

**Normen en richtlijnen voor
methodvalidatie in chemische laboratoria**

Wetenschappelijk ondersteuningsprogramma
voor de normalisatie

deel II

Eindverslag

Federale Diensten voor
WETENSCHAPPELIJKE, TECHNISCHE
EN CULTURELE AANGELEGENHEDEN

**Wetenschappelijk
ondersteuningsprogramma voor de
normalisatie**

**Onderzoeksovereenkomst Nr. NO/03/003
Normen en richtlijnen voor
methodvalidatie in chemische
laboratoria**

Eindverslag

Deelnemende laboratoria :

- **VUB : prof. D.L. Massart**
- **ULg : prof. J. Crommen**
- **KUL : prof. J. Hoogmartens en prof. E. Roets**

Wetenschappelijk ondersteuningsprogramma voor de normalisatie

Onderzoeksovereenkomst Nr. NO/03/003

Normen en richtlijnen voor methodenvalidatie in chemische laboratoria

Eindverslag

1. Inleiding

Overeenstemming tussen meetresultaten en hun onderlinge erkenning voor het nagaan van het naleven van de productspecificaties is een probleem dat belangrijk is voor iedereen die betrokken is in een consument-producent relatie. Dit is zowel een probleem voor de industrie als voor nationale en internationale beleidsorganen. Een dergelijke overeenkomst vereist dat de meetmethoden gevalideerd zijn, dit wil zeggen dat aangetoond moet worden dat hun precisie en andere performantiekarakteristieken voldoen aan de eisen van alle belanghebbende partijen. Er is een nood aan afspraken over de wijze waarop deze validatie moet uitgevoerd worden, of met andere woorden : men heeft standaarden en richtlijnen nodig voor de validatie van meetmethoden.

De technische en wetenschappelijke doelstelling van het project was om gestandaardiseerde benaderingen en richtlijnen te ontwikkelen voor de validatie van meetmethoden in chemische, farmaceutische, agro-voeding en aanverwante industrieën. Het is duidelijk dat standaardisatie niet enkel nodig is op nationaal vlak maar nog meer op internationaal vlak en dat de doelstelling daarom het definiëren van standaarden voor Europees gebruik of zelfs voor een wereldwijd gebruik is. Op organisatorisch vlak bracht dit project de bekwaamheid van een aantal Belgische onderzoekers en een aantal industrieën op het gebied van methodenvalidatie samen, teneinde de Belgische wetenschappers een prominente rol te laten spelen in de ontwikkeling van normen en richtlijnen.

Methodenvalidatie is dus een belangrijk regulatory probleem voor vele industrieën en is nodig om de validiteit aan te tonen van de methoden die aangewend worden om product specificaties na te gaan. De voorgestelde normen voor methodenvalidatie moeten kost-efficiënt zijn zeker wat de industrie betreft.

Bovendien is methodenvalidatie nodig voor de bescherming van de veiligheid van de consument en om de kwaliteit van consumptieproducten te garanderen. De Farmacopee, bijvoorbeeld, beschrijft standaardmethoden om de kwaliteits- en veiligheidskarakteristieken van geneesmiddelen te controleren. Er is echter nood aan een betere validatie van deze analysemethoden. Daarom werden in het kader van dit project een aantal case studies uitgevoerd die zich concentreerden op de veiligheid van geneesmiddelen. Hierbij handelt het

enerzijds over methodenvalidatie van farmakopee-analyses op bulkpreparaten en anderzijds van bioanalysemethodes van geneesmiddelen voor farmacokinetsiche doeleinden.

De ontwikkeling van snelle, multivariate technieken, zoals o.a. de nabij-infrarood analyse, voor het meten van de kwaliteit tijdens of in onmiddellijke relatie met een proces, is een belangrijke trend in de industrie. Pre-normatief onderzoek om guidelines te ontwikkelen voor de procesanalyse was gepland en werd uiteindelijk uitgevoerd behandeld.

Het hier uitgevoerde project sloot nauw aan bij een internationaal project genaamd "*Robust process analytical methods for industrial practice and their validation*" uit het Standards, Measurements and Testing programma van de Europese Unie. Dit project had als partners de Universiteiten van Brussel, Tarragona (Spanje) en Genua (Italië) en de volgende bedrijven Elf (Frankrijk), Rhone Poulenc (Frankrijk), Pfizer (Verenigd Koninkrijk), Novartis (Zwitserland) en Unilever (Nederland). Het was een sterk industrie-gericht project met de volgende doelstellingen :

- a) het voorstellen van werkmiddelen voor de methodenvalidatie in een industriële context, welke dan uitgetest werden door bovenvermelde belangrijke industriële partners, zodat bijgedragen wordt tot de Europese standardisatie in dit gebied,

- b) het vooruit helpen van methodenvalidatie in de procesanalyse.

De twee projecten zijn complementair aan elkaar, zowel wat de wetenschappelijke inhoud als de samenstelling van de deelnemers betreft. Terwijl het SMT project slechts de Vrije Universiteit Brussel als partner heeft, liet dit project toe een samenwerking op te zetten op nationaal niveau tussen de industrie, wetenschappelijke instituten (als leden van het begeleidingscomite) en de universiteiten van Leuven, Luik en Brussel. Wat de wetenschappelijke inhoud betreft legt het SMT project meer de nadruk op het ontwikkelen robuuste methodes, terwijl in dit project het accent lag op de kosten en de efficiëntie in het kader van methodenvalidatie. Alhoewel de projecten complementair waren aan elkaar werd gehoopt dat er tussen beide synergieën ontwikkeld konden worden, wat in de loop van beide projecten ook het geval gebleken is.

2. Methodologie

Binnen het project werden verschillende taken uitgevoerd en verschillende aspecten van de methodenvalidatie bestudeerd. De verschillende onderdelen waarvoor getracht werd richtlijnen te formuleren of bestaande richtlijnen te evalueren, zijn :

- a) de evaluatie van de robuustheid van analytische methodes,
- b) de validatie van chromatografische bioanalysemethodes, d.w.z. methodes die bepalingen uitvoeren in biologische milieus,
- c) bepaling van de detectielimiet,

- d) de vergelijking van alternatieve meetmethodes,
- e) methodenvalidatiedesigns die kostenefficiënt zijn.
- f) bepaling van de reproduceerbaarheid van een methode,
- g) robuuste procesanalyse met multivariate calibratie,

Voor de verschillende bovenvermelde onderdelen was de aanpak meestal gelijklopend. In eerste instantie werd de reeds bestaande situatie, d.w.z. de uit de literatuur voorhanden zijnde gegevens, dieper onderzocht. In een tweede stap werden enkele specifieke applicatiegebieden geselecteerd en werden hierin een aantal case studies uitgevoerd. Vervolgens werd tenslotte geprobeerd om een algemene benadering en gefinaliseerde standaardprocedures te definiëren.

2.1. Bestudering van de bestaande situatie

Verschillende standaardisatie-organisaties hebben richtlijnen voorgesteld voor de terminologie die gebruikt moet worden om de validatie-parameters (zoals repeteerbaarheid en bias) van een analysemethode te beschrijven. De belangrijkste wereldorganisaties die hierbij betrokken zijn, zijn o.a. ISO (International Organisation for Standardization), IUPAC (International Union of Pure and Applied Chemistry) en AOAC (Association of Official Analytical Chemists). Naast deze wereldwijde organisaties zijn er nog vele actoren op Europees (zoals Eurachem en CEN), nationaal (zoals de Duitse DIN, de Franse AFNOR en de Nederlandse Keuringsdienst voor Waren) en regionaal niveau. Er zijn vroeger enkele pogingen tot harmonisatie geweest die geleid hebben tot een paar geharmoniseerde protocols tussen ISO, AOAC en IUPAC. In sommige gevallen hebben bepaalde industriële sectoren, zoals die van de farmaceutische analyse, via het ICH comité (International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use) een overeenkomst bereikt over de terminologie. In dit specifieke geval werd er ook een overeenkomst bereikt over validatie-parameters die geanalyseerd moeten worden naargelang de aard van de toepassing. Men was daarbij echter nog ver van het bereiken van een overeenkomst over hoe deze parameters gemeten en statistisch behandeld moeten worden. In sommige sectoren heeft men getracht een stap verder te gaan. Op het gebied van de bioanalyse van geneesmiddelen voor farmacokinetische studies, specificeert de zogenaamde "Washington consensus guideline" hoe goed de repeteerbaarheid moet zijn en hoeveel replicaten er minimaal nodig zijn om dit te bepalen.

Voor elk van de onderdelen en aspecten van de methodenvalidatie waarvoor criteria en normen opgesteld moesten worden, werd in eerste instantie het uit de literatuur beschikbare materiaal verzameld en bestudeerd.

Systematische, algemene en overeengekomen benaderingen die preciese experimentele richtlijnen bevatten en advies geven over hoe de validatie-parameters statistisch berekend

moeten worden, waren oorspronkelijk nog niet beschikbaar en werden gedurende dit project gedefinieerd.

2.2 Specifieke applicatiegebieden

Traditionele methodenvalidatie

Twee van de specifiek te bestuderen applicatiegebieden werden gekozen omdat het proces voor de ontwikkeling van richtlijnen op Europees- en Wereldniveau bezig is. Deze applicatiegebieden werden bestudeerd teneinde ook Belgische wetenschappers de gelegenheid te geven een prominente rol op dit terrein te spelen. De twee applicatiegebieden zijn (i) richtlijnen voor methodenvalidatie in farmacopee-analyses en (ii) de bioanalyse van geneesmiddelen voor farmacokinetische doeleinden.

Het eerste wordt bestudeerd door de Belgische Pharmacopee-commissie op nationaal niveau. Validatie of voortgezette validatie van bestaande officiële methoden voor de analyse van geneesmiddelen, gepubliceerd in de farmacopees, wordt meer en meer belangrijk. De ontwikkeling van voldoende gevalideerde nieuwe of verbeterde methoden voor de analyse van geneesmiddelen is eveneens noodzakelijk. Ook binnen dit project werden in dit verband een aantal case studies uitgevoerd.

Bioanalyse van geneesmiddelen werd, voor de aanvang van dit project, bestudeerd door een informeel comité met leden van o.a. de FDA, de Canadese tegenhanger ervan en leden van de industrie. Dit resulteerde in de zogenaamde 'Washington consensus guidelines'. De VUB (dienst prof. Massart) publiceerde, eveneens voor de aanvang van dit project, een kritiek op deze richtlijn en was eveneens aanwezig op de follow-up conferentie in Munchen. Gedurende dit project werden richtlijnen opgesteld voor de validatie van chromatografische methodes voor de dosage van geneesmiddelen in biologische milieus. Dit gebeurde in samenwerking met een commissie van de Franse SFSTP (Société Française de Sciences et Techniques Pharmaceutiques).

Vooraleer algemene richtlijnen voor methodenvalidatie gedefinieerd kunnen worden, moesten eerst een aantal specifieke problemen in detail bestudeerd worden. Het betreft hier validatie-criteria die gekend waren als knelpunten in de ontwikkeling van een methodenvalidatiestrategie, zoals de bepaling van de robuustheid en de detectielimiet van een methode.

Het eerste criterium dat aan een grondige studie onderworpen werd, is de bepaling van de robuustheid van een methode. Volgens de meest gangbare definitie bestudeert een robuustheidstest het effect van kleine veranderingen in een analysemethode op het gevonden resultaat. De bedoeling van dergelijke test is om de lange termijn precisie of de reproduceerbaarheid van een methode te voorspellen en zonodig te verbeteren teneinde uitwisselbaarheid van methodes tussen laboratoria of instrumenten te verhogen. Binnen dit

project werden een aantal case studies uitgevoerd, werd een strategie vooropgesteld welke een statistische interpretatie van de resultaten toelaat evenals een strategie voor minimale evaluatie van de robuustheid, werd een softwareprogramma geschreven en werden richtlijnen opgesteld voor het opstellen en interpreteren van een robuustheidstest.

Het tweede validatiecriterium dat bestudeerd werd, is de detectielimiet. Twee belangrijke benaderingen zijn mogelijk : (i) de klassieke, gebaseerd op de meting van blancostalen, en (ii) deze die gebaseerd is op informatie van de calibratielij. De tweede is statistisch nogal complex, maar is diegene die IUPAC voorstelt in een (concept)richtlijn. Praktische ervaring ermee was echter beperkt en het was noodzakelijk ze in verschillende situaties uit te testen voordat ze in praktijk gebracht kan worden.

Een situatie die zich bij het ontwikkelen en valideren van een methode frequent voordoet is dat men de resultaten van deze methode wil vergelijken met deze van een bestaande referentiemethode. Richtlijnen werden voorgesteld voor de vergelijking van twee analysemethoden, nl. voor de bias, de verschillende precisiecomponenten en de detectielimieten.

Proces analyse

Proces analytische methoden brengen de meting dicht bij het proces of binnen het proces. De gebruikelijke benadering van het ontwikkelen van selectieve analysemethoden om de aanwezige substanties of de karakteristieken van een produkt te meten, levert dan vaak methoden op die te langzaam zijn. De chemisch selectieve methodes moeten dan vervangen worden door een aantal nieuwe methoden, die gebruik maken van minder selectieve informatie (een spectrum, een chromatogram waarin niet alle pieken zijn gescheiden of geïdentificeerd, een set van resultaten van sensoren) en essentieel een multivariaat karakter hebben. Deze methoden zijn minder tijdsrovend en laten toe om de meting dicht bij het proces te brengen. Het economisch voordeel is evident.

Er zijn twee typische problemen die gebruik kunnen maken van dergelijke methoden. In het eerste, dat mengselanalyse wordt genoemd, is men meer geïnteresseerd in de kennis van het aantal componenten dat aanwezig is en met welke concentratieprofielen. In het tweede, multivariate calibratie genaamd, heeft men referentiestalen voor de te analyseren bestanddelen, en is men geïnteresseerd in hun concentraties.

In de twee gevallen worden de metingen direct op een vloeistofstroom uitgevoerd of op het produkt als dusdanig met een minimale voorbehandeling. Dergelijke methoden worden het meest gebruikt in de agro-voedingsindustrie, maar toepassingen in de farmaceutische industrie (bijvoorbeeld, analyse van tabletten zonder deze op te lossen) worden frekwenter en in de petrochemische industrie meet men op deze manier bijvoorbeeld het octaangetal.

Om informatie te extraheren uit de multivariate gegevens vertrouwt men op mathematische (chemometrische) methoden zoals evolving factor analyse, partial least squares en vele andere. De introductie van deze methoden wordt echter bemoeilijkt omdat er problemen zijn met de acceptatie ervan als standaardmethode.

Dit komt omdat men pas recent begonnen is met het onderzoeken van de beperkingen van deze methodes. De grootste zorg is hun betrouwbaarheid. Bijvoorbeeld, in multivariate calibratie is het in essentie de robuustheid van de calibratiemodellen en soms hun transfer die aandacht vereist. Het was de bedoeling om binnen dit project pre-normatief onderzoek uit te voeren met betrekking tot de manier waarop dit type van methoden gevalideerd moet worden. Uiteindelijk werd de problematiek uitgebreid onderzocht en leidde dit tot de formulering van een aantal richtlijnen terzake.

3. Resultaten

De verschillende onderdelen waarvoor getracht werd richtlijnen te formuleren of bestaande richtlijnen te evalueren, zijn :

- a) de evaluatie van de robuustheid van analytische methodes,
- b) de validatie van chromatografische bioanalysemethodes, d.w.z. methodes die bepalingen uitvoeren in biologische milieus,
- c) bepaling van de detectielimiet,
- d) de vergelijking van alternatieve meetmethodes,
- e) methodenvalidatiedesigns die kostenefficiënt zijn.
- f) bepaling van de reproduceerbaarheid van een methode,
- g) robuuste procesanalyse met multivariate calibratie,

a) Robuustheid (robustness/ruggedness) van analytische methodes

I.v.m. de studie van de robuustheid werd een filosofie opgesteld wat betreft de keuze van de experimentele designs en de statistische interpretatie van de effecten. Deze filosofie is beschreven in het artikel " *Y. Vander Heyden, F. Questier and D.L. Massart, A ruggedness test strategy for procedure related factors : experimental set-up and interpretation; Journal of Pharmaceutical and Biomedical Analysis, 17 (1998) 153-168*". Ze werd eveneens geïmplementeerd in een computerprogramma dat beschreven is in het manuscript "*F. Questier, Y. Vander Heyden and D.L. Massart; RTS, a computer program for the*

experimental set-up and interpretation of ruggedness tests; Journal of Pharmaceutical and Biomedical Analysis 18 (1998) 287-303".

Gebaseerd op discussies met de leden van het begeleidingscomité kon een lijst met discussiepunten opgesteld worden die in de loop van het project bestudeerd dienden te worden. Het gaat om de volgende punten :

- Wat is de bij voorkeur te gebruiken definitie voor ruggedness/robustness? Worden in een robuustheidstest voornamelijk procedure-gebonden factoren onderzocht in screening designs of volgt men meer de United States Pharmacopeia methode waar voornamelijk niet-procedure gebonden factoren onderzocht worden in nested designs? Nog een andere mogelijkheid welke door bepaalde bedrijven toegepast wordt, is het testen van de robuustheid m.b.v. een interlaboratoriumtest.

Case studies op voornamelijk procedure-gebonden factoren werden uitgevoerd in volgende manuscripten : (i) "Y. Vander Heyden and D.L. Massart; Y. Zhu and J. Hoogmartens; J. De Beer; *Ruggedness tests on the HPLC assay of the United States Pharmacopeia XXIII for tetracycline hydrochloride : comparison of different columns in an interlaboratory approach; Journal of Pharmaceutical and Biomedical Analysis 14 (1996) 1313-1326*", (ii) "Y. Vander Heyden, A. Bourgeois and D.L. Massart, *Influence of the sequence of experiments in a ruggedness test when drift occurs; Analytica Chimica Acta 347 (1997) 369-384*". Nested designs werden bestudeerd in "Y. Vander Heyden, K. De Braekeleer, Y. Zhu, J. Hoogmartens, J. De Beer and D.L. Massart, *Nested designs in ruggedness testing, aanvaard voor publicatie in Journal of Pharmaceutical and Biomedical Analysis*"

- Wanneer voert men een ruggedness test uit? Alleen voor methodes die onderworpen worden aan een interlaboratoriumtest of ook voor methodes die een meer beperkt gebruik kennen, b.v. alleen in één lab. Past men voor deze laatste methodes ook steeds een ruggedness test toe of alleen wanneer de tussen-dagen precisie slecht is?

- In de huidige filosofie worden de designs zodanig gekozen dat een bepaalde statistische interpretatie mogelijk is. Soms zijn daarvoor relatief veel experimenten nodig, meer dan de analyst verkiest uit te voeren. Daarom wordt eraan gedacht een strategie in stappen uit te bouwen die de analyst moet toelaten toch bepaalde conclusies te trekken na het uitvoeren van een minder uitgebreid design (invoeren van minimale designs). Hierdoor zal bepaalde

informatie ingeleverd moeten worden en zal er bijvoorbeeld geen of slechts een beperkte statistische analyse mogelijk zijn.

Daarom werden minimale designs aan de strategie toegevoegd. Deze laten wel het berekenen van effecten toe maar niet de oorspronkelijke statistische interpretatie. Een aantal onder hen kunnen na een eerste evaluatie wel uitgebreid worden tot de designs uit de oorspronkelijke strategie. Voor deze minimale designs werd toch getracht een significantie aan te geven voor de berekende effecten m.b.v. randomisatietesten.

De randomisatietesten gaven problemen wanneer er meer dan één significant effect was. Een sterk significant effect kan namelijk verantwoordelijk zijn voor de onderschatting van de significantie van de andere effecten. Na herhaaldelijke correcties waarbij elk significant bevonden effect geëlimineerd wordt, bleek dat in het algemeen dezelfde effecten als significant aangeduid worden door de randomisatietesten als door de t-testen uitgevoerd bij de "recommended designs" en waarbij de standard error berekend wordt gebaseerd op interactie-effecten of dummy factor effecten.

Uiteindelijk werden in de minimale designs de significante effecten geschat, gebruik makend van de gegevens van de niet-significante (algorithme van Dong).

- Kunnen supersaturated designs gebruikt worden in een ruggedness test?

Het gebruik van supersaturated designs als minimale designs in een robuustheidstest werd onderzocht. De gebruikte supersaturated designs werden geconstrueerd uit Plackett-Burman designs en voor een aantal case studies werd de variantie van de supersaturated designs vergeleken met deze van de Plackett-Burman designs (de berekening van effecten is nl. niet meer mogelijk voor de supersaturated designs). De verkregen varianties voor de supersaturated designs waren over het algemeen niet verschillend van deze van de Plackett-Burman designs wat erop wijst dat de supersaturated designs toelaten de variantie veroorzaakt door de factoren op een geschikte manier te schatten. Er zijn echter nog meer case studies vereist om dit te kunnen veralgemenen.

In een volgende stap diende bepaald te worden of de variantie gemeten voor een supersaturated design, veroorzaakt wordt door een al dan niet robuuste methode, m.a.w. een statistische test is vereist. In deze context werden een aantal voorstellen gedaan waarbij een kritische waarde berekend wordt uitgaande zowel van een theoretische benadering als van

practische resultaten. De designvariantie moet nl. vergeleken worden met een variantie die een normale experimentele fout voorstelt. De praktische bruikbaarheid van de voorgestelde criteria dient nog nagegaan te worden. Vooraleer daar echter mee te starten werd er nagegaan of er statistisch aanvaardbare resultaten gevonden kunnen worden. Dit blijkt echter niet het geval te zijn door het geringe aantal vrijheidsgraden te wijten aan het lage aantal uitgevoerde experimenten. Afhankelijk van de gebruikte hypothese zal men ofwel een grote probabiliteit hebben om niet-robuste methodes als robuust te beschouwen, ofwel robuuste methodes te verwerpen als niet-robust. Daarom dienen er andere interpretatiemethodes voor dergelijke designs gezocht te worden.

- Moet een factor op 2 (extreme) of op 3 (2 extreme en 1 nominaal) niveaus onderzocht worden?

- Wenst de gemiddelde gebruiker in staat te zijn de effecten statistisch te interpreteren? Het antwoord lijkt te zijn dat nogal wat gebruikers dat niet wensen. Ook de Eurachem voorstellen lijken in die richting te gaan.

Vaak is een analyst meer geïnteresseerd in chemisch relevante effecten dan in statistisch significante. Hoe kunnen deze chemisch relevante effecten bepaald worden?

In het artikel "*Y. Vander Heyden, C. Hartmann and D.L. Massart; P. Nuyten, A.M. Hollands and P. Schoenmakers; Ruggedness testing of a size exclusion chromatographic assay for low molecular mass polymers; Journal of Chromatography A, 756 (1996) 89-106*" is getracht daar een antwoord op te geven.

- Kan men bij de bepaling van statistisch significante of chemisch relevante effecten er van uitgaan dat men reeds de repeteerbaarheid of de tijdsafhankelijke precisie kent?

- Hoe moet men bepaalde factoren includeren in de screening designs die uitgevoerd worden? B.v. (i) de componenten van een mengsel, (ii) de zure en basische component van een buffermengsel.

Het onderzoeken van deze factoren evenals van enkele andere in screening designs werd bestudeerd en is beschreven in een artikel "*Y. Vander Heyden, F. Questier and D.L. Massart; Ruggedness testing of chromatographic methods : selection of factors and levels; Journal of Pharmaceutical and Biomedical Analysis 18 (1998) 43-56*". In deze referentie

wordt er beschreven hoe voor een bepaalde set van proces- en mengselvariabelen men de meest economische experimentele set-up kan definiëren. Er wordt eveneens aangetoond dat afhankelijk van hun formulering, voor bepaalde factoren (buffercomponenten) men meer of minder fysisch relevante effecten kan berekenen.

In de tekst wordt tevens een algemene richtlijn gegeven voor de keuze van de extreme niveaus in een robuustheidstest, alsook worden enkele speciale gevallen besproken (keuze van asymmetrische intervallen rond het nominale niveau bijvoorbeeld).

- Een robuustheidstest moet aangeven in welke mate fouten/afwijkingen in de factoren aanleiding geven tot wijzigingen in de respons. Het doel van een robuustheidstest moet zijn dat de gevonden resultaten en de getrokken conclusies leiden tot een verbetering van de resultaten van een interlaboratoriumtest d.w.z. de reproduceerbaarheid verbetert. Daarom zouden de resultaten van een robuustheidstest moeten leiden tot het creëren van system suitability criteria voor de (kritieke) factoren. Een voorstel terzake is gedaan in "*Y. Vander Heyden, M. Jimidar, E. Hund, N. Niemeijer, R. Peeters, J. Smeyers-Verbeke, D.L. Massart and J. Hoogmartens; Determination of system suitability limits with a robustness test, aanvaard voor publicatie in J. of Chromatography A*".

- Stel dat men een methode niet zelf ontwikkeld heeft, maar men er toch een robuustheidstest op uitvoert. De nominale waarden voor de onderzochte factoren worden in dit geval overgenomen van een bepaalde bron. Wanneer nu bepaalde effecten groot zijn d.w.z. de methode is niet robuust, hoe kan men gebaseerd op de ruggedness test resultaten, niveaus definiëren voor deze factoren om het responsoppervlak te bepalen en tot een nieuwe optimalisatie te komen?

