keuringsresultaten, waarbij dan y de bekende keuringsresultaten zijn en p de onbekende frequenties in de populatie;

daarvoor moet men het stelsel vergelijkingen in p oplossen, cq. de matrix H inverteren: $p = H^{-1} y$.

Uit dit eenvoudige model is het duidelijk dat het noodzakelijk is de kansen h_{ij} in H nauwkeurig te kennen: er wordt immers gehandeld alsof H bekend is. Voorts kan men gemakkelijk zien dat $y_0 + y_1 + y_2 = 1$ als geldt dat $p_0 + p_1 + p_2 = 1$ en $h_{i0} + h_{i1} + h_{i2} = 1$ voor elke index i . Het komt er op neer, dat elke fractie in de populatie wordt herverdeeld over de fracties in de waarnemingen. Omgekeerd geldt ook dat $p_0 + p_1 + p_2 = 1$ volgt uit $y_0 + y_1 + y_2 = 1$ en de condities voor H . Door toevalsfluctuatie is het echter mogelijk dat de waarnemingen y afwijken van hun theoretische waarde en dit kan resulteren in schattingen buiten het interval $(0,1)$ voor een of meer p_i . Daarom is de bovengeschetste primitieve aanpak in de praktijk niet bruikbaar. Het probleem is echter in principe oplosbaar met behulp van een meer realistisch model en bijbehorende statistische schattingsprocedures mits men tevens de eigenschappen van de meet- en beslissingsprocedure op bevredigende wijze kan kwantificeren. Hierover gaan de volgende paragrafen.

