

Een drietal voorbeelden waarbij chemometrie als hulpmiddel voor de evaluatie van klinisch-chemische en hematologische gegevens wordt gebruikt

M. VOLMER, B. G. WOLTHERS, F. A. J. MUSKIET en W. van der SLIK

Chemometrie maakt gebruik van mathematische methoden voor de interpretatie van analytische gegevens. **Patroonherkenning**, gebruikt als een evaluatie hulpmiddel, is één van de chemometrische methoden. Naast het oplossen van classificatie problemen kan patroonherkenning gebruikt worden om het inzicht in complexe biochemische processen te verbeteren. Met deze methode zoekt men naar samenvattende indexwaarden met een grotere informatieve inhoud dan de resultaten van univariate analyses. Laatstgenoemde wordt geïllustreerd aan de hand van een studie naar klinisch-chemische en hematologische gegevens van personen met heterozygote sikkelcelanemie (HbAS). In tegenstelling tot univariate analyses identificeerde de "Lineaire Discriminant Analyse" methode verscheidene nieuwe parameters die een bijdrage leveren aan de scheiding tussen controles en personen met HbAS en twee HbAS subgroepen die verschillen in percentage HbS. Twee andere studies omvatten het gebruik van multivariate kalibratietechnieken voor het voorspellen van de samenstelling van urinewegstenen volgend op analyses met behulp van Fourier-transformatie infrarood spectroscopie. Het werd aangetoond dat, vergeleken met gecomputeriseerd zoeken in een bibliotheek gevolgd door visuele beoordeling van de infrarood spectra, de analyses met de multivariate kalibratietechnieken de juistheid verbeteren en de noodzaak van expertise voor visuele evaluatie minimaliseren.

Trefwoorden: chemometrie; patroonherkenning; discriminant analyse; heterozygote sikkelcelanemie; urinewegstenen

Chemometrie is een discipline waarbij mathematische methoden worden ontwikkeld en toegepast, ten einde maximale informatie te ontlenen aan (klinisch) chemisch onderzoek (1). Twee technieken die hierbij worden toegepast zijn "patroonherkenning" en "multivariate kalibratie voor multicomponent-mengsels" (2). Van patroonherkenning zal één voorbeeld gegeven worden en van multivariate kalibratie methoden voor multicomponent-mengsels een tweetal voorbeelden.

Centraal Klinisch Chemisch Laboratorium, Academisch Ziekenhuis Groningen

Correspondentie: Ing. M. Volmer, Centraal Klinisch Chemisch Laboratorium, Academisch Ziekenhuis Groningen, Postbus 30.001, 9700 RB Groningen.
Ingekomen: 14.10.94

Patroonherkenning

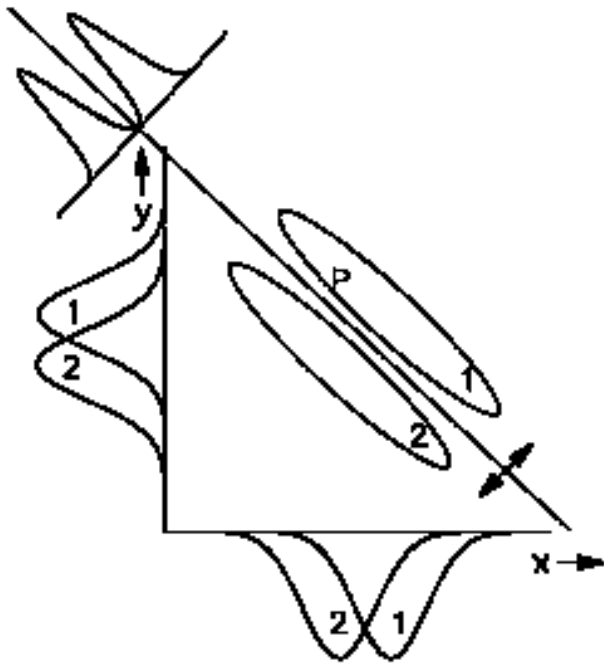
Patroonherkenning heeft tot doel het bewerken en herkennen van patronen (profielen) van overeenkomstige waarnemingsresultaten en het detecteren van groepen (klassen) van objecten. Patroonherkenningstechnieken zijn goed in staat om de relatie tussen de in- en uitvoer zichtbaar te maken in onoverzichtelijke tabellen waarin van een groot aantal objecten (b.v. personen) de resultaten (scores) van een groot aantal laboratorium bepalingen (variabelen) zijn opgenomen. Een patroon wordt voorgesteld door de scores van meer dan één variabele van één object. Toepassing van patroonherkenning in het laboratorium is vooral van belang voor de interpretatie van complexe analyse patronen, zoals celclassificatie op grond van bijvoorbeeld densitometrische en morfologische kenmerken (3), signaalanalyse (chromatogrammen) en diagnostiek (4-6).