Uit de opgedane kennis werden er tenslotte richtlijnen gedefinieerd voor het opstellen en interpreteren van een robuustheidstest welke de definitie voor robuustheid van de ICH volgen. Deze richtlijnen bevatten eveneens de mogelijkheid om system suitability limieten te bepalen uit de resultaten van een robuustheidstest. Dit laatste wordt eveneens aanbevolen door de ICH. De verschillende stappen in de set-up en interpretatie van een robuustheidstest worden in de richtlijnen gegeven. De richtlijnen werden opgesteld in samenwerking met Unilever (Vlaardingse, Nederland), terwijl er een coöperatie bestond met Novartis (Basel,

Zwitserland), Bayer (Dormagen, Duitsland) en Shell (Amsterdam, Nederland) tijdens het uitwerken van een aantal case studies die het uiteindelijk opstellen van de richtlijnen toelieten. De richtlijn bevat twee delen, namelijk de selectie van de experimentele set-up (definitie van factoren en hun niveaus, en selectie van de designs) en de behandeling van de resultaten van een uitgevoerd design. Ze bevat bijvoorbeeld voorstellen voor het uitvoeren design (Plackett-Burman of fractional factorial design) in functie van het aantal te bestuderen factoren. De behandeling van de resultaten bestaat in de berekening van effecten en hun grafische en statistische interpretatie teneinde hun significantie te bepalen. Wanneer een methode dan als robuust beschouwd wordt wat zijn kwantitatief aspect betreft, kunnen voor een aantal responsen, zoals resolutie, analysetijd, asymmetriefactoren in chromatografie, system suitability limieten gedefinieerd worden. Deze richtlijn (*Y. Vander Heyden, A.Nijhuis, B.G.M. Vandeginste and D.L. Massart; Robustness/Ruggedness Tests in Method Validation*) werd toegevoegd in bijlage.

Daarnaast werken een aantal leden van dit onderzoeksprogramma, nl. P. Chiap (ULg), Ph. Hubert (ULg) en Y. Vander Heyden (VUB), en van het begeleidingscomité (B. Boulanger (E. Lilly) samen binnen een commissie "Robustesse" van de Franse SFSTP aan een artikel dat eveneens een aantal richtlijnen zal voorstellen voor het opstellen en het interpreteren van een robuustheidstest en dus gedeeltelijk overlapt met de hogervermelde richtlijnen. Het artikel zal twee grote delen omvatten, nl. (i) het implementeren van robuustheid en robuustheidscriteria in methodeoptimalisatie, en (ii) het uitvoeren van robuustheidstesten in methodenvalidatie. Het gedeelte over de robuustheid in methodenvalidatie zal grotendeels gebaseerd zijn op de resultaten en richtlijnen beschreven in de doctoraatsthesis van Y. Vander Heyden. Hieruit blijkt duidelijk de invloed, binnen deze commissie, van de Belgische leden van dit normalisatieprogramma op het Franse niveau.

Een onderwerp dat naar het einde van het project toe onderzocht werd, is het gebruik van zogenaamde asymmetrische designs in robuustheidstesten. Deze designs laten toe sommige factoren op meer dan twee niveaus te onderzoeken. Dit laatste is voor kwalitatieve factoren in feite een vereiste om een min of meer representatief idee te krijgen over hun invloed op de robuustheid van een analyseresultaat. De resultaten van dit onderzoek zijn beschreven in de volgende manuscripten "*E. Hund, Y. Vander Heyden, M. Haustein, D.L. Massart and J. Smeyers-Verbeke, Comparison of several criteria to decide on the significance of effects in*

a robustness test with an asymmetrical factorial design, document in voorbereiding” en “*E. Hund, Y. Vander Heyden, M. Haustein, D.L. Massart and J. Smeyers-Verbeke, Application of fractional factorial and asymmetrical factorial designs in the robustness testing of a RP-HPLC method, document in voorbereiding*”.

b) Validatie van chromatografische bioanalysemethodes, d.w.z. methodes die bepalingen uitvoeren in biologische milieus

In samenwerking met de SFSTP werd een strategie besproken voor de validatie van chromatografische methodes voor bepalingen in biologische milieus. Deze besprekingen leidden enerzijds tot de publicatie van een aantal richtlijnen i.v.m. validatie in biologische milieus (*S. Bervoas-Martin, B. Boulanger, P. Chevalier, P. Chiap, D. Grandjean, P. Hubert, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, E. Chapuzet et N. Mercier; Méthodes chromatographiques de dosage dans les milieux biologiques : stratégie de validation; S.T.P. Pharma Pratiques 7, 3 (1997) 169-194*) en tot een aantal case studies ter evaluatie van de voorgestelde richtlijnen ((i) *S. Bervoas-Martin, B. Boulanger, P. Chevalier, P. Chiap, D. Grandjean, P. Hubert, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, E. Chapuzet et N. Mercier; Méthodes chromatographiques de dosage dans les milieux biologiques. Exemple d'application de la stratégie de validation; S.T.P. Pharma Pratiques 8, 2 (1998) 81-107*, (ii) *Ph. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie et J.C. Nivet; The SFSTP Guide on the validation of chromatographic methods for drug bioanalysis: from the Washington Conference to the laboratory, ingediend voor publicatie*, (iii) *P. Chiap, Ph. Hubert, B. Boulanger, J. Crommen, Validation of an automated method for the liquid chromatographic determination of atenolol in plasma: application of a new validation protocol, ingediend voor publicatie*, en (iv) *B. Boulanger, Ph. Hubert, P. Chiap, J. Crommen; An analysis of the SFSTP guide on validation of chromatographic bioanalytical methods: progress and limitations, ingediend voor publicatie*) en anderzijds tot de ontwikkeling van een softwaresysteem voor methodevalidatie in deze biologische milieus.

c) Bepaling van de detectielimiet

Een aanvang werd genomen met het vergelijken van twee benaderingen voor de evaluatie van de detectielimiet : *i)* gebaseerd op de meting van blancostalen; *ii)* gebaseerd op de informatie van de calibratielijn. Ze werden toegepast op NIR metingen gebaseerd op univariate calibratie. Verdere evaluatie is echter noodzakelijk.

d) De vergelijking van alternatieve meetmethodes

Een aantal bestaande standaarden en richtlijnen, zoals ISO 5725, AFNOR T90-120 en Eurachem werden verzameld en bestudeerd met het oog op hun bruikbaarheid voor het vergelijken van analysemethoden. Het doel van deze studie is richtlijnen te hebben die toelaten aan te tonen dat (i) een nieuwe methode even goed of beter presteert dan een bestaande referentiemethode, of (ii) een bepaalde methode zonder problemen getransfereerd kan worden van een researchcentrum naar een industrieel labo of van het ene naar het andere industrieel labo. De ISO-normen zijn beperkt tot het gebruik van een methode in een bepaalde matrix en voor een bepaalde concentratie en zijn bedoeld voor interlaboratoriumstudies. Ze moesten geherdefinieerd worden voor het bestuderen van meerdere matrices en meerdere concentratieniveaus en voor intralaboratoriumstudies. Het voordeel van de ISO-norm tegenover andere validatiestandaarden is dat zowel de β -fout als interval hypothesetesten geïnccludeerd zijn. In verband met het vergelijken van methodes werd dus een protocol opgesteld uitgaande van de bestaande standaarden. De resultaten zijn beschreven in het artikel "*S. Kuttatharmakul, D.L. Massart and J. Smeyers-Verbeke, Comparison of alternative measurement methods, accepted for publication in Analytica Chimica Acta*".

Op basis van de ISO 5725 (deel 6-sectie 8) norm werd een voorstel van richtlijnen uitgewerkt voor het vergelijken van analysemethoden. Daartoe werd de ISO norm, die gebaseerd is op interlaboratoriumstudies, aangepast voor het vergelijken van twee analysemethoden binnen eenzelfde organisatie (intralaboratoriumvergelijking). Dit betekent b.v. dat de precisie niet bestudeerd wordt onder reproduceerheidscondities maar dat intermediaire precisie condities (v.b. verschillende dagen en operatoren) moeten beschouwd worden. De voorgestelde richtlijnen houden nu rekening met de "time-different intermediate" precisie en dienen uitgebreid te worden tot de "[time+operator]-different intermediate" precisie.

Een eerste voorstel van richtlijn start zoals de ISO norm met een bepaling van het aantal metingen dat nodig is om de β -fout te controleren (de β -fout in deze context is de probabiliteit dat een relevante bias niet gedetecteerd wordt of dat een aanvaardbare precisie niet gehaald wordt). Deze benadering kan echter aanleiding geven tot een aantal uit te voeren metingen dat in praktijk niet of moeilijk haalbaar is. Daarom werd een tweede richtlijn ontwikkeld die start met een door de gebruiker bepaald aantal metingen en die de β -fout evalueert. Op die manier krijgt de analyst ten minste een idee van de waarschijnlijkheid dat een methode die eigenlijk niet aan de voorwaarden voldoet, aanvaard wordt.

De eerder voorgestelde richtlijnen voor het vergelijken van analysemethoden houden enkel rekening met de "time-different intermediate" precisie ($s_{I(T)}^2$). Nagegaan werd hoe een

uitbreiding naar de "[time+instrument+operator]-different intermediate" precisie ($s_{I(OIT)}^2$) mogelijk is. Het probleem bij een intralaboratoriumvergelijking van methoden is namelijk dat meestal het aantal instrumenten gebruikt voor de analyse alsook het aantal analisten dat de analyses uitvoert beperkt is. Dit leidt tot een slechte schatting van de variantiecomponenten voor instrumenten en operatoren. Daarom wordt voorgesteld $s_{I(OIT)}^2$ niet als de som van de individuele variantiecomponenten te berekenen maar deze af te leiden uit de kwadraatsommen van de verschillen tussen de daggemiddelden en het algemeen gemiddelde. Alhoewel de op deze manier bekomen $s_{I(OIT)}^2$, die afhankelijk kan zijn van het aantal beschouwde instrumenten en operatoren, misschien geen goede schatter is voor de "[time+instrument+operator]-intermediate" precisie is het bruikbaar in de vergelijking van methodes op voorwaarde dat eenzelfde aantal operatoren en instrumenten voor beide methodes beschouwd worden.

De twee hoger voorgestelde richtlijnen voor het vergelijken van analysemethoden (de eerste gebaseerd op een bepaling van het aantal metingen nodig om de β -fout tot een gespecificeerde waarde te beperken, de tweede vertrekkend van een door de gebruiker gespecificeerd aantal metingen en een evaluatie van de β -fout) werden uitgewerkt en aangepast. Daarbij werd rekening gehouden met opmerkingen gemaakt door o.a. Elf en Novartis (industriële partners in het hoger vermelde Europese SMT project gecoördineerd door de VUB).

Het was daarbij noodzakelijk om in de eerste richtlijn een beslissing te nemen i.v.m. de aanvaardbare β -fout (hier de probabiliteit dat een relevante bias niet gedetecteerd wordt of een gespecificeerde precisie niet gehaald wordt) voor het berekenen van het aantal uit te voeren metingen (n). Om praktische redenen (n neemt toe naarmate men meer zekerheid wil dat b.v. een relevante bias inderdaad gedetecteerd zal worden, dit betekent naarmate men een kleinere β -fout wil) werd de voorkeur gegeven om $\beta=20\%$ te specificeren in plaats van 5% zoals in de eerste versie voorgesteld. Vermits de verschillende statistische testen op het klassieke 5% significantieniveau ($\alpha=5\%$) uitgevoerd worden, betekent dit dat nog altijd meer belang gehecht wordt aan de α - dan aan de β -fout. Met andere woorden men beschouwt het, louter om praktische redenen, belangrijker om een methode die eigenlijk aanvaardbaar is niet te verwerpen dan om een methode die eigenlijk niet geschikt is te aanvaarden.

Een betere balans tussen de α - en de β -fout zou verkregen kunnen worden door $\alpha=10\%$ en $\beta=10\%$ te specificeren. Het aantal uit te voeren metingen met deze waarden voor α en β is

vergelijkbaar met deze voor $\alpha=5\%$ en $\beta=20\%$. Alhoewel dit logischer is, werd het in de richtlijn niet voorgesteld omdat dit zou betekenen dat de statistische testen uitgevoerd worden op het 10% in plaats van op het klassieke 5% significantieniveau. Dit laatste is zo standaard dat afwijken hiervan momenteel waarschijnlijk moeilijk aanvaard zou worden.

Vermits voorgesteld werd om de richtlijnen in te dienen bij de CEN werden ze samengevoegd in één document dat verspreid werd onder de partners in dit onderzoeksprogramma en in het SMT project.

In de hoger vermelde richtlijnen voor het vergelijken van analysemethodes wordt enkel rekening gehouden met de "time-different intermediate" precisie. Zoals beschreven in het vorig verslag werd nagegaan hoe een uitbreiding naar de "[time+instrument+operator]-different intermediate" precisie mogelijk is. Een artikel in dat verband is aanvaard voor publicatie in *Analytica Chimica Acta*, nl. "S. Kuttatharmakul, D.L. Massart and J. Smeyers-Verbeke, *Comparison of alternative measurement methods, accepted for publication in Analytica Chimica Acta*"

Vergelijking van analysemethodes : de voorgestelde richtlijnen voor de vergelijking van twee analysemethodes, beperken zich tot een vergelijking van de bias en van verschillende precisiecomponenten. De voorgestelde richtlijnen zijn ofwel gebaseerd op een bepaling van het aantal metingen nodig om de β -fout tot een gespecificeerde waarde te beperken, ofwel vertrekkend van een door de gebruiker gespecificeerd aantal metingen en een evaluatie van de β -fout. Nagegaan werd ook hoe de detectielimieten van beide methodes vergeleken konden worden. Dit is uiteraard belangrijk wanneer het methoden voor sporenanalyse betreft.

e) Methodvalidatiedesigns die kostenefficiënt zijn : lineariteit ijklijn

Wat betreft de studie van methodevalidatie designs die kosten-efficiënt zijn, werden verschillende designs vergeleken voor de evaluatie van de lineariteit van de ijklijn. Een evaluatie van lineariteitstesten (Anova lack-of-fit test en testen van de significantie van de kwadratische coëfficiënt b_2) werd uitgevoerd aan de hand van gesimuleerde niet-lineaire calibratielijnen met een verschillende graad van afbuiging die gebaseerd zijn op een beperkt aantal metingen ($6 < n \leq 10$). Verschillende designs waarin de n metingen verdeeld zijn over 3, 4 of 5 concentratieniveaus werden daarbij bestudeerd. De waarschijnlijkheid dat een lack-of-

fit niet gedetecteerd wordt (= de β -fout), is het kleinst wanneer slechts 3 concentratieniveaus beschouwd worden. Heteroscedastiteit (niet constante variantie over het beschouwde concentratiegebied), die gezien het gering aantal metingen, niet aan de hand van de experimentele gegevens kan onderzocht worden, geeft aanleiding tot een aanzienlijke toename van de β -fout. Vermits de β -fout ook afhankelijk is van de precisie werd de invloed van deze laatste ook nagegaan. Dit is belangrijk omdat dit toelaat te bepalen welke precisie noodzakelijk is om een relevante afwijking van de lineariteit te detecteren. In praktijk is het echter moeilijk te specificeren welke afwijking van lineariteit als belangrijk beschouwd wordt. Vermits een lack-of-fit echter aanleiding geeft tot fouten in de voorspelde concentraties (en dus tot een bias) wordt nu de β -fout van de lineariteitstesten in functie van de predictiefout bestudeerd. Bovenvermelde resultaten zijn beschreven in het manuscript "S. Kuttatharmakul, D.L. Massart and J. Smeyers-Verbeke, Influence of precision, sample size and design on the beta error of linearity tests; *Journal of Analytical Atomic Spectroscopy (JAAS)* 13 (1998) 109-118"

f) Bepaling van de reproduceerbaarheid van een methode en methodenvalidatie in farmacopee-analyses

Een van de applicatiegebieden waarbinnen het uittesten van richtlijnen en normen gepland waren, is de methodenvalidatie in farmacopee-analyses.

De toepassing van standaarden en richtlijnen in verband met meetmethoden uit de farmacopee voor de analyse van geneesmiddelen maakte dus een belangrijk onderdeel uit van het project. De bestaande voorschriften en richtlijnen, moeten daarom eerst toegepast worden op enkele case studies. Op punt stellen van nieuwe of verbeterde methoden blijft eveneens noodzakelijk.

Na toetsing werd gekozen om de analysemethodes voor enkele belangrijke antibiotica als model te nemen. De biotechnologie van de industriële antibiotica is op dit ogenblik trouwens in volle expansie, wat ongetwijfeld een diepgaande invloed zal hebben op de aard en samenstelling van deze producten.

Onze keuze kan als volgt toegelicht worden. Spiramycine en tylosine, beide behorend tot de groep van macrolide antibiotica, worden niet alleen aangewend respectievelijk in de humane

en veterinaire geneeskunde, maar hebben ook toepassing gevonden in de agro-voeding als groeibevorderende stoffen.

Bacitracine, een polypeptide antibioticum, wordt veelal aangewend als zinkcomplex ter bevordering van de voedselopname efficiëntie bij sommige diersoorten en ziekte controle. Het is in deze sector het economisch belangrijkste produkt.

Amoxicilline, een belangrijke vertegenwoordiger van de β -lactam groep werd eveneens als model gekozen. Het is een semi-synthetische penicilline met activiteit tegen zowel gram-positieve als gram-negatieve bacteria. Het wordt gebruikt in verschillende farmaceutische vormen zoals capsules, orale suspensies als injectievloeistoffen.

De 3 eerst vermelde antibiotica kunnen alleen als fermentatieprodukt verkregen worden door klassieke biosynthese. Dit houdt in dat het isoleren gebeurt uitgaande van complexe voedingsmilieus ontstaan tijdens de gecontroleerde groei van veelal gemodificeerde micro-organismen.

De produktievoorwaarden en isoleringsschema's, bekomen na een uitgebreide reeks van experimenten, leiden tot economisch rendabele opbrengsten en behoren tot de confidentiële kennis van de producenten. De bekomen actieve grondstoffen bezitten meestal een complexe samenstelling met een groep actieve componenten naast een wisselend aantal inactieve bestanddelen. Hun chemische aard is de oorzaak dat gemakkelijk ongewenste nevenprodukten kunnen ontstaan door chemische omzetting tijdens het isoleren of de opzuivering. Deze onzuiverheden bezitten meestal geen activiteit en sommige kunnen zelfs toxisch zijn.

De oorspronkelijke microbiologische procedure die tot op heden voor heel wat antibiotica wordt toegepast, is gebaseerd op de activiteit ten opzichte van gevoelige micro-organismen. Deze methoden zijn zeer bewerkelijk, langdurig en onderhevig aan tal van interferenties. De wisselende samenstelling van de monsters ten opzichte van de gebruikte standaard geeft aanleiding tot problemen bij deze activiteitsbepalingen, wat ook zal blijken bij de bespreking van de bekomen resultaten! Het gebruik van alternatieve methoden, zoals de scheidingsmethoden HPLC en CE, dringt zich dan ook op.

Het gebruik van fysico-chemische methoden om de kwaliteitscontrole van antibiotica en analoge farmaceutische produkten uit te voeren, is reeds enkele jaren in opgang. Er moet

naar gestreefd worden om de vereiste selectiviteit te bereiken. De signalen die gemeten worden moeten selectief zijn voor de te bepalen stoffen. De methode die tegenwoordig meest gebruikt wordt, is een combinatie van een chromatografische scheiding gekoppeld aan een spectrometrische bepaling. Voor de onderzochte modellen is vloeistofchromatografie of capillaire elektroforese met ultraviolet detectie de voor de hand liggende keuze.

Interlaboratoriumstudies werden opgesteld op twee analysemethodes (farmacopee methodes) voor de bepaling van het antibioticum tylosine. Deze studies werden opgesteld volgens de ISO-normen terzake (ISO 5725-2). De eigenlijke interlaboratoriumexperimenten werden voorafgegaan door een zogenaamde trainingsronde. Dergelijke trainingsronde laat toe een aantal problemen die zich voordoen in de deelnemende labo's op te lossen zodanig dat men in de uiteindelijke studies deze problemen kan vermijden en tot een maximaal aantal bruikbare resultaten komt.

De behandeling van de resultaten van de interlaboratoriumstudies gebeurde volgens de ISO 5725-2 richtlijn. De gegevens werden eerst getest op outliers (zowel outliers op het niveau van de binnen-laboratoriumvariantie als van de tussen-laboratoriumvariantie). Na verwijderen van de outliers wordt de repeteerbaarheid en de reproduceerbaarheid van de methode berekend en wordt er nagegaan of er een verband bestaat (lineair of logaritmisch) tussen de variantie en de concentratie te bepalen bestanddeel (homo/heteroscedasticiteit).

Aangezien de studie uitgevoerd werd m.b.v. twee bepalingsmethodes op dezelfde stalen, konden de resultaten van beide methodes vergeleken worden. Er werden zowel kwantitatieve parameters (varianties, gemiddelde resultaten) als kwalitatieve (resolutie, analysetijd, enz.) beschouwd.

De resultaten van dit onderzoek werden opgenomen in een tutorial artikel "*Y. Vander Heyden, J. Saevels, E. Roets, J. Hoogmartens, D. Decolin, M.G. Quaglia, W. Van den Bossche, R. Leemans, O. Smeets, F. Van de Vaart, B. Mason, G.C. Taylor, W. Underberg, A. Bult, P. Chiap, J. Crommen, J. De Beer, S.H. Hansen and D.L. Massart, Interlaboratory studies on two HPLC assays for tylosin (tartrate), Journal of Chromatography A, 830 (1999) 3-28*".

Wat de andere substanties (spiramycine, bacitracine, amoxicilline) betreft, werd zowel methodenoptimalisatie als methodenvalidatie uitgevoerd. De resultaten ervan zijn in meer detail beschreven in (i) *Y.M. Li, A. Van Schepdael, Y. Zhu, E. Roets, J. Hoogmartens; Development and validation of amoxicillin determination by micellar electrokinetic capillary chromatography, Journal of Chromatography A, 812 (1998) 227-236*, en (ii) *L. Liu, J. Saevels, P. Louis, H. Nelis, S. Rico, K. Dierick, S. Guyomard, E. Roets, J. Hoogmartens; Interlaboratory study comparing the microbiological potency of spiramycins I, II and III; in druk in Journal of Pharmaceutical and Biomedical Analysis.*

g) Robuuste procesanalyse met multivariate calibratie

Wat de ontwikkeling van een filosofie voor procesanalyse betreft, werden twee problemen bestudeerd. Een eerste was het nagaan of de gegevens afwijkende metingen bevatten. Als oplossing van dit probleem werd een strategie vooropgesteld, voor NIR metingen, gebaseerd op het meten van leverage punten en op een evolutieprogramma, lijkend op een genetisch algoritme.

Het tweede probleem was het bepalen van de factoren welke een invloed hebben op de precisie. Hiervoor werd gevonden dat een aantal duplicate of replicate metingen van een beperkt aantal standaarden een even goed model en een even goede precisie oplevert als een groot aantal standaarden die slechts éénmaal gemeten worden.

In het laatste deel van het project werkte de VUB partner eveneens mee aan het opstellen van validatierichtlijnen voor Nabij Infrarood (NIR) spectrofotometrie. Deze richtlijnen worden opgesteld op Europees niveau en zijn bedoeld voor toepassing in de industrie. Andere partners in dit project zijn o.a. de multinationals Pfizer, Shell en Elf.

Veel methoden voor multivariate calibratie zijn in de literatuur voorgesteld. Het bleek nu dat meerdere van deze methoden een analoge performantie hebben. Om verwarring te vermijden, stellen we voor dat slecht de volgende methodes beschouwd zullen worden

- Multiple linear regression (MLR) : multi-pele lineaire regressie
- Principal component regression (PCR) : principale componentenregressie
- Partial least squares (PLS)

- Neural networks (NN) : neurale netwerken
- Locally weighted regression (LWR) : lokaal gewogen regressie
- Radial basis functions combined with PLS (RBF-PLS) : radial basis functies gecombineerd met PLS

Ze werden opgenomen op basis van theoretische overwegingen, die bevestigd werden door onderlinge vergelijking van hun performantie uitgevoerd op NIR data sets met verschillende data structuren.

Tijdens bovenvermelde vergelijkingen werden nog vele andere methodes uitgetest, zoals niet-lineaire varianten van PLS, PCR en MLR, nearest-neighbour methodes, methodes waarin de regressie is uitgevoerd op Fourier coëfficiënten, en wavelets. Deze werden tenslotte niet weerhouden omdat sommige geen goede resultaten opleverden, terwijl andere vergelijkbare resultaten met de bovenvermelde methodes gaven maar veel moeilijker te begrijpen of minder algemeen zijn. Om deze redenen is hun gebruik niet aangewezen door gebruikers die niet vertrouwd zijn met chemometrie.