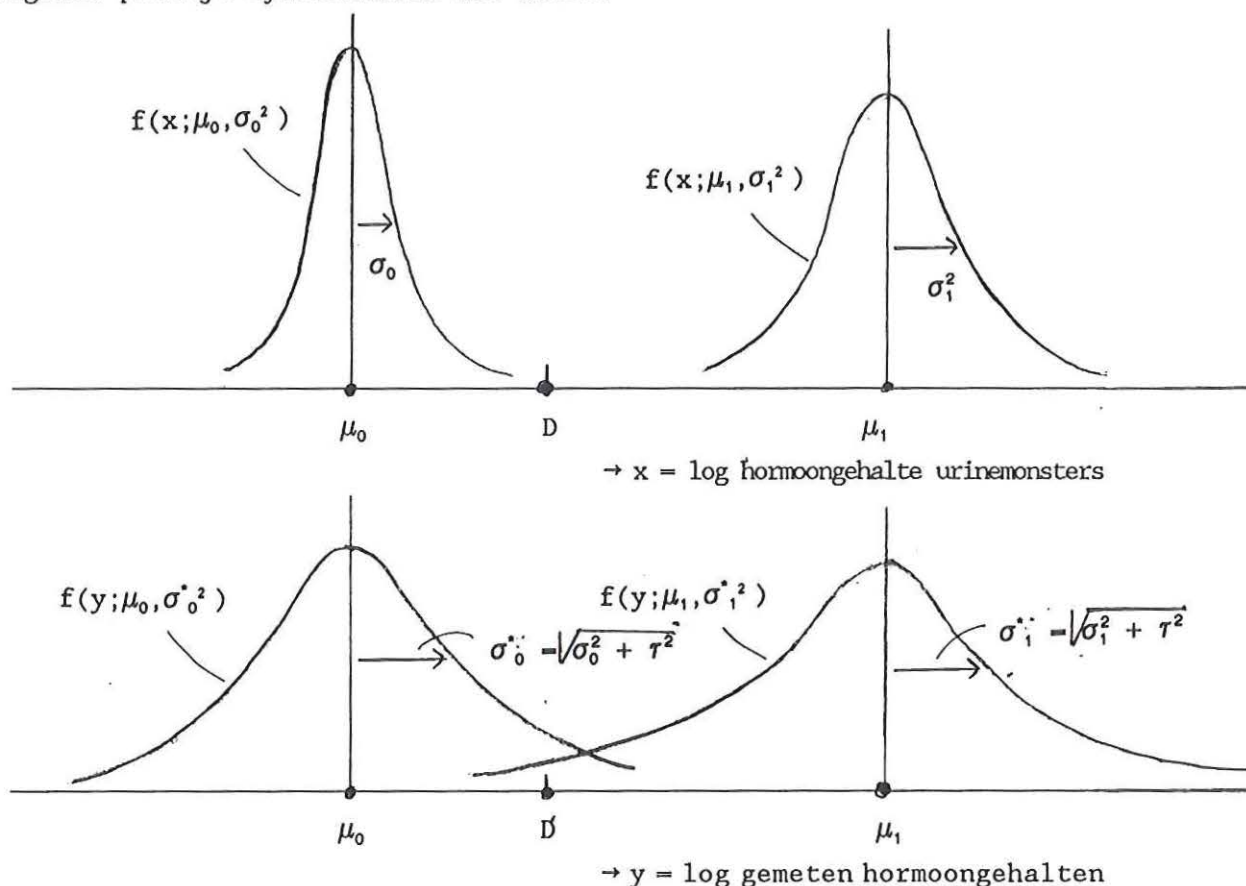
5.3 Een realistisch model

We zullen veronderstellen dat de verdeling van de hormoongehalten in de urinemonsters een mengsel is van 2 verdelingen, namelijk een fractie p_0 urinemonsters met een natuurlijk hormoongehalte en een fractie $1-p_0$ met een verhoogd hormoongehalte; de (logaritmen van de) hormoongehalten x in deze fracties volgen normale verdelingen met gemiddelden μ_0 resp. μ_1 en standaardafwijkingen σ_0 en σ_1 en dichtheidsfuncties $f(x; \mu_0, \sigma_0^2)$ en $f(x; \mu_1, \sigma_1^2)$. Het gaat er hierbij niet zozeer om of de normaliteitsveronderstelling opgaat; in een later stadium kan zonodig ook een ander type verdeling worden genomen.

Voorts wordt verondersteld dat ook de keuringsuitslagen y met een dergelijk type verdeling worden beschreven. De parameters daarvan, die we zullen aangeven met μ_0^* , μ_1^* , σ_0^* en σ_1^* , zijn echter verschillend. De kansverdeling van de keuringsuitslagen wordt bepaald door de eigenschappen van de keuringsprocedure. Als eenvoudig model hiervoor wordt verondersteld dat de meetprocedure de variantie van de log (hormoongehalten) met τ^2 verhoogt, maar de verwachtingswaarden gelijk laat; derhalve

$$\begin{array}{lll} \mu_0 \rightarrow \mu_0^* = \mu_0 & \sigma_0^2 \rightarrow \sigma_0^{*2} = \sigma_0^2 + \tau^2 & f(x; \mu_0, \sigma_0^2) \rightarrow f(y; \mu_0, \sigma_0^{*2}) \\ \mu_1 \rightarrow \mu_1^* = \mu_1 & \sigma_1^2 \rightarrow \sigma_1^{*2} = \sigma_1^2 + \tau^2 & f(x; \mu_1, \sigma_1^2) \rightarrow f(y; \mu_1, \sigma_1^{*2}) \end{array}$$

Als men verder een drempelwaarde D kiest waar beneden het hormoongehalte wordt geacht van natuurlijke oorsprong te zijn is het model compleet. Het volgende plaatje symboliseert het model:



Vermoedelijk moet het model nog worden uitgebreid met een parameter die beschrijft dat een fractie ϕ van de urinemonsters, al dan niet met verhoogd hormoongehalte, niet kan worden gemeten als gevolg van matrix-effecten. Als men beschikt over het model met de parameters p_0 , μ_0 , μ_1 , σ_0 en σ_1 kan men berekenen wat de consequenties zijn van andere instellingen voor D , ϕ en τ . Met behulp hiervan kan men een optimale strategie kiezen met betrekking tot de kansen op ten onrechte goedkeuren en afkeuren en de gevoeligheid van deze kansen voor methodenverbetering aangeven. Het is daarbij uiteraard gewenst te letten op de financiële en/of maatschappelijke risico's van de te kiezen strategie en niet uitsluitend op de kansen van goed- of afkeuren. Ook de kosten van eventuele methodenverbeteringen moeten daartegen worden afgewogen.

In Bijlage 2 wordt een mathematische beschrijving gegeven van het hier besproken model.

5.4 Vaststelling van de eigenschappen van de meet- en keuringsprocedure

In het in de vorige paragraaf beschreven model worden de eigenschappen van de meet- en keuringsprocedure gekarakteriseerd door de bijdrage van de meetfoutvariantie τ^2 aan de variantie van de waarnemingen en door de fractie ϕ van meetuitkomsten die niet aan de voorwaarden voldoen. Deze grootheden kunnen in principe worden vastgesteld door routinematige urine-monsters met bekende hormoongehalten in combinatie met verschillende maten van "vuilheid" te laten meelopen.

