Maastricht University

De berekening van referentiewaarden in de klinische chemie uit analyseresultaten van een patiëntenpopulatie

Citation for published version (APA):

Naus, A. J. M. (1982). *De berekening van referentiewaarden in de klinische chemie uit analyseresultaten van een patiëntenpopulatie*. [Doctoral Thesis, Maastricht University]. Rijksuniversiteit Limburg. https://doi.org/10.26481/dis.19820513an

Document status and date:
Published: 01/01/1982

DOI:
10.26481/dis.19820513an

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 20 Apr. 2024

Voor Marijke en Emma.

Voorwoord.

Velen hebben een onmisbare bijdrage geleverd aan de tot standkoming van dit proefschrift. Op de eerste plaats bedank ik mijn ouders die mijn studie mogelijk hebben gemaakt en mijn vrouw, zonder wie dit onderzoek niet tot een goed einde zou zijn gebracht.

Verder wil ik mijn promotoren, prof.dr. P.J. Brombacher en prof. dr. R. Doornbos bedanken voor de prettige en stimulerende samenwerking. Prof.dr. H.H. Greiling en prof.dr. H.C. Hemker dank ik voor hun bereidheid als referent op te treden.

Ir. J.Th.M. Wijnen heeft met zijn opbouwende kritiek een wezenlijke bijdrage geleverd aan de tot standkoming van dit proefschrift.

Mijn opleider, dr. ir. P.S.H. Kuppens dank ik niet alleen voor zijn niet aflatende belangstelling voor dit onderzoek, maar ook voor de prettige wijze waarop hij in de afgelopen jaren mijn opleiding tot klinisch chemicus heeft begeleid.

Het manuscript is getypt door Margriet Crins-Crins en Trinette Heynen. Het verzamelen van de zeer vele gegevens is uitgevoerd door o.a. Wilma Heynders-Rijkers, Tiny Willekens-Reyven en Nelly Verdonschot.

Drs. A. Borst heeft waardevolle bijdragen geleverd bij de opzet en uitvoering van dit onderzoek.

Drs. B. Lahaye verzorgde de correctie van de Zusammenfassung en dr. B. Davies de correctie van de summary.

Door de stichting dr. ir. J.H.J. van de Laar werd een financiële bijdrage in de drukkosten geleverd.

Bestuur en directie van het St. Laurentiusziekenhuis ben ik zeer erkentelijk dat ze mij in de gelegenheid hebben gesteld dit onderzoek te verrichten.

Tot slot dank ik alle personeelsleden van het klinisch chemisch, hematologisch en isotopen laboratorium van het St. Laurentiusziekenhuis voor de prettige samenwerking in de afgelopen jaren.

Inhoud.

1. Inleiding. 1
   1.1. Enkele aspecten van het aandachtsveld van de klinische chemie. 1
   1.2. Definitie en begripsbepaling. 3
   1.3. Toelichting van enkele punten uit de definitie. 7
      1.3.1. Grenzen van het referentiewaarden gebied. 7
      1.3.2. Keuze van de referentiepopulatie. 8
   1.4. Factoren die van invloed (kunnen) zijn op de referentiewaarden. 13
   1.5. Meerdimensionale referentiewaarden. 15
   1.6. Samenvatting en probleemstelling. 19
2. Methoden voor de berekening van referentiewaarden. 21
   2.1. Inleiding. 21
   2.2. Methoden die uitgaan van een homogene populatie. 22
      2.2.1. De frequentieverdeling is Gaussisch. 22
      2.2.2. Er worden geen veronderstellingen gedaan over de vorm van de frequentieverdeling. 24
         2.2.2.1. De range van de waarnemingen. 24
         2.2.2.2. De percentielenmethode. 25
   2.3. Methoden, uitgaande van een inhomogene populatie. 27
      2.3.1. De methode van Hoffmann. 27
      2.3.2. De methode van Neumann. 29
      2.3.3. De Pryce conventie. 33
      2.3.4. De methode van Becktel. 34
      2.3.5. De methode van Bhattacharya. 37
3. Vergelijking van methoden die uitgaan van een inhomogene populatie. 41
   3.1. Keuze van de methoden voor de vergelijking. 41
   3.2. De wijze van vergelijking. 42
   3.3. Het genereren van verdelingen. 45
   3.4. Het verwerken van de verdelingen. 49
   3.5. Resultaten en discussie. 52