Twee belangrijke methodieken voor patroonherkenning zijn "multivariate classificatie methoden" en "neurale netwerken". Multivariate classificatie methoden (MCM) zijn doorgaans gebaseerd op technieken afkomstig uit de lineaire algebra (7). Zoals de naam reeds doet vermoeden vereisen ze gegevens met een overwegend lineaire datastructuur. Neurale netwerken zijn computer programma's waarvan de werking gedeeltelijk is afgeleid van het veronderstelde werkingsmechanisme van de hersenen. Kenmerkend voor neurale netwerken is hun associatieve vermogen en hun capaciteit om te leren van voorbeelden. Netwerken kunnen toegepast worden voor zowel lineaire als niet-lineaire datastructuren (8). MCM's en neurale netwerken hebben elk hun specifieke voor- en nadelen.

Een voorbeeld van patroonherkenning wordt geïllustreerd aan de hand van een studie van de klinisch-chemische en hematologische parameters van personen met heterozygote sikkelcelanemie. Het doel was om een beter inzicht te krijgen in de onderliggende patho-fysiologische processen. Hierbij is gebruik gemaakt van een MCM (9).

Lineaire Discriminant Analyse

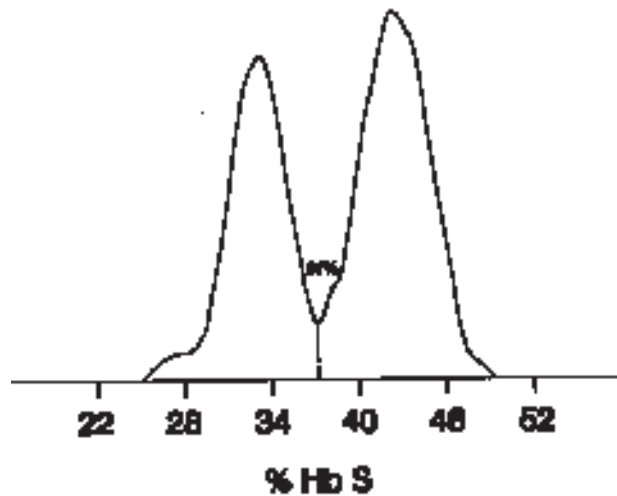
De meeste laboratoriumbepalingen vertonen overlappende verdelingen van de patiënten en de referentie groep. De diagnostische bruikbaarheid is afhankelijk van de mate van overlap. Met de gebruikelijke univariate aanpak is het doorgaans niet mogelijk om maximale informatie aan de gegevens te ontlenen. De reden hiervan is dat geen rekening wordt gehouden met mogelijke correlaties tussen de diverse variabelen. Multivariate statistische methoden verschaffen



Figuur 1. Het verband tussen de variabelen X en Y voor twee hypothetische groepen 1 en 2. De groepen kunnen niet volledig van elkaar worden gescheiden op grond van de afzonderlijke variabelen X of Y. Combinatie van X en Y leidt voor beide groepen tot hyperelliptische intervallen, die volledig gescheiden zijn in een nieuw coördinatensysteem.

ons deze mogelijkheid wel, zodat de informatie-waarde van klinisch-chemisch onderzoek vergroot kan worden (figuur 1). De figuur illustreert de puntenwolken van meetresultaten, omvat door een tweetal ellipsen, van de analyses X en Y van 2 hypothetische groepen (b.v. 1=patiëntengroep en 2=referentiegroep). Aan de hand van de frequentieverdelingen langs de assen is te zien dat noch variabele X, noch variabele Y, een complete scheiding te zien geven tussen de twee hypothetische groepen. Met de hier geïllustreerde Lineaire Discriminant Analyse (LDA) techniek wordt met een mathematische methode gezocht naar een lineaire combinatie van de oorspronkelijke variabelen in de vorm van een producten som van X en Y. LDA zoekt een optimale scheiding tussen de twee groepen door een nieuwe projectie van het oorspronkelijke assenstelsel toe te passen. In dit voorbeeld van 2-groeps-LDA is de beslissingsas (gezond-ziek) zodanig berekend dat de afstand tussen de 2 groepen maximaal is. Uit deze illustratie blijkt, dat rekening houdend met de relatie tussen de verschillende variabelen, soms meer informatie verkregen kan worden dan bij een univariate aanpak.

Van 18 vrouwelijke controles (HbAA) en 37 vrouwen met heterozygote sikkelcelanemie (HbAS) werden de resultaten van standaard hematologische en klinisch-chemische variabelen vergeleken. Dit onderzoek werd zowel met univariate als multivariate classificatie methoden uitgevoerd. Voor het percentage HbS werd een bimodale verdeling vastgesteld (figuur 2).



