

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

Rapport 91.21

Juni 1991

**Chemometrisch onderzoekmodel voor borging van
het keuren op de aanwezigheid van residuen**

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

* Groep Landbouwwiskunde/DLO

Afdeling: Coördinatie Chemometrie

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

INHOUD

ABSTRACT	1
SAMENVATTING	5
1 INLEIDING	7
2 RESULTATEN KEURINGSONDERZOEK	7
2.1 Geval 1: Keuring in de boerderijfase	8
2.2 Geval 2: Keuring in de slachtfase	10
3 EFFECT VAN VERLAGING VAN DE AANTOONBAARHEIDSGRENS VAN DE SCREENINGSMETHODE	10
4 MATHEMATISCHE BESCHRIJVING VAN DE KEURINGSPROCEDURE	12
5 BORGING VAN DE KWALITEIT VAN DE KEURINGSPROCEDURE	15
5.1 Toetsen keuringsresultaten	15
5.2 Permanente borging	19
5.3 Dynamisch model	19
5.4 Verdeling analytgehalten	19
6 SLOTOPMERKINGEN	20

SAMENVATTING

Wettelijke regelingen betreffende samenstelling en veiligheid van voedingsmiddelen vereisen een doelmatige controle. Dit geschiedt door een keuring, die bestaat uit het nemen van geschikte monsters volgens een vastgesteld bemonsteringsschema, gevolgd door analyse van de monsters.

Wanneer de meeste te onderzoeken eenheden aan de gestelde eisen voldoen, kan een tweetraps onderzoek aantrekkelijk zijn. Eerst worden de monsters onderzocht met een goedkope en snelle 'screeningsmethode', om het grote aantal negatieve monsters (= die aan de eisen voldoen) uit te zeven. Het kleine gedeelte dat positief (= voldoet niet aan de eisen) bevonden wordt, wordt verder onderzocht met een 'bevestigingsmethode' voor het verkrijgen van het uiteindelijke keuringsresultaat.

Voor een dergelijke keuring wordt een model gegeven, waarmee de fractie fout-negatieve resultaten in de keuring kan worden berekend. Er wordt aangenomen dat, als aan de bevestigingsmethode geschikte criteria zijn gesteld, de fractie fout-positieve resultaten verwaarloosbaar klein is.

1 INLEIDING

Wettelijke regelingen betreffende samenstelling en veiligheid van voedingsmiddelen vereisen een doelmatige controle. Dit geschiedt door een keuring, die bestaat uit het nemen van geschikte monsters volgens een vastgesteld bemonsteringsschema, gevolgd door analyse van de monsters.

Wanneer de meeste te onderzoeken eenheden aan de gestelde eisen voldoen, kan een tweetraps onderzoek aantrekkelijk zijn. Eerst worden de monsters onderzocht met een goedkope en snelle 'screeningmethode', om het grote aantal negatieve monsters (= die aan de eisen voldoen) uit te zeven. Het kleine gedeelte dat positief (= voldoet niet aan de eisen) bevonden wordt, wordt verder onderzocht met een 'bevestigingsmethode' voor het verkrijgen van het uiteindelijke keuringsresultaat.

Een dergelijk model wordt in dit rapport nader uitgewerkt en gekwantificeerd.

2 RESULTATEN KEURINGSONDERZOEK

Voor het evalueren van keuringsresultaten gaan we uit van het volgende model.

De keuring bestaat uit een eerste onderzoek met een screeningmethode, gevolgd door een bevestigingsonderzoek van de bij de screening positief bevonden monsters. De populatie bestaat uit een fractie p_0 afkomstig van onbehandelde dieren en een fractie $(1-p_0)$, afkomstig van behandelde dieren.

In geval van een niet-natuurlijk analyt is in de fractie van de onbehandelde dieren het analytgehalte 0. Voor de analytgehalten x in de fractie 'behandeld' nemen we een normale verdeling aan met gemiddelde μ_1 en standaardafwijking σ_1 en dichtheidsfunctie $f(x; \mu_1, \sigma_1^2)$.