4. Meer gedetailleerde bespreking van de Bhattacharya plot. 56
4.1. Wiskundige achtergrond van de plot. 56
4.1.1. De breedte van de klassen. 60
4.2. Berekening van de Bhattacharya plot. 61
4.2.1. Berekening van een rechte lijn met behulp van de kleinste kwadraten methode. 63
4.2.2. De gewogen kleinste kwadraten berekening. 64
4.3. Manuele berekening na smoothen van de frequenties. 70
4.4. Berekening van meerdimensionale referentiewaarden gebieden met behulp van de Bhattacharya plot. 74
5. Niet Gaussische frequentieverdelingen. 78
5.1. Inleiding. 78
5.2. Toetsen of een frequentieverdeling Gaussisch is. 78
5.3. Methoden om de frequentieverdeling mathematisch te beschrijven,wanneer de vorm ervan niet Gaussisch is 79
5.3.1. Inleiding. 79
5.3.2. De benadering van Pearson. 80
5.3.3. Het beschrijven van een waargenomen frequentieverdeling met een reeksontwikkeling. 81
5.3.4. Het transformeren van de waarnemingen. 83
5.4. Keuze van de rekenprocedure, wanneer de frequentieverdeling niet Gaussisch is. 85
5.4.1. Inleiding. 85
5.4.2. De Gamma verdeling. 85
5.4.3. Log-Gauss verdelingen, benaderd met een gamma verdeling. 89
5.4.4. Keuze van de benadering, wanneer er twijfel is over de vorm van de frequentieverdeling. 91
6. Toepassingen. 93
6.1. Inleiding. 93
6.2. Enkele voorbeelden van bepalingen waarvan de frequentieverdeling Gaussisch is. 94
6.2.1. Kalium. 94
6.2.2. Calcium. 97
6.2.3. Thyroxine. 100

6.3. Enkele voorbeelden van bepalingen waarvan de frequentieverdeling niet Gaussisch is. 103
6.3.1. De transaminases ASAT (SGOT) en ALAT (SGPT). 103
6.3.1.1. ASAT (SGOT ; EC 2.6.1.1.). 104
6.3.1.2. ALAT (SGPT ; EC 2.6.1.2.). 111
6.3.2. Urinezuur. 113
6.4. The use of patient data for the calculation of reference values for some haematological parameters. 119
6.5. Determination of n-dimensional reference ellipsoids using patient data. 125
Samenvatting. 131
Summary. 134
Zusammenfassung. 137
Literatuur. 140
Levensbeschrijving. 154

DE BEREKENING VAN REFERENTIEWAARDEN IN DE KLINISCHE CHEMIE UIT ANALYSERESULTATEN VAN EEN PATIENTENPOPULATIE

DE BEREKENING VAN REFERENTIEWAARDEN IN DE KLINISCHE CHEMIE UIT ANALYSERESULTATEN VAN EEN PATIENTENPOPULATIE

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE GENEESKUNDE
AAN DE RIJKSUNIVERSITEIT LIMBURG TE MAASTRICHT,
OP GEZAG VAN DE RECTOR MAGNIFICUS PROF. DR. H.C. HEMKER,
VOLGENS BESLUIT VAN HET COLLEGE VAN DECANEN
IN HET OPENBAAR TE VERDEDIGEN OP
DONDERDAG 13 MEI 1982 DES NAMIDDAGS TE 4 UUR

DOOR

ANDREAS JOHANNES MARIA NAUS

GEBOREN TE GRATHEM.