- MLR is vanuit statistisch oogpunt gezien de best gekende methode. De methode is gebaseerd op de selectie van variabelen. Ze verdient de voorkeur wanneer de selectie van variabelen eenvoudig is, d.w.z. wanneer bepaalde variabelen relatief selectief zijn voor de componenten of eigenschappen die bepaald worden.
- PCR is eveneens een gekende methode die, van de beschreven methodes, het dichtst aanleunt bij PCR. Het is een tweestapsprocedure. In de eerste stap worden principale componenten bepaald. Dit zijn lineaire combinaties van de originele variabelen. Ze kunnen als nieuwe variabelen beschouwd worden die op een optimale manier de variantie bundelen welke in de spectra aanwezig is. Dergelijke variabelen die de originele variabelen combineren op een dusdanige manier dat deze combinaties een bepaald criterium optimaliseren (hier de verklaarde variantie maximaliseren) worden ook latente variabelen genoemd. In de tweede stap wordt dan MLR toegepast op de nieuw verkregen latente variabelen.
- PLS is eveneens een methode die vaak gebruikt wordt in multivariate calibratie. Het lijkt op PCR maar werkt in één stap. De methode bepaalt eveneens latente variabelen,

maar het gebruikte criterium hier is de maximale covariantie tussen de y-waarden en de spectrale variabelen.

- Neurale netwerken (NN) : slechts backpropagation netwerken werden gebruikt. Neurale netwerken starten met het maken van combinaties van de originele variabelen. Deze combinaties worden dan non-lineair gemaakt door het gebruik van transferfuncties.
- Lokaal gewogen regressie (LWR) gebruikt PCR of PLS, in combinatie met een wegingssysteem zodat de calibratiestandaarden die het dichtst staan tot het te voorspellen staal, ook het grootste gewicht krijgen.
- RBF-PLS : dit is een lokale methode die zeer goed blijkt te werken voor moeilijke datastructuren

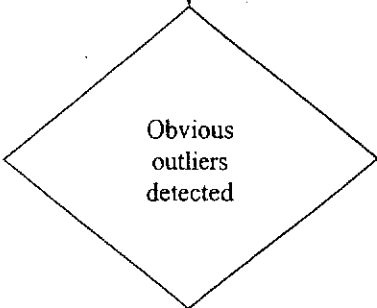
De publicaties gemaakt in het kader van dit onderzoek zijn opgenomen in de bijlage.

Op basis van de opgedane kennis kon volgend beslissingsschema opgesteld worden.

Data set acquisition and pre-processing



Data inspection :
- PC score plots
- Plot of most correlated PCs vs y
- Histogram of y-values

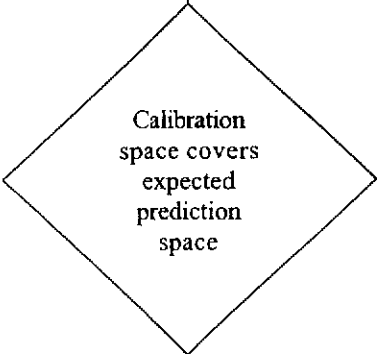


yes

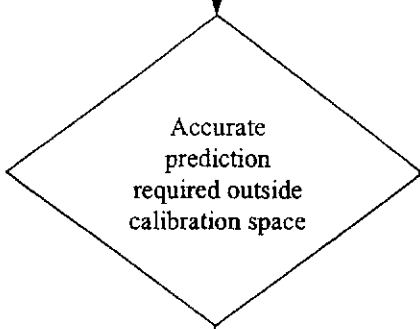
Remove outliers



no



no



yes

Predicted y-values outside the calibration space will indicate that a different model must be applied

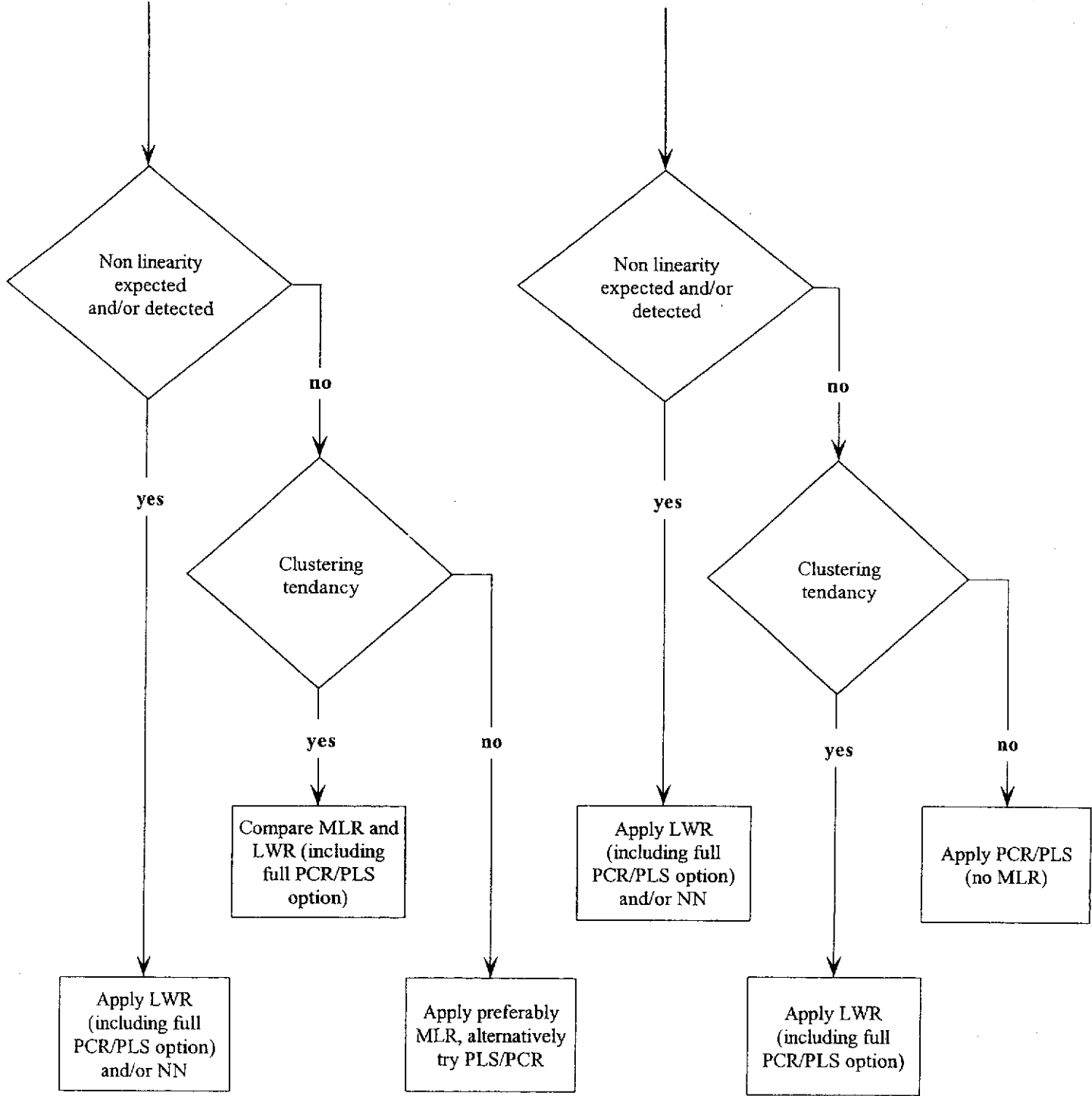


yes



no





4. Besluiten en aanbevelingen

De verschillende onderdelen waarvoor in dit project richtlijnen geformuleerd werden of bestaande richtlijnen geëvalueerd, waren :

- a) de evaluatie van de robuustheid van analytische methodes,
- b) de validatie van chromatografische bioanalysemethodes, d.w.z. methodes die bepalingen uitvoeren in biologische milieus,
- c) bepaling van de detectielimiet,
- d) de vergelijking van alternatieve meetmethodes,
- e) methodenvalidatiedesigns die kostenefficiënt zijn, met name voor de lineariteit van een ijklijn,
- f) bepaling van de reproduceerbaarheid van een methode,
- g) robuuste procesanalyse met multivariate calibratie,

Methodenvalidatie is noodzakelijk voor de industrie, maar ook duur. Om die reden werd er bij het opstellen van bepaalde normen en richtlijnen, eveneens getracht om volgende zaken te definiëren : (i) methodenvalidatiedesigns die kostenefficiënt zijn, en (ii) minimale vereisten voor methodenvalidatie.

- (i) Een van de meest belangrijke vragen in methodenvalidatie is welke en hoeveel experimenten uit te voeren. Hoe meer experimenten er uitgevoerd worden, hoe beter de prestatiekenmerken bepaald kunnen worden, maar des te duider ook de validatie. Goede experimentele designs laten toe het aantal experimenten te beperken en toch genoeg informatie te verkrijgen. Het experimentele design hangt af van de situatie. Het is bijvoorbeeld verschillend voor een methode toegepast voor het testen van sommige productspecificaties, waar de concentratierange die wordt getest beperkt is, en voor situaties waar de range groot kan zijn, zoals in de voedingsmiddelen- of bioanalyse.
- (ii) Methoden kunnen in verschillende mate gevalideerd worden. In veel praktische situaties is een uitgebreide validatie niet haalbaar, noch noodzakelijk. De vraag is dan of overeenstemming kan worden gevonden over at het minimum is, afhankelijk van de context. Een benadering voor het definiëren van dergelijke minimum vereisten moet ook diagnostica of regels bevatten die toelaten te beslissen wanneer het minimum niet genoeg is en er meer gedaan moet worden.

Er dient opgemerkt te worden dat in dit project voor een aantal methodenvalidatieonderdelen wel rekening gehouden werd met het opstellen van kostenefficiënte designs (b.v. het nagaan van de lineariteit van een ijklijn) en met minimale vereisten (b.v. gebruik van minimale designs in robuustheidstesten), maar dat dit niet op een veralgemeende basis het geval is. In

de toekomst is er dus nog zeker werk weggelegd voor het ontwikkelen van verdere minimale en kostenefficiënte methodenvalidatieonderdelen, zodat men uiteindelijk tot een situatie komt waarbij men validatierichtlijnen heeft zowel voor een volledige als voor een minimale validatie.

Alhoewel de gebruikelijke statistische pakketten kunnen helpen in methodenvalidatie, is er geen enkel beschikbaar dat voldoende algemeen is en expliciet gericht op methodenvalidatie. De mogelijkheden van de bestaande methodenvalidatieprogramma's daarentegen zijn echter sterk gelimiteerd in vergelijking tot het hier onderzochte aantal onderwerpen. Inderdaad moet men rekening houden met de chemische context, die testen en tools vereisen die niet beschikbaar zijn in de meeste klassieke statistische pakketten. Een eenvoudig voorbeeld is het vergelijken van de helling van twee calibratielijnen (is er een statistisch significant verschil tussen beide?), hetgeen vaak noodzakelijk is voor het detecteren van interferenties met de zogenaamde standaardadditiemethode. Omdat statistische conclusies vaak beperkt zijn tot fouten van de eerste soort is het niet mogelijk te besluiten hoeveel experimenten men moet uitvoeren om twee methoden of twee laboratoria te vergelijken, zoals vaak vereist is in methodenvalidatie. Bovendien zijn de meeste chemici niet voldoende getraind in statistiek en in de terminologie, zodat een klassiek geconstrueerd pakket niet geschikt is.

Om die reden is het nodig specifieke software te produceren. Naast de voorgestelde richtlijnen is software dus één van de eindproducten van het project. Er werden programma's ontwikkeld voor de validatie van bioanalytische methodes, voor het opstellen en interpreteren van robuustheidstesten, voor het vergelijken van methodes en voor multivariate calibratie. Omdat het maken van professionele, gebruiksvriendelijke en gevalideerde software moeilijk en kostbaar is, werd geopteerd voor het ontwikkelen van demonstratiesoftware.

Het is de bedoeling om in de toekomst enerzijds de door ons opgestelde richtlijnen te laten officialiseren en anderzijds om ze te promoten via verspreiding langs het Internet. De richtlijnen opgesteld i.v.m. robuustheidstesten zullen b.v. voorgelegd worden aan de ICH teneinde ze te laten officialiseren terwijl andere aan ISO zullen voorgesteld worden. De opgestelde richtlijnen zullen eveneens beschikbaar gesteld worden door ze op het Internet te publiceren, in combinatie met tutorials, praktische voorbeelden en de eventueel reeds beschikbare software. Dit moet de verspreiding van de in het project opgestelde richtlijnen bevorderen.

5. Synthese van het onderzoek

5.1. Inleiding

Overeenstemming tussen meetresultaten en hun onderlinge erkenning voor het nagaan van het naleven van de productspecificaties is een probleem dat belangrijk is voor iedereen die betrokken is in een consument-producent relatie. Dit is zowel een probleem voor de industrie als voor nationale en internationale beleidsorganen. Een dergelijke overeenkomst vereist dat de meetmethoden gevalideerd zijn, dit wil zeggen dat aangetoond moet worden dat hun precisie en andere performantiekenmerken voldoen aan de eisen van alle belanghebbende partijen. Er is een nood aan afspraken over de wijze waarop deze validatie moet uitgevoerd worden, of met andere woorden : men heeft standaarden en richtlijnen nodig voor de validatie van meetmethoden.

De technische en wetenschappelijke doelstelling van het project was om gestandaardiseerde benaderingen en richtlijnen te ontwikkelen voor de validatie van meetmethoden in chemische, farmaceutische, agro-voeding en aanverwante industrieën. Op organisatorisch vlak bracht dit project de bekwaamheid van een aantal Belgische onderzoekers en een aantal industrieën op het gebied van methodenvalidatie samen, teneinde de Belgische wetenschappers een prominente rol te laten spelen in de ontwikkeling van normen en richtlijnen.

Methodenvalidatie is dus een belangrijk regulatory probleem voor vele industrieën en is nodig om de validiteit aan te tonen van de methoden die aangewend worden om product specificaties na te gaan. Bovendien is methodenvalidatie nodig voor de bescherming van de veiligheid van de consument en om de kwaliteit van consumptieproducten te garanderen.

De ontwikkeling van snelle, multivariate technieken, zoals o.a. de nabij-infrarood analyse, voor het meten van de kwaliteit tijdens of in onmiddellijke relatie met een proces, is een belangrijke trend in de industrie. Pre-normatief onderzoek om guidelines te ontwikkelen voor de procesanalyse was gepland en werd uiteindelijk uitgevoerd behandeld.

5.2. Methodologie

Binnen het project werden verschillende taken uitgevoerd en verschillende aspecten van de methodenvalidatie bestudeerd. De verschillende onderdelen waarvoor getracht werd richtlijnen te formuleren of bestaande richtlijnen te evalueren, zijn : a) de evaluatie van de robuustheid van analytische methodes, b) de validatie van chromatografische bioanalysemethodes, d.w.z. methodes die bepalingen uitvoeren in biologische milieus, c) bepaling van de detectielimiet, d) de vergelijking van alternatieve meetmethodes, e) methodenvalidatiedesigns die kostenefficiënt zijn, met name wat betreft de lineariteit van ijklijnen, f) bepaling van de reproduceerbaarheid van een methode, g) robuuste procesanalyse met multivariate calibratie.

Voor de verschillende bovenvermelde onderdelen was de aanpak meestal gelijklopend. In eerste instantie werd de reeds bestaande situatie, d.w.z. de uit de literatuur voorhanden zijnde gegevens, dieper onderzocht. In een tweede stap werden enkele specifieke applicatiegebieden geselecteerd en werden hierin een aantal case studies uitgevoerd. Vervolgens werd tenslotte geprobeerd om een algemene benadering en gefinaliseerde standaardprocedures te definiëren.

Twee van de specifiek te bestuderen applicatiegebieden werden gekozen omdat het proces voor de ontwikkeling van richtlijnen op Europees- en Wereldniveau bezig is. Deze applicatiegebieden werden bestudeerd teneinde ook Belgische wetenschappers de gelegenheid te geven een prominente rol op dit terrein te spelen. De twee applicatiegebieden zijn (i) richtlijnen voor methodenvalidatie in farmacopee-analyses en (ii) de bioanalyse van geneesmiddelen voor farmacokinetische doeleinden.

Bioanalyse van geneesmiddelen werd, voor de aanvang van dit project, bestudeerd door een informeel comité met leden van o.a. de FDA, de Canadese tegenhanger ervan en leden van de industrie. Dit resulteerde in de zogenaamde 'Washington consensus guidelines'. De VUB (dienst prof. Massart) publiceerde, eveneens voor de aanvang van dit project, een kritiek op deze richtlijn en was eveneens aanwezig op de follow-up conferentie in Munchen. Gedurende dit project werden richtlijnen opgesteld voor de validatie van chromatografische methodes voor de dosage van geneesmiddelen in biologische milieus. Dit gebeurde in samenwerking met een commissie van de Franse SFSTP (Société Française de Sciences et Techniques Pharmaceutiques).

Vooraleer algemene richtlijnen voor methodenvalidatie gedefinieerd kunnen worden, moesten eerst een aantal specifieke problemen in detail bestudeerd worden. Het betreft hier validatie-criteria die gekend waren als knelpunten in de ontwikkeling van een methodenvalidatiestrategie, zoals de bepaling van de robuustheid en de detectielimiet van een methode.

Het eerste criterium dat aan een grondige studie onderworpen werd, is de bepaling van de robuustheid van een methode. Binnen dit project werden een aantal case studies uitgevoerd, werd een strategie vooropgesteld welke een statistische interpretatie van de resultaten toelaat evenals een strategie voor minimale evaluatie van de robuustheid, werd een softwareprogramma geschreven en werden richtlijnen opgesteld voor het opstellen en interpreteren van een robuustheidstest.

Het tweede validatiecriterium dat bestudeerd werd, is de detectielimiet. Twee belangrijke benaderingen zijn mogelijk : (i) de klassieke, gebaseerd op de meting van blancostalen, en (ii) deze die gebaseerd is op informatie van de calibratielijn.

Een situatie die zich bij het ontwikkelen en valideren van een methode frequent voordoet is dat men de resultaten van deze methode wil vergelijken met deze van een bestaande referentiemethode. Richtlijnen werden voorgesteld voor de vergelijking van twee analysemethoden, nl. voor de bias, de verschillende precisiecomponenten en de detectielimieten.

Proces analytische methoden brengen de meting dicht bij het proces of binnen het proces. De gebruikelijke benadering van het ontwikkelen van selectieve analysemethoden om de aanwezige substanties of de karakteristieken van een produkt te meten, levert dan vaak methoden op die te langzaam zijn. De chemisch selectieve methodes moeten dan vervangen worden door een aantal nieuwe methoden, die gebruik maken van minder selectieve informatie (een spectrum, een chromatogram waarin niet alle pieken zijn gescheiden of geïdentificeerd, een set van resultaten van sensoren) en essentieel een multivariaat karakter hebben.

Om informatie te extraheren uit de multivariate gegevens vertrouwt men op mathematische (chemometrische) methoden zoals evolving factor analyse, partial least squares en vele andere. De introductie van deze methoden wordt echter bemoeilijkt omdat er problemen zijn met de acceptatie ervan als standaardmethode. Dit komt omdat men pas recent begonnen is met het onderzoeken van de beperkingen van deze methodes. De grootste zorg is hun betrouwbaarheid. Bijvoorbeeld, in multivariate calibratie is het in essentie de robuustheid van de calibratiemodellen en soms hun transfer die aandacht vereist. Het was de bedoeling om binnen dit project pre-normatief onderzoek uit te voeren met betrekking tot de manier waarop dit type van methoden gevalideerd moet worden. Uiteindelijk werd de problematiek uitgebreid onderzocht en leidde dit tot de formulering van een aantal richtlijnen terzake.

5.3. Resultaten

a) Robuustheid (robustness/ruggedness) van analytische methodes

I.v.m. de studie van de robuustheid werd een filosofie opgesteld wat betreft de keuze van de experimentele designs en de statistische interpretatie van de effecten. Ze werd eveneens geïmplementeerd in een computerprogramma. Deze strategie was gebaseerd op een aantal case studies waarin voornamelijk procedure-gebonden factoren onderzocht werden. De mogelijkheid om tot chemisch relevante conclusies te kunnen i.p.v. tot statistisch significante werd voorgesteld. Een algemene richtlijn wordt eveneens gegeven voor de keuze van de extreme niveaus in een robuustheidstest. Een voorstel werd gedaan voor de definiëring van system suitability limieten uit de resultaten van een robuustheidstest.

Het gebruik van nested designs voor de evaluatie van niet-procedure-gebonden factoren werd eveneens bestudeerd.

Minimale designs werden aan de oorspronkelijk gedefinieerde strategie toegevoegd. Deze laten wel het berekenen van effecten toe maar niet de oorspronkelijke statistische interpretatie. Een aantal onder hen kunnen na een eerste evaluatie uitgebreid worden tot de designs uit de oorspronkelijke strategie. Voor deze minimale designs werd getracht een significatie aan te geven voor de berekende effecten m.b.v. randomisatietesten. Uiteindelijk werden in de minimale designs de significante effecten geschat, gebruik makend van de gegevens van de niet-significante (algorithme van Dong).

Het gebruik van supersaturated designs als minimale designs in een robuustheidstest werd onderzocht. De gebruikte supersaturated designs werden geconstrueerd uit Plackett-Burman designs en voor een aantal case studies werd de variantie van de supersaturated designs vergeleken met deze van de Plackett-Burman designs (de berekening van effecten is nl. niet meer mogelijk). De verkregen varianties voor de supersaturated designs waren over het algemeen niet verschillend van deze van de Plackett-Burman designs wat erop wijst dat de supersaturated designs toelaten de variantie veroorzaakt door de factoren te schatten. In een volgende stap diende bepaald te worden of de variantie gemeten voor een supersaturated design, veroorzaakt wordt door een al dan niet robuuste methode, m.a.w. een statistische test is vereist. Vooraleer daar echter mee te starten werd er nagegaan of er statistisch aanvaardbare resultaten gevonden kunnen worden. Dit blijkt echter niet het geval te zijn door

het geringe aantal vrijheidsgraden te wijten aan het lage aantal experimenten. Afhankelijk van de gebruikte hypothese zal men ofwel een grote probabiteit hebben om niet-robuste methodes als robuust te beschouwen, ofwel robuuste methodes te verwerpen als niet-robust.

Uit de opgedane kennis werden er tenslotte richtlijnen gedefinieerd voor het opstellen en interpreteren van een robuustheidstest welke de definitie voor robuustheid van de ICH volgen. Deze richtlijnen bevatten eveneens de mogelijkheid om system suitability limieten te bepalen. Dit laatste wordt eveneens aanbevolen door de ICH. De verschillende stappen in de set-up en interpretatie van een robuustheidstest worden in de richtlijnen gegeven. De richtlijnen werden opgesteld in samenwerking met Unilever (Vlaardingen, Nederland), terwijl er een coöperatie bestond met Novartis (Basel, Zwitserland), Bayer (Dormagen, Duitsland) en Shell (Amsterdam, Nederland) tijdens het uitwerken van een aantal case studies.

Een onderwerp dat naar het einde van het project toe onderzocht werd, is het gebruik van zogenaamde asymmetrische designs in robuustheidstesten. Deze designs laten toe sommige factoren op meer dan twee niveaus te onderzoeken.

b) Validatie van chromatografische bioanalysemethodes, d.w.z. methodes die bepalingen uitvoeren in biologische milieus

In samenwerking met de SFSTP werd een strategie besproken voor de validatie van chromatografische methodes voor bepalingen in biologische milieus. Deze besprekingen leidden enerzijds tot de publicatie van een aantal richtlijnen i.v.m. validatie in biologische milieus en tot een aantal case studies ter evaluatie van de voorgestelde richtlijnen en anderzijds tot de ontwikkeling van een softwaresysteem voor methodevalidatie in deze biologische milieus.

c) Bepaling van de detectielimiet

Een aanvang werd genomen met het vergelijken van twee benaderingen voor de evaluatie van de detectielimiet : *i)* gebaseerd op de meting van blancostalen; *ii)* gebaseerd op de informatie van de calibratielijn. Ze werden toegepast op NIR metingen gebaseerd op univariate calibratie. Verdere evaluatie is echter noodzakelijk.

d) De vergelijking van alternatieve meetmethodes

Een aantal bestaande standaarden en richtlijnen, zoals ISO 5725, AFNOR T90-120 en Eurachem werden verzameld en bestudeerd met het oog op hun bruikbaarheid voor het vergelijken van analysemethodes. Het doel van deze studie is richtlijnen te hebben die toelaten aan te tonen dat *(i)* een nieuwe methode even goed of beter presteert dan een bestaande referentiemethode, of *(ii)* een bepaalde methode zonder problemen getransfereerd kan worden van een researchcentrum naar een industrieel labo of van het ene naar het andere industrieel labo.

Twee richtlijnen werden voorgesteld voor het vergelijken van analysemethoden: de eerste gebaseerd op een bepaling van het aantal metingen nodig om de β -fout tot een gespecificeerde

waarde te beperken, de tweede vertrekkend van een door de gebruiker gespecificeerd aantal metingen en een evaluatie van de β -fout. Daarbij werd rekening gehouden met opmerkingen gemaakt door o.a. Elf en Novartis. Het was daarbij noodzakelijk om in de eerste richtlijn een beslissing te nemen i.v.m. de aanvaardbare β -fout (hier de probabiliteit dat een relevante bias niet gedetecteerd wordt of een gespecificeerde precisie niet gehaald wordt) voor het berekenen van het aantal uit te voeren metingen.