Voorts wordt verondersteld dat voor deze monsters ook de keuringsuitslagen y , resp. z met de screenings-, resp. bevestigingsmethode met een dergelijk type verdeling kunnen worden beschreven, met parameters μ_1^* , μ_1^{**} , σ_1^* en σ_1^{**} . We nemen aan dat bij de meetprocedure wel een meetonnauwkeurigheid optreedt, maar dat er geen systematische afwijking is. Dus de varianties van de analytgehalten worden verhoogd, neem aan met τ^2 , en de verwachtingswaarden blijven gelijk:

Screening:

$$\mu_1 \rightarrow \mu_1^* = \mu_1 \qquad \sigma_1^2 \rightarrow \sigma_1^{*2} = \sigma_1^2 + \tau^2 \qquad f(x; \mu_1, \sigma_1^2) \rightarrow f(y; \mu_1, \sigma_1^{*2})$$

Bevestiging:

$$\mu_1 \rightarrow \mu_1^{**} = \mu_1 \qquad \sigma_1^2 \rightarrow \sigma_1^{**2} = \sigma_1^2 + \tau^2 \qquad f(x; \mu_1, \sigma_1^2) \rightarrow f(z; \mu_1, \sigma_1^{**2})$$

Voor blanco monsters, dus monsters die de matrix bevatten, maar niet de analyt, is per definitie $\mu_B = 0$ en $\sigma_B = 0$. Doordat de screeningsmethode niet selectief genoeg is, zullen

daar toch een aantal negatieve monsters positief bevonden worden. Wanneer de criteria voldoende scherp gesteld zijn, zullen deze monsters bij de bevestiging door de mand vallen en alsnog als negatief worden aangemerkt.

Zowel bij de screening als bij de bevestiging zullen gehalten beneden een bepaald niveau niet meer aangetoond kunnen worden. In dit model stellen we de grens waaronder aantonen niet mogelijk is op 1 ppb voor beide methoden.

Dit algemere model wordt hierna voor twee concrete gevallen uitgewerkt.

2.1 Geval 1: Keuring in boerderijfase

Voor de keuring in de boerderijfase veronderstellen we de volgende omstandigheden.

De populatie bestaat uit 40000 onderzochte dieren. Neem aan dat 15% behandeld is, dus $p_0 = 0,15$. Van de 40000 monsters zijn er dus 34000 afkomstig van onbehandelde en 6000 van behandelde dieren.

Eerst worden alle monsters onderworpen aan het screeningsonderzoek.

Neem aan dat in de verzameling monsters, afkomstig van de behandelde dieren de gehalten normaal verdeeld zijn, met een gemiddelde gehalte $\mu = 5$ ppb en een spreiding $\sigma = 2,5$ ppb, d.w.z. 95 % van de ware gehalten ligt tussen 0 en 10 ppb.

Stel dat voor de meeton nauwkeurigheid geldt: $\tau = 0,33$ ppb, zowel voor de screeningsmethode als voor de bevestigingsmethode. De aantoonbaarheidsgrens is gesteld op 1 ppb.

In geval van onbehandelde dieren is de analyt dan in alle gevallen afwezig, dus $\mu_0 = 0$ en $\sigma_0 = 0$. Bij de meting zullen door toevallige afwijkingen waarden ongelijk aan 0 gevonden worden. De kans dat waarden boven 1 ppb gevonden worden is 0,0013; voor 34000 negatieve monsters kunnen hierdoor dus 44 fout positieve monsters verwacht worden. Verder kunnen door matrixinvloeden systematisch positieve monsters worden gevonden. Stel, dat we hierdoor in ongeveer 0,5 % van de gevallen een waarde boven de aantoonbaarheidsgrens van 1 ppb zullen vinden. Voor 34000 monsters zijn er dus $44 + 170 = 214$ fout positief en 33786 terecht negatief.

Voor analyse van de monsters van behandelde dieren is $\sigma = \sqrt{(2,5^2 + 0,33^2)} = 2,52$, hetgeen slechts weinig hoger is dan de spreiding in de werkelijke gehalten. De meeton nauwkeurigheid heeft dus maar weinig invloed op het resultaat. De verwachting blijft $\mu = 5$ ppb. De afstand tussen de verwachting en de aantoonbaarheidsgrens is $5 - 1 = 4$ ppb, dus ongeveer $1,6 \sigma$. Van deze verdeling ligt dan ongeveer 5,5 % beneden de aan-

toonbaarheidsgrens van 1 ppb. Van de 6000 monsters zullen bij de analyse dus 330 monsters fout negatief zijn en 5670 terecht positief.