Promotores:	prof. dr. P.J. Brombacher
	prof. dr. R. Doornbos
Referenten:	prof. dr. H.H. Greiling
	prof. dr. H.C. Hemker

1. Inleiding.

1.1. Enige aspecten van het aandachtsveld van de klinische chemie.

De klinische chemie is de wetenschap die zich bezighoudt met de chemische analyse van menselijk materiaal zoals bloed, urine, faeces, liquor cerebrospinalis, zweet etc. Ook de analyse van stenen (urineweg- en galstenen) behoort tot het terrein van de klinische chemie.
Doel van deze analyses is een bijdrage te leveren tot de diagnostiek en de behandeling van ziekten, waarbij chemische en biochemische processen een rol spelen.
De uitkomst van een klinisch chemisch onderzoek kan kwantitatief zijn (bijv. de glucose concentratie in serum = 5,4 mmol/l) of kwalitatief. Een kwalitatief resultaat geeft alleen maar de aan- of afwezigheid van een bepaalde stof aan. Ook tussenvormen, de semi-kwantitatieve uitkomsten, zijn uiteraard mogelijk; hierbij wordt aangegeven of een bepaalde stof in hoge of lage concentratie voorkomt. De interpretatie van een kwalitatieve uitkomst is meestal éénduidig: twijfel over de betekenis van het resultaat zal meestal niet optreden. Een positieve zwangerschapstest bijv., waarbij in urine de aanwezigheid van het hormoon HCG wordt aangetoond, geeft zelden aanleiding tot een verkeerde interpretatie, hoewel een verhoogde HCG concentratie niet hoeft te betekenen dat er sprake is van zwangerschap. (Door bepaalde testistumoren worden bijv. ook grote hoeveelheden HCG geproduceerd.)
Kwantitatieve bepalingsuitkomsten leveren echter meestal wél interpretatieproblemen op. De aanvrager van het onderzoek is nl. in het algemeen niet geïnteresseerd in de absolute getalgrootte van de uitkomst, maar in de vraag of deze uitkomst wijst op pathologie. Met andere woorden: een kwantitatieve uitkomst wordt door de aanvragend arts getransformeerd tot een kwalitatieve uitspraak.

Om deze transformatie met enige mate van betrouwbaarheid te kunnen doen, moet bekend zijn, welke bepalingsuitkomsten gevonden worden wanneer er géén sprake is van pathologie en ook, welke uitkomsten gevonden worden wanneer dit wél het geval is.
Wil een klinisch chemische bepaling enig nut hebben bij een bepaalde ziekte, dan moeten de uitkomsten van die bepaling duidelijk anders zijn, wanneer de ziekte al dan niet aanwezig is. Grofweg kunnen "normaalwaarden" gedefinieerd worden als de grenzen, waarbinnen de meeste bepalingsuitkomsten liggen van de mensen, die geen ziekte hebben waarvoor de bepaling differentieert. Een bepalingsuitkomst binnen deze grenzen kan nu met enige statistische zekerheid de kwalificatie "normaal" krijgen, terwijl een uitkomst buiten deze grenzen de kwalificatie "pathologisch" meekrijgt.

Er zijn echter ook bepalingen, waarvan de uitkomsten kwantitatief zijn en waarvoor toch het begrip "normaalwaarde" geen rol speelt. Dit zijn de bepalingen waarvan de uitkomsten tijdens een ziekteproces vervolgd worden. Zo is CEA (carcino embryonic antigen) een stof die door tal van maligne tumoren geproduceerd wordt. De bepaling van de CEA concentratie in serum wordt nu bijv. gebruikt, om te vervolgen of chemotherapie het beoogde effect heeft op het kwaadaardig proces. Een daling wijst met grote zekerheid op "aanslaan" van de therapie, terwijl een stijging wijst op verdere uitbreiding van het ziekteproces. Hier is dus duidelijk niet de absolute uitkomst van belang, maar de verandering van de uitkomst in de tijd.
Een tweede voorbeeld van een kwantitatieve analyse, waarbij niet alleen de absolute grootte van de uitkomst een rol speelt, is de bepaling van een steroid-spectrum in urine. Hier is men vooral geinteresseerd in de verhouding van de concentraties van de afzonderlijke steroiden. Een verstoring van deze verhouding kan bijv. wijzen op een erfelijk bepaalde afwijking in het metabolisme van steroiden in de bijnier.