Om de gedachten te bepalen kan men denken aan het volgende experiment:

Aan 5 verschillende urinesoorten als achtergrondmatrix wordt een range van 5 concentraties van een hormoon (op logschaal c_1, \dots, c_5) toegevoegd. De 25 monsters worden in een aselechte volgorde onder praktijkomstandigheden gemeten. Dit wordt een aantal keren, bijvoorbeeld 5 keer, herhaald. Bij elke meting komen de gebruikelijke waarnemingen beschikbaar en kan derhalve worden vastgesteld:

- of de resultaten aan de criteria voldoen voor de detectie van het desbetreffende hormoon (retentietijd en verhouding tussen de massafragmenten); hiervoor introduceren we een variabele z : als aan de criteria is voldaan, dan is $z=0$, anders $z=1$,
- wat de logaritme van het gemeten hormoongehalte y is als aan de criteria wordt voldaan, dus als $z=0$.

De relatie tussen de waarden van z , $y-c$, $\log(y-c)^2$ en de ingestelde hormoongehalten en urinesoorten kan worden onderzocht met behulp van generaliseerde lineaire modellen of met variantie-analyse. Hierbij is $y-c$ een maat voor een systematische afwijking en $(y-c)^2$ voor de variabiliteit van de uitkomst. Dit levert informatie over de grootte van ϕ ($\phi = Ez$) en τ en over de stabiliteit van deze grootheden en over eventuele systematische afwijkingen. Systematische afwijkingen kunnen wellicht worden vermeden. Als mocht blijken dat de parameters ϕ en τ in belangrijke mate afhangen van de urinesoorten dient men te inventariseren welke urinesoorten in de praktijkmonsters voorkomen en in welke frequenties, en de parameterwaarden over de urinesoorten "gewogen" te middelen. Een alternatief is de eigenschappen per urinesoort te beschrijven, met als consequentie dat in de routinematige controle voor ieder monster ook de urinesoort dient te worden vastgesteld en bij de keuring te worden betrokken. Als op logschaal wordt gewerkt zijn hopelijk ϕ en τ niet afhankelijk van de ingestelde hormoonconcentraties; indien dit wel zo is wordt het model veel gecompliceerder.

Omdat de waarden van ϕ en τ en de eventuele relatie van deze parameters met urinesoorten en concentraties in het model van de vorige paragraaf bekend worden verondersteld is het gewenst deze zeer nauwkeurig vast te stellen. Dit vraagt vermoedelijk aanmerkelijk meer waarnemingen dan in bovenstaand experiment voorzien. Bovendien is het niet uitgesloten dat de parameters een zeker verloop met de tijd vertonen. Wellicht is het daarom zinvol de informatie over ϕ en τ te vergroten en zonedig bij te stellen door routinematig meenemen van controlemonsters bij de praktijkbepalingen.

In bovenstaande is uitgegaan van controlemonsters met 1 hormoon. De procedure zal echter moeten worden uitgevoerd voor meer hormonen tegelijk. Zolang de meting van afzonderlijke hormonen niet wordt beïnvloed door de simultane aanwezigheid van andere hormonen kunnen zonder bezwaar willekeurige cocktails worden gemaakt van alle hormonen in de urinemonsters. Als echter onderlinge beïnvloeding niet is uitgesloten is het noodzakelijk nader te bezien of en in hoeverre cocktails kunnen worden gebruikt. Steeds zal een zorgvuldige planning noodzakelijk zijn om te garanderen dat alle gewenste combinaties evenwichtig worden verdeeld over de beschikbare tijd en ruimte.

5.5 Multivariatie aspecten

In 5.4 is reeds de mogelijkheid aan de orde gesteld om cocktails van hormonen te gebruiken bij het vaststellen van de eigenschappen van de meet- en keuringsprocedure. Daarbij ontstaan problemen als de metingen aan de verschillende hormonen niet onafhankelijk zijn. Als dit laatste het geval is zijn er ook problemen bij de keuringsprocedure omdat de eigenschappen daarvan dan niet alleen voor de afzonderlijke hormonen maar ook voor alle mogelijke combinaties van hormonen die in de praktijk voorkomen moeten worden vastgesteld. Dit zou tot welhaast onoverkomelijke problemen voeren m.b.t. de hoeveelheid informatie die nodig is om de procedure te beheersen.