Bij het screenen worden in totaal $33786 + 330 = 34116$ monsters negatief en $214 + 5670 = 5884$ monsters positief bevonden.

Nu volgt het bevestigingsonderzoek, van de (terecht of niet terecht) bij de screening positief bevonden monsters. De terecht en fout negatieve monsters worden niet verder onderzocht.

Er waren 5884 positieve monsters, waarvan 214 fout positief en 5670 terecht positief. Een aanvaardbare schatting is, dat van de 5670 terecht positieven nu ongeveer 1 % fout negatief wordt gevonden, dat is dus ongeveer 60 fout-negatieven. De 214 fout positieve monsters uit het screeningsonderzoek worden nu alle negatief bevonden.

Bij de bevestigingsonderzoek worden dus 5610 monsters positief en $60 + 214 = 274$ monsters negatief bevonden.

Het eindresultaat van de hele keuringsprocedure is, dat er $33786 + 44 + 170 + 60 + 330 = 34390$ monsters = 86 % negatief en 5610 monsters = 14 % positief worden bevonden. Van de negatieve monsters is ongeveer 1 % fout negatief. Er zijn geen fout positieve resultaten.

De bovenstaande redenering is in het volgende schema samengevat.

Werkelijk	G e v o n d e n			
	Screening		Bevestiging	Keuring
34000 -	33786 - 0 0,5% = 170 + 0 _{sys} . 0,0013% = 44 + 0 _{toev} .		170 - 0 44 - 0	33786 - 0 170 - 0 44 - 0
6000 +	5670 + 0 330 - 0		5610 + 0 60 - 0	5610 + 0 60 - 0 330 - 0

Eindresultaat van de keuring:

Aantal uitgevoerde screeningsonderzoeken: 40000
Aantal uitgevoerde bevestigingsonderzoeken: 5884

Positief: 5610 14 %

Terecht + : 5610 14 %
Fout + : 0 0 %

Negatief: 34390 86 %

Terecht - : 34000 85 %
Fout - : 390 1 %

0 = terecht

⊖ = fout

2.2 Geval 2: Keuring in de slachtfase

Bij deze uitwerking gaan we uit van 4000 monsters, waarvan 5 % positief is. Een dergelijke situatie zou zich mogelijk voor kunnen doen bij een keuring in de slachtfase.

Toepassing van dezelfde redenering als boven leidt tot het volgende schema.

Werkelijk	G e v o n d e n		
	Screening	Bevestiging	Keuring
3800 -	0,5 % = 19 syst fout + 0,13 % = 5 toev fout + 3781 terecht -	19 - 0 5 - 0	19 - 0 5 - 0 3781 - 0
200 +	5,5 % = 10 fout - 190 terecht +	1 % = 2 fout - 188 terecht +	10 - 0 2 - 0 188 + 0
Eindresultaat van de keuring:			
Aantal uitgevoerde screeningsonderzoeken: 4000			
Aantal uitgevoerde bevestigingsonderzoeken: 214			
Positief: 188 4,7 %		Terecht + : 188 4,7 % Fout + : 0 0 %	
Negatief: 3812 95,35 %		Terecht - : 3800 95 % Fout - : 12 0,3 %	

3 EFFECT VAN VERLAGING VAN DE AANTOONBAARHEIDSGRENS VAN DE SCREENINGSMETHODE

In de voorgaande berekeningen is uitgegaan van de veronderstelling dat de aantoonbaarheidsgrens = afkeurgrens voor de screeningsmethode en voor de bevestigingsmethode in beide gevallen 1 ppb bedraagt. In werkelijkheid kan voor een ELISA-methode, die voor de screening gebruikt wordt een lagere aantoonbaarheidsgrens en dus ook een lagere afkeurgrens worden bereikt dan voor een GC-MS methode, die voor de bevestiging wordt gebruikt.

Laten we nagaan wat het effect is, als de aantoonbaarheidsgrens voor de screening verlaagd wordt tot 0,1 ppb. In de praktijk zou dit misschien net haalbaar zijn.

Geval 1

De afstand tussen aantoonbaarheidsgrens en meetverwachting is nu verhoogd tot $5 - 0,1 = 4,9$ ppb, ofwel $4,9 / 2,52 = 1,994$ s. Dit betekent, dat het aantal te verwachten fout negatieve resultaten daalt van 5,5 naar 2,3 %; voor 6000 positieve monsters dus tot 138 monsters. Er slippen dus $330 - 138 = 192$ minder positieve monsters door de screeningsprocedure.