Uit het voorgaande is duidelijk, dat de klinische chemie zeer nauw verweven is met de medische wetenschap. Zonder klinische chemie is moderne medische diagnostiek en behandeling dan ook vrijwel ondenkbaar. De klinische chemie heeft echter ook raakvlakken met de technische wetenschappen (analytische technieken) en met de wiskunde (verwerking van meetresultaten; statistische bewerking van grote hoeveelheden gegevens).
Vooral de gegevensverwerking zal in de nabije toekomst een nog belangrijker rol in de klinische chemie gaan spelen. Aan deze ontwikkeling wil dit proefschrift een bijdrage leveren.

1.2. Definitie en begripsbepaling.

Referentiewaarden worden in de dagelijkse klinisch chemische routine meestal gebruikt om bepalingsuitkomsten te verdelen in twee kategorieën : uitkomsten die géén aanwijzing geven voor bestaande of in de nabije toekomst optredende ziekte en uitkomsten die dat wél doen. Een definitie van referentiewaarden kan dan ook zijn :
Dié getallen die aangeven binnen welke grenzen de uitkomsten van een nader te specificeren percentage van de referentiepopulatie voor die bepaling terecht komen.
Bij het vaststellen van referentiewaarden in deze zin wordt slechts rekening gehouden met de analyse uitkomsten verkregen uit materiaal van mensen die geen ziekte hebben waarvoor de bepaling differentieert (de referentiepopulatie). Wanneer ook de frequentieverdeling van de bepalingsuitkomsten, wanneer er wél sprake is van ziekte, bekend is, kan bij iedere individuele uitkomst de (a posteriori) kans berekend worden, dat de patiënt waarvan deze uitkomst afkomstig is, lijdt aan de betreffende ziekte (13, 34, 48-52, 59, 67, 111). Referentiewaarden in bovenstaande betekenis hebben dan geen zin meer.

De term referentiewaarden verdient de voorkeur boven de term normaalwaarden, aangezien het begrip normaal vele dingen kan impliceren die niet gedekt worden door het begrip normaalwaarde (25, 31).
Zo betekent normaal in medische zin meestal "niet pathologisch". Een bepalingsuitkomst die valt buiten het normaalwaardengebied krijgt dan vrij automatisch het predikaat afwijkend mee, waarmee bedoeld wordt dat een ziekteproces aan het vinden van de afwijkende uitslag ten grondslag ligt.
Deze conclusie is echter niet zonder meer te trekken. Binnen de grenzen van het normaalwaardengebied vallen de bepalingsuitkomsten van een arbitrair bepaalde fractie (α) van de referentiepopulatie. Een uitkomst buiten deze grenzen kan ook betekenen, dat de gevonden waarde is verkregen door analyse van materiaal afkomstig van een persoon die behoort tot dat gedeelte van de referentiepopulatie (1- α) waar, zonder dat er sprake is van pathologie, toch bepalingsuitkomsten worden gevonden die vrij sterk afwijken van het gemiddelde. De term referentiewaarden geeft veel beter aan dat er sprake is van getallen waarmee een bepalingsuitkomst vergeleken moet worden, zonder de strikte tweedeling pathologie - geen pathologie op te dringen.
Normaal heeft verder nog de betekenis "ideaal". Het gebruik van de term normaalwaarden impliceert dan ook enigszins, dat er een ideale uitkomst voor de betreffende bepaling bestaat, nl. de mediaan van het gebied, en dat iedere afwijking van deze ideale uitkomst minder gunstig is.
In statistische zin tenslotte is normaal synoniem met "Gaussisch verdeeld". Wanneer de factoren die de uitkomst van een bepaling beïnvloeden, voldoen aan bepaalde voorwaarden, kan beredeneerd worden dat, wanneer een voldoende groot aantal resultaten van die bepaling verzameld wordt, de frequentieverdeling van deze resultaten een Gauss curve zal vormen. Of de factoren die een bepalingsuitkomst beïnvloeden inderdaad voldoen aan deze voorwaarden, kan niet a priori vastgesteld worden. Bovendien geldt voor veel bepalingen, dat niet alle factoren bekend zijn.