Figuur 2. Frequentieverdeling van het percentage HbS (x-as) voor de gehele groep personen met heterozygote sikkelcelanemie (n=37). De niet getekende y-as representeert de frequentiedichtheid. Het percentage HbS blijkt bimodaal verdeeld te zijn met toppen (modi) bij 33,3% en 42,4% en een optimale scheiding bij 37% HbS.

Een percentage HbS $\leq 37\%$ werd toegeschreven aan het gelijktijdig vóórkomen van heterozygote α -thalassemie-2 (3 actieve α -genen). Personen met HbS $> 37\%$ werden verondersteld 4 actieve α -genen te bezitten. Verder onderzoek werd uitgevoerd met de controle groep, de gehele HbAS groep en beide HbAS subgroepen.

Met de Student t-test werd voor de gehele HbAS groep afwijkingen gevonden in natrium, log(ASAT) en MCH. HbAS personen met HbS $\leq 37\%$ (n=16) hadden verlaagd MCH en log(ASAT) en personen met HbS $> 37\%$ (n=21) hadden verhoogd natrium en urinezuur en verlaagd log(ASAT). Bij het gebruik van de "Lineaire Discriminant Analyse" (LDA) werden additionele afwijkingen gevonden in kalium, ureum, urinezuur en ALAT (gehele HbAS groep), natrium en ALAT (subgroep met HbS $\leq 37\%$) en kalium, ureum, ALAT (subgroep met HbS $> 37\%$) (tabel 1). De beste univariate scheiding tussen de gehele HbAS groep en de controle groep werd gevonden voor de MCH met een "non error rate" (NER) van 69,1%, terwijl bij het gebruik van LDA met de variabelen natrium, kalium, ureum, urinezuur, log(ASAT), ALAT en MCH een NER gevonden werd van 85,5%. De aan sikkelcelnefropathie gerelateerde variabelen natrium, kalium, ureum en urinezuur zijn tot nu toe slechts afwijkend bevonden bij patiënten met homozygote sikkelcelanemie. De tussen de HbAS subgroepen gevonden verschillen in parameters voor sikkelcelnefropathie zijn waarschijnlijk gerelateerd aan de verminderde tendens van erythrocyten om te sikkelen naarmate het percentage HbS lager is. De gegevens van de LDA tonen aan dat zelfs schijnbaar gezonde personen met HbAS heterogeen zijn in de ontwikkeling van klinisch-chemisch aantoonbare abnormaliteiten.

Tabel 1. Resultaten van "Lineaire Discriminant Analyse", zoals verkregen uit de vergelijkingen van de gehele HbAS groep en twee HbAS subgroepen met de controle groep (HbAA)

Geselecteerde variabelen	Gehele HbAS groep (n=37)		HbAS sub-groep met HbS \leq 37% (n=16)	HbAS sub-groep met HbS \geq 37% (n=21)
	MCH*	natrium* kalium ureum urinezuur log(ASAT)* ALAT MCH*	natrium log(ASAT)* ALAT MCH*	natrium* kalium ureum urinezuur* log(ASAT)* ALAT
Standard distance D	0,77	2,23	1,91	2,71
Non Error Rate (%)	69,1	85,5	82,4	89,7
Sensitiviteit (%)	70,3	86,5	81,3	90,5
Specificiteit (%)	66,7	83,3	83,3	88,9

D: aantal standaard deviaties tussen de gemiddelden van de aangegeven groep en de controle groep; Non Error Rate: fractie correct geclassificeerde objecten; Sensitiviteit: ratio tussen het aantal werkelijk positieve resultaten en de som van de werkelijk positieve en vals negatieve resultaten; Specificiteit: ratio tussen de werkelijk negatieve resultaten en de som van de werkelijk negatieve en vals negatieve resultaten; *: ook afwijkend bevonden in de univariate analyse.