Deze winst wordt betaald met een aanzienlijk hoger aantal fout positieve monsters.

Door toevallige afwijkingen wordt nu ($0,1$ ppb = $0,1/0,33 = 0,3$ s, fractie die hierboven ligt is 0,38) 38 % of 12920 van de 34000 negatieve monsters fout positief bevonden.

Ook door matrixeffecten zullen nu meer monsters positief worden geduid. Stel dat dit 2 % is ofwel 680 monsters. In totaal worden dan nu $12920 + 680 = 13600$ monsters fout positief worden. Bij de bevestiging zullen deze monsters alsnog negatief worden, maar ze moeten wel allemaal onderzocht worden. Het komt er op neer, dat bijna de helft van de monsters ook met de bevestigingsmethode onderzocht moeten worden.

Samengevat in het schema:

Werkelijk	----- G e v o n d e n -----		Keuring
	Screening	Bevestiging	
34000 -	20400 - 0 2% = 680 + 0 syst. 38% = 12920 + 0 toev	680 - 0 12920 - 0	20400 - 0 680 - 0 12920 - 0
6000 +	5860 + 0	5800 + 0 60 - 0	5800 + 0 60 - 0
	140 - 0		140 - 0

Eindresultaat van de keuring:

Aantal uitgevoerde screeningsonderzoeken: 40000

Aantal uitgevoerde bevestigingsonderzoeken: 19460

Positief: 5800 14,5 %

Terecht + : 5800 14,5 %
Fout + : 0 0 %

Negatief: 34200 85,5 %

Terecht - : 34000 85 %
Fout - : 200 0,5 %

Door verlaging van de aantoonbaarheidsgrens voor de screeningsmethode daalt het percentage fout-negatieve resultaten van 1 naar 0,5 % en stijgt het aantal terecht positieven van 14 naar 14,5 %. Het aantal monsters dat met de bevestigingsmethode onderzocht moet worden stijgt hiervoor van 5884 naar 19270.

Geval 2

Voor Geval 2 wordt de situatie nu als volgt.

Werkelijk	G e v o n d e n		
	Screening	Bevestiging	Keuring
3800 -	2 % = 76 syst fout + 38% = 1444 toev fout + 2280 terecht -	76 - 0 1444 - 0	76 - 0 1444 - 0 2280 - 0
200 +	2,3 % = 5 fout - 195 terecht +	1 % = 2 fout - 193 terecht +	5 - 0 2 - 0 193 + 0
Eindresultaat van de keuring: Aantal uitgevoerde screeningsonderzoeken: 4000 Aantal uitgevoerde bevestigingsonderzoeken: 1710			
Positief: 193 4,8 %		Terecht + : 193 4,8 % Fout + : 0 0 %	
Negatief: 3812 95,35 %	Terecht - : 3800 95 %	Fout - : 7 0,2 %	

Het percentage fout-negatieve resultaten daalt van 0,3 naar 0,2 %.

Het percentage terecht positieve resultaten stijgt van 4,7 naar 4,8 %.

Het aantal met de bevestigingsmethode uit te voeren analyses stijgt van 214 naar 1710.

4 MATHEMATISCHE BESCHRIJVING VAN DE KEURINGSPROCEDURE

Stel het ware gehalte x van een analyt in de populatie analytbevattende monsters is normaal verdeeld met gemiddelde $\mu=5$ en standaardafwijking $\sigma=2.5$. Dit gehalte wordt gemeten met een meetfout die een standaardafwijking heeft van $\tau=0.33$. Bij een afkeurgrens van $D=1$ zullen bij deze waarde van τ monsters met analytgehalten $x=1$ en $x=2$ afkeurkansen hebben van 50% en 99.87% respectievelijk; dit is in redelijke overeenstemming met de praktijkervaring. De populatie monsters wordt in het hier gehanteerde voorbeeld geacht te bestaan uit 85% analytvrije en 15% analytbevattende monsters.

Bij de screening worden de volgende resultaten verkregen:

- van de analytvrije monsters wordt een fractie $\phi_{0\alpha}$ systematisch positief als gevolg van matrixeffecten.
- van de analytbevattende monsters wordt een fractie ϕ_{α} positief als gevolg van toevallige

afwijkingen; voor $\tau = 0,33$ en $D = 1$ is $\phi_{\alpha} = 0,0013$.