Hoewel de frequentieverdeling van veel klinisch chemische en hematologische bepalingen goed beschreven kan worden met een Gausscurve, is er dus geen enkele principiële reden waarom dit zo zou zijn (43, 44, 78). In dit verband is de door Elveback et al (29) gemaakte opmerking over de door Gauss opgestelde "law of errors", zeer illustratief :

"This law states that if repeated measures are made on the same physical object, the distribution of the random component on the errors can be well approximated by the Gaussian, or normal, distribution. This law implies that repeated measurements of the length of the same desk would follow the normal distribution ; but the law certainly does not imply that measurements of the length of all of the different desks to be found in a medical school would do so. The law applies to repeated measurements of one's own height, but the difference between that height and that of one's neighbour is something quite different from a random error of measurement."

De laatste zin van bovenstaand citaat kan enige bevreemding wekken, omdat de lengte verdeling van de Nederlandse bevolking bijna een perfecte Gauss curve vormt. Een populatie echter die bestaat uit een mengsel van een klein en een groot ras (bijv. Pygmeeën en Bantoes) zal echter vrijwel zeker een lengte verdeling vertonen die niet Gaussisch is. Wanneer dus geen nadere gegevens over de populatie voorhanden zijn, kan geen uitspraak gedaan worden over de vorm van de frequentieverdeling.

Ook in statistische betekenis is het gebruik van het woord normaal dus verwarrend.

Het gebruik van de term referentiewaarden in bovenstaande betekenis is niet in overeenstemming met de aanbevelingen van de Scandinavische vereniging van klinische chemie (24). In deze

aanbevelingen wordt de term referentie interval aanbevolen en wordt met referentiewaarden bedoeld de individuele uitkomsten verkregen uit de referentiepopulatie. De term referentiewaarden is echter in Nederland meer algemeen gebruikelijk dan de term referentie interval.
Elion-Gerritzen (27) geeft de voorkeur aan het gebruik van de term normaalwaarden omdat de term referentiewaarden ook gebruikt wordt voor de betekenis : uitkomsten verkregen met een referentiemethode. Het gebruik in deze betekenis is echter zeker niet algemeen. Een tweede reden om de term referentiewaarden niet te gebruiken, is voor haar gelegen in de definitie van het begrip door Dybkaer (26) :

"A set of values of a certain type of quantity obtainable from a single individual or a group of individuals corresponding to a stated description. This description must be spelled out if others are to use the reference values. For each type of quantity a series of reference groups will be necessary taking into consideration : age, sex, race, menstruation, previous diet and exercise, posture etc."

Het gebruiken van de term referentiewaarden houdt echter niet altijd in, dat bij het bepalen van deze waarden een onderscheid gemaakt is op alle punten die door Dybkaer genoemd zijn (geslacht, leeftijd etc.). Op praktische gronden zal dit veelal ook niet mogeljk zijn. Een referentiegroep moet, wil er statistische waarde aan gehecht kunnen worden, voldoende groot zijn. Wanneer te veel of te strenge critéria worden aangelegd, zal het in de praktijk vaak onmogelijk zijn een representatieve en tevens voldoende grote groep te formeren (5). Het houdt wel in dat de samenstelling en de eigenschappen van de referentiepopulatie bekend moeten zijn, tenminste op die punten die geacht kunnen worden van invloed te zijn op het resultaat van het te verrichten onderzoek. Dit laatste is meestal wel het geval.