Multivariate kalibratie technieken

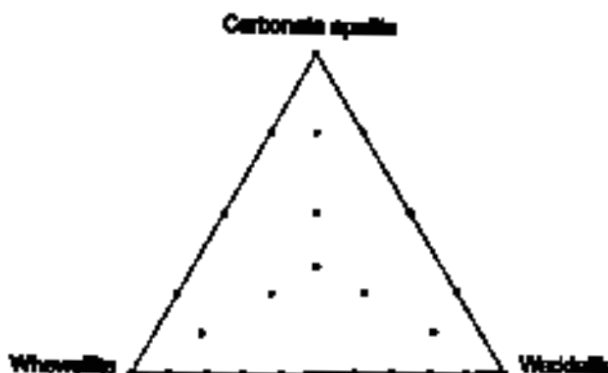
Multivariate kalibratietechnieken hebben als doel de kwantificering van één of meerdere stoffen in een complex mengsel. Deze mengsels veroorzaken vaak een moeilijk te interpreteren analyse patroon. Een voorbeeld hiervan is de spectroscopische bepaling van niet functionele hemoglobines. Voor multivariate kalibratie kan gebruik gemaakt worden van "multivariate en multi-pele lineaire-regressie methoden" en "neurale netwerken". Multivariate en multi-pele lineaire-regressie methoden (MLR) omvatten methoden zoals de klassieke multi-pele lineaire regressie (CLS), de inverse lineaire regressie (ILS), de principale componenten regressie (PCR) en de "Partial Least-Squares" (PLS) regressie (10). PLS is van deze de meest robuuste (11). Al deze technieken vereisen een niet al te grote afwijking van de lineariteit. In het algemeen geven MLR methoden en neurale netwerken vergelijkbare resultaten. Indien echter sprake is van grote afwijkingen van lineariteit door bijvoorbeeld interferenties (de wet van Lambert-Beer gaat vaak niet meer op bij infrarood spectroscopie van multi-component mengsels) dan is het gebruik van neurale netwerken te prefereren. Een voorbeeld van multivariate kalibratie wordt gegeven aan de hand van een studie naar kwantificering van urinewegstenen met zowel PLS (12) als een neurale netwerk (13).

Partial Least-Squares regressie

Kwantitatieve analyse van urinewegstenen (nierstenen) met behulp van Fourier-transformatie infrarood (FT-IR) spectroscopie wordt bemoeilijkt door de benodigde expertise bij de interpretatie van het spectrum. In ons laboratorium werd routinematig met behulp van een computer gezocht in verschillende bibliotheken. De bibliotheekbestanden bevatten 235 referentiespectra van kunstmatige mengsels met verschillende samenstellingen. Na geautomatiseerd zoeken werd voor een meer exacte semi-kwantitatieve analyse een visuele interpretatie van de bandinten-

siteiten uitgevoerd. De waarde van "Partial Least-Squares" (PLS) regressie werd uitgetest voor de meest frequent voorkomende samenstellingen van urinewegstenen.

Als kalibratieset voor PLS regressie werd een "gesloten" mengsel-ontwerp (figuur 3) gemaakt, bestaande uit 25 verschillende monsters met de componenten whewelliet, weddelliet and carbonaat-apatiet. De voorspellende waarde van de PLS analyse werd gevalideerd aan de hand van 17 kunstmatige testmengsels en 20 geselecteerde urinewegstenen. De semi-kwantitatieve samenstellingen van deze urinewegstenen werden bepaald door middel van gecomputeriseerd zoeken in een bibliotheek, gevolgd door een visuele interpretatie. Twee urinewegstenen bevatten brushiet en struviet als extra componenten. Met behulp van een FT-IR spectrofotometer werd van alle monsters een transmissiespectrum opgenomen van 4000-400 cm^{-1} met een resolutie van 4 cm^{-1} . Voorafgaand aan de PLS regressie werden de spectra gereduceerd tot



Figuur 3. Ternair mengsel-ontwerp van whewelliet, weddelliet en carbonaat-apatiet, zoals gebruikt voor de "Partial Least-Squares" trainingsset. Het ontwerp is een gesloten factor ruimte waarbinnen alle componenten altijd sommeren tot 100%.