- van de analytbevattende monsters vindt men een fractie ϕ_1 negatief en een fractie $(1-\phi_1)$ positief. Hier is ϕ_1 de kans op een waarneming kleiner dan $D=1$ bij de verdeling van de meetuitkomsten met $\mu=5$ en $\sigma_s = \sqrt{(\sigma^2 + \tau^2)} = 2,522$. Uit een tabel van de normale verdeling leest men af $\phi_1 = 0,05637$. Van alle monsters is dus $\phi_1 * 15\%$ fout negatief en $(1-\phi_1) * 15\%$ terecht positief.

Bij het bevestigingsonderzoek (GC-MS) dat op de positieve monsters uit het screeningsonderzoek wordt toegepast en waarvoor dezelfde meetprecisie verondersteld wordt, vindt men de volgende resultaten:

- de analytvrije monsters, die door de screeningsmethode ten onrechte positief worden bevonden, worden met zeer grote zekerheid negatief als gevolg van de strenge selectiviteitscriteria die worden aangelegd bij de bevestigingsmethode.
- van de fractie terecht positieven $(1-\phi_1) * 15\%$ wordt een fractie ϕ_2 negatief bij de bevestiging. Dit laat in de totale populatie een extra bijdrage ter grootte van $\phi_2 * (1-\phi_1) * 15\%$ ten onrechte negatief.

De grootte van ϕ_2 kan worden berekend als volgt:

Stel y is het resultaat van de screening met $\mu=5$ en $\sigma_s = 2,522$ en z is het resultaat van de bevestiging met $\mu=5$ en $\sigma_s = 2,522$. Als y en z beide voor alle werkelijk positieve monsters worden bepaald zijn y en z biviaat normaal verdeeld met bovenstaande gemiddelden en standaardafwijking en met een correlatie-coëfficiënt die gelijk is aan $\rho = \sigma^2 / \sigma_s^2 = 0,9825$. Voor deze verdeling moet men berekenen de kans dat $z \leq 1$, gegeven dat $y > 1$, dus

$$\begin{aligned} \phi_2 &= P(z \leq 1 \mid y > 1) \\ &= P(z \leq 1 \text{ en } y > 1) / P(y > 1) \\ &= [P(z \leq 1) - P(z \leq 1 \text{ en } y \leq 1)] / [1 - P(y \leq 1)] = \\ &= [P(\chi \leq -1,586) - P(\chi_1 \leq -1,586 \text{ en } \chi_2 \leq -1,586)] / [1 - P(\chi \leq -1,586)] \end{aligned}$$

waarin χ_1 en χ_2 gecorreleerde standaard normale variabelen zijn.

De kans in de teller vindt men met behulp van een daarvoor geschikt rekenprogramma en blijkt te zijn 0,00845; de kans in de noemer is gelijk aan de eerder gevonden $(1-\phi_1) = 0,94363$, zodat

$$\phi_2 = 0,00845 / 0,94363 = 0,0090.$$

Het eindresultaat van de keuring is derhalve:

* terecht negatief		85%
* terecht positief $(1-\phi_2)(1-\phi_1)*15\%$	(= B)	14%
* fout-negatief $[1-(1-\phi_2)(1-\phi_1)]*15\%$	(= A)	1%
* fout-positief		0%

5:

Geschikte criteria om de eigenschappen van de keuringsprocedure vast te leggen zijn de fractie fout-negatieve keuringsuitslagen f van het totaal aantal positieve monsters C , en de fractie monsters F uit het totale aantal monsters N , dat bevestigingsonderzoek vergt.

Met f kan men ondermeer beschrijven hoe de kwaliteit van de keuring afhangt van de meetonnauwkeurigheid τ . Met F hangen de totale kosten van de keuring samen; immers het aantal monsters in het ódureò bevestigingsonderzoek is gelijk aan $F \times N$. Voor f en F gelden de volgende uitdrukkingen

De normaliteitsveronderstelling is niet essentieel maar voor de uitvoerbaarheid van dergelijke globale berekeningen wel handig; uiteraard kunnen de resultaten gevoelig zijn voor afwijkingen van deze veronderstelling in de werkelijke situatie. In principe kunnen de berekeningen ook worden uitgevoerd voor andere veronderstellingen t.a.v. de verdeling in de populatie monsters en de verdeling van de meetfouten van de screening en van de bevestiging; wellicht is het realistischer uit te gaan van bijvoorbeeld een lognormale verdeling van y en z omdat deze alleen positieve uitkomsten toelaat. Hiernaar is geen verder onderzoek verricht.