Vergelijking van analysemethoden : de voorgestelde richtlijnen voor de vergelijking van twee analysemethoden, beperken zich tot een vergelijking van de bias en van verschillende precisiecomponenten. Nagegaan werd ook hoe de detectielimieten van beide methodes vergeleken konden worden. Dit is uiteraard belangrijk wanneer het methoden voor sporenanalyse betreft.

e) Methodenvalidatiedesigns die kostenefficiënt zijn : lineariteit ijklijn

Wat betreft de studie van methodevalidatie designs die kosten-efficiënt zijn, werden verschillende designs vergeleken voor de evaluatie van de lineariteit van de ijklijn. Een evaluatie van lineariteitstesten (Anova lack-of-fit test en testen van de significantie van de kwadratische coëfficiënt) werd uitgevoerd aan de hand van gesimuleerde niet-lineaire calibratielijnen die gebaseerd zijn op een beperkt aantal metingen ($6 < n \leq 10$). Verschillende designs waarin de metingen verdeeld zijn over 3, 4 of 5 concentratieniveaus werden daarbij bestudeerd. De waarschijnlijkheid dat een lack-of-fit niet gedetecteerd wordt (= de β -fout), is het kleinst wanneer slechts 3 concentratieniveaus beschouwd worden. Heteroscedastiteit, die gezien het gering aantal metingen, niet aan de hand van de experimentele gegevens kan onderzocht worden, geeft aanleiding tot een aanzienlijke toename van de β -fout. Vermits de β -fout ook afhankelijk is van de precisie werd de invloed van deze laatste ook nagegaan. Dit is belangrijk omdat dit toelaat te bepalen welke precisie noodzakelijk is om een relevante afwijking van de lineariteit te detecteren. In praktijk is het echter moeilijk te specificeren welke afwijking van lineariteit als belangrijk beschouwd wordt. Vermits een lack-of-fit echter aanleiding geeft tot fouten in de voorspelde concentraties (en dus tot een bias) wordt nu de β -fout van de lineariteitstesten in functie van de predictiefout bestudeerd.

f) Bepaling van de reproduceerbaarheid van een methode en methodenvalidatie in farmacopee-analyses

Een van de applicatiegebieden waarbinnen het uittesten van richtlijnen en normen gepland waren, is de methodenvalidatie in farmacopee-analyses. De bestaande voorschriften en richtlijnen, moeten daarom eerst toegepast worden op enkele case studies. Op punt stellen van nieuwe of verbeterde methoden blijft eveneens noodzakelijk. Na toetsing werd gekozen om de analysemethoden voor enkele belangrijke antibiotica als model te nemen.

Interlaboratoriumstudies werden opgesteld op twee analysemethoden (farmacopee methodes) voor de bepaling van het antibioticum tylosine. Deze studies werden opgesteld en de resultaten behandeld volgens de ISO-normen terzake (ISO 5725-2). De eigenlijke interlaboratoriumexperimenten werden voorafgegaan door een trainingsronde. De eigenlijke

gegevens werden eerst getest op outliers (zowel outliers op het niveau van de binnen-laboratoriumvariantie als van de tussen-laboratoriumvariantie). Na verwijderen ervan wordt de repeeteerbaarheid en de reproduceerbaarheid van de methode berekend en wordt er nagegaan of er een verband bestaat (lineair of logaritmisch) tussen de variantie en de concentratie te bepalen bestanddeel (homo/heteroscedasticiteit).

Wat de andere substanties (spiramycine, bacitracine, amoxicilline) betreft, werd zowel methodenoptimalisatie als methodenvalidatie uitgevoerd.

g) Robuuste procesanalyse met multivariate calibratie

Wat de ontwikkeling van een filosofie voor procesanalyse betreft, werden twee problemen bestudeerd. Een eerste was het nagaan of de gegevens afwijkende metingen bevatten. Als oplossing van dit probleem werd een strategie vooropgesteld, voor NIR metingen, gebaseerd op het meten van leverage punten en op een evolutieprogramma, lijkend op een genetisch algoritme.

Het tweede probleem was het bepalen van de factoren welke een invloed hebben op de precisie. Hiervoor werd gevonden dat een aantal duplicate of replicate metingen van een beperkt aantal standaarden een even goed model en een even goede precisie oplevert als een groot aantal standaarden die slechts éénmaal gemeten worden.

In het laatste deel van het project werkte de VUB partner eveneens mee aan het opstellen van validatierichtlijnen voor Nabij Infrarood (NIR) spectrofotometrie. Deze richtlijnen worden opgesteld op Europees niveau en zijn bedoeld voor toepassing in de industrie. Andere partners in dit project zijn o.a. de multinationals Pfizer, Shell en Elf.

Veel methoden voor multivariate calibratie zijn in de literatuur voorgesteld. Het bleek nu dat meerdere van deze methoden een analoge performantie hebben. Om verwarring te vermijden, stellen we voor dat slecht de volgende methodes beschouwd zullen worden: Multiple linear regression (MLR) : multipole lineaire regressie, Principal component regression (PCR) : principale componentenregressie, Partial least squares (PLS), Neural networks (NN) : neurale netwerken, Locally weighted regression (LWR) : lokaal gewogen regressie, Radial basis functions combined with PLS (RBF-PLS) : radial basis functies gecombineerd met PLS. Ze werden opgenomen op basis van theoretische overwegingen, die bevestigd werden door onderlinge vergelijking van hun performantie uitgevoerd op NIR data sets met verschillende data structuren. Op basis van de opgedane kennis kon een beslissingsschema opgesteld worden.

5.4. Besluiten

Methodenvalidatie is noodzakelijk voor de industrie, maar ook duur. Om die reden werd er bij het opstellen van bepaalde normen en richtlijnen, eveneens getracht om volgende zaken te definiëren : (i) methodenvalidatiedesigns die kostenefficiënt zijn, en (ii) minimale vereisten voor methodenvalidatie.

Er dient opgemerkt te worden dat in dit project voor een aantal methodevalidatieonderdelen wel rekening gehouden werd met het opstellen van kostenefficiënte designs (b.v. het nagaan

van de lineariteit van een ijklijn) en met minimale vereisten (b.v. gebruik van minimale designs in robuustheidstesten), maar dat dit niet op een veralgemeende basis het geval is. In de toekomst is er dus nog zeker werk weggelegd voor het ontwikkelen van verdere minimale en kostenefficiënte methodenvalidatieonderdelen, zodat men uiteindelijk tot een situatie komt waarbij men validatierichtlijnen heeft zowel voor een volledige als voor een minimale validatie.

Alhoewel de gebruikelijke statistische pakketten kunnen helpen in methodenvalidatie, is er geen enkel beschikbaar dat voldoende algemeen is en expliciet gericht op methodenvalidatie. De mogelijkheden van de bestaande methodenvalidatieprogramma's daarentegen zijn echter sterk gelimiteerd in vergelijking tot het hier onderzochte aantal onderwerpen. Inderdaad moet men rekening houden met de chemische context, die testen en tools vereisen die niet beschikbaar zijn in de meeste klassieke statistische pakketten. Omdat statistische conclusies vaak beperkt zijn tot fouten van de eerste soort is het niet mogelijk te besluiten hoeveel experimenten men moet uitvoeren om twee methoden of twee laboratoria te vergelijken, zoals vaak vereist is in methodenvalidatie.

Om die reden is het nodig specifieke software te produceren. Er werden programma's ontwikkeld voor de validatie van bioanalytische methodes, voor het opstellen en interpreteren van robuustheidstesten, voor het vergelijken van methodes en voor multivariate calibratie. Omdat het maken van professionele, gebruiksvriendelijke en gevalideerde software moeilijk en kostbaar is, werd geopteerd voor het ontwikkelen van demonstratiesoftware.

Het is de bedoeling om in de toekomst enerzijds de door ons opgestelde richtlijnen te laten officialiseren en anderzijds om ze te promoten via verspreiding langs het Internet. De richtlijnen opgesteld i.v.m. robuustheidstesten zullen b.v. voorgelegd worden aan de ICH teneinde ze te laten officialiseren terwijl andere aan ISO zullen voorgesteld worden. De opgestelde richtlijnen zullen in de toekomst eveneens beschikbaar gesteld worden door ze op het Internet te publiceren, in combinatie met tutorials, praktische voorbeelden en de eventueel reeds beschikbare software. Dit moet de verspreiding van de in het project opgestelde richtlijnen bevorderen.

Gedetailleerde resultaten

Gedetailleerde resultaten

In dit deel zijn de gedetailleerde resultaten van het project opgenomen. Het betreft ofwel de geformuleerde richtlijnen en normen, ofwel de meest relevante publicaties, ofwel representatieve samenvattingen van het werk uitgevoerd binnen een bepaald labo dat deel uitmaakte van dit project.

Robuustheid
(robustness/ruggedness)
van analytische methodes

6. Bijlagen

a) Referentielijst

Gezien de uitgebreidheid van de onderzoeksgebieden, de resultaten en de geraadpleegde referenties lijkt het ons niet opportuun hier een opsomming van de gebruikte referenties te geven. Ze kunnen op een veel praktischer manier teruggevonden worden in de referentielijsten van de verdere bijlagen.

b) Lijst van publicaties voortvloeiend uit het onderzoek

Robuustheid (robustness/ruggedness) van analytische methodes

- [1] Y. Vander Heyden, F. Questier and D.L. Massart, A ruggedness test strategy for procedure related factors : experimental set-up and interpretation; **Journal of Pharmaceutical and Biomedical Analysis**, 17 (1998) 153-168.
- [2] F. Questier, Y. Vander Heyden and D.L. Massart; RTS, a computer program for the experimental set-up and interpretation of ruggedness tests; **Journal of Pharmaceutical and Biomedical Analysis** 18 (1998) 287-303.
- [3] Y. Vander Heyden and D.L. Massart; Y. Zhu and J. Hoogmartens; J. De Beer; Ruggedness tests on the HPLC assay of the United States Pharmacopeia XXIII for tetracycline hydrochloride : comparison of different columns in an interlaboratory approach; **Journal of Pharmaceutical and Biomedical Analysis** 14 (1996) 1313-1326.
- [4] Y. Vander Heyden, A. Bourgeois and D.L. Massart, Influence of the sequence of experiments in a ruggedness test when drift occurs; **Analytica Chimica Acta** 347 (1997) 369-384.
- [5] Y. Vander Heyden, K. De Braekeleer, Y. Zhu, J. Hoogmartens, J. De Beer and D.L. Massart, Nested designs in ruggedness testing, aanvaard voor publicatie in **Journal of Pharmaceutical and Biomedical Analysis**
- [6] Y. Vander Heyden, S. Kuttatharmmakul, J. Smeyers-Verbeke and D.L. Massart; Supersaturated designs for robustness testing; in preparation
- [7] Y. Vander Heyden, C. Hartmann and D.L. Massart; P. Nuyten, A.M. Hollands and P. Schoenmakers; Ruggedness testing of a size exclusion chromatographic assay for low molecular mass polymers; **Journal of Chromatography A**, 756 (1996) 89-106.
- [8] Y. Vander Heyden, F. Questier and D.L. Massart; Ruggedness testing of chromatographic methods : selection of factors and levels; **Journal of Pharmaceutical and Biomedical Analysis** 18 (1998) 43-56.

- [9] Y. Vander Heyden, M. Jimidar, E. Hund, N. Niemeijer, R. Peeters, J. Smeyers-Verbeke, D.L. Massart and J. Hoogmartens; Determination of system suitability limits with a robustness test, aanvaard voor publicatie in **J. of Chromatography A**.
- [10] E. Hund, Y. Vander Heyden, M. Haustein, D.L. Massart and J. Smeyers-Verbeke, Comparison of several criteria to decide on the significance of effects in a robustness test with an asymmetrical factorial design, **document in voorbereiding**.
- [11] E. Hund, Y. Vander Heyden, M. Haustein, D.L. Massart and J. Smeyers-Verbeke, Application of fractional factorial and asymmetrical factorial designs in the robustness testing of a RP-HPLC method, **document in voorbereiding**.

Validatie van chromatografische bioanalysemethodes

- [12] S. Bervoas-Martin, B. Boulanger, P. Chevalier, P. Chiap, D. Grandjean, P. Hubert, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, E. Chapuzet et N. Mercier; Méthodes chromatographiques de dosage dans les milieux biologiques : stratégie de validation; **S.T.P. Pharma Pratiques** 7, 3 (1997) 169-194.
- [13] S. Bervoas-Martin, B. Boulanger, P. Chevalier, P. Chiap, D. Grandjean, P. Hubert, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, E. Chapuzet et N. Mercier; Méthodes chromatographiques de dosage dans les milieux biologiques. Exemple d'application de la stratégie de validation; **S.T.P. Pharma Pratiques** 8, 2 (1998) 81-107
- [14] Ph. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie en J.C. Nivet; The SFSTP Guide on the validation of chromatographic methods for drug bioanalysis: from the Washington Conference to the laboratory, **ingediend voor publicatie**
- [15] P. Chiap, Ph. Hubert, B. Boulanger, J. Crommen, Validation of an automated method for the liquid chromatographic determination of atenolol in plasma: application of a new validation protocol, **ingediend voor publicatie**
- [16] B. Boulanger, Ph. Hubert, P. Chiap, J. Crommen; An analysis of the SFSTP guide on validation of chromatographic bioanalytical methods: progress and limitations, **ingediend voor publicatie**

De vergelijking van alternatieve meetmethodes

- [17] S. Kuttatharmakul, D.L. Massart and J. Smeyers-Verbeke, Comparison of alternative measurement methods, accepted for publication in **Analytica Chimica Acta**.

Methodenvalidatiedesigns die kostenefficiënt zijn : lineariteit ijklijn

- [18] S. Kuttatharmakul, D.L. Massart and J. Smeyers-Verbeke, Influence of precision, sample size and design on the beta error of linearity tests; **Journal of Analytical Atomic Spectroscopy (JAAS)** 13 (1998) 109-118.

Bepaling van de reproduceerbaarheid van een methode en methodenvalidatie in farmacopee-analyses

- [19] Y. Vander Heyden, J. Saevels, E. Roets, J. Hoogmartens, D. Decolin, M.G. Quaglia, W. Van den Bossche, R. Leemans, O. Smeets, F. Van de Vaart, B. Mason, G.C. Taylor, W. Underberg, A. Bult, P. Chiap, J. Crommen, J. De Beer, S.H. Hansen and D.L. Massart, Interlaboratory studies on two HPLC assays for tylosin (tartrate), **Journal of Chromatography A**, 830 (1999) 3-28.
- [20] Y.M. Li, A. Van Schepdael, Y. Zhu, E. Roets, J. Hoogmartens; Development and validation of amoxicillin determination by micellar electrokinetic capillary chromatography, **Journal of Chromatography A**, 812 (1998) 227-236.
- [21] L. Liu, J. Saevels, P. Louis, H. Nelis, S. Rico, K. Dierick, S. Guyomard, E. Roets, J. Hoogmartens; Interlaboratory study comparing the microbiological potency of spiramycins I, II and III, in druk in **Journal of Pharmaceutical and Biomedical Analysis**.

Robuuste procesanalyse met multivariate calibratie

- [22] V. Centner, D.L. Massart, O.E. de Noord, "Detection of inhomogeneities in sets of NIR spectra", *Anal. Chim. Acta*, **330** (1996) 1-17.
- [23] V. Centner, O.E. Noord, D.L. Massart, "Detection of nonlinearities in multivariate calibration", *Anal. Chim. Acta*, **376** (1998) 153-168.
- [24] D. Jouan-Rimbaud, D.L. Massart, C.A. Saby, C. Puel, "Characterisation of the representativity of selected sets of samples in multivariate calibration and pattern recognition", *Anal. Chim. Acta*, **350** (1997) 149-161.
- [25] D. Jouan-Rimbaud, D.L. Massart, C.A. Saby, C. Puel, "Determination of the representativity between two multidimensional data sets by a comparison of their structure", *Chemom. Intelligent Lab. Syst.*, **40** (1998) 129-144.
- [26] D. Jouan-Rimbaud, E. Bouveresse, D.L. Massart, O.E. de Noord, "Detection of prediction outliers and inliers in multivariate calibration", *Anal. Chim. Acta*, in press.
- [27] V. Centner, D.L. Massart, S. de Jong, "Inverse calibration predicts better than classical calibration", *Fresenius' J. Anal. Chem.*, **361** (1998) 2-9.
- [28] V. Centner, D.L. Massart, S. de Jong, "Accuracy of prediction from multivariate calibration models in the errors-in-variables situation", in preparation.

- [29] F. Despagne, D.L. Massart, "Variable selection for neural networks in multivariate calibration", *Chemom. Intell. Lab. Syst.*, **40** (1998) 145-163.
- [30] V. Centner, D. L. Massart, "Optimisation in locally weighted regression", *Anal. Chem.*, in press.
- [31] D. Jouan-Rimbaud, D.L. Massart, O.E. de Noord, "Random correlation in variable selection for multivariate calibration with a genetic algorithm", *Chemom. Intell. Lab. Syst.*, **35** (1996) 213-220..
- [32] F. Despagne, D.L. Massart, O.E. de Noord, "Optimization of PLS calibration models complexity by simulation of instrumental perturbations", *Anal. Chem.*, **69** (1998) 3391-3399.
- [33] V. Centner, D.L. Massart, O.E. de Noord, S. de Jong, B.M. Vandeginste, C. Sterna, "Elimination of uninformative variables for multivariate calibration", *Anal. Chem.*, **68** (1996) 3851-3858.
- [34] D. Jouan-Rimbaud, B. Walczak, R.J. Poppi, O. E. de Noord, D.L. Massart, "Application of wavelet transform to extract the relevant component from spectra data for multivariate calibration", *Anal. Chem.*, **69** (1997) 4317-4323.
- [35] L. Pasti, D. Jouan-Rimbaud, D.L. Massart, O.E. de Noord, "Application of Fourier transform to multivariate calibration of Near-Infrared data", *Anal. Chim. Acta*, **364** (1998) 253-263.
- [36] F. Despagne, B. Walczak, D.L. Massart, "Transfer of calibrations of Near-Infrared spectra using neural networks", *Appl. Spectrosc.*, **52** (1998) 732-745.
- [37] F. Despagne, D.L. Massart, M. Jansen, H.van Daalen, "Intersite transfer of industrial calibration models", in preparation.
- [38] F. Despagne, D.L. Massart, "Tutorial review: neural networks in multivariate calibration", *The Analyst*, **123** (1998) 157R-178R.
- [39] V. Centner, J. Verdú-Andrés, B. Walczak, D. Jouan-Rimbaud, F. Despagne, L. Pasti, R. Poppi, D.L. Massart, O.E. de Noord, "A comparison of multivariate calibration techniques applied to experimental NIR data sets", submitted.
- [40] F. Estienne, L. Pasti, V. Centner, B. Walczak, F. Despagne, D. Jouan Rimbaud, D. L. Massart, O. E. de Noord, "A comparison of multivariate calibration techniques applied to experimental NIR data sets. Part II: evaluation of the predictive ability of the models in extrapolation conditions", in preparation.
- [41] F. Despagne, F. Estienne, L. Pasti, B. Walczak, D. L. Massart, O. E. de Noord, "A comparison of multivariate calibration techniques applied to experimental NIR data sets. Part III: evaluation of the predictive ability of the models in the presence of instrumental perturbations in the test set", in preparation.

c) Gedetailleerde resultaten

cf. verder.

d) Overzicht voornaamste normen waarnaar verwezen wordt

- [1] International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use, Draft Consensus Text, Validation on Analytical Procedures, Released for Consultation, 26 October 1993, p. 1-7.
- [2] International Organisation for Standardisation (ISO); Accuracy (trueness and precision) of measurement methods and results - Part 3 : Intermediate measures of the precision of a standard measurement method; International Standard ISO 5725-3:1994(E), First edition.
- [3] International Organisation for Standardisation (ISO); Accuracy (trueness and precision) of measurement methods and results - Part 2 : Basic method for the determination of repeatability and reproducibility of a standard measurement method; *International Standard* ISO 5725-2:1994(E), First edition.
- [4] International Organisation for Standardisation (ISO); Accuracy (trueness and precision) of measurement methods and results - Part 1; International Standard ISO 5725-1:1994(E), First edition.
- [5] Eurachem, A focus for Analytical Chemistry in Europe; Quantifying Uncertainty in Analytical Measurement, version 5; Eurachem Workshop Draft, September 1994.
- [6] ICH Harmonised Tripartite Guideline prepared within the Third International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), Text on Validation of Analytical Procedures, 1994, (<http://www.ifpma.org/ich1.html>).
- [7] ICH Harmonised Tripartite Guideline prepared within the Third International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), Validation of Analytical Procedures : Methodology, 1996, 1-8 (<http://www.ifpma.org/ich1.html>).
- [8] *Drugs Directorate Guidelines, Acceptable Methods*; Health Protection Branch - Health and Welfare Canada; 1992; 20-22.

- [9] Eurachem, A focus for Analytical Chemistry in Europe; *Quantifying Uncertainty in Analytical Measurement*, First Edition 1995.
- [10] Current concepts for the validation of compendial assays – Pharmacopeial Forum, The United States Pharmacoeia Inc., Rockville 1986, p.1241
- [11] Validation of compendial methods – Guidelines, Pharmacopeial Forum, The United States Pharmacoeia Inc., Rockville 1986, p.4129.
- [12] Note explicative III/844/87-FR FINAL – Commission des Communautés européennes, Groupe de travail du Comité des spécialités pharmaceutiques, août 1989.
- [13] International Standard, Accuracy (Trueness and Precision) of Measurement methods and results, ISO 5725-6, Geneva, 1994.
- [14] International Standard, Statistics (Vocabulary and Symbols): Design of experiments , ISO 5725-3, Geneva, 1985.
- [15] An American National Standard ASTM, E 1655-94, Standard Practices for Infrared, Multivariate, Quantitative Analysis (1994).



**Validatie van
chromatografische
bioanalysemethodes**

Scientific Impulse program Normalisation (Part II)**Contribution of the Department of Pharmaceutical Analytical Chemistry
(University of Liège)**

The objective of the project was to develop generally accepted and standardised approaches as well as guidelines for the validation of measurement methods in chemical laboratories. Among these measurement methods, the department of Pharmaceutical Analytical Chemistry at the University of Liège focused mainly on the validation of methods for the determination of drugs and their metabolites in biological fluids and particularly in plasma. The validation of such methods is necessary in order to ensure their reliability and hence the confidence that can be placed on the results generated in bioavailability, bioequivalence and pharmacokinetic studies. Very often important decisions are taken which are based on the interpretation of data obtained from bioanalytical results.

For about ten years, several articles have been published about the validation of bioanalytical methods in the international literature. Even if the currently available documents have contributed to significant improvements in the validation of analytical methods in biological matrices, they are still too general. The Washington Conference report which is now being utilised as a basis for bioanalytical method validation gives the minimum requirements (e.g. at least 5 concentration levels covering the range of the expected concentrations, the minimum number of experiments to be performed at each concentration level, the acceptance criteria of 15% and 20% for precision and accuracy, etc.), but does not provide any validation strategy. Under these conditions, the validation work performed in bioanalysis depends mainly on the analyst's experience, on his personal point of view and on the laboratory strategy.

Several methods were developed for the determination of drugs and their metabolites in plasma using solid phase extraction on disposable cartridges in combination to liquid chromatography (LC). The LC determination of benzodiazepines (diazepam, nordazepam and camazepam), acetylsalicylic acid used as anti-thrombotic drug and two of its metabolites, salicylic and salicyluric acids as well as molsidomine, used as a prodrug in the treatment of angor and its active metabolite, 3-morpholino-sydnonimine or SIN-1 can be quoted as an example. All these methods were then validated according to the Washington Conference report.

Although this document can assist the analysts in validating their bioanalytical procedures, it does not provide neither experimental design nor validation protocol. Therefore, the "Société Française des Sciences et Techniques Pharmaceutiques (SFSTP)" created in 1995 a Commission in the bioanalytical validation field. An experimental and statistical approach was proposed to obtain good estimates of the validation parameters in accordance with the acceptance criteria defined in the Washington Conference report. This validation guide illustrates the consensus between both experimental and statistical points of view and describes the recommended strategy, allowing the analyst to perform the most profitable experimental work possible [1-3]. This is achieved by drawing maximum

information from the results and in routine use by reducing the re-analysis risk due to a lack of agreement with the validation criteria.

Taking into account the statistical requirements, the original strategy proposed by the SFSTP for the validation of a bioanalytical method comprises two steps : a pre-validation phase and a validation phase. The different experiments proposed in both steps were designed and optimised in order to minimise the risks to accept a method which is effectively not valid or to reject a method even though it is actually valid.

The aim of the prevalidation step is to prepare the validation step. Consequently, before starting the experiments, the configuration of the equipment, the preparation of the stock solution and its dilutions, the validation standards being the QC samples used during the validation step must be well defined in the validation protocol. The prevalidation step allows the analyst to identify the response function of the calibration curve by means of a decision tree (linear, non-linear, mathematical transformation, weighting). For the analysis of the response function, it is foreseen an adjustment of the responses on the mean concentrations, a visual analysis of the residuals allowing the detection of potential aberrant values or a lack of linearity, a statistical analysis in order to test the hypothesis of homoscedacity (Levene's and Cochran's tests) and to confirm the adequacy of the regression model. The prevalidation step also allows the estimation of the limit of detection, the determination of the extraction efficiency and the optimal number of experiments to be performed in the validation step. The estimation of the limit(s) of quantitation and the calibration range is also possible by means of an accuracy profile. After the regrouping of the found concentrations for the calibration samples, intermediate precision and accuracy are estimated at each concentration level. The one-sided upper and lower confidence limits for a measure of accuracy (e.g. recovery) at the 95% level are then calculated by introducing the estimation of the standard deviation for the intermediate precision and plotted as a function of concentration in order to obtain an accuracy profile.