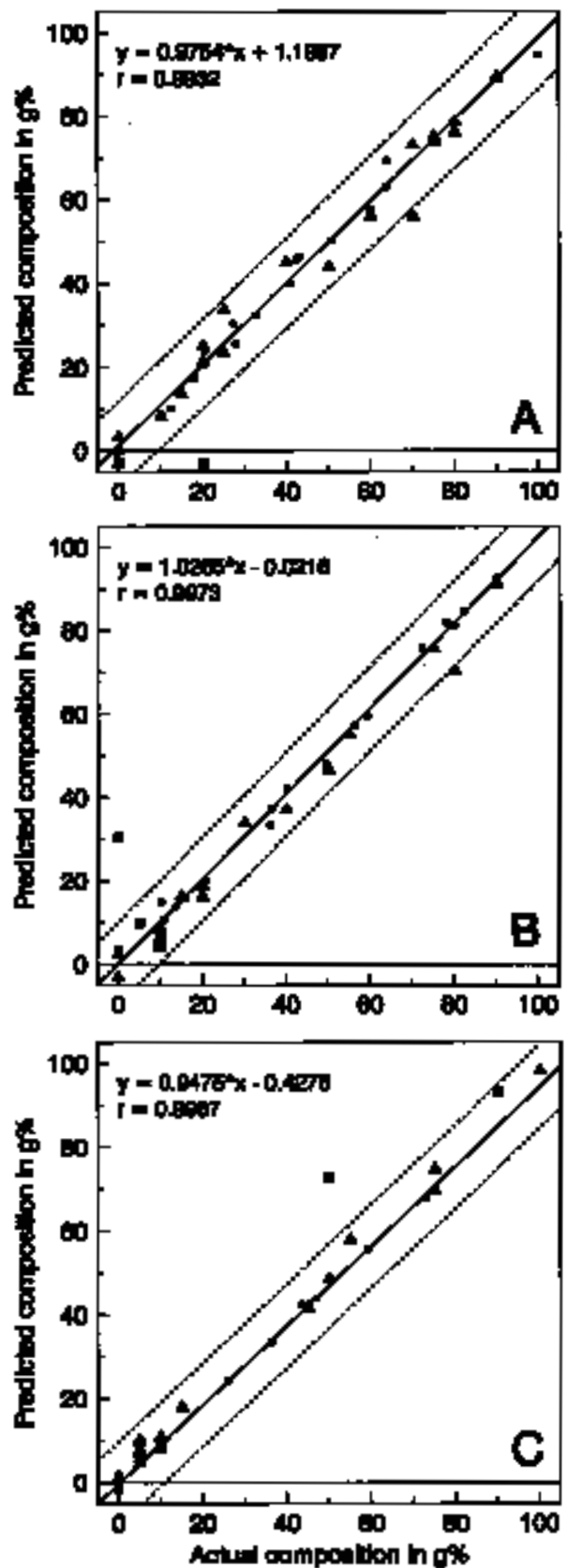
16 cm⁻¹. Na selectie van 33 relevante golfgetallen werden de spectra genormaliseerd.

We vonden dat afwijkingen tot 20% van de werkelijke samenstelling niet ongebruikelijk zijn bij gebruik van het zoeken in een bibliotheek, ook indien dit gevolgd wordt door een visuele inspectie van de spectrale bandintensiteiten. De regressielijnen van de werkelijke- en door de PLS-voorspelde- samenstellingen van de 25 kalibratiemonsters hadden correlatie coëfficiënten van 0,9944 (whewelliet), 0,9925 (weddelliet) en 0,9976 (carbonaat-apatiet). De regressielijnen van de validatiemonsters (figuur 4) tonen dat PLS regressie goed bruikbaar is voor het vaststellen van de kwantitatieve samenstelling van mengsels. Slechts twee kunstmatige mengsels hadden voorspelde waarden voor whewelliet die 5-6% afweken van de werkelijke compositie. Slechts twee urinewegstenen hadden afwijkingen van iets meer dan 10% van de "werkelijke" samenstelling. Het brushiet bevattende validatiemonster toonde aan dat composities die sterk afwijken van de samenstelling van de kalibratieset grote afwijkingen van de werkelijke samenstelling kunnen geven.

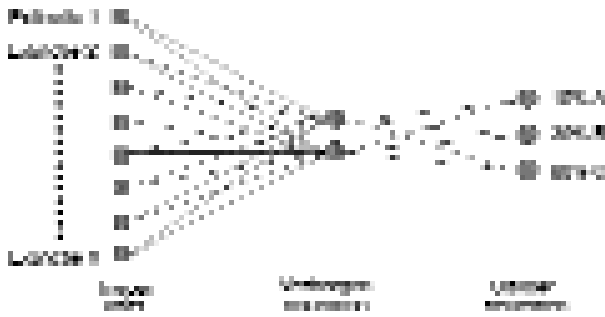
Vergeleken met geautomatiseerd zoeken in een bibliotheek gevolgd door een visuele beoordeling van de spectra, werd de PLS analyse superieur bevonden met betrekking tot de juistheid van de uitslag. Bovendien wordt de noodzaak van een visuele beoordeling van het spectrum verminderd. Afgezien van enkele praktische beperkingen in de faciliteiten voor gegevensverwerking zijn we ervan overtuigd dat PLS regressie een veelbelovend gereedschap is voor de routine kwantificering van niet slechts whewelliet, weddelliet en carbonaat-apatiet mengsels, maar eveneens voor andere toepassingsgebieden dan de analyse van urinewegstenen.