Het aantal monsters dat in het bevestigingsonderzoek moet worden onderzocht bedraagt

$$\{ 0,85 (\phi_{0a} + \phi_{0c}) + 0,15 (1 - \phi_1) \} N,$$

waarin N = de totale populatie.

Een criterium voor de kwaliteit van de keuring is, dat het aantal fout-negatieve resultaten ten opzichte van het aantal positieve monsters in de steekproef niet te groot mag worden als τ groter wordt, dus

$$f = A / C \text{ mag niet te groot worden als } \tau \text{ groter wordt,}$$

waarin

A = de fractie fout-positieve resultaten in de keuring

C = de fractie totaal aantal positieve monsters in de steekproef.

5 BORGING VAN DE KWALITEIT VAN DE KEURINGSPROCEDURE

5.1 Toetsen keuringsresultaten

De eigenschappen van een keuringsprocedure worden volledig bepaald door:

- de afgesproken keuringsregels, waaronder de wijze van steekproeftrekking, de te hanteren meetvoorschriften en de grenzen van goed- of afkeuren,
- de meetfouten die optreden bij het vaststellen van de grootheden waarop wordt gekeurd.

Nadat de procedure eenmaal is vastgesteld moet de naleving ervan worden verzekerd door goede afspraken te maken over de wijze waarop dit wordt verantwoord en kan worden gecontroleerd - afspraken over 'good laboratory practice' dus. Gegeven dat dit goed is geregeld is alleen de grootte en de verdeling van de meetfouten - die bij een keuring op analytgehalten altijd een rol zullen spelen - bepalend voor de kwaliteit van de keuring. De borging daarvan zal zich derhalve moeten richten op niveau- en reproduceerbaarheidscontrole van de meetresultaten i.c. dus op controle van de laboratoriumcalibratie. Deze kan in principe worden uitgevoerd door regelmatig monsters met bekende of onbekende gehalten te laten meten door zowel de keuringsinstantie als de borgingsinstantie. Deze metingen dienen in de normale productieroutine te worden opgenomen, bij voorkeur zodanig, dat duplicaten van deze monsters op verschillende tijdstippen worden geanalyseerd. Uiteraard dienen deze monsters bij de analyse-uitvoering niet als controlemonsters of als duplicaten herkenbaar te zijn. De keuze van de monsters dient het relevante meettraject te bestrijken. Gegeven de selectiviteit van de GC-MS bevestiging heeft het geen enkele zin bij deze controle ook negatieve monsters mee te nemen; deze leveren bijna per definitie steeds negatieve resultaten. Men denke derhalve aan bijvoorbeeld een aantal monsters per combinatie van:

- 2 gehaltetrajecten.
- 2 matrixsoorten.

In geval er weinig verschil in de matrixinvloeden is te verwachten, kan men zich beperken tot alleen verschillen in gehalten. Voor kalverurine zou dit het geval kunnen zijn.

Van elk monster worden 4 gelijkwaardige duplicaten gemaakt die als afzonderlijke monsters worden toegezonden aan beide laboratoria op 2 willekeurig gekozen tijdstippen.

Of een dergelijke werkwijze zou kunnen worden gerealiseerd en hoe deze moet worden uitgevoerd zal ongetwijfeld nader moeten worden bezien; zonodig dienen ook alternatieven te worden bedacht en overwogen, bijvoorbeeld minder monsters met meer duplicaten per monster. Het is echter noodzakelijk de controle zo in te richten dat zowel een niveau-controle als een vaststelling van de reproduceerbaarheid mogelijk is. De analyse van de resultaten van een aantal monsters krijgt onder het voorgestelde schema de volgende vorm (bijvoorbeeld; een en ander hangt af van de precieze uitvoering!).