Ervan uitgaande dat de metingen van verschillende hormonen bij simultane aanwezigheid onderling onafhankelijk zijn kan men zich de vraag stellen of de simultane aanwezigheid van hormonen zelf aanleiding kan zijn tot bijstelling van de keuringsprocedures. Bijvoorbeeld in de vorm van een drempelwaarde D die een (lineaire of kwadratische) functie is van de concentraties van alle hormonen gezamenlijk. Een dergelijke discriminerende functie kan in sommige gevallen veel effectiever zijn bij het detecteren van afwijkingen van de natuurlijke situatie. Om een dergelijke functie optimaal te kiezen moet men beschikken over de multivariate verdeling van de hormoongehalten in de natuurlijke en de afwijkende monsters. In theorie

kunnen deze worden verkregen met een multivariate formulering van het in 5.3 gestelde model en daarbij aangepaste schattingsprocedure. De uitvoering in de praktijk zal niet zo eenvoudig blijken, mede omdat de gevoeligheid voor de keuze van de vorm van de multivariate verdeling zowel belangrijker als moeilijker is dan in het univariate geval.

6. KWALITEITSBORGING

Bij een werkwijze voor de analyse hoort ook een werkwijze voor de kwaliteitsborging.

Voor het aantonen van hormonen met GCMS is dan ook een uitgebreid borgingsprogramma voorgeschreven. Er worden in routine door het CLRVV per dag per produktielijn ongeveer 12 monsters en 11 controles gemeten. Aangezien het CLRVV over 3 GCMS apparaten beschikt, dus 3 produktielijnen, betekent dit 33 controles op een produktie van 36 monsters.

Echter, ook deze borging is vooral gericht op het voorkomen van fout-positieve resultaten. Bijvoorbeeld: geen besmetting van een monster door het voorgaande (cross-over). De huidige controlemonsters verschaffen dan ook weinig informatie over het optreden van fout-negatieve resultaten (monster bevat hormoon, maar wordt niet als zodanig aangemerkt).

Het is echter, door een andere keuze van controlemonsters, mogelijk met dezelfde inspanning deze informatie wel te verkrijgen.

Het doel van de borging is een evenwicht te verzekeren tussen de gewenste zekerheid en de inspanning die hiervoor nodig is.

7. VOORSTEL ONDERZOEK OM INZICHT IN DE KEURINGSKARAKTERISTIEK TE VERKRIJGEN

Om inzicht te krijgen in de keuringskarakteristiek en daarmee op de kans op fout-negatieve resultaten, zal het nodig zijn een onderzoek uit te voeren voor een aantal verschillende soorten urine, ieder met verschillende, bekende, gehalten. Om een betrouwbaar beeld te krijgen onder routineomstandigheden moet dit onderzoek wordt uitgevoerd op het CLRVV, waarbij de onderzoeksmonsters tussen de normale monsters gespreid worden. Het is de taak van RIKILT en RIVM de monsters te verschaffen.

Onderzoek reeds aanwezig materiaal

Er is reeds veel materiaal voorhanden aan gegevens van uitgevoerde analyses. Hierin kan onderzocht worden wat de spreiding is in relatieve retentietijden en in de verhoudingen van de intensiteiten van de massafragmenten. Een en ander is van belang in verband met de revisie van de EEG Beschikkingen 87/410 en 89/610.

Oriënterend onderzoek

5 soorten urine zonder hormonen, van 'schoon' tot 'vuil'

Analyten: DES DE HEX ED α E2 β E2 E1 EE2 α NT β NT α T β T MT

Ook α TB β TB α Zer β Zer ?

Mengsel van analyten toevoegen aan urines, 5 gehalten

Elk verschillend monster 10 x analyseren, gespreid in de tijd.

Voor elke soort gewenste gehalten bepalen, van 'goed aantoonbaar' tot 'niet of nauwelijks aantoonbaar', monsteronderzoek door RIKILT, RIVM en/of CLRVV, i.s.m. RIKILT Coörd. Chemometrie.

Daarnaast ook standaardenoplossingen met verschillende gehalten.

Definitief onderzoek

In totaal 250 monsters urine, gekozen op grond van oriënterend onderzoek. Monsteronderzoek door RIKILT, RIVM en/of CLRVV, evaluatie door RIKILT Coörd. Chemometrie.

Chemometrische evaluatie

Uit de meetresultaten kunnen de volgende gegevens worden verkregen.