Kunstmatig neurale netwerk

FT-IR spectroscopie is een uitstekende methode voor de bepaling van de samenstelling van urinewegstenen. Zoals reeds vermeld zijn kunstmatige neurale netwerken uitermate geschikt voor het extraheren van kwantitatieve (en kwalitatieve) informatie uit gegevens met niet-lineaire structuren. We onderzochten of neurale netwerken geschikt zijn voor het kwantificeren van de 8 meest voorkomende componenten in urinewegstenen met behulp van FT-IR. De gebruikte neurale netwerken bestaan uit een invoerlaag, één of meer verborgen lagen en een uitvoerlaag (figuur 5). Iedere laag bestaat uit één of meerdere neuronen die verbonden zijn met alle neuronen in de volgende laag. De neuronen zijn simpele rekeneenheden die invoersignalen ontvangen. Het netwerk wordt getraind door het aanbieden van een bestand van trainingsmonsters met bijbehorende uitvoerwaarden. Tijdens de leer- of trainingsfase worden de sterktes van de verbindingen (gewichten) tussen de neuronen in de opeenvolgende lagen aangepast. Deze gewichten geven de stimulerende of remmende invloeden van de invoersignalen weer. Het doel van de gewichtsaanpassingen is om de fout tussen de berekende uitkomst



Figuur 4. Vergelijkingen van de met behulp van "Partial Least-Squares" voorspelde samenstellingen van urinewegstenen (y) en de werkelijke composities (x). Weergegeven is de vergelijking van de componenten whewelliet (A), weddelliet (B) en carbonaat-apatiet (C) in validatie monsters van kunstmatige mengsels (●), urinewegstenen (▲) en urinewegstenen die qua samenstelling afwijken van de ijkset (■). De regressielijnen (—) en de lijnen die 10% hiervan afwijken (---) gelden voor de kunstmatige validatie monsters (●).



Figuur 5. Vereenvoudigd schema van een kunstmatig neurale netwerk voor het vaststellen van de percentages A, B en C in een mengsel, aan de hand van spectroscopische invoersignalen.

van het netwerk en de werkelijke waarde te verkleinen tot een vooraf ingestelde tolerantie. In zekere zin extraheert het netwerk zelf een geschikt model aan de hand van de in- en uitvoergegevens. De procedure van aanpassen aan de gegevens van de trainingsset resulteert uiteindelijk in een combinatie van gewichten die gebruikt kan worden voor een snelle berekening van monsters met een onbekende uitvoer.

Het netwerk werd getraind met 160 binaire en ternaire mengsels van ammonium-uraat, brushiet, carbonaat-apatiet, cystine, struviet, urinezuur, weddelliet en whewelliet. Om de kwaliteit van het netwerk tijdens de training te testen werden er 57 extra kunstmatige mengsels gemaakt met samenstellingen die overeen kwamen met die van de trainingsset. Om de kwaliteit van het netwerk na de training te testen werd gebruik gemaakt van een validatieset van 36 geselecteerde urinewegstenen, die verschilden in de onderlinge verhoudingen van de bovengenoemde componenten. De samenstelling van deze urinewegstenen werd op de traditionele wijze vastgesteld aan de hand van gecomputeriseerd zoeken in bibliotheek bestanden, gevolgd door een visuele beoordeling van het spectrum. Drie van de 36 urinewegstenen bevatten één of meerdere extra component(en) die niet in de trainingsset (ijkset) voorkwam(en), te weten urinezuur-dihydraat, amorf calciumfosfaat en chlooramfenicol. Er werden spectra opgenomen van 4000-400 cm^{-1} met een resolutie van 16 cm^{-1} . Uit elk spectrum werd een continue reeks van golfgetallen geselecteerd, die vervolgens werden genormaliseerd op het totale oppervlak onder de curve. Deze gegevens werden aangeboden aan het netwerk. Het aantal invoerunits van het netwerk was gelijk aan het aantal geselecteerde golfgetallen. Aangezien er 8 componenten in de trainingsset waren opgenomen werd een overeenkomstig aantal uitvoer-neuronen gebruikt. Na de training werd de voorspellende waarde van het netwerk beoordeeld door een vergelijking te maken tussen de samenstellingen van de 36 urinewegstenen zoals die bepaald werden door het neurale netwerk en de overeenkomstige samenstellingen zoals die bepaald werden via gecomputeriseerd zoeken in de bibliotheek, gevolgd door visuele interpretatie.

Een continue reeks van golfgetallen van 1936-400 cm^{-1} bleek voldoende voor de identificatie van de 8 componenten. Op grond van de evaluatie-procedure met de 57 extra kunstmatige mengsels in de trainingsfase bleek dat één verborgen laag van 8 neuronen optimaal was voor het voorspellen van de samenstelling. Tabel 2 toont een goede overeenkomst tussen de werkelijke- (geschatte) en de door het netwerk voorspelde- samenstellingen van de 36 urinewegstenen. Retrospectief onderzoek in ons laboratorium heeft aangetoond dat ongeveer 80% van alle urinewegstenen calcium bevatten. Hoewel de frequentie van voorkomen, zoals in tabel 2 is weergegeven, niet representatief is voor de dagelijkse praktijk hadden alle urinewegstenen samenstellingen zoals die kunnen voorkomen. De overgrote meerderheid van de analyses vertoonden afwijkingen die kleiner waren dan 10% ten opzichte van de via de bibliotheek zoekacties vastgestelde samenstellingen. Monster 15 vertoonde afwijkende resultaten voor carbonaat-apatiet en struviet. Monster 31 gaf afwijkende resultaten voor de calciumoxalaten whewelliet en weddelliet. De niet met de trainingsset overeenkomende urinewegstenen 34, 35 en 36 werden aan de validatieset toegevoegd om de robuustheid van het netwerkmodel te testen voor samenstellingen die geheel of gedeeltelijk afwijken van de trainingsset. Een zeer ernstige afwijking werd gevonden voor de chlooramphenicol bevattende urinewegsteen (nummer 36), met zelfs een sterk negatieve waarde van -120 % voor de component whewelliet.