Som van duplicaten	g.v.v.	Verwachtingswaarde gem. kwadraat
Niveau	1	$\tau_1^2 + \tau_2^2 + 4\sigma^2$ + niveau bijdragen
Monsters A gehalte	1	$\tau_1^2 + \tau_2^2 + 4\sigma^2$ + bijdrage A-effect
B matrix	1	$\tau_1^2 + \tau_2^2 + 4\sigma^2$ + bijdrage B-effect
A × B	1	$\tau_1^2 + \tau_2^2 + 4\sigma^2$ + bijdrage A × B
Rest	n-4	$\tau_1^2 + \tau_2^2 + 4\sigma^2$
Labs	1	$\tau_1^2 + \tau_2^2$ + niveauverschil labs
Labs × A	1	$\tau_1^2 + \tau_2^2$ + labverschil A-effect
Labs × B	1	$\tau_1^2 + \tau_2^2$ + labverschil B-effect
Labs × A × B	1	$\tau_1^2 + \tau_2^2$ + labverschil A × B
Labs × rest	n-4	$\tau_1^2 + \tau_2^2$
Totaal	2n	

Verschil van duplicaten	g.v.v.	Verwachtingswaarde gem. kwadraat
Lab 1	n	$2\tau_1^2$
Lab 2	n	$2\tau_2^2$
Totaal	2n	

Uit de analyse van de sommen van de duplicaten kan men informatie verkrijgen over de niveauverschillen tussen de labs en of deze afhangen van gehalte en/of matrix; de

verschillen tussen de monsters zelf, inclusief σ^2 , zijn kunstmatig. Uit de analyse van de verschillen tussen de duplicaten kan men informatie halen over de reproduceerbaarheid van de labs (dit betreft een nader te definiëren intermediaire reproduceerbaarheidsmaat) en de eventuele afhankelijkheid daarvan van gehalte en/of matrix. In dit laatste geval moeten de varianties van de duplicaatverschillen voor elke combinatie matrix \times gehalte worden berekend.

Bepalend voor de omvang van dit onderzoek wordt geacht het aantal monsters dat nodig is om de reproduceerbaarheid en de afhankelijkheid daarvan van gehalte en/of matrix voldoende nauwkeurig vast te stellen. Wat voldoende nauwkeurig is kan in principe worden afgeleid van een eis te stellen aan de fractie f fout-negatieven onder de analytbevattende monsters. In het voorbeeld is deze gelijk aan $f=1/15=0.067$. Als men bijvoorbeeld stelt dat deze niet groter dan $(1+\epsilon)*f$ mag worden als gevolg van grotere meetfouten dan overeengekomen (met ϵ een klein getal, bijvoorbeeld $\epsilon=0.1$) dan volgen hieruit eisen voor de nauwkeurigheid waarmee de reproduceerbaarheid moet worden vastgesteld.

Met behulp van de volgende functie in τ en f is het mogelijk de gevoeligheid van f voor τ vast te stellen:

$$\begin{aligned} f &= \{1 - [1-\phi_1(\tau)][1-\phi_2(\tau)]\} \\ &= \phi_1(\tau) + \phi_2(\tau) - \phi_1(\tau)*\phi_2(\tau) \end{aligned}$$

Deze functie kan men voor een aantal waarden van τ berekenen. De berekening verloopt als volgt (voor de berekening van ϕ_1 en ϕ_2 wordt het programma BINORM.EXE misbruikt resp. gebruikt):

$$\begin{aligned} \sigma_s^2 &= \sigma^2 + \tau^2 \\ \rho &= \sigma^2/\sigma_s^2 \\ h,k &= (D-\mu)/\sigma_s \\ \phi_1 &= B(h, k, \rho=1) \\ \phi_2 &= [B(h, k, 1) - B(h, k, \rho)]/[1 - B(h, k, 1)] \\ f &= \phi_1 + \phi_2 - \phi_1\phi_2 \end{aligned}$$

De output van BINORM.EXE is de kans $B(h, k, \rho)$ op een uitkomst in het kwadrant $x_1 \leq h$, $x_2 \leq k$ bij een bivariate standaard normale verdeling met correlatiecoëfficiënt ρ .

Enkele resultaten zijn:

$\mu = 5 \quad \sigma = 2.5 \quad D = 1$

τ	σ	ρ	h	ϕ_1	ϕ_2	f
0	2.5	1	-1.6	0.05480	0	0.055
0.17	2.506	0.995	-1.596	0.05524	0.00471	0.060
0.33	2.522	0.983	-1.586	0.05637	0.00883	0.066
0.50	2.550	0.962	-1.569	0.05832	0.01354	0.073
0.67	2.588	0.933	-1.546	0.06105	0.01863	0.081

Bij dit voorbeeld blijkt in het beschouwde traject f met ongeveer 10% toe te nemen bij een verhoging van τ met 0.17.