The objective of the validation step is to demonstrate the method selectivity by means of 6 independent sources of the same biological matrix, to determine precision (repeatability and intermediate precision) and accuracy at each concentration level of the validation standards, to confirm the limit(s) of quantitation and to verify the linearity of the relationship between the estimated concentrations for the validation standards and the introduced concentrations in order to estimate the global accuracy of the method. The experiments of the validation step should be performed on different days under conditions as close as possible to the routine ones. A method is considered acceptable with respect to accuracy if the one-sided lower and upper confidence limits of a measure of accuracy (e.g. recovery) at the 95% level are comprised between 80 % and 120 %. These confidence limits take into account the estimation of the standard deviation for intermediate precision. This approach is in conformity with the recommendations given in the Washington Conference report. Indeed, if the confidence limits are not greater than $\pm 20\%$, the estimates for accuracy are obviously smaller than $\pm 20\%$ at the limit of quantitation and $\pm 15\%$ for the other concentration levels. Moreover, it is not examined whether the 100% value is contained within the confidence limits. The reason for this is that when poorly precise

methods are used or when the sample size is too small, it is likely that unacceptable analytical methods may actually be accepted.

In order to perform the statistical calculations of the proposed protocol for the validation of a bioanalytical method, a software was elaborated by means of Microsoft Excel® for Windows. The analysis of the response function has been developed particularly and takes into account of the problems related to the heterogeneity of variances and the lack of fit for the selected model. For a bioanalytical method, the application of the simple regression model based on the least squares method may be inappropriate. In this case, some weighed regression models or transformations of data (square roots, natural logarithm, etc.) are proposed. This software (Excel® worksheets) was then checked by means of the JMP® statistical software (S.A.S.) (collaboration with Dr. B. Boulanger from E. Lilly, member of the Accompanying Committee).

In addition, in order to illustrate the experimental and statistical strategy proposed by the SFSTP Commission, several methods were developed for the LC determination in plasma of drugs and their metabolites (e.g. sotalol, atenolol used as β -blockers as well as albendazole, an antihelminthic agent, and its main metabolites (sulfoxide and sulfone derivatives)). These methods, using dialysis and trace enrichment as sample preparation technique, were then validated according to the validation guide. The validation results of the methods for the determination of atenolol and sotalol were presented in two articles [4,5]. The experimental design used as well as the results of statistical calculations were also described.

All the successful applications of this validation protocol were also presented in several international symposiums. Moreover, the validation guide was described at a study sitting of the SFSTP in Paris.

On the other hand, only the results obtained from the validation of the LC method for the determination of atenolol in plasma are given in this report. In the prevalidation step, the analysis of the response function was first performed over a concentration range from 20 ng/ml to 1000 ng/ml. A simple regression model (least squares method) was selected. However, a transformation of data (square root) was needed due to the variance heterogeneity. After the estimation of the concentrations obtained from calibration samples, a new calibration range was determined. Indeed, the one-sided confidence limits of the recovery at the 95 % level exceeded the norms of 80% and 120% at the concentration equivalent to 20 ng/ml. Consequently, this concentration level had to be eliminated from the calibration range and the analysis of the response function was again performed over a range from 25 ng/ml to 1000 ng/ml. The same regression model was selected in order to describe the relationship between the responses and the concentrations. The limit of quantitation was then estimated at 26 ng/ml by means of the accuracy profile. Two other validation criteria were also tested : the limit of detection was estimated at 10 ng/ml and the mean absolute recovery was calculated (about 65%). The validation step permitted the assessment of method precision and accuracy. Indeed, the confidence limits of the measures of accuracy did not exceed the acceptable norms irrespective of the concentration level of the validation standards. The limit of quantitation was also confirmed. Moreover, the linearity of the

relationship between the found concentrations of the validation standards and the introduced concentrations was verified. Finally, the method selectivity was demonstrated by means of 6 independent sources of the same biological matrix. No interfering peak from endogenous components of plasma was present at the retention times of atenolol and its internal standard, sotalol.

This validation guide has been himself validated in several real cases before being published and some practical applications provide the analyst on one hand a better understanding on the way to proceed and on the other hand real data for qualifying his own computations that he could perform using a commercial spreadsheet. It should be noted that this validation strategy is principally meant for chromatographic methods, but it could also be applied to other bioanalytical methods based on techniques such as capillary electrophoresis, fluorometry, atomic absorption spectrometry, etc. Moreover, even though the strategy was originally developed for pharmaceutical applications, it can reasonably be applied to other fields with similar specifications (environment, food products, etc.).

The SFSTP guide does not constitute a final end point, but on the contrary was thought as a large basis to pave the way for developments that are expected from readers and analysts that will practice the guide. On one hand, since the publication of the guide in 1997, some members of the SFSTP Commission already have some modifications or warnings to propose in order to initiate a continuous process of improvements. On the other hand, many choices and decisions that have been taken in this guide constitute major progresses compared to traditional ways to proceed.

1. E. Chapuzet, N. Mercier, S. Bervoas-Martin, B. Boulanger, P. Chevalier, P. Chiap, D. Granjean, P. Hubert, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie et J.C. Nivet, Méthodes chromatographiques de dosage dans les milieux biologiques : stratégie de validation. Rapport d'une Commission SFSTP, *S.T.P. Pharma Pratiques*, 7, 169-194 (1997).
2. Ph. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie and J.C. Nivet, The SFSTP Guide on the validation of chromatographic methods for drug bioanalysis : from the Washington Conference to the laboratory, *Analytica Chimica Acta*, 391, 135-148 (1999).
3. B. Boulanger, Ph. Hubert, P. Chiap and J. Crommen, An analysis of the SFSTP Guide on validation of chromatographic bioanalytical methods : progresses and limitations, *Analytica Chimica Acta*, accepté pour publication (Réf. ACAVAL-011) (1999).
4. E. Chapuzet, N. Mercier, S. Bervoas-Martin, B. Boulanger, P. Chevalier, P. Chiap, D. Granjean, P. Hubert, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie et J.C. Nivet, Méthodes chromatographiques de dosage dans les milieux biologiques : exemple d'application de la stratégie de validation. Rapport d'une Commission SFSTP, *S.T.P. Pharma Pratiques*, 8, 81-107 (1998).
5. P. Chiap, Ph. Hubert, B. Boulanger and J. Crommen, Validation of an automated method for the liquid chromatographic determination of atenolol in plasma : application of a new validation strategy, *Analytica Chimica Acta*, 391, 227-238 (1999).

Programme d'appui scientifique à la Normalisation (volet II)**Rapport final d'activité du Service d'Analyse des
Médicaments de l'Université de Liège****I. Objectifs**

Dans le cadre du programme d'appui scientifique à la normalisation (volet II), l'objectif principal de notre projet est de développer des normes et recommandations (notes de guidance) communément acceptées, tant au niveau national qu'international, pour la validation de procédures de mesure dans les laboratoires d'analyse. A l'instigation du coordinateur du projet, un consortium regroupant des compétences académiques et industrielles dans le domaine a été créé afin de permettre aux scientifiques et industriels belges de jouer un rôle prépondérant dans ce domaine. C'est dans cet esprit que le laboratoire d'Analyse des Médicaments de l'Université de Liège collaborera aux différents aspects de ce projet.

Parmi les procédures de mesure envisagées, nous nous intéresserons tout d'abord aux méthodes de dosage de médicaments et de leurs métabolites dans les fluides biologiques et en particulier dans le plasma. Nous participerons, de manière plus spécifique, à l'évaluation des recommandations (par exemple, celles issues de la conférence de Washington) existant dans le domaine de la validation des méthodes bioanalytiques utilisées dans le cadre des études de bio-équivalence, de biodisponibilité et de pharmacocinétique. A cette fin, nous mettrons au point différentes méthodes de dosage de médicaments et de leurs métabolites dans le plasma par association de l'extraction en phase solide comme technique de préparation de l'échantillon à la chromatographie liquide. Nous nous efforcerons ensuite de développer des notes de guidance adaptées à la validation de méthodes bioanalytiques. Notre participation aux travaux d'une Commission de la Société Française des Sciences et Techniques Pharmaceutiques (SFSTP) intitulée "Validation analytique dans les milieux biologiques" sera déterminante. En effet, l'objectif de cette commission est de faire le point sur la méthodologie utilisée par chacun de ses membres et de définir sur cette base et en fonction des recommandations internationales des lignes directrices devant aboutir à l'élaboration d'un protocole expérimental minimal. Ce guide de validation devra notamment tenir compte de certaines particularités liées aux milieux biologiques telles que l'étendue de la gamme de calibration, le choix du modèle de régression, l'établissement du ou des seuil(s) de quantification, l'évaluation de la sélectivité, la détermination de la précision et de l'exactitude de la méthode, etc. Le protocole expérimental proposé

par la SFSTP sera ensuite évalué au cours d'une session spéciale d'étude élargie au monde industriel français et européen.

L'étape suivante consistera à tester la version finale de cette stratégie de validation à l'aide d'exemples concrets. A cette fin, nous mettrons au point plusieurs méthodes de dosage de médicaments par chromatographie liquide, que nous validerons selon la stratégie proposée par la Commission. Au terme de cette étude, nous soumettrons ce guide de validation pour publication dans une revue scientifique afin de permettre sa diffusion au niveau international et présenterons le résultat de nos recherches lors de symposiums nationaux et internationaux.

D'autre part, nous nous préoccupons de la validation de méthodes développées pour l'analyse pharmaceutique. Lors de cette étude, nous aurons l'opportunité de tester les performances de deux logiciels d'aide à la validation analytique, basés sur la méthodologie proposée voici quelques années par une commission de la SFSTP. Nous pourrions ainsi tester cette stratégie de validation et formuler éventuellement quelques remarques.

Par ailleurs, nous nous intéresserons à la problématique de la robustesse des méthodes, critère important pour leur transfert vers d'autres laboratoires. A cette fin, nous participerons à une autre Commission de la SFSTP dont l'objectif sera de développer des notes de guidance pour la robustesse des méthodes analytiques. Avant d'envisager la robustesse en validation, nous nous focaliserons sur la robustesse lors du développement d'une méthode. C'est dans cet esprit que nous considérerons l'optimisation de méthodes analytiques à l'aide de plans d'expérience. Nous développerons une méthode de dosage d'un médicament et ses principaux métabolites dans le plasma par association de la dialyse et de l'enrichissement de traces à la chromatographie liquide. Lors de cette mise au point, nous tenterons d'appliquer la méthodologie des plans d'expérience pour le criblage et l'optimisation de facteurs influençant les processus de dialyse et d'enrichissement. Ensuite, nous testerons les performances d'un logiciel d'optimisation des séparations par chromatographie liquide, notamment lors de l'application d'une élution graduée.

Enfin, nous apporterons également notre contribution lors d'une étude collaborative dont l'objectif sera, d'une part, de déterminer la répétabilité et la reproductibilité de deux méthodes chromatographiques de dosage des différents composants d'un antibiotique et, d'autre part, de comparer les performances des deux méthodes proposées. Dans le cadre de cette collaboration avec nos partenaires de la VUB et de la KUL, nous pourrions étudier les problèmes de "transposabilité" des méthodes et de validation "inter-laboratoire".

II. Méthodologie

Nous allons à présent décrire la méthodologie que nous avons suivie afin de mener à bien les objectifs de notre projet avant d'envisager les résultats que nous avons obtenus dans le cadre du programme d'appui à la normalisation.

Nous nous sommes tout d'abord focalisés sur le développement et la validation de méthodes de dosage de médicaments et de leurs métabolites dans le plasma par chromatographie liquide. Ces méthodes d'analyse jouent un rôle primordial dans l'évaluation et l'interprétation des paramètres relatifs à la biodisponibilité, la bioéquivalence ou la pharmacocinétique d'un médicament. Il est essentiel d'utiliser des méthodes bien caractérisées et entièrement validées afin de générer des résultats crédibles, qui puissent être valablement interprétés. Depuis une dizaine d'années, plusieurs articles traitant de la validation des méthodes analytiques dans les milieux biologiques ont été publiés dans la littérature internationale. Même si ces documents contribuent à faire progresser de manière significative la validation dans le domaine de la bioanalyse, ils se limitent, le plus souvent, aux préceptes généraux et ne proposent pas d'approche expérimentale. Le travail effectué dans ce domaine dépend donc largement de l'expérience de l'analyste, de son jugement et de la stratégie du laboratoire d'analyse.

L'injection directe d'un échantillon biologique n'étant généralement pas possible en chromatographie liquide, nous avons envisagé un traitement préalable de l'échantillon et avons sélectionné, comme technique de préparation, une extraction en phase solide entièrement automatisée sur cartouches à usage unique. Cette automatisation s'avère nécessaire, lorsque le nombre d'échantillons à traiter est élevé, comme c'est souvent le cas lors d'études de biodisponibilité, de bioéquivalence ou de pharmacocinétique. Tous les échantillons ont été traités et analysés au moyen du système ASPEC (Automatic Sample Preparation with Extraction Cartridges), qui associe l'extraction sur cartouches à usage unique à la chromatographie liquide.

A l'aide de cette technique, nous avons développé plusieurs méthodes chromatographiques pour le dosage de médicaments et de leurs métabolites dans le plasma dans le cadre d'études de biodisponibilité et de bioéquivalence. A titre d'exemples, nous pouvons citer le dosage par chromatographie liquide de benzodiazépines (diazépam, nordazépam et camazépam) ou d'antiagrégants plaquettaires (l'acide acétylsalicylique et deux de ses métabolites, les acides salicylique et salicylurique) ou encore d'antiangoreux (la molsidomine et son principal métabolite actif, le 3-morpholino-sydnominine ou SIN-1) en milieu plasmatique. Ces méthodes ont ensuite été validées sur la base du rapport intitulé "Guidelines of the Washington Conference" qui reprend les conclusions d'une conférence organisée au début des années nonante à Washington afin de traiter de la validation des procédures bioanalytiques. Ce document pré-

sente une série de recommandations qu'il convient de suivre lors d'une validation d'une procédure bioanalytique et définit, outre les paramètres de validation, des exigences minimales. D'autres conférences traitant du même sujet ont ensuite été organisées à Barcelone et plus récemment à Munich sous le patronage de diverses organisations du monde pharmaceutique.

Bien que le document publié à l'issue de la Conférence de Washington puisse aider le bioanalyste à valider ses procédures de dosage, il n'apporte pas d'approche expérimentale, ni de plan standard de validation. C'est la raison pour laquelle, la Société Française des Sciences et Techniques Pharmaceutique (SFSTP) a mis en place en 1995 une commission à laquelle nous avons participé activement afin de rédiger un guide de validation tenant compte des particularités liées aux milieux biologiques. Le souci de la commission a été de rationaliser chaque étape et chaque prise de décision afin d'améliorer, à terme, la qualité sans accroître pour autant le coût global de la validation.

Ce guide de validation a été élaboré afin d'aider les industriels à valider leurs méthodes de dosage dans les milieux biologiques. Il représente un consensus de plusieurs professionnels de la bioanalyse et de la biostatistique. Se basant sur les normes communément admises, il permet de rentabiliser les essais effectués et de tirer le maximum d'informations des résultats obtenus afin d'éviter qu'une méthode soit déclarée fiable alors qu'elle ne l'est pas et inversement. Tout au long de nos réflexions, nous nous sommes efforcés de concevoir une stratégie de validation rationnelle, optimisée, pratique et statistiquement crédible.

Ensuite, afin de tester cette nouvelle stratégie de validation, nous avons mis au point des méthodes entièrement automatisées pour le dosage en milieu plasmatique de β -bloquants, le sotalol et l'aténolol ainsi que d'un agent thérapeutique à activités anthelminthiques, l'albendazole et ses principaux métabolites (les dérivés sulfoxyde et sulfone). Ces méthodes originales sont basées sur l'association de la dialyse comme technique de préparation de l'échantillon à la chromatographie liquide couplée à une détection par spectrophotométrie d'absorption dans l'ultra violet. La dialyse est basée sur la diffusion passive de molécules à travers une membrane de porosité sélective, caractérisée par un seuil de coupure (cut-off) déterminé et constituant une barrière entre deux liquides. Le gradient de concentration qui s'établit entre la solution échantillon (liquide donneur) et le liquide de dialyse (liquide accepteur) constitue la seule force de passage des molécules à travers la membrane jusqu'à ce qu'un équilibre soit atteint. Seules les molécules présentant une taille inférieure au seuil de coupure de la membrane pourront être dialysées. Dans le cas des échantillons plasmatiques où l'analyte se trouve en partie sous forme liée aux protéines, seule la fraction libre peut diffuser à travers la membrane. Afin de compenser la dilution des composés à analyser qui est inévitable lors de tout processus de dialyse et d'apporter une certaine sélectivité chimique dont est dépourvu ce mode de traitement de l'échantillon, un enrichissement du dialysat préalable à l'analyse

chromatographique est effectué à l'aide soit d'une pré-colonne échangeuse de cations (contenant des groupements sulfonate greffés sur des phases silice ou polymérique), soit d'une pré-colonne renfermant un support de silice modifiée chimiquement au moyen de groupements octadécyle. Toutes les expériences ont été réalisées au moyen d'un système entièrement automatisé permettant d'effectuer une dialyse à haut rendement couplée à la chromatographie liquide. Il porte le nom de ASTED (Automated Sequential Trace Enrichment of Dialysates). Ces différentes méthodes ont ensuite été validées sur la base de la stratégie à la fois expérimentale et statistique proposée par la Commission de la SFSTP.

Afin de pouvoir réaliser l'ensemble des calculs statistiques prévus dans le protocole proposé pour la validation d'une procédure bioanalytique, nous avons conçu un logiciel à l'aide du programme Microsoft Excel[®] pour Windows. Les calculs statistiques décrits dans ce logiciel ont ensuite été vérifiés à l'aide du programme statistique JMP[®] (S.A.S.). Cette étape a été réalisée en étroite collaboration avec Monsieur Bruno Boulanger, Statisticien au sein de la Société Pharmaceutique E. Lilly et membre du Comité d'Accompagnement. Nous avons également testé notre logiciel lors du traitement statistique de données issues de la validation de plusieurs méthodes de dosage de médicaments dans les fluides biologiques proposées par la Société Pharmaceutique Innothera. L'exploitation statistique des résultats obtenus nous a permis de valider ces méthodes de dosage.

Par ailleurs, le protocole expérimental proposé par la SFSTP a été évalué au cours d'une session d'étude organisée par la SFSTP à Paris. Au cours de cette réunion présidée par le Professeur Massart et élargie au monde industriel français et européen, plusieurs membres de la Commission " Validation analytique dans les milieux biologiques " ont présenté cette démarche de validation.

De plus, afin d'assurer une diffusion de nos travaux, nous avons publié deux articles dans la revue scientifique S.T.P. Pharma Pratiques. Le premier est intitulé " Méthodes chromatographiques de dosage dans les milieux biologiques : stratégie de validation. Rapport d'une Commission SFSTP " et décrit le protocole expérimental caractéristique de cette stratégie de validation ainsi que l'exploitation statistique des données obtenues. Quant au second article publié sous le titre " Méthodes chromatographiques de dosage dans les milieux biologiques : exemple d'application de la stratégie de validation. Rapport d'une Commission SFSTP ", il présente un exemple d'application de cette stratégie développé dans le cadre de la validation d'une méthode originale de dosage du sotalol en milieu plasmatique par chromatographie liquide. Trois autres articles vont également être publiés très prochainement dans la revue scientifique " Analytica Chimica Acta ". Leur parution dans un numéro spécial de cette revue consacré à la validation assurera la diffusion de nos travaux au niveau international. Le premier article présente une description de la nouvelle stratégie de validation des méthodes

bioanalytiques et insiste sur l'analyse de la fonction de réponse et la détermination du ou des seuil(s) de quantification. Le deuxième article propose une analyse essentiellement statistique de ce guide de validation et indique les progrès apportés par ce guide dans le domaine de la validation en bioanalyse ainsi que ses limitations. Quant au troisième article, il décrit un exemple d'application de cette stratégie et présente les résultats obtenus lors de la validation d'une méthode, que nous avons développée au sein de notre laboratoire, pour le dosage de l'aténolol dans le plasma. Nous avons également pu présenter cette nouvelle stratégie de validation ainsi que les exemples d'application lors de symposiums nationaux et internationaux.

Conformément aux autres objectifs de notre projet, nous nous sommes ensuite intéressés à la validation de méthodes développées pour l'analyse pharmaceutique. A cette fin, nous avons mis au point et validé des méthodes chromatographiques pour le dosage d'un anti-inflammatoire non stéroïdien, l'ibuprofène et de son principal produit de dégradation, le 4-isobutylacétophénone dans des lidoses ainsi que l'analyse d'antibiotiques, la bacitracine et la néomycine dans une spécialité pharmaceutique à usage ophtalmique, nasal ou otique.

La stratégie de validation que nous avons appliquée était basée sur le guide de validation analytique publié par une Commission de la SFSTP en 1992 dans S.T.P. Pharma Pratiques. Nous avons ensuite comparé les notes de guidance proposées dans cet article avec les recommandations internationales reprises dans le document Q2B "Validation of analytical procedure : methodology" publié à l'issue de la réunion ICH3 (International Conference on Harmonization). Ces normes ICH sont à l'heure actuelle communément admises pour la validation de méthodes dans le domaine pharmaceutique.

Ensuite, nous avons testé les performances de deux logiciels commercialisés d'aide à la réalisation de l'ensemble des calculs et des analyses statistiques nécessaires à la vérification de chacun des critères de validation analytique (AVA[®] (Qualilab) et ELSA[®] (Waters)). Tous deux sont basés sur la méthodologie proposée par la SFSTP.

Parmi les critères testés lors de la validation des méthodes d'analyse pharmaceutique ou biopharmaceutique, le test de robustesse représente certainement le critère de validation le moins fréquemment appliqué, bien qu'il soit important à considérer. Il permet d'identifier les paramètres qui légèrement modifiés sont susceptibles d'influencer significativement les performances d'une méthode et permet d'anticiper les problèmes qui pourraient apparaître lors de son transfert vers d'autres laboratoires. Il est également requis lors de la détermination de la reproductibilité (fidélité inter-laboratoire) d'une méthode. A l'heure actuelle, à notre connaissance, la seule autorité exigeant que la robustesse des méthodes analytiques destinées à un dossier d'enregistrement soit vérifiée est la "Food and Drug Administration" (FDA). Toutefois, conformément aux normes ICH, ce critère de validation sera, à notre sens, requis prochainement dans tout dossier d'autorisation de mise sur le marché de médicaments. Conscients de

l'importance de ce critère de validation, nous avons participé aux travaux d'une nouvelle Commission de la SFSTP créée dans le but de développer des notes de guidance pour la robustesse des méthodes analytiques. Toutefois, nous ne nous sommes pas seulement préoccupés de la robustesse en validation, mais nous avons également envisagé la robustesse lors du développement de la méthode.

C'est dans cette optique que nous avons considéré la planification des expériences afin d'optimiser les méthodes analytiques. Schématiquement, lors de l'optimisation d'une méthode, les deux grandes étapes qui encadrent l'expérimentation sont, en amont, la planification des expériences et, en aval, l'analyse des données. Par rapport à l'approche classique qui consiste à étudier de manière empirique et/ou systématique les paramètres susceptibles d'influencer une méthode, la méthodologie des plans d'expérience permet, en un minimum d'expériences, de sélectionner les facteurs influents et ensuite de les optimiser. Cette approche garantit l'acquisition d'un maximum d'informations en un minimum de temps.

Dans le cadre de cette étude, nous avons appliqué la méthodologie des plans d'expériences pour la mise au point d'une méthode de dosage d'anesthésiques locaux (mépivacaïne, bupivacaïne et prilocaïne) dans le plasma. Cette méthode associait la dialyse et l'enrichissement du dialysat sur une cartouche échangeuse de cations à la chromatographie liquide.

Nous avons tout d'abord recherché les facteurs qui influençaient significativement les processus de dialyse et d'enrichissement. A cette fin, nous avons réalisé un test de criblage. Nous avons ensuite recherché les conditions à la fois optimales et robustes pour le dosage de ces anesthésiques locaux dans le plasma. Nous souhaitions obtenir une récupération des analytes la plus élevée possible avec un temps de dialyse n'excédant pas 20 minutes et dans des conditions telles que la largeur des pics chromatographiques soit la plus faible possible afin d'accroître la sensibilité de la méthode. Pour ce faire, nous avons sélectionné 4 facteurs à optimiser. Une fois les expériences réalisées, nous avons appliqué une fonction de désirabilité de Derringer afin d'obtenir des conditions à la fois optimales et robustes. Le traitement statistique des résultats a été réalisé à l'aide du programme JMP® (S.A.S.).