- Spreiding van de criteria-parameters (retentietijd, intensiteit 4 mas safragmenten)

per gehalte en

per urinesoort en standaardenoplossing.

- De goedkeurkansen

bij hantering van de huidige criteria en

bij wijziging van de criteria.

Bij 'wijzigen van de criteria' kan men denken aan een modificatie van het begrip 'intensiteitsverhoudingen moeten kloppen', of ook aan het aantal in de beschouwing te betrekken massafragmenten: in plaats van de huidige 4 een aantal van 3, 2 of 1. In alle gevallen wordt van de beschikbare informatie slechts een (klein) deel gebruikt. Een andere aanpak is, geen informatie weg te gooien, maar *alle* massafragmenten te beschouwen. Uit chemometrisch oogpunt is dit aantrekkelijk; bij gebruik van de huidige MS-apparatuur gaat dit echter ten koste van de gevoeligheid.

Welke criteria ook gekozen worden, er dient rekening gehouden te worden met de nodige moeilijkheden in de dataverwerking in verband met incompatibiliteit van computerprogramma's.

8. OPTIMALISERING INFORMATIEINHOUD

Het in §7 voorgestelde onderzoek dient om inzicht te verkrijgen in de parameters van de keuringskarakteristiek, en op grond daarvan keuzes te

kunnen maken over de gewenste bemonsterings- en keuringsprocedures. Daarnaast is het mogelijk de bij de analyse verkregen informatie beter te benutten.

Er zijn een aantal mogelijkheden.

'Grijs' gebied van verdachte monsters

Door naast de huidige scherpe criteria ook minder scherpe criteria te formuleren verkrijgt men een 'grijs' gebied van monsters die niet aan de scherpe, maar wel aan de minder scherpe criteria voldoen. Deze monsters zou men 'verdacht' kunnen noemen. Ze bieden geen grond tot afkeuring, maar kunnen wel aanleiding geven tot vervolgonderzoek, monsterneming op de boerderij, selectieve bemonstering in geval van dezelfde herkomst, enz.

Criteria optimaliseren

De huidige criteria zijn gedeeltelijk arbitrair opgesteld. Men zou, mede op grond van de inmiddels verkregen rijke ervaring beter onderbouwde criteria kunnen opstellen, o.a. voor de toegelaten spreiding in retentietijd en intensiteitsverhouding van de massafragmenten.

Opsporen hormoonpatronen

In §4 is reeds opgemerkt, dat patronen van hormoongehalten wellicht meer informatie over illegaal gebruik verschaffen dan de gehalten op zich. Uit de schat van informatie die intussen door het CLRVV is verkregen zouden, in samenhang met nader onderzoek, dergelijke patronen zijn op te sporen.

Normale procedure CLRVV

Toegepaste gehalten in standaardoplossing door CLRVV.

DES DE HEX	0,5 x 10 ⁻⁹
E1 αE2 βE2	1
ED αNT βNT αT βT MT EE2	2,5
αZer βZer Zeralenon Zeralanon	4
αTB βTB	5
E2-D3	0,1
PCB	2

In de monsters 'urine + toev.' zijn de gehalten 2x zo hoog.

Procedure

Gedurende de dag: GC-MSD apparatuur schoonmaken en tunen.

Intussen worden de monsters opgewerkt; deze worden vanaf ongeveer 16.00 uur gemeten en zijn de volgende ochtend om ongeveer 8.00 uur gereed. 's Morgens worden de resultaten geëvalueerd.

Normale meetprocedure:

Volgnr.

C1	Standaardenoplossing)retentietijd en piekverhoudingen	
C2	Standaardenoplossing)moeten kloppen	*
C3	Standaardenoplossing)) hieruit	
C4	Blanco urine + toevoeging #) terugvinding	**
C5	Oplosmiddel + PCB	rechte basislijn?	***
		controle op contaminatie	
M1	Monster #		
M2	Monster #		
M3	Monster #		
C6	Blanco urine #	controle op contaminatie	
M4	Monster #		
M5	Monster #		
M6	Monster #		
C7	Standaardenoplossing		
M7	Monster #		
M8	Monster #		
M9	Monster #		
C8	Blanco chemicaliën #	controle op contaminatie	
M10	Monster #		
M11	Monster #		
M12	Monster #		
C9	Blanco chemicaliën + toevoeging #	uitkomst in orde?	
C10	Standaardenoplossing) weer terugvinding	
C11	Blanco urine +toevoeging #) klopt met eerste?*	**

= doorloopt de gehele procedure.