Een redelijke eis aan de omvang van het borgingsonderzoek is dat per lab, per combinatie van gehalte en matrix, m duplicaatverschillen beschikbaar zijn, waarbij m volgt uit de eis:

Als voor het keuringslab $\tau^2 \geq 0.25$ moet de kans om een significant verschil tussen keuringslab en borgingsinstantie (met $\tau^2 = 0.11$) te krijgen gelijk zijn aan tenminste 90% bij een onbetrouwbaarheid van 5%. Dus er moet voldaan worden aan:

$$\text{onder } H_0: P((\hat{\tau}_2^2)/(\hat{\tau}_1^2) > F_m^m) \leq 0.05$$

$$\text{onder } H_1: P((\hat{\tau}_2^2)/(\hat{\tau}_1^2) > F_m^m) = P((\hat{\tau}_2^2/0.25)/(\hat{\tau}_1^2/0.11) > (0.11/0.25) F_m^m) \geq 0.90$$

Door proberen vindt men m.b.v. F-tabellen $m \approx 60$. Dit getal moet niet worden opgevat als een absolute eis, maar als een orde van grootte waarmee het naar verwachting mogelijk zal zijn relevante verschillen te detecteren.

In het voorgaande is verondersteld dat de niveaucontrole geen extra eisen stelt aan het borgingsonderzoek. Dit is mede ingegeven door de interne niveaucontrole die bij de meetprocedure is inbegrepen.

5.2 Permanente borging

Bovenstaande berekeningen hebben betrekking op een eenmalige controle en zijn bedoeld om een steekproefomvang vast te stellen waarmee de keuringsresultaten kunnen worden getoetst aan een norm. Het is ook mogelijk een permanente borging te realiseren door routinematig controlemonsters te laten analyseren en telkens nadat een voldoende grote verzameling resultaten is verkregen een analyse uit te voeren en de resultaten te monitoren in een controlediagram. De definitie van de reproduceerbaarheidsmaat beïnvloedt de controleintensiteit: als men hiervoor bijvoorbeeld kiest de reproduceerbaarheid van een meting verkregen door een laboratorium op een aselekt gekozen analysetijdstip binnen een jaar, dan zal men de genoemde steekproefgrootte binnen een jaar moeten realiseren. Het lijkt aantrekkelijk om met een voortschrijdende werkwijze te werken waarbij op elk analysemoment zowel de resultaten over het afgelopen jaar als de resultaten over de periode sinds het laatste analysemoment worden gezien; de laatste uiteraard met minder onderscheidingsvermogen, maar met de mogelijkheid sneller indicaties te krijgen over instabiliteiten in de procedure. Welke analysemomenten men moet kiezen hangt mede af van de mate waarin de keuring onder controle blijkt te zijn. Men kan in principe naar behoefte de frequentie daarvan laten toe- of afnemen in afhankelijkheid van de gerealiseerde resultaten.

5.3 Dynamisch model

Het is mogelijk deze gehele procedure gelijktijdig voor een mengsel van analyten uit te voeren, met een dynamische keuze van de samenstelling van het mengsel die afhankelijk is van de actualiteit van analyten en/of van verschillen in controleintensiteit.

5.4 Verdeling analytgehalten

De controlemonsters geven geen goed beeld van de in de populatie voorkomende verdeling van analytgehalten omdat deze immers gericht zijn gekozen. Als het wenselijk is deze verdeling te schatten moet een daarop gerichte steekproefprocedure worden gekozen. Het is dan waarschijnlijk beter zich te baseren op de gerealiseerde resultaten van de routinematige keuring.

6 SLOTOPMERKINGEN

De verdelingsveronderstellingen in 1 en 3 zijn niet realistisch en moeten in het kader van een model dat voor de gehele meetprocedure moet worden opgesteld nader worden gezien. Hetzelfde geldt voor het begrip meetfout. In het model is geen rekening gehouden met de criteria die ten behoeve van de selectiviteitseisen worden aangelegd.

Het verdient daarom aanbeveling met behulp van monsters met bekende gehalten en aan de hand van historische gegevens hierover nader onderzoek te verrichten.