D'autre part, nous avons envisagé la mise au point de méthodes chromatographiques à l'aide d'un logiciel d'optimisation des séparations chromatographiques intitulé "DryLab®". Il permet la mise au point rapide de méthodes optimales et robustes en chromatographie liquide ou gazeuse en procédant à des prédictions reposant sur des simulations obtenues après simple introduction de données expérimentales. Ces données préliminaires sont issues de l'analyse de deux, trois ou quatre chromatogrammes obtenus dans des conditions expérimentales bien déterminées. Le logiciel DryLab® est conçu pour modéliser les paramètres expérimentaux dans les séparations par chromatographie liquide, avec utilisation d'éluion isocratique et graduée, et

par chromatographie gazeuse sur la base de la résolution d'une paire critiques de pics et du temps d'analyse. Il se charge des calculs complexes générés par l'introduction des temps de rétention des pics correspondant aux différents constituants à séparer obtenus dans les deux, trois ou quatre conditions opératoires initiales. Les paramètres pouvant être modélisés en chromatographie liquide à polarité de phases inversée sont d'une part, pour une élution isocratique, le pourcentage de phase organique, le pH, la force ionique, la composition d'une phase mobile ternaire, l'utilisation d'additifs dans la phase mobile (agents d'appariement d'ions, amines compétitrices et tampons) et d'autre part, pour une élution graduée, les conditions de gradient et la température.

Nous avons utilisé ce logiciel pour l'optimisation de trois méthodes chromatographiques associées à une détection par spectrophotométrie d'absorption dans l'ultra violet. Il s'agissait tout d'abord d'une méthode de dosage au sein d'une formulation pharmaceutique du paclitaxel, médicament indiqué dans le traitement de divers carcinomes, et de ses nombreuses substances apparentées ; ensuite, d'une méthode d'analyse quantitative en milieu plasmatique de l'albendazole, substance thérapeutique à activités anthelminthiques, et de ses principaux métabolites (les dérivés sulfoxyde et sulfone) et enfin d'une méthode de dosage d'un antibiotique, la bacitracine, au sein d'une formulation pharmaceutique à usage ophtalmique, nasal ou otique. La bacitracine est un antibiotique polypeptidique constitué d'un mélange de plusieurs composés ne différant entre eux que par la présence de certains acides aminés. Ceux possédant une activité microbiologique appartiennent aux groupes A (constituant principal), B et C. Leurs produits de dégradation résultant d'une désamination oxydative appartiennent au groupe F. Pour le dosage du paclitaxel et de la bacitracine, la principale difficulté résidait dans la séparation des différents constituants en un temps d'analyse raisonnable, tandis que pour la méthode de dosage de l'albendazole, il s'agissait de déterminer quantitativement l'albendazole et ses deux principaux métabolites présentant des polarités différentes de la molécule mère par rapport autres constituants endogènes du plasma susceptibles d'interférer dans le dosage de ces composés.

Parmi les paramètres expérimentaux pouvant être modélisés à l'aide de ce logiciel, nous avons sélectionné les conditions de gradient (temps, intervalle et forme) ainsi que la température lors de la mise au point de ces méthodes de dosage par élution graduée. Il a, en effet, été démontré que la modification simultanée du temps de gradient et de la température constituait une approche similaire à l'utilisation d'une phase mobile ternaire dans le but d'accroître la résolution. Nous avons été contraints à appliquer le mode d'élution graduée afin de séparer, au cours d'une même analyse, tous les constituants présentant des polarités différentes.

Enfin, lors de la dernière phase de notre projet, nous avons envisagé les problèmes liés au transfert de méthodes et à la validation inter-laboratoire. C'est la raison pour laquelle nous

avons participé, sous la responsabilité des Professeurs Massart et Hoogmartens, à une étude collaborative ayant pour but, d'une part, de déterminer la répétabilité et la reproductibilité de deux méthodes chromatographiques pour le dosage des différents composants d'un antibiotique, la tylosine et, d'autre part, de comparer ces deux méthodes afin de déterminer celle qui serait la plus appropriée pour l'analyse quantitative de la composition d'échantillons de tylosine. La tylosine est un mélange d'antibiotiques macrolides dont le principal composant est la tylosine A. La première méthode utilisait une phase stationnaire de silice greffée au moyen de groupements octadécyle, tandis que pour la seconde, la séparation chromatographique s'effectuait à l'aide d'une phase stationnaire polymérique (co-polymère de styrène-divinylbenzène). Douze laboratoires ont participé à cette étude et nous avons procédé, à l'instar de ces laboratoires, à l'analyse de plusieurs échantillons de tylosine selon un protocole élaboré selon les normes ISO 5725-2. Les expériences à réaliser ont également été déterminées sur la base de ces normes. Puisque le traitement statistique des résultats a été effectué par le Laboratoire du Professeur Massart, nous laisserons à nos collègues le soin de présenter les résultats de cette étude collaborative et d'en tirer les conclusions.

III. Résultats

Dans le cadre de notre projet lié au programme d'appui scientifique à la Normalisation, notre première préoccupation a été de valider plusieurs méthodes de dosage de médicaments et de leurs métabolites, que nous avons développées au sein de notre laboratoire dans le cadre d'études de biodisponibilité ou de bioéquivalence. Pour ces validations, nous avons suivi les recommandations publiées à l'issue de la Conférence de Washington dans un document intitulé "Guidelines of the Washington Conference". Parmi ces recommandations, nous pouvons citer :

- a) le contrôle de la stabilité des analytes dans des conditions bien déterminées ;
- b) l'évaluation de la sélectivité de la méthode au moyen de six sources indépendantes de la même matrice biologique ;
- c) l'établissement d'une courbe de calibration constituée d'au moins 5 niveaux de concentration ;
- d) la sélection de la fonction de réponse qui doit être la plus simple possible et dont l'ajustement a été testé statistiquement ;
- e) la détermination de la fidélité et de l'exactitude de la méthode à trois niveaux de concentration couvrant l'entièreté de l'intervalle de dosage avec un minimum de 5 répétitions par niveau de concentration. La fidélité ne peut excéder 15 % en termes de coefficient de variation, excepté au seuil de quantification pour lequel on tolère 20%. Quant à l'exactitude, les valeurs estimées doivent être comprises entre $\pm 20\%$ de la valeur supposée vraie au seuil de quantification et entre $\pm 15\%$ pour les autres niveaux de concentration.

Au terme de leur validation, les méthodes que nous avons développées au sein de notre laboratoire répondaient aux exigences reprises dans le document émanant de la Conférence de Washington. Bien que ce dernier puisse aider le bioanalyste à valider ses procédures de dosage, il n'apporte aucune approche expérimentale et statistique.

C'est pourquoi, conformément au principal objectif de notre projet, nous avons mis au point, sur la base des normes communément admises, des notes de guidance minimales adaptées à la validation de procédures bioanalytiques à la suite des travaux de la Commission "Validation dans les milieux biologiques" de la SFSTP.

Le guide de validation que nous avons élaboré au sein de la Commission SFSTP définit tout d'abord les différents critères de validation et présente ensuite une approche à la fois expérimentale et statistique pour la validation d'une procédure bioanalytique. La stratégie propo-

sée comporte deux étapes : une phase de prévalidation et une phase de validation. Pour chacune d'elles, le protocole expérimental ainsi que l'exploitation statistique des données brutes obtenues sont décrits.

L'objectif de la phase de prévalidation est de préparer tous les éléments nécessaires à la validation formelle de la méthode, c'est-à-dire la configuration précise du matériel à utiliser, la préparation des différentes solutions et des standards de validation, lesquels sont choisis en fonction des concentrations retenues pour les contrôles de qualité utilisés en routine. Cette première étape doit permettre d'identifier, au moyen d'un arbre de décision, la fonction de réponse la plus appropriée (relation linéaire ou non-linéaire, transformation mathématique, pondération). Au cours de l'analyse de la fonction de réponse, nous avons prévu un alignement des réponses sur les concentrations moyennes en fonction des pesées respectives de chaque solution mère, une visualisation graphique des résidus permettant de déceler une éventuelle valeur aberrante ou des problèmes de linéarité, une analyse statistique afin de tester l'hypothèse d'homoscédasticité, c'est-à-dire d'homogénéité des variances au moyen des tests de Cochran et de Levene et de confirmer l'adéquation du modèle de régression choisi par l'application d'un test de manque d'ajustement, une méthode de sélection du facteur de pondération, etc. La phase de prévalidation permet également de définir le seuil de détection par la méthode statistique de Miller et Miller, de déterminer le rendement d'extraction absolu et de déterminer le nombre optimal d'expériences (répétitions et séries) à réaliser en phase de validation en fonction des résultats obtenus lors de la prévalidation. L'estimation du ou des seuil(s) de quantification ainsi que de l'intervalle de dosage est également possible grâce au tracé d'un profil d'exactitude. Pour ce faire, après le regroupement des concentrations calculées en retour, la fidélité intermédiaire et l'exactitude sont estimées pour chaque niveau de concentration. Les limites de confiance unilatérales à 95% des mesures d'exactitude (par exemple, le pourcentage de recouvrement) sont ensuite calculées en y incluant les estimations de la fidélité intermédiaire et représentées graphiquement en fonction de la concentration afin d'obtenir le profil d'exactitude.

Quant à la phase de validation proprement dite, son objectif est de démontrer la sélectivité de la méthode de dosage à l'aide de six sources indépendantes de la matrice biologique, de vérifier la linéarité de la relation existant entre les concentrations estimées à partir des standards de validation et les concentrations introduites afin de déterminer l'exactitude globale de la procédure, de déterminer la fidélité de la méthode (répétabilité et fidélité intermédiaire) ainsi que son exactitude à chaque niveau de concentration des standards de validation et, enfin, de confirmer le ou les seuil(s) de quantification. Les expériences de la phase de validation doivent être réalisées sur plusieurs jours (pas nécessairement consécutifs mais sans rejet non documenté de résultats intermédiaires prévus au protocole) et dans les conditions les plus proches de celles qui seront rencontrées lors de l'analyse en routine (appareils, opérateurs, ...).

L'approche qui permet de confirmer le ou les seuil(s) de quantification et d'attester de l'exactitude et de la fidélité de la méthode respecte les recommandations communément admises lors de la Conférence de Washington. En effet, dans la mesure où les limites de confiance ne dépassent pas les critères d'acceptation de $\pm 20\%$, les valeurs estimées sont d'office inférieures à $\pm 20\%$ pour le seuil de quantification et à $\pm 15\%$ pour les autres niveaux de concentration de l'intervalle de dosage. Par ailleurs, l'appartenance de la valeur de 100 % pour le pourcentage de recouvrement dans les limites de confiance n'est pas considérée. La raison étant que dans le cas de méthodes peu fidèles ou lorsque le nombre d'échantillons est insuffisant, le risque de considérer une méthode comme valide alors qu'elle ne l'est pas en réalité est trop important.

Afin de pouvoir réaliser l'ensemble des calculs statistiques prévus dans ce protocole de validation, nous avons créé un logiciel à l'aide du programme Microsoft Excel[®] pour Windows. L'analyse de la fonction de réponse a été particulièrement développée et tient compte des problèmes liés à l'hétérogénéité des variances ou au manque d'ajustement du modèle de régression sélectionné. Lors de l'étude de la fonction de réponse d'une méthode bioanalytique, l'utilisation du modèle simple de régression des moindres carrés peut s'avérer inadéquate. Notre logiciel prévoit, dans ce cas de figure, le recours à des modèles de régression pondérée ou à des méthodes faisant appel à des transformations préalables des données (racine carrée, logarithme népérien, ...).

Par ailleurs, afin de tester cette stratégie de validation à la fois expérimentale et statistique, nous avons validé, selon le protocole établi par la Commission de la SFSTP, plusieurs méthodes chromatographiques pour le dosage de médicaments (aténolol, sotalol, albendazole, anesthésiques locaux) et éventuellement de leurs métabolites dans le plasma. Lors de l'application de ce protocole de validation, nous n'avons rencontré aucun problème majeur et nous avons pu valider ces méthodes de dosage dans un domaine de concentration relativement large.

Nous n'allons pas présenter les résultats de la validation de chacune de ces méthodes. Nous insisterons simplement sur la validation de la méthode de dosage de l'aténolol dans le plasma par association de la dialyse à la chromatographie liquide. L'ensemble des résultats générés au cours de cette validation a d'ailleurs constitué l'exemple d'application qui sera prochainement publié dans le numéro spécial consacré à la validation de la revue scientifique *Analytica Chimica Acta*. Lors de la phase de prévalidation, l'analyse de la fonction de réponse, telle qu'elle est proposée dans le guide de validation, a tout d'abord été réalisée dans un domaine de concentration s'étendant de 20 à 1000 ng/ml. Un modèle simple de régression linéaire (méthode des moindres carrés) a été sélectionné. Toutefois, une transformation mathématique des données (racine carrée) s'est avérée nécessaire afin de pallier un manque d'homogénéité

des variances. L'hypothèse d'homogénéité avait en effet été rejetée lors de l'application des tests statistiques de Cochran et de Levene. Après l'estimation des concentrations des échantillons de calibration à partir des différentes droites de régression obtenues, l'établissement d'un profil d'exactitude a permis de définir un nouvel intervalle de dosage. En effet, les limites de confiance unilatérales à 95% des mesures d'exactitude obtenues à 20 ng/ml dépassaient les normes d'acceptation de 80% et 120%. Puisque ce niveau de concentration a été éliminé, nous avons évalué à nouveau la fonction de réponse aux niveaux de concentration restants et avons sélectionné le même modèle de régression afin d'exprimer, de façon optimale, la relation existant entre les réponses et les concentrations. Nous avons également estimé par cette procédure le seuil de quantification à 26 ng/ml. Au terme des expériences de prévalidation, deux autres critères de validation ont encore été évalués : le seuil de détection, estimé à 10 ng/ml par la méthode statistique de Miller et Miller, et la récupération absolue moyenne correspondant à 65 %.

En phase de validation, les résultats obtenus à partir des standards de validation ont permis de vérifier la fidélité et l'exactitude de la méthode. En effet, les limites de confiance des mesures d'exactitude ne dépassaient pas les normes d'acceptation fixées quel que soit le niveau de concentration choisi dans l'intervalle de dosage. Le seuil de quantification estimé lors de la prévalidation a donc pu être confirmé. En outre, nous avons vérifié la linéarité de la relation existant entre les concentrations mesurées et les concentrations effectivement ajoutées et avons contrôlé la sélectivité de la méthode en analysant six sources indépendantes de la même matrice biologique. Aucune interférence provenant des composés endogènes du plasma n'était présente au temps de rétention des pics correspondant à l'aténolol et à son étalon interne, le sotalol.

Après avoir envisagé la problématique de la validation dans les milieux biologiques, nous nous sommes intéressés à la validation de méthodes, que nous avons développées pour l'analyse de médicaments (ibuprofène, bacitracine, néomycine) dans des formes pharmaceutiques. A cette fin, nous avons utilisé deux logiciels d'aide à la validation analytique (AVA[®] (Qualilab) et ELSA[®] (Waters)). Tous deux sont basés sur la méthodologie proposée dans un guide de validation publié par une Commission de la SFSTP en 1992, mais qui reste conforme aux normes ICH actuelles.

Lors de cette étude, nous avons pu tester les performances de ces deux logiciels. Ils ont recours à un protocole expérimental approprié permettant d'optimiser le nombre d'essais à réaliser pour satisfaire à l'étude des différents critères de validation. Ils assurent également l'ensemble des calculs mathématiques et statistiques ainsi que des représentations graphiques proposés dans le guide de validation. Les principales différences entre ces logiciels se situent au niveau des fonctionnalités. A titre d'exemples, nous pouvons citer l'intégration du logiciel Elsa[®] dans un environnement Windows, la possibilité de saisie automatique des données ainsi

que la réalisation d'un test de robustesse à l'aide de ce logiciel. Quant au programme AVA[®], il permet, entre autres, une traçabilité des données (suivi des modifications) en fonction du contexte d'utilisation, un suivi dans l'avancement de la vérification des critères, un affichage dynamique de messages spécifiques renseignant l'opérateur sur l'évolution des calculs et signalant les événements particuliers (valeur aberrante, effet de matrice, ...) et l'édition d'un rapport de validation plus complet.

Au terme de cette étude, nous pouvons formuler quelques remarques sur la stratégie proposée par la SFSTP pour la validation de méthodes analytiques dans le domaine pharmaceutique. Afin de sélectionner de façon adéquate le modèle de régression, il conviendrait peut-être d'effectuer, sur un même jour, des répétitions à chaque niveau de concentration lors de l'étude de la fonction de réponse. La variance intra-jour pourrait ainsi être estimée. Par ailleurs, le protocole appliqué pour déterminer l'exactitude, en l'occurrence l'utilisation de la valeur cible de 100 % comprise dans l'intervalle de confiance du recouvrement moyen, pourrait être modifié. Le risque de considérer une méthode comme valide alors qu'elle ne l'est pas en réalité est, en effet, trop important, notamment lorsqu'il s'agit de méthodes peu fidèles. Pour cette raison, nous préférons envisager les normes d'acceptation canadiennes. Toutefois, nous ne considérons pas que la valeur du biais doit être égal ou inférieur à 2 %, ainsi que l'indiquent ces normes, pour attester de l'exactitude d'une méthode, mais nous préférons que les limites de confiance d'une mesure d'exactitude (par exemple du pourcentage de recouvrement) soient comprises entre 98 et 102 % pour affirmer qu'une méthode est exacte.

L'étape suivante de notre projet lié au programme de normalisation a consisté à étudier la problématique de la robustesse d'une méthode lors de son développement. Pour ce faire, nous avons appliqué la méthodologie des plans d'expériences pour la mise au point d'une méthode de dosage d'anesthésiques locaux (bupivacaine, mépivacaine et prilocaïne) dans le plasma. Cette méthode associait la dialyse et l'enrichissement du dialysat sur une cartouche échangeuse de cations à la chromatographie liquide.

Afin de sélectionner les facteurs qui influençaient le plus significativement la récupération des analytes en milieu plasmatique, nous avons sélectionné un plan d'expériences D-optimal algorithmique, c'est-à-dire un plan qui, en maximisant le déterminant de la matrice, permet de calculer les effets avec la meilleure précision possible. Seize expériences ont été programmées afin de tester l'influence de sept facteurs. Avant chaque expérience, un blanc plasmatique était analysé afin de démontrer la sélectivité de la méthode dans chacune des conditions opératoires. Comme il s'agissait d'un test de criblage, nous avons seulement considéré les effets principaux. Nous avons ainsi sélectionné la stratégie, qui nous paraissait la plus appropriée, avant d'exécuter les mesures. Les facteurs les plus significatifs à considérer étaient le volume de dialyse et le débit de dispensation du dialysat sur la cartouche d'enrichissement.

Nous avons ensuite optimisé les conditions de dialyse et d'enrichissement à l'aide d'un autre plan d'expériences D-optimal algorithmique basé sur quatre facteurs (volume de dialyse, débit de dispensation du dialysat, concentration du liquide de dialyse et volume d'agent de dénaturation des protéines plasmatiques). Ce plan de 24 expériences permettait d'évaluer les effets principaux et quadratiques de ces facteurs ainsi que leurs effets d'interaction tout en imposant une contrainte : le temps de dialyse devait être inférieur à 20 minutes. Avant chaque expérience, un blanc plasmatique a été analysé afin d'attester de la sélectivité de la méthode. La récupération relative des analytes ainsi qu'un paramètre chromatographique (la largeur des pics à mi-hauteur) ont été ensuite modélisés. Grâce à cette modélisation, nous avons pu tracer des profils de prédiction et d'interaction facilitant l'interprétation de l'influence des facteurs sélectionnés. Ensuite, l'application d'une fonction de désirabilité de Derringer englobant la récupération relative des analytes, la largeur des pics à mi-hauteur ainsi que leurs dérivées locales nous a permis de déduire des conditions à la fois optimales et robustes à l'intérieur du domaine expérimental. Les valeurs prédites pour la récupération des analytes dans ces conditions opératoires étaient comparables à celles obtenues expérimentalement.

Dans le cadre de l'optimisation de méthodes analytiques, nous avons ensuite envisagé l'application d'un logiciel d'optimisation des séparations chromatographiques intitulé "DryLab[®]" pour la mise au point d'une méthode de dosage de l'albendazole et de ses principaux métabolites dans le plasma ovin et de procédures d'analyse de la bacitracine ainsi que du paclitaxel et de ses substances apparentées dans des formes pharmaceutiques.

Comme les différents constituants à séparer présentaient des polarités différentes, nous avons préféré utiliser une élution graduée. C'est la raison pour laquelle, parmi les paramètres expérimentaux pouvant être modélisés à l'aide de ce logiciel, nous avons sélectionné les conditions de gradient (temps, intervalle et forme) et éventuellement la température lors de la mise au point de ces méthodes de dosage par élution graduée.

Pour la méthode de dosage du paclitaxel et de ses substances apparentées, nous sommes parvenus, grâce à ce logiciel, à sélectionner le solvant organique le plus approprié pour la constitution de la phase mobile, à déterminer les conditions de gradient optimales à l'aide de cartes de résolution décrivant la relation entre le temps de gradient et la résolution d'une paire critique et à obtenir une séparation de douze composés suffisamment résolus en un temps d'analyse raisonnable. D'autre part, nous avons pu attester de l'excellente prédiction des temps de rétention par le logiciel DryLab. En effet, les chromatogrammes prédits et expérimentaux étaient comparables.

Quant aux méthodes de dosage de l'albendazole et de ses deux principaux métabolites et d'analyse des différents composants de la bacitracine, nous avons optimisé à la fois les conditions de gradient et la température, ce qui a nécessité quatre chromatogrammes préliminaires.

Il a, en effet, été démontré que la modification simultanée du temps de gradient et de la température constituait une approche similaire à l'utilisation d'une phase mobile ternaire afin d'accroître la résolution. Nous pouvons ainsi nous limiter à la seule utilisation de l'acétonitrile comme modificateur organique, celui-ci étant le solvant de choix en élution graduée pour de nombreuses raisons. Quant à la fraction aqueuse, il est conseillé d'utiliser un tampon de pH 3,5 lors de la séparation de composés ionisables, qu'ils soient acides ou basiques.

Lors du développement de la procédure de dosage de l'albendazole et de ses métabolites, basée sur l'utilisation de la dialyse comme technique de préparation de l'échantillon associée à la chromatographie liquide, une des difficultés auxquelles nous étions confrontés était l'obtention d'une séparation optimale des trois composés et d'une sélectivité suffisante par rapport aux constituants endogènes du plasma ovin. En considérant les pics correspondant non seulement aux analytes à séparer, mais également aux principaux composés du plasma détectables dans les conditions opératoires initiales, nous sommes parvenus, à l'aide du logiciel DryLab, à obtenir des conditions chromatographiques optimales ainsi qu'une sélectivité suffisante. De plus, les chromatogrammes prédits et expérimentaux étaient tout à fait similaires. Nous avons ensuite entrepris la validation de cette méthode de dosage selon le protocole établi par la Commission de la SFSTP à laquelle nous avons participé activement et les résultats obtenus étaient excellents. Cette méthode a permis ensuite la détermination des paramètres pharmacocinétiques chez l'ovin après administration par voie intraveineuse d'une nouvelle formulation galénique d'albendazole, une solution aqueuse.

En ce qui concerne la méthode d'analyse de la bacitracine, nous sommes également parvenus à optimiser à la fois les conditions de gradient et la température. Nous avons pu obtenir une séparation satisfaisante des constituants de la bacitracine. Toutefois, nous souhaitons poursuivre l'optimisation de la méthode en étudiant l'influence du pH de la phase mobile sur la résolution des différents composés. Pour ce faire, nous avons optimisé le pH par une approche classique, c'est-à-dire que nous avons fait varier uniquement le pH, les autres paramètres restant constants. Nous avons appliqué cette approche, car, dans sa version actuelle, le logiciel DryLab ne permet pas de modéliser, en élution graduée, ni le pH, ni d'autres paramètres, tels que la force ionique du tampon ou l'addition d'additifs dans la phase mobile (agents d'appariement d'ions, amines compétitrices, ...). Une très bonne séparation des différents composés de la bacitracine a pu être obtenue avec un tampon de pH 3,0 en un temps d'analyse de 30 minutes seulement. D'excellents résultats ont également été obtenus lors de la validation de cette méthode. Actuellement, nous souhaiterions tester d'autres types de phase stationnaire ainsi que l'addition d'agents d'appariement d'ions. Il nous semble, par ailleurs, préférable d'étudier ces paramètres, en répétant les quatre expériences préliminaires pour chaque condition opératoire.

D'autre part, les résultats de ces deux dernières études (albendazole et bacitracine) seront présentés prochainement lors d'une forum national et d'un symposium international et seront soumis pour publication dans la revue scientifique " Journal of Chromatography A ".

IV. Conclusions

L'objectif principal de notre projet était de développer des normes et recommandations communément acceptées pour la validation de procédures de mesure dans les laboratoires d'analyse. Dans le cadre de ce programme de normalisation, nous nous sommes plus particulièrement intéressés à la validation de méthodes pour le dosage de médicaments et de leurs métabolites dans les fluides biologiques.

Après avoir évalué les recommandations existant dans le domaine de la validation des méthodes bioanalytiques, nous avons mis au point des notes de guidance adaptées à la validation de telles procédures de dosage. Celles-ci sont le résultat des travaux d'une Commission intitulée " Validation dans les milieux biologiques " de la SFSTP à laquelle nous avons participé activement. Le guide de validation proposé par cette commission tient compte des particularités liées aux milieux biologiques et décrit une nouvelle stratégie à la fois expérimentale et statistique pour la validation d'une procédure bioanalytique. Le protocole retenu prévoit la séparation des expériences en deux phases : l'une de prévalidation et l'autre de validation. La prévalidation permet de déterminer statistiquement le modèle de régression par l'analyse de la fonction de réponse et d'estimer le ou les seuil(s) de quantification à valider au moyen d'un profil d'exactitude. La validation reprend en les complétant, notamment par l'utilisation généralisée de critères statistiques, les modes opératoires généralement décrits dans la littérature. De plus, le modèle de régression proposé tient compte des nouvelles techniques de plus en plus utilisées en bioanalyse, en l'occurrence les techniques chromatographiques couplées à une détection par spectrométrie de masse, qui ne sont généralement pas décrites par un modèle de régression linéaire. Toutefois, pour que cette démarche soit applicable au niveau du laboratoire, il a également été pris en compte l'aspect pratique des manipulations proposées. Les protocoles aboutissent ainsi à un nombre suffisant mais réaliste d'expériences. Le gain de qualité obtenu ne se fait donc pas au détriment du coût global de la validation.