Gebruik van gecomputeriseerde zoekacties is zinvol voor de routinematige vaststelling van de kwalitatieve en globale kwantitatieve samenstellingen van urinewegstenen. Voor een exactere kwantitatieve samenstelling blijft een visuele beoordeling van het spectrum noodzakelijk. Met name als een mengsel bestaat uit meer dan twee componenten wordt visuele interpretatie complex. In de voorgaande studie werd gevonden dat de PLS regressie een betrouwbare methode is voor het vaststellen van de kwantitatieve samenstelling van mengsels bestaande uit carbonaat-apatiet, whewelliet en weddelliet. Tabel 2 laat zien dat met een neurale netwerk slechts twee urinewegstenen afwijkingen vertoonden van iets meer dan 10% van de werkelijke compositie. De "werkelijke" samenstellingen van de urinewegstenen zijn echter ook schattingen, waardoor een objectieve analyse van de juistheid van de neurale netwerk resultaten niet te geven is. Aan de hand van tabel 2 is tevens te zien dat urinewegstenen die afwijken van de kwalitatieve samenstellingen van de trainingsset soms grote afwijkingen vertonen van de werkelijke composities. Dit heeft als consequentie dat voor de dagelijkse praktijk voldoende kennis van het herkennen van afwijkende spectrale patronen noodzakelijk blijft. Voorafgaande aan de netwerk-analyse blijft een gecomputeriseerde zoekactie in een bibliotheek met bijzondere componenten dan ook zinvol. Op grond van onze bevindingen concluderen we dat kunstmatige neurale netwerken goed bruikbaar zijn voor het vaststellen van de

Tabel 2. Vergelijking van de werkelijke- en de met een neurale netwerk voorspelde- samenstelling van urinewegstenen^a

Monster	AMUR	BRUS	CARB	CYST	STRU	URIC	WEDD	WHEW
1							20, 10	80, 90
2							20, 20	80, 80
3							85, 75	15, 25
4			100, 95				0, 5	
5			40, 40				60, 60	
6			70, 75				10, 10	20, 15
7			30, 30				40, 40	30, 30
8			65, 70				5, 5	30, 25
9			15, 15				20, 15	65, 75
10			5, 5				30, 20	65, 75
11			5, 5				15, 15	80, 80
12			60, 60				10, 10	30, 30
13			90, 95		5, 5			5, 0
14			80, 80		15, 15		5, 5	
15			85, 70		10, 25		5, 5	
16			80, 90		20, 10			
17			95, 95		5, 5			
18			50, 55		50, 45			
19			65, 75		35, 25			
20			55, 60		40, 40			
21			5, 5			80, 80	5, 10	10, 5
22						50, 65	20, 10	30, 25
23						30, 30	10, 5	60, 55
24		20, 20	80, 80					
25		90, 90	10, 10					
26		50, 45	50, 55					
27		30, 20	70, 80					
28		40, 40	50, 50				10, 10	
29		55, 55	35, 35				10, 10	
30		40, 40	45, 45				15, 15	
31		70, 65					0, 20	30, 15
32		40, 45		30, 40		30, 15		
33	65, 65					10, 10	15, 20	10, 5
34 ^b	0, 10					0, 80	0, 10	
35 ^b			0, 90				0, 5	0, 5
36 ^b			0, 40	0, 40			0, 140	0, -120

AMUR: ammoniumuraat; BRUS: brushiet; CARB: carbonaat-apatiet; CYST: cystine; STRUV: struviet; URIC: urinezuur; WEDD: weddelliet; WHEW: whewelliet; ^a: het linker getal is het percentage zoals geschat door middel van een geautomatiseerde zoekactie in een bibliotheek, gevolgd door een visuele inspectie van de bandintensiteiten en het rechter getal is het percentage zoals voorspeld door het neurale netwerk; ^b: monster 34 was samengesteld uit urinezuur en urinezuur-dihydraat (50:50 g/g), monster 35 uit amorf calcium-fosfaat en carbonaat-apatiet (80:20 g/g) en monster 36 uit chlooramfenicol (100%).

kwantitatieve samenstelling van urinewegstenen. Verfijning van het neurale netwerk door uitbreiding met meerdere componenten die in urinewegstenen worden aangetroffen kan het toepassingsgebied vergroten. Voor de volautomatische verwerking van FT-IR spectra die opgenomen zijn van urinewegstenen ontwikkelden we een op een neurale netwerk gebaseerd computer programma voor de analyse van bovengenoemde componenten.