Totaal: 12 monsters (M1 -M12) en 11 controles (C1 - C11) per GC-MSD meetsysteem; er zijn drie van deze systemen in gebruik. Er worden dus in totaal 36 monsters gemeten en 33 controlemetingen uitgevoerd.

Er worden monsters gemeten op dinsdag, woensdag, donderdag en vrijdag.

Kwaliteitscontrole en criteria

* RT standaardopl. klopt onderling vrijwel altijd

Piekverhoudingen in standaardopl. kloppen vrijwel altijd

** Terugvinding toevoegingen in urine t.o.v. standaardopl. klopt vrijwel altijd

*** De rechte basislijn klopt vrijwel altijd

Het criterium voor de retentietijd

òf: retentietijd in monster = retentietijd in standaard-oplossing \pm 5 sec,

òf:

$$\frac{\{RT\ \text{analyt}\}_{\text{monster}}}{\{RT\ \text{int.st}\}} = \frac{\{RT\ \text{analyt}\}_{\text{standaard}}}{\{RT\ \text{int.st}\}} \pm 5/RT\ \text{int. stand. in sec}$$

klopt vrijwel altijd. Is dit niet het geval, dan is dit bij een nieuwe injectie vaak alsnog wel het geval.

Ook het criterium voor de aanwezigheid van de 4 juiste massafragmenten klopt meestal. Soms ontbreken één of twee massafragmenten; dit kan gevangen worden onder het criterium voor de piekverhoudingen.

Wanneer er iets niet klopt, is het vrijwel steeds, dat de intensiteitsverhoudingen voor één piek of voor twee pieken niet kloppen. Dit is de ervaring van zowel CLRVV als van RIKILT en wordt bevestigd door het turven van 113 praktijkmonsters van het CLRVV.

De berekening van de piekverhoudingen vindt in eerste instantie plaats ten opzichte van de voorafgaande standaardoplossing. Wanneer het resultaat niet aan de criteria voldoet kan de berekening ook ten opzichte van één van de andere standaardoplossingen worden uitgevoerd. In de criteria is niet voorgeschreven, welke standaardoplossing moet worden gebruikt.

Mathematische beschrijving van het in 5.3 beschreven model

Voor verdeling van de meetuitkomsten geldt:

$$F(Y) \equiv P(Y \leq y) = p_0 \int_{-\infty}^y f(Y; \mu_0, \sigma_0^2) dY + (1-p_0) \int_{-\infty}^y f(Y; \mu_1, \sigma_1^2) dY$$

De meetuitkomsten $y \leq D$ zullen voor een deel beneden de detectiegrens liggen zodat de praktijkgegevens wellicht met de volgende verdeling moeten worden gekarakteriseerd:

$$p_0^* \equiv F(D) = P(Y \leq D) = p_0 \int_{-\infty}^D f(Y; \mu_0, \sigma_0^2) dY + (1-p_0) \int_{-\infty}^D f(Y; \mu_1, \sigma_1^2) dY$$

$$f_*(Y | Y > D) \equiv f(Y) / (1-p_0^*)$$

Hierin is $f(y)$ de kansdichtheid van y , dus $f(y) = F'(y)$. Zowel p_0^* als f_* zijn afhankelijk van de parameters p_0 , μ_0 , μ_1 , σ_0 en σ_1 ; τ en D worden bekend verondersteld. Voor hormonen die niet natuurlijk voorkomen kunnen wellicht voor μ_0 en σ_0 zodanig kleine waarden worden gekozen dat deze in combinatie met τ de vastgestelde eigenschappen van de meet- en keuringsprocedure van blanco's simuleren. Op basis van de waarnemingen kunnen in principe p_0 , μ_1 en σ_1 worden geschat.

Als het model nog moet worden uitgebreid met een parameter die beschrijft dat een fractie ϕ van de urinemonsters, al dan niet met verhoogd hormoongehalte, niet kan worden gemeten als gevolg van matrix-effecten, dan kan hier als volgt rekening mee worden gehouden.

De fractie ϕ komt in p_0^* terecht zodat het model voor p_0^* wordt:

$$p_0^* = F(D) + \phi[1-F(D)] = \phi + (1-\phi)F(D)$$

Als ϕ niet los van de andere parameters kan worden geschat moet informatie over ϕ worden verkregen uit nader onderzoek van de eigenschappen van de meet- en keuringsprocedure.