Nous avons ensuite testé cette stratégie de validation à l'aide d'exemples concrets. Nous avons ainsi mis au point plusieurs méthodes de dosage de médicaments et éventuellement de leurs métabolites dans le plasma, que nous avons ensuite validées selon le protocole présenté. Afin de faciliter l'exploitation statistique des résultats obtenus, nous avons développé un logiciel à l'aide du programme Excel[®] pour Windows et avons insisté particulièrement sur l'analyse de la fonction de réponse. Après l'évaluation de cette stratégie de validation au cours d'une session spéciale d'étude élargie au monde industriel français et européen, le guide de validation ainsi que des exemples d'application ont fait l'objet de plusieurs présentations lors de symposiums et de publications dans des revues scientifiques internationales.

Ce guide de validation est déjà utilisé dans plusieurs sociétés pharmaceutiques (Sanofi-Winthrop, E. Lilly, etc.) et par l'Agence Nationale du Médicament Vétérinaire en France. Il a également été repris dans d'autres secteurs que le secteur pharmaceutique. C'est notamment le cas pour la Société Wallonne de distribution des eaux, qui l'utilise pour la validation de ses méthodes de dosage.

Par ailleurs, nous nous sommes également intéressés à la validation de méthodes analytiques dans le domaine pharmaceutique. Nous avons mis au point plusieurs méthodes de dosage de médicaments dans des spécialités pharmaceutiques. Nous les avons ensuite validées selon une stratégie de validation proposée en 1992 par une Commission de la SFSTP. Lors de cette étude, nous avons pu formuler quelques remarques sur cette stratégie, qui reste toutefois conforme aux normes ICH actuelles. Nous avons également comparé les performances de deux logiciels d'aide à la validation analytique (AVA[®] et ELSA[®]), basés tous deux sur la méthodologie proposée par la SFSTP.

Nous nous sommes ensuite focalisés sur l'optimisation de méthodes analytiques à l'aide de plans d'expériences. Nous avons appliqué cette méthodologie pour le criblage et l'optimisation de facteurs influençant les processus de dialyse et d'enrichissement lors de la mise au point d'une méthode entièrement automatisée pour le dosage d'anesthésiques locaux dans le plasma. Nous avons également envisagé l'optimisation de séparations chromatographiques, essentiellement en élution graduée, au moyen d'un logiciel d'optimisation appelé "DryLab[®]". A partir d'un nombre limité d'expériences (2 ou 4), nous sommes parvenus à optimiser les conditions de gradient et éventuellement de température lors de la mise au point de méthodes de dosage de médicaments tant dans le plasma qu'au sein de formes pharmaceutiques. D'excellentes séparations ont été obtenues grâce à l'application de ce logiciel d'optimisation.

Enfin, afin d'étudier les problèmes liés au transfert de méthodes et à la validation "inter-laboratoire", nous avons participé à une étude collaborative qui a permis la comparaison sous différents aspects de deux méthodes de dosage d'un antibiotique, la tylosine.

Même si nous avons pu présenter et publiés nos travaux dans le cadre de ce programme de normalisation, il nous semble opportun de développer à l'avenir un site sur le réseau Internet afin de promouvoir davantage nos travaux et en particulier ceux liés à la validation des méthodes bioanalytiques. La diffusion du guide de validation sur ce réseau permettra d'accroître l'impact de cette stratégie de validation au niveau international et d'élargir son domaine d'application.

***Bepaling van de
reproduceerbaarheid van
een methode en
methodvalidatie in
farmacopee-analyses***

**Wetenschappelijk Ondersteuningsprogramma
voor de Normalisatie, deel II**

**K.U.Leuven
Laboratorium voor farmaceutische
chemie en analyse van
geneesmiddelen
Van Evenstraat 4
3000 Leuven**

1. Inleiding

De ontwikkeling en validatie van meetmethoden in de chemische, farmaceutische, agro-voeding en aanverwante industrieën is een belangrijk gegeven.

Het opstellen van standaarden en richtlijnen in verband met deze meetmethoden is een belangrijke opdracht.

De toepassing daarvan voor de farmacopee, bij de analyse van geneesmiddelen en bioanalyse, maakt een belangrijk onderdeel daarvan uit. Op punt stellen van nieuwe of verbeterde methoden blijft noodzakelijk. De voorschriften en richtlijnen die zullen ontwikkeld worden, moeten daarom eerst toegepast worden op enkele gevallen.

Na toetsing werd gekozen, gezien de opgedane kennis en ervaring door het jarenlang onderzoek in ons laboratorium, om enkele belangrijke antibiotica als model te nemen.

De biotechnologie van de industriële antibiotica is op dit ogenblik trouwens in volle expansie, wat ongetwijfeld een diepgaande invloed zal hebben op de aard en samenstelling van deze producten.

De keuze kan als volgt toegelicht worden.

Spiramycine en tylosine, beide behorend tot de groep van macrolide antibiotica, worden niet alleen aangewend respectievelijk in de humane en veterinaire geneeskunde, maar hebben ook toepassing gevonden in de agro-voeding als groeibevorderende stoffen.

Bacitracine, een polypeptide antibioticum, wordt veelal aangewend als zinkcomplex ter bevordering van de voedselopname efficiëntie bij sommige diersoorten en ziekte controle. Het is in deze sector het economisch belangrijkste produkt.

Amoxicilline, een belangrijke vertegenwoordiger van de β -lactam groep werd eveneens als model gekozen. Het is een semi-synthetische penicilline met activiteit tegen zowel gram-positieve als gram-negatieve bacteria. Het wordt gebruikt in verschillende farmaceutische vormen zoals capsules, orale suspensies als injectievloeistoffen.

De 3 eerst vermelde antibiotica kunnen alleen als fermentatieproduct verkregen worden door klassieke biosynthese. Dit houdt in dat het isoleren gebeurt uitgaande van complexe voedingsmilieus ontstaan tijdens de gecontroleerde groei van veelal gemodificeerde micro-organismen.

De produktievoorwaarden en isoleringeschema's, bekomen na een uitgebreide reeks van experimenten, leiden tot economisch rendabele opbrengsten en behoren tot de confidentiële kennis van de producenten. Deze bekomen actieve grondstoffen bezitten meestal een complexe samenstelling met een groep actieve componenten naast een wisselend aantal inactieve bestanddelen. Hun chemische aard is de oorzaak dat gemakkelijk ongewenste nevenprodukten kunnen ontstaan door chemische omzetting tijdens het isoleren of de opzuivering. Deze onzuiverheden bezitten meestal geen activiteit en sommige kunnen zelfs toxisch zijn.

De oorspronkelijke microbiologische procedure die tot op heden voor heel wat antibiotica wordt toegepast, is gebaseerd op de activiteit ten opzichte van gevoelige micro-organismen. Deze methoden zijn zeer bewerkelijk, langdurig en onderhevig aan tal van interferenties. De wisselende samenstelling van de monsters ten opzichte van de gebruikte standaard geeft aanleiding tot problemen bij deze activiteitsbepalingen, wat ook zal blijken bij de bespreking van de bekomen resultaten! Het gebruik van alternatieve methoden dringt zich dan ook op.

2. Methodologie

Het gebruik van fysico-chemische methoden om de kwaliteitscontrole van antibiotica en analoge farmaceutische produkten uit te voeren, is reeds enkele jaren in opgang. Er moet naar gestreefd worden om de vereiste selectiviteit te bereiken. De signalen die gemeten worden moeten selectief zijn voor de te bepalen stoffen. De methode die tegenwoordig meest gebruikt wordt, is een combinatie van een chromatografische scheiding gekoppeld aan een spectrometrische bepaling.

Voor de onderzochte modellen is vloeistofchromatografie of capillaire electroforese met ultraviolet detectie de voor de hand liggende keuze.

Voor tylosine werden bestaande methoden vergeleken en een collaboratieve studie uitgevoerd. Om tot een vergelijking van de vloeistofchromatografische analyse methoden te kunnen overgaan, moet aan een aantal voorafgaande vereisten voldaan zijn. De methodologie is immers gesteund op een vergelijking van signaalresponsen

van de verschillende componenten aanwezig in te onderzoeken monsters met standaarden met dezelfde identiteit en gekende zuiverheid als de te bepalen componenten. Het was daarom voor de studie van tylosine nodig te beschikken over de nodige referentiestoffen en dit in voldoende hoeveelheid en met gekende zuiverheid.

Er werd uitgegaan van een commercieel tylosine monster dat gezuiverd werd om het hoofdbestanddeel tylosine A te verkrijgen. Deze zuivering vergt een arbeidsintensieve procedure aangezien de andere componenten die aanwezig zijn in commerciële monsters, verwante verbindingen zijn met zeer gelijkende chemische structuur en dus ook zeer gelijkende fysische eigenschappen. Dit bemoeilijkt aanzienlijk de scheiding tussen tylosine A en deze verwante verbindingen.

Uit tylosine A werd dan door chemische omzetting tylosine B en tylosine D verkregen. De drie produkten werden door kristallisatie omgevormd tot een fysische vorm die hun stabiliteit gedurende een voldoende lange periode waarborgt. Ze werden onderzocht op identiteit, zuiverheid en gehalte met behulp van spectroscopische, fysische en chemische methoden, en geschikt bevonden voor deze studie.

Voor de vergelijking van deze beide analyse methoden was het ook nuttig te beschikken over andere referentiestoffen zoals tylosine C en demycinosyltylosine A. Deze produkten kunnen in het laboratorium niet bereid worden, maar moeten uit fermentatiekulturen afgezonderd worden. Een producent was bereid om zeer kleine hoeveelheden van deze stoffen te leveren.

De eerste methode gebruikt een omgekeerde fase silicagel kolom met een zure mobiele fase en natriumperchloraat als ionenpaar reagens. De tweede methode is gebaseerd op een scheiding met behulp van een poly(styreen)divinylbenzeen stationaire fase en een alkalische mobiele fase zonder toevoeging van tegen-ionen. Een voordeel van deze keuze is tevens dat er nagenoeg geen correlatie bestaat wat betreft scheidingsmechanismen voor beide methoden. Het protocol werd opgesteld in samenwerking met de V.U.B. en een voorbereidende analyse ronde werd uitgevoerd. Na bespreking van de problemen die optraden en de opmerkingen van de deelnemers werden de resultaten verwerkt. Uiteindelijk werd de definitieve studie aangevat. Ze werd uitgevoerd door 12 binnen- en buitenlandse laboratoria. Deze studie werd opgezet en geïnterpreteerd volgens de geldende ISO-normen. De vooruitgang van de werkzaamheden werd besproken in groep 7 van de antibiotica experts van de Europese Farmacopeecommissie tijdens haar vergaderingen in Straatsburg.

Voor spiramycine werd als voorstel weerhouden om de microbiologische activiteit van de verschillende componenten te vergelijken alvorens een vergelijking van vloeistofchromatografische methoden aan te vangen. Het laboratorium heeft reeds ondervinding met dergelijk onderzoek door de uitvoering van een collaboratieve studie van het antibioticum neomycine in 6 laboratoria voor de microbiologische waardebepaling en een vergelijking met vloeistofchromatografie. Zoals voor tylosine was het daarom ook noodzakelijk de zuivere referentiestoffen spiramycine I, II, III en IV te bereiden of af te zonderen. Het vergt het gebruik van scheiding en zuivering met behulp van preparatieve kolomchromatografie op silicagel.

Daar spiramycine II en III in kleine hoeveelheden aanwezig zijn in commerciële monsters, waren meerdere scheidingen nodig om voldoende hoeveelheden van deze mineure componenten te bekomen om deze studie uit te voeren.

Spiramycine IV en de neospiramycines I, II en III werden in kleine hoeveelheden bekomen door chemische omzettingen.

De referentiestoffen werden eveneens met behulp van spectroscopische, fysische en chemische methoden onderzocht op hun identiteit, zuiverheid en gehalte, en geschikt bevonden voor hun toepassing.

Een model waarbij de individuele actieve componenten van een antibioticum aan een microbiologische waardebepaling met verschillende varianten onderworpen worden, werd uitgewerkt.

Het protocol werd door ons laboratorium opgesteld in overeenstemming met de richtlijnen beschreven in de Europese Farmacopee onder de rubriek microbiologische waardebepaling van antibiotica. Een voorafgaande studie werd in het eigen laboratorium uitgevoerd om eventuele problemen te kunnen opsporen. Het bereik van de te gebruiken concentraties voor de 3 componenten om tot een aanvaardbaar respons te komen, werd vastgelegd. Aan de uiteindelijke studie werd deelgenomen door 5 binnen- en buitenlandse laboratoria en door 1 buitenlandse producent van spiramycine. Er werd gebruik gemaakt van de diffusie en van de turbidimetrische methoden en van verschillende indicatormicro-organismen. Verschillende proefopzetten werden aangewend om de gebruikelijke varianten voor deze bepalingen na te bootsen. De resultaten werden eveneens geanalyseerd volgens de geldende ISO-richtlijnen. De activiteiten tussen de produkten werden vergeleken per methode alsmede de verschillen voor hetzelfde produkt tussen de twee methoden. De interlaboratorium relatieve standaardafwijkingen werden berekend. De resultaten voor de

herhaalbaarheid, de varianties tussen de laboratoria en de reproduceerbaarheid werden eveneens bepaald.

Voor amoxicilline, één van de belangrijkste semi-synthetische β -lactam antibiotica op de markt, werd een capillaire electroforetische scheidingsmethode ontwikkeld. Omdat er een groot aantal vloeistofchromatografische methoden beschreven zijn, werd een micellaire elektrokinetische chromatografische variant ontworpen. Deze nieuwe technologie bezit een zeer hoog scheidingsvermogen wat veelal resulteert in betere resolutie en kortere analyse tijden. De operationele kostprijs is ook lager dan voor vloeistofchromatografie. Er is een gepaste, zuivere en stabiele chemische referentiestof van amoxicilline beschikbaar (Ph. Eur., CRS). Zoals alle β -lactam antibiotica is deze stof vrij onstabiel in oplossing en tijdens het bewaren. Er ontstaat een grote variëteit aan ontbindingsprodukten waaronder ook polymeren. Precursoren van de synthese kunnen eveneens als onzuiverheden in monsters voorkomen.

In ons laboratorium waren reeds een groot aantal van deze ontbindingsprodukten vroeger bereid. Bepaalde synthese-onzuiverheden werden door een fabrikant ter beschikking gesteld.

De ontwikkeling van deze elektroforetische scheidingsmethoden omvatte het onderzoek van de verschillende parameters op het bekomen resultaat. Er volgde een validatie van al de variabele parameters die de scheiding merkkelijk beïnvloeden.

De toepassing van een factoriële proefopstelling, alsmede de analyse van de gemeten responsvariabelen en multivariate regressieberekeningen worden daarbij uitgevoerd met behulp van een klassiek statistische grafische software systeem "statgraphics".

Het onderzoek naar een geschikte methode voor de kwaliteitscontrole van bacitracine en zijn zinkcomplex werd bemoeilijkt door een aantal factoren, eigen aan dit produkt.

Het commercieel product bestaat uit een relatief onzuiver mengsel van meerdere onstabiele, verwante verbindingen met complexe polypeptide structuur die een vrij hoge molecuulmassa bezitten.

Tijdens de preliminaire controle van de beschreven methoden uit de literatuur bleek dat bij alle onderzochte monsters, er steeds een belangrijke interferentie optrad tijdens de scheiding tussen onbekende matrixbestanddelen en de biologisch actieve componenten.

Een variatie van de pH, een zeer belangrijke parameter wat betreft invloed op de scheiding, gaf geen verbetering.

Een methode recent voorgesteld aan de Farmacopeecommissie door een Europese fabrikant van bacitracine, werd eveneens aan controle onderworpen. De scheiding tussen de actieve componenten onderling was vollediger, maar de interferentie van de matrixbestanddelen blijft aanwezig.

Daar al deze beschreven vloeistofchromatografische methoden die gebruik maken van isocratische elutie niet voldeden, werd overgeschakeld op gradiëntelutie. De invloed van verschillende typen organische modificatoren en de invloed van verschillende typen organische zuren en vluchtige buffers werd onderzocht. Een volledige scheiding werd nog niet bereikt, maar deze voorlopige methode voldoet het best aan het gestelde criterium. Het gebruik van vluchtige componenten in de mobiele fase bezit ook een ander voordeel. Een koppeling van de chromatografische scheiding met massaspectrometrie zal toelaten de identiteit van de geëluëerde pieken vast te stellen.

Dit is des te meer belangrijk omdat geen referentiestoffen voor de individuele componenten beschikbaar zijn. Het optimaliseren van de monstervoorbehandeling is eveneens noodzakelijk om degradatie tijdens het oplossen of bewaren van de oplossingen vóór de scheiding te verhinderen.

De bedekkingsgraad en de aard van het gebruikte kolom materiaal kunnen een significante invloed hebben op het resultaat. Het aanwijzen van andere factoren met invloed moeten eveneens onderzocht worden.

Een chromatografische scheiding bekomen met gradiëntelutie kan wel meer problemen opleveren bij methodetransfer. Er zijn meer voorzorgen nodig tijdens de validatie van deze methode. Een voorafgaandelijke uitvoering van een uitgebreide robuustheidstest met gebruik van statistisch programma dringt zich dan ook op vooraleer tot een collaboratieve studie kan besloten worden.

Terzelfdertijd is ook prelininair onderzoek verricht om een alternatieve methode te ontwikkelen met behulp van capillaire elektroforese. Omdat doorgaans betere selectiviteit bekomen wordt na toevoegen van micellen aan de bufferoplossing, werd dan ook geopteerd voor deze modus. Een micellen vormend zwitterionische ionenpaar reagens werd geschikt bevonden. Zoals bij grotere polypeptiden is de invloed van de pH en de aard en de concentratie van de gebruikte buffer van zeer groot belang bij de optimalisering van de scheiding.

Deze methode voor de scheiding van bacitracine moet nog onderworpen worden aan een gedetailleerde experimentele proefopstelling om een optimale scheiding te bekomen binnen de kortst mogelijke tijdsduur.

Richtlijnen die de aanvaardingscriteria vastleggen, zijn hier eveneens noodzakelijk om problemen inherent aan alle technologische transfers te vermijden.

3. Resultaten

De resultaten van interlaboratorium studies uitgevoerd op twee hoge-druk vloeistofchromatografische bepalingen voor tylosine en zijn tartraatzout werden verzameld en verwerkt in een dit jaar gepubliceerd artikel. Het beschrijft tevens op verklarende wijze de principes en methoden van de gevolgde richtlijnen. Een van de onderzochte methoden is nu toegevoegd als annex voor de nieuwe monografie ontwerp tekst. Daarin wordt deze bepaling gebruikt om de samenstelling van tylosine mengsels te verifiëren en wordt gevolgd door een microbiologische waarde bepaling om de totale activiteit te meten.

De resoluties tussen kritische paren werden bestudeerd in samenhang met de kolom afhankelijke selectiviteit. Alle monsters en standaarden werden door ons laboratorium ter beschikking gesteld. De te bepalen chromatografische parameters en de te bereiden oplossingen werden beschreven. In totaal werden 7 verschillende omgekeerde fase silicagel kolommaterialen, met dezelfde lengte en licht verschillende diameters gebruikt voor methode I door in totaal 12 deelnemende laboratoria, industrieën en officiële controle-organisaties.

Voor methode II werd identiek polymeer kolom materiaal met gelijke afmetingen aangewend.

In de ganse opstelling werd een panel van personeelstypen met verschillende functies gedefinieerd. Het betrof, uitvoerende, superviserende, operationele en statistische projectmedewerkers. Om uniformiteit in rapportering te waarborgen werden standaardformulieren ontworpen. Het uitgewerkt proefmodel bevat uitgebalanceerde uniforme meetniveaus. Een vijftal monsters werden onderzocht in de 12 laboratoria volgens een bepaalde volgorde van dubbel herhaalde metingen voor elk van de monsters. De statistische analyse van de resultaten werd uitgevoerd volgens 3 op elkaar volgende stappen. Er werd een kritisch onderzoek uitgevoerd op logische samenhang en het opzoeken van afwijkende resultaten. Er werden berekeningen uitgevoerd om de algemene gemiddelden en varianties te bepalen alsmede om de belangrijkste functionele relaties vast te stellen.

Veranderingen in de verhouding acetonitril-waterige fase voor methode I om de resolutie tussen tylosine A (TA) en tylosine D (TD) waren toegelaten wanneer de vooropgestelde limietwaarde niet werd bereikt. Hetzelfde gebeurde voor de verhouding tetrahydrofuraan-water voor methode II. Hoge (110%) en lage (80%) concentraties voor een monster, ten opzichte van de gekozen nominale waarde werden gemeten om variabele zuiverheid van de monsters te simuleren. De apparatuur en de opstelling werden eveneens gecontroleerd. Referentie- en blanco-oplossingen werden geïnjecteerd, wat toeliet de bekomen pieken te identificeren.

De resultaten van de herhaalbaarheid en de signaal-ruis verhouding werden berekend. De eisen gesteld voor de resolutie waren voor methode I tussen TD en TA > 2.0 en voor methode II > 4.0 . De signaal-ruis verhouding was > 10 voor beide methoden.

De herhaalbaarheid van injectie moest een relatieve standaarddeviatie bezitten (RSD) $< 1.0\%$. De omschrijving van de kwaliteit van de te gebruiken tetrahydrofuraan (THF) alsmede de bewaaromstandigheden van dit oplosmiddel waren kritische parameters. Een modelchromatogram van elke methode werd toegevoegd. Een opmerkelijk feit werd vastgesteld. De elutie volgorde van demycinosylytylosine (DMT) en tylosine C (TC) was omgekeerd bij het gebruik van een bepaald kolom materiaal. Verschillen in elutievolgorde voor een bepaalde groep verwante stoffen op nochtans gelijkende kolommen is niet zeldzaam. Dit kan problemen veroorzaken bij methodetransfer indien de benodigde referentiestoffen niet beschikbaar zijn.

Het gehalte aan TA werden in tweevoud bepaald voor alle monsters door alle deelnemende laboratoria met de 2 methoden. De waarden van de hoge en lage concentraties daarbij gevoegd, werden een totaal aantal van 336 resultaten bekomen.

De originele data van de studie werden vooraf gecontroleerd op eventuele fouten en op afwijkende resultaten. De volledige statistische analyse werd vervolledigd en samengevat in tabellen en grafieken. Er werd een lineair verband aangetoond tussen standaardafwijkingen en het gemiddelde gehalte aan sommige verwante verbindingen. Er werd ook vastgesteld dat wanneer de verhouding tussen de laagste en hoogste hoeveelheid van een bepaalde component voldoende groot is in verschillende monsters, de standaarddeviatie afhandelbaar is van de concentratie. De verhouding blijft lineair bij intermediaire concentraties en exponentieel bij grote verschillen.

Behalve de kwantitatieve resultaten, werden ook de andere parameters van beide methoden vergeleken zoals de gevoeligheid, de analyseduur en de technische toepasbaarheid van de methode. Vergelijking van de variaties en de gemiddelden

worden niet in de ISO-richtlijnen beschreven omdat deze opgesteld zijn voor de evaluatie van slechts één methode. Deze vergelijkingen lieten echter de evaluatie toe van beide methoden ten opzichte van elkaar.

De berekende en bepaalde varianties voor de bekende tylosines werden eveneens berekend en samengevoegd in tabellen. De gegevens voor andere, onbekende componenten die in de verschillende monsters voorkwamen, werden vermeld maar niet verder geanalyseerd.

De beide methoden leverden aanvaardbare resultaten op zowel voor de hoofdcomponent TA als voor de verwante verbindingen TB, TC, TD en DMT. Methode I gaf wel systematisch iets hogere resultaten voor de verwante verbindingen dan methode II. Volgens opmerkingen van een deelnemer zijn de absorptiemaxima voor alle verbindingen gelegen bij 288 nm voor de beide methoden. Dit werd bepaald door koppeling van de kolom aan een meer gesofistikeerde diode-array detector. Het verschil in golflengte bij de uitgevoerde metingen, 290 μm voor methode I en 280 μm voor methode II, kan wellicht de verklaring geven voor de bekomen afwijkingen. De voor- en nadelen van beide methoden werden eveneens behandeld. Methode I is niet zeer apparaatvriendelijk door het gebruik van een lage pH en een uitzonderlijk hoge zoutconcentratie in de mobiele fase met gevaar van uitkristallisatie, alsmede door de aanwezigheid van corrosieve chloride-ionen. Er treedt ook een onvolledige scheiding op van bepaalde onbekende producten, aanwezig in sommige monsters, van tylosine D of tylosine A. De selectiviteit voor een bepaald paar is kolom afhankelijk wat de identificatie van deze pieken kan bemoeilijken. Er ontstaan soms artefacten gedurende de chromatografie en tenslotte is de stabiliteit van de referentieoplossingen niet optimaal. Methode II kan alleen toegepast worden met behulp van één soort kolom materiaal en de analysetijd is beduidend langer dan bij de vorige methode. De scheiding tussen DMT en TB is soms zeer onvolledig en tevens afhankelijk van de onderlinge relatieve concentraties. Kleine variaties van de THF concentratie in de mobiele fase veroorzaakt grote variaties in de bekomen retentietijden, wat tevens verklaart waarom verschillende methoden om te ontgassen tot andere waarden kunnen leiden.