Algemene conclusies

Met name de praktische bruikbaarheid van de chemometrische technieken voor het kwantificeren van complexe mengsels heeft ons doen besluiten om hier in toekomst verder aan te werken. Toekomstige studies zullen betrekking hebben op het aanwenden van een neurale netwerk voor de analyse van aminozuur-

profielen van patiënten met aangeboren stofwisselingsziekten en de ontwikkeling van een expertsysteem dat een optimaal aanvraagpatroon ten behoeve van patiënten met een myocard infarct beoogt.

Literatuur

1. Massart DL, Vandeginste S, Demming Y, Michotte Y, Kaufman L. Chemometrics a textbook. Elsevier: New York, 1988.
2. Kowalski BR, Seasholtz MB. Recent developments in multivariate calibration. J Chemometrics 1991; 5: 129-145.
3. Smit JW, Gelsema ES, Huiges W, Nawrath RF, Halie MR. A commercially available interactive pattern recognition system for the characterization of blood cells: description of the system, extraction and evaluation of simple geometrical parameters of normal white cells. Clin Lab Haemat 1979; 1: 109-119.

4. Bommel JH van, Willems JL. Handboek medische informatica. Bohn Scheltema & Holkema: Utrecht/Antwerpen, 1989.
5. Goldschmidt HMJ. The application of multivariate statistical analysis in Clinical Chemistry and Haematology Medicometrical applications. ICG Printing Dordrecht, ISBN 909001828T, 1987.
6. Voet H van der, Hemel JB. Multivariate classification methods and their evaluation in applications. Proefschrift, Rijksuniversiteit Groningen, 1988.
7. Derde MP, Massart DL. Extraction of information from large sets by pattern recognition. Fresen Z Anal Chem 1982; 313: 484-495.
8. Smits JRM, Melssen WJ, Buydens LMC, Kateman G. Using artificial neural networks for solving chemical problems. Part I. Multi-layer feed-forward networks. Chemometr Intell Lab 1994; 22: 165-189.
9. Volmer M, Muskiet FAJ, Hindriks FR, van der Slik W. Potential of descriptive linear discriminant analysis for studying clinical and haematological data from persons with heterozygous sickle cell disease. Ann Clin Biochem 1991; 28: 379-385.
10. Haaland DM, Thomas EV. Partial least-squares methods for spectral analysis. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. Anal Chem 1988; 80: 1193-1202.
11. Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. Anal Chim Acta 1986; 185: 1-17.
12. Volmer M, Bolck A, Wolthers GB, Ruiter AJ, Doornbos DA, Slik W van der. Partial Least-Squares regression for routine analysis of urinary calculus composition with Fourier Transform Infrared Analysis. Clin Chem 1993; 39: 948-954.
13. Volmer M, Wolthers BG, Metting HJ, Haan THY, Coenegracht PMJ, Slik W van der. Artificial neural network predictions of urinary calculus compositions analyzed with infrared spectroscopy. Clin Chem 1994; 40: 1692-1697.

Summary

Chemometrics as an aid for the evaluation of clinical chemical and haematological data. Volmer M, Wolthers BG, Muskiet FAJ and Slik W van der. Ned Tijdschr Klin Chem 1995; 20: 239-245.

Chemometrics makes use of mathematical methods for the interpretation of analytical data. Pattern recognition, used as a tool for evaluation, is one of the chemometrical methods. Apart from solving classification problems, pattern recognition can be used to improve insight into complex biochemical processes. With this method one searches for summing index values with greater informative content than results from univariate analyses. The latter is illustrated by a study of clinical chemical and haematological data from subjects with heterozygous sickle cell anaemia (HbAS). In contrast to univariate analyses "Linear Discriminant Analysis" identified various new parameters that contribute to separation between controls and subjects with HbAS and two HbAS subgroups that differ in percentage HbS. Two other studies comprised the use of multivariate calibration techniques for the prediction of urinary calculus compositions following analyses with Fourier-transform infrared spectroscopy. It was shown that, compared with computerized library search followed by visual judgement of infrared spectra, analyses with multivariate calibration techniques improve accuracy and minimize the need of expertise for visual evaluation of infrared spectra.

Key-words: Chemometrics; pattern recognition; Linear Discriminant Analysis; heterozygous sickle cell anaemia; Partial Least-Squares regression; artificial neural network; urinary calculus composition