Aan de collaboratieve studie om de relatieve microbiologische activiteit van spiramycine I, II en III te bepalen namen 6 laboratoria van 3 landen deel. De studie is

aanvaard voor publicatie en in druk. Spiramycine IV en de beschikbare degradatieproducten neospiramycine I, II, en III werden niet in de studie opgenomen omdat deze mineure componenten een lagere activiteit bezitten en de beschikbare hoeveelheden zeer klein waren.

Er werd gebruik gemaakt van goed gekarakteriseerde standaarden van spiramycine I, II en III. Het gehalte in deze monsters voor de hoofdcomponent bedroeg respectievelijk 95.0; 93.3 en 94.2 % m/m en de som van de totale massa die kan verklaard worden, bedroeg 99.2; 99.0 en 98.9 % respectievelijk. Het base gehalte werd bekomen met behulp van water vrije titratie, het watergehalte met Karl-Fischer bepaling, de hoeveelheid residueel oplosmiddel met behulp van gaschromatografie en de verwante verbindingen en onzuiverheden met behulp van vloeistofchromatografie. De overeenkomstige hoeveelheden andere spiramycine was in elke standaard kleiner dan 1.0 %.

Voor de diffusiemethode werd steeds gebruik gemaakt van hetzelfde milieu en een identiek gevoelig indicator micro-organisme, *Bacillus subtilis*. De dosisverhoudingen varieerden tussen 1.33 en 2 en de opgebrachte volumina tussen 50 en 200 μ l.

De gebruikte proefopstellingen waren latijns vierkant of willekeurig gekozen blokken. Het aantal herhaalde metingen bedroeg respectievelijk 6 of 9. De incubatietemperatuur was vastgesteld tussen 30 ° en 37 °C met een tijdsduur van 18 uur. Het aantal onafhankelijke bepalingen bedroeg 3 tot 5 met precisiezonemetingen tot op 0.02 tot 0.2 mm.

Voor de turbidimetrische methode werd gebruik gemaakt van steeds hetzelfde vloeibare milieu met identieke pH instelling en identiek indicator micro-organisme, *Staphylococcus aureus*.

De dosisverhoudingen varieerden van 1.3 tot 2. Het aantal verdunningen die in de proeven gebruikt werden, bedroegen 3 of 4, terwijl de concentraties van de te onderzoeken oplossingen 10 tot 40 IE/ml bedroegen in laboratorium I, 3.2 tot 7.5 IE/ml in laboratorium II en 10.5 tot 17.7 IE/ml voor spiramycine I; 22.9 tot 38.7 IE/ml voor spiramycine II en 21.0 tot 35.5 IE/ml voor spiramycine II in laboratorium III. De gebruikte proefopstellingen waren steeds willekeurig gekozen blokken. Het aantal herhaalde metingen was 3 tot 6. De incubatietemperatuur was steeds 37 °C met een tijdsduur van 3.30 uur tot 4.15 uur. Het aantal onafhankelijke bepalingen bedroeg 4 of 5. De groei van de micro-organismen werd beëindigd door toevoegen van

formaldehydoplossing. De optische densiteit van de bekomen suspensies werd gemeten bij 532 μm .

De microbiologische activiteit bekomen voor spiramycine I is beduidend hoger dan deze voor spiramycine II en III. Er is tevens een verschil tussen de 2 gebruikte methoden. Bij de diffusie zijn de relatieve activiteiten voor spiramycine II en III respectievelijk 52 en 72 % ten opzichte van spiramycine I. Bij turbidimetrische bepaling werd respectievelijk 45 en 52 % bekomen. De relatieve interlaboratorium standaarddeviaties varieerden voor de waarden bekomen met diffusie van 3.6 tot 16.3 %, terwijl dit met turbidimetrische bepaling van 2.6 tot 7.7 % bedroeg. De resultaten van deze studie werden eveneens geanalyseerd volgens de geldende ISO-richtlijnen. De herhaalbaarheid tussen de laboratoria en de reproduceerbaarheid voor beide methoden werd bepaald.

Zoals bij biologische waardebepalingen veel voorkomt, hebben de variaties op de resultaten en de onderling bekomen waarden meestal een beduidend grotere spreiding dan bij fysico-chemische methoden, zelfs indien het de toepassing betreft op zuivere standaardstoffen. Het laboratorium dat bij de diffusiemethode dunnere agarbodems gebruikte, bekwam systematisch een hogere variantie binnen het eigen laboratorium.

Individuele waarden voor de activiteit van de verschillende componenten werden berekend. In totaal werden 144 onafhankelijke bepalingen uitgevoerd. De vergelijking van spiramycine I standaard met identiek spiramycine I als monster gaf confidentielimieten ($P = 0.95$) tussen 89.2 en 115 % voor de diffusie en tussen 95.6 en 102.8 % voor de turbidimetrische bepaling.

Voor spiramycine II waren de confidentielimieten begrepen tussen 86.6 en 116.6 % voor diffusie en tussen 92.5 en 105.3 % voor turbidimetrische bepaling.

Voor spiramycine III waren de limieten voor diffusie gelegen tussen 88.8 en 118.4 % en tussen 98.0 en 108.4 % voor turbidimetrie.

De methode ontwikkeld en gevalideerd, voor de bepaling van amoxicilline en zijn potentiële onzuiverheid met behulp van micellaire elektrokinetische capillaire chromatografie, werd uitgevoerd onder volgende experimentele omstandigheden.

Er werd gebruik gemaakt van gesmolten-silica capillair met 36 cm effectieve lengte en 50 μm inwendige diameter bij een temperatuur van 25 °C onder een elektrische hoogspanning van 15 kilovolt. De gebruikte migratiebuffer was samengesteld uit een

mensel van 70 millimolair natriumdiwaterstoffosfaat, 125 millimolair natriumdodecylsulfaat en 5 procent acetonitril, gebracht op pH 6.0. De hydrodynamische injectietijd bedroeg 4 seconden. De gebruikte golflengte was 230 nm.

De invloed van volgende parameters, pH, bufferconcentratie, natriumdodecylsulfaatconcentratie, en van de organische modifier, werden onderzocht. De robuustheid van de methode werd bepaald door middel van een volledige fractie factoriële proefopstelling.

De relatieve standaarddeviatie, lineariteit, precisie, detectie en bepalinglimieten werden bepaald voor de hoofdpijk.

De ontwikkelde methode laat toe amoxicilline te scheiden van 15 potentiële onzuiverheden waaronder dimeer en trimeer.

Dit is merkkelijk beter dan een vroeger gepubliceerde methode met behulp van capillaire elektroforese zonder toevoeging van ionisch oppervlakte-actieve stoffen.

Standaard paretokaarten alsmede respons-oppervlakken die de invloed op de migratietijden van de 2 kritische paren aantonen, werden opgesteld. Er werd nergens overlapping tussen L- en D-amoxicilline en tussen D-amoxicilline en 4-hydroxyfenylglycyl-amoxicilline vastgesteld binnen de bestudeerde grenzen. De methode is dus robuust. De scheiding tussen 2 mineure componenten, amoxicillinedimeer en 4-hydroxyfenylglycylamoxicilline, is gevoelig aan kleine veranderingen in de achtergrondelektrolyet.

Kwalitatieve gegevens bekomen voor het systeem waren: de herhaalbaarheid binnen één dag voor de migratietijd 1.4 % (n = 6) en voor de gecorrigeerde piekoppervlakte 0.84 (n = 6).

De herhaalbaarheid tussen verschillende dagen voor de migratietijd was 3.5 % (n = 3) en voor de gecorrigeerde piekoppervlakte 1.7 % (n = 3). De regressielijn voor de lineariteit werd bepaald binnen de grenzen 0.25 tot 1.5 mg/ml, waarbij de gecorrigeerde piekoppervlakte (Y) = 256282 X (amoxicillineconcentratie in mg/ml) - 3644.

De correlatiecoëfficiënt (r) bedroeg 0.9992. De detectielimiet, met signaal-ruisverhouding (= 3) bedroeg bij een hydrodynamische injectie van 9.9 seconden 3.3 pg (0.02 %) en de bepalinglimiet bedroeg 13.3 pg (0.08 %) voor (n = 9; RSD 9.3 %).

Een merkwaardige vaststelling was tevens dat de sterke afhankelijkheid van de pH voor de migratietijden alleen optrad bij de beide amoxicilline (5R en 5S) - piperazinediones. Dit fenomeen kon verklaard worden door het bepalen van de pK_a -waarden van dit piperazinedionemengsel in vergelijking met dit van de gelijkende amoxicilloinezuren. De pK_a -waarden van respectievelijk 3.3 en 7.7 toonden aan dat de piperazinediones veel sterkere zuren zijn waardoor de distributies in de micellen sterk verminderd worden en dus ook de migratietijden sterk verminderden.

De bekomen methode is superieur wat betreft selectiviteit en snelheid, maar vloeistofchromatografie geeft betere kwantitatieve gegevens.

De resultaten bekomen voor de kwaliteitscontrole van bacitracine zijn nog onvolledig. Dit onderzoek wordt bemoeilijkt door de complexiteit en instabiliteit van dit product en door het ontbreken van referentiestoffen.

De gepubliceerde methoden werden onderzocht, maar voldoen niet aan één van de belangrijkste scheidingsvereisten. De belangrijkste en meest actieve componenten moeten niet alleen onderling, maar ook van de inactieve onzuiverheden gescheiden zijn. De belangrijke hoeveelheid matrixbestanddelen met onbekende chemische structuur interfereert steeds met de bepaling van de actieve componenten.

Een vloeistofchromatografische methode in ontwikkeling, met gradiëntelutie vertoont een goede scheiding met minimale interferentie.

Verder onderzoek op commerciële monsters alsmede een robuustheidstest is noodzakelijk.

De voorlopige resultaten bekomen met behulp van micellaire elektrokinetische capillaire chromatografie lijken veelbelovend. De matrixbestanddelen migreren als een afzonderlijke groep gescheiden van de actieve componenten. De methode dient echter nog verder geoptimaliseerd te worden en daarop aan validatie en robuustheid onderworpen te worden.

4. Besluiten en aanbevelingen

Het onderzoek leverde nuttige informatie op nopens de minimum richtlijnen die vereist zijn bij technologische transfer in verband met chromatografische scheidingstechnieken, één van de belangrijkste hedendaags aangewende, analytische methoden.

De minimum criteria waaraan dergelijke methoden dienen te voldoen, werden op nauwkeurige wijze gedefinieerd.

Interlaboratoriumstudies opgezet met zorgvuldig gekozen doelstellingen hebben daarbij een belangrijke bijdrage geleverd.

De uitvoering van 2 verschillende bepalingmethoden voor samenstelling en gehaltebepaling van 1 antibioticum leverde ook aanvullende gegevens op om de voor- en nadelen van elke methode vast te stellen.

Chromatografische parameters met minimale vereisten werden gebruikt als voorafgaande controle om te bepalen dat de gebruikte apparatuur voldeed aan de voorwaarden voor het onderzoek.

Resolutie tussen kritische paren werd weerhouden om het scheidingsvermogen te controleren en de signaal-ruis verhouding werd bepaald om de vereiste meetgevoeligheid van het meetsysteem te waarborgen.

Het gebruik van goed gedefinieerde standaarden, zoals in de studie van tylosine en de injectie van referentie- en blanco-oplossingen, lieten toe voldoende resultaten te bekomen om tot een grondige statistische vergelijking over te gaan nopens de resultaten verkregen in de verschillende laboratoria. Daarbij kon een verandering in het elutiepatroon bij bepaald kolom materiaal ontdekt worden. Als besluit werd geformuleerd dat de resolutie tussen het kritische paar, zoals voorgesteld in het ontwerp van Europese monografie, waarschijnlijk te laag is.

De stabiliteit van de referentieoplossingen is eveneens onvoldoende.

De interlaboratoriumstudie om de microbiologische activiteit van verschillende spiramycines te vergelijken toonde aan dat het noodzakelijk is om de vrijheid in proefopstelling te beperken tot ofwel diffusie- of turbidimetrische metingen.

Gebruikte doses, de dosisverhoudingen in verdunningen en volumes alsmede incubatietemperaturen dienen binnen nauwere grenzen te variëren.

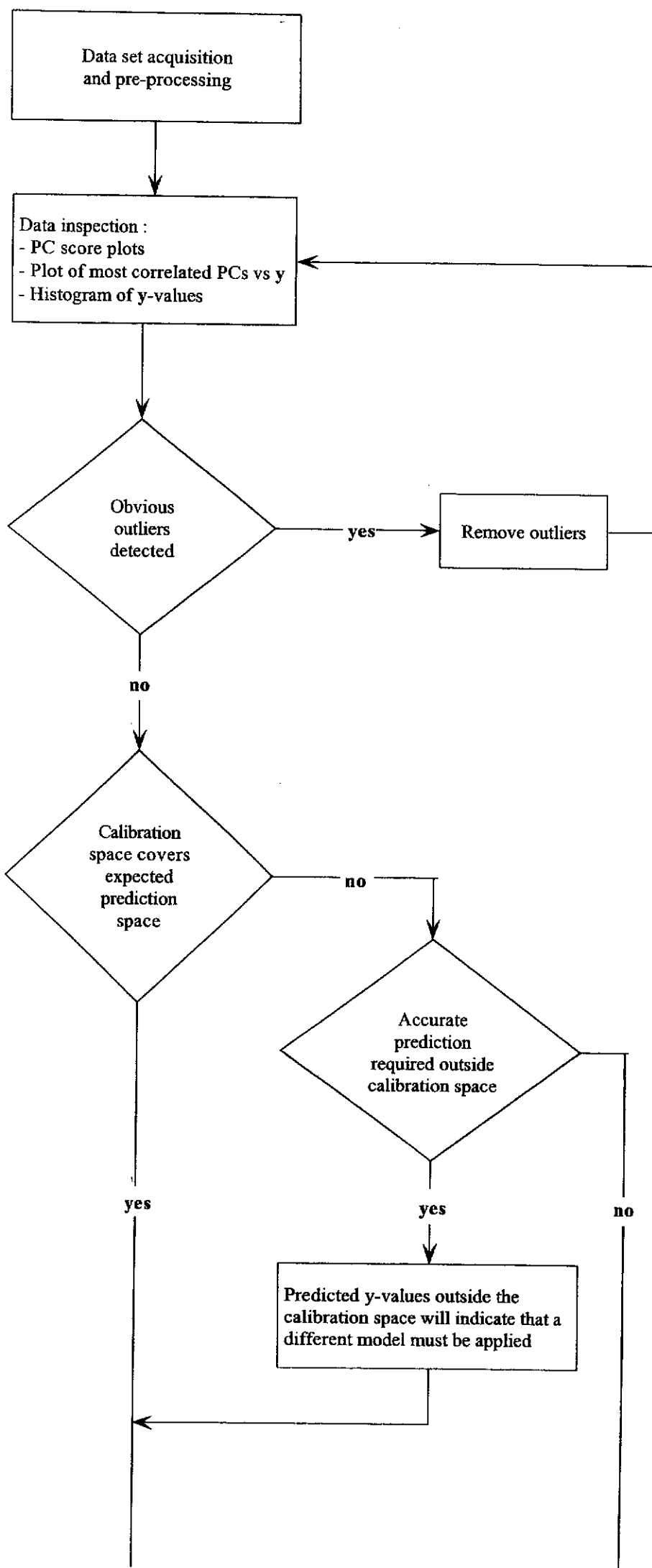
De gekende problemen voor kwantitatieve biologische metingen zullen blijven bestaan, zolang de samenstelling van de commerciële monsters aanzienlijk varieert ten opzichte van de gebruikte standaard.

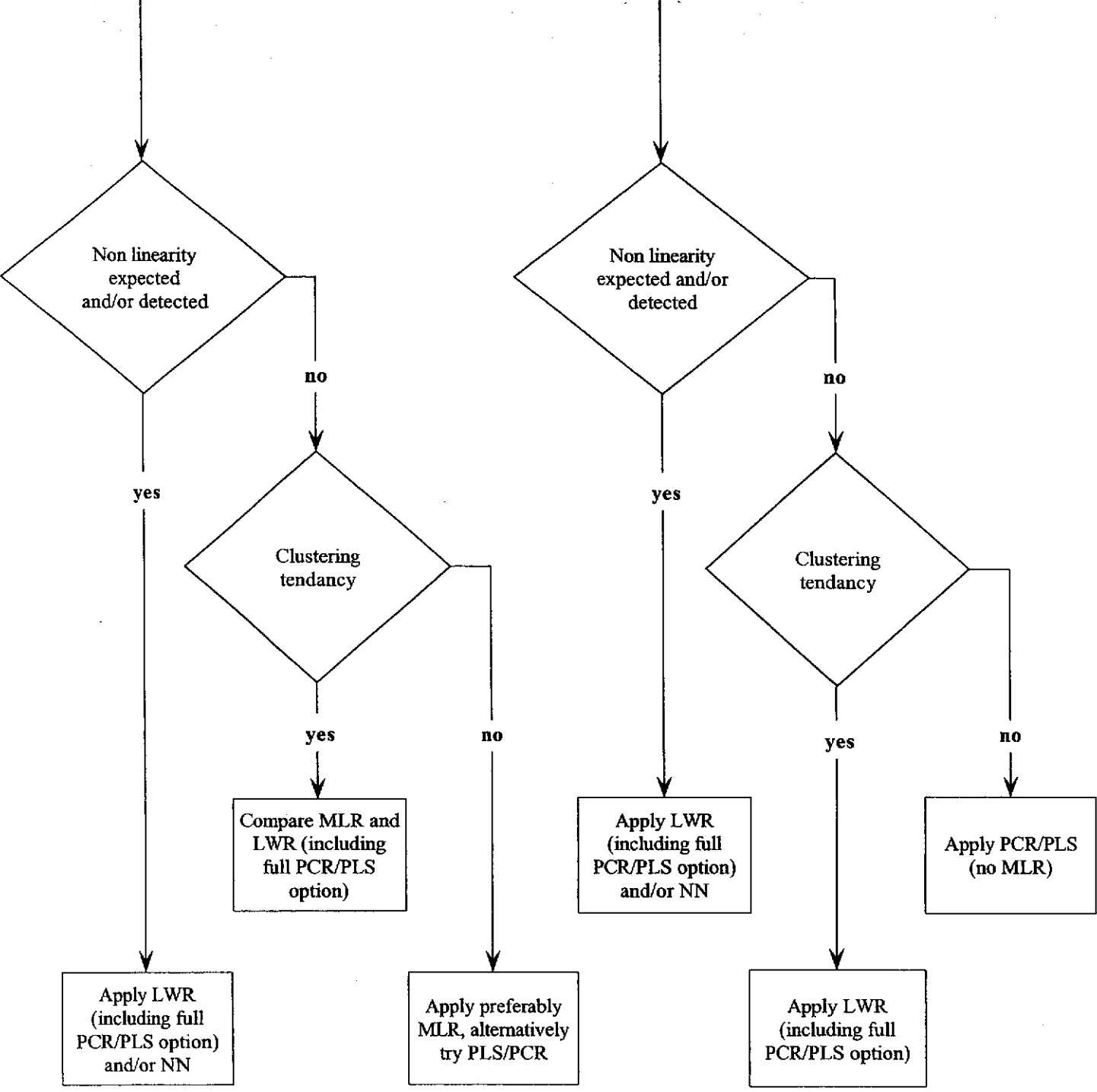
Het blijkt aanbevolen de bemerkelijke biologische waardebepalingen, waar mogelijk te vervangen door meer nauwkeurige selectieve fysico-chemische metingen.

Het uitvoeren van een aantal collaboratieve studies waarbij chromatografische en elektroforetische analysemethoden vergeleken worden, zou het sneller invoeren van laatst vermelde techniek, kunnen bevorderen.

***Robuuste procesanalyse met
multivariate calibratie***

Op basis van de opgedane ervaring werd i.v.m. multivariate calibratie volgend beslissingsschema opgesteld (cfr. volgende pagina's)





Voor meer gedetailleerde resultaten verwijzen we U graag naar de volgende referenties welke in de loop van dit project opgesteld werden.

- [1] V. Centner, D.L. Massart, O.E. de Noord, "Detection of inhomogeneities in sets of NIR spectra", *Anal. Chim. Acta*, **330** (1996) 1-17.
- [2] V. Centner, O.E. Noord, D.L. Massart, "Detection of nonlinearities in multivariate calibration", *Anal. Chim. Acta*, **376** (1998) 153-168.
- [3] D. Jouan-Rimbaud, D.L. Massart, C.A. Saby, C. Puel, "Characterisation of the representativity of selected sets of samples in multivariate calibration and pattern recognition", *Anal. Chim. Acta*, **350** (1997) 149-161.
- [4] D. Jouan-Rimbaud, D.L. Massart, C.A. Saby, C. Puel, "Determination of the representativity between two multidimensional data sets by a comparison of their structure", *Chemom. Intelligent Lab. Syst.*, **40** (1998) 129-144.
- [5] D. Jouan-Rimbaud, E. Bouveresse, D.L. Massart, O.E. de Noord, "Detection of prediction outliers and inliers in multivariate calibration", *Anal. Chim. Acta*, in press.
- [6] V. Centner, D.L. Massart, S. de Jong, "Inverse calibration predicts better than classical calibration", *Fresenius' J. Anal. Chem.*, **361** (1998) 2-9.
- [7] V. Centner, D.L. Massart, S. de Jong, "Accuracy of prediction from multivariate calibration models in the errors-in-variables situation", in preparation.
- [8] F. Despagne, D.L. Massart, "Variable selection for neural networks in multivariate calibration", *Chemom. Intell. Lab. Syst.*, **40** (1998) 145-163.
- [9] V. Centner, D. L. Massart, "Optimisation in locally weighted regression", *Anal. Chem.*, in press.
- [10] D. Jouan-Rimbaud, D.L. Massart, O.E. de Noord, "Random correlation in variable selection for multivariate calibration with a genetic algorithm", *Chemom. Intell. Lab. Syst.*, **35** (1996) 213-220.
- [11] F. Despagne, D.L. Massart, O.E. de Noord, "Optimization of PLS calibration models complexity by simulation of instrumental perturbations", *Anal. Chem.*, **69** (1998) 3391-3399.
- [12] V. Centner, D.L. Massart, O.E. de Noord, S. de Jong, B.M. Vandeginste, C. Sterna, "Elimination of uninformative variables for multivariate calibration", *Anal. Chem.*, **68** (1996) 3851-3858.
- [13] D. Jouan-Rimbaud, B. Walczak, R.J. Poppi, O. E. de Noord, D.L. Massart, "Application of wavelet transform to extract the relevant component from spectra data for multivariate calibration", *Anal. Chem.*, **69** (1997) 4317-4323.

- [14] L. Pasti, D. Jouan-Rimbaud, D.L. Massart, O.E. de Noord, "Application of Fourier transform to multivariate calibration of Near-Infrared data", *Anal. Chim. Acta*, **364** (1998) 253-263.
- [15] F. Despagne, B. Walczak, D.L. Massart, "Transfer of calibrations of Near-Infrared spectra using neural networks", *Appl. Spectrosc.*, **52** (1998) 732-745.
- [16] F. Despagne, D.L. Massart, M. Jansen, H.van Daalen, "Intersite transfer of industrial calibration models", in preparation.
- [17] F. Despagne, D.L. Massart, "Tutorial review: neural networks in multivariate calibration", *The Analyst*, **123** (1998) 157R-178R.
- [18] V. Centner, J. Verdú-Andrés, B. Walczak, D. Jouan-Rimbaud, F. Despagne, L. Pasti, R. Poppi, D.L. Massart, O.E. de Noord, "A comparison of multivariate calibration techniques applied to experimental NIR data sets", submitted.
- [19] F. Estienne, L. Pasti, V. Centner, B. Walczak, F. Despagne, D. Jouan Rimbaud, D. L. Massart, O. E. de Noord, "A comparison of multivariate calibration techniques applied to experimental NIR data sets. Part II: evaluation of the predictive ability of the models in extrapolation conditions", in preparation.
- [20] F. Despagne, F. Estienne, L. Pasti, B. Walczak, D. L. Massart, O. E. de Noord, "A comparison of multivariate calibration techniques applied to experimental NIR data sets. Part III: evaluation of the predictive ability of the models in the presence of instrumental perturbations in the test set", in preparation.

De wetenschappelijke verantwoordelijkheid over de inhoud van dit eindverslag berust volledig bij de auteurs.
Voor verdere inlichtingen betreffende het Wetenschappelijk ondersteuningsprogramma voor de Normalisatie,
gelieve contact op te nemen met de DWTC-verantwoordelijke van dit programma:

Anna CALDERONE
Tel.: (02) 238 34 40
Fax: (02) 230 59 12
E-mail: cald@belspo.be

DWTC homepage: <http://www.belspo.be>