1.3. Toelichting van enkele punten uit de definitie.

1.3.1. Grenzen van het referentiewaardengebied.

In de definitie van referentiewaarden wordt gesproken over grenzen waarbinnen de uitslagen van "een bepaald percentage" van de referentiepopulatie liggen. Voor dit "bepaalde percentage" wordt meestal 95% gekozen en wel zodanig dat 2,5% van de referentiepopulatie uitslagen heeft lager dan de ondergrens en 2,5% uitslagen heeft hoger dan de bovengrens van het gebied.
Deze keuze van 95% is echter arbitrair en komt waarschijnlijk voort uit het feit dat bij de Gauss verdeling 95% van de waarnemingen ligt tussen $\mu \pm 1{,}96\,\sigma$. De factor 1,96 wordt in de praktijk afgerond tot 2 en dit is een mooi rond, gemakkelijk hanteerbaar getal. Wanneer 95% van de uitkomsten afkomstig van "gezonden" binnen de referentiewaarden ligt, houdt dit ook in dat 5% buiten deze grenzen terecht komt (fout van de eerste soort ; vals positieven). Hoe groot het percentage vals negatieven is (fout van de tweede soort ; bepalingsuitkomsten binnen de referentiewaarden afkomstig van "zieken") kan alleen vastgesteld worden wanneer de frequentieverdeling van uitkomsten afkomstig van mensen die lijden aan de ziekte, waarvoor de bepaling differentieert, bekend is.
Bij de vaststelling van klinisch meer bruikbare referentiewaarden moet met beide soorten fouten rekening worden gehouden en bovendien met de consequenties ervan. Een voorbeeld, waarbij de gevolgen van een vals negatieve uitkomst zeer groot zijn, vormen de bepalingen die uitgevoerd worden bij alle pasgeborenen om bepaalde ernstige aangeboren stofwisselingsziekten, zoals phenylketonurie (PKU), op te sporen. Het niet ontdekken van een kind met PKU heeft enorme consequenties, terwijl de kosten van de bepaling gering zijn.
Het is verstandig in een dergelijke situatie de grenzen van het referentiewaardengebied aan de veilige (nauwe) kant te kiezen.

Ook het leggen van de linker en rechter referentiewaarde op evengrote afstanden van het gemiddelde (er voor het gemak vanuit gaande dat de frequentieverdeling Gaussisch is) is arbitrair en hoeft niet altijd juist te zijn.

De uitkomsten van een klinisch chemische bepaling die in de dagelijkse routine geproduceerd worden, zijn afkomstig van zowel "gezonden" als "zieken". De frequentieverdeling van resultaten afkomstig van zo'n inhomogene populatie, kan voorgesteld worden als een verdeling afkomstig van "gezonden", die gecontamineerd is door "pathologische" uitkomsten. Doel van dit proefschrift is nu, een mathematische beschrijving te geven van de frequentieverdeling van de resultaten afkomstig van "gezonden", uitgaande van de resultaten uit de dagelijkse routine. Wanneer een frequentieverdeling mathematisch vast ligt, kan ieder percentiel bepaald worden. In dit proefschrift wordt niet ingegaan op de vraag, welke percentielwaarden in bepaalde situaties het beste gebruikt kunnen worden als referentiewaarden. Op deze kwestie wordt uitgebreid ingegaan door Bezemer (9).

1.3.2. Keuze van de referentiepopulatie.

De samenstelling van een referentiepopulatie moet zodanig zijn dat :

1. alle leden van de populatie voldoen aan bepaalde welomschreven (gezondheids)kenmerken en eventueel aan een aantal objectief vast te stellen kenmerken, zoals geslacht, leeftijd, e.d.
2. wat betreft alle andere kenmerken gesproken kan worden van een aselecte steekproef uit de totale populatie.

Punt 1 kan bijv. inhouden dat alle leden van de referentiepopulatie gezond moeten zijn. De eis kan echter ook zijn de aanwezigheid van leverziektes, andere dan obstructieve galwegaandoeningen. De formulering van deze eis hangt af enerzijds van

de bepaling zelf en anderzijds van het doel waarvoor ze gebruikt wordt. Wat betreft dit laatste punt kan het bijv. zo zijn, dat de samenstelling van de referentiepopulatie anders moet zijn, wanneer de bepaling gebruikt gaat worden voor screeningsdoeleinden, dan wanneer ze gebruikt wordt voor differentiatie tussen bepaalde verwante ziektebeelden. Ook maakt het nog verschil of met screening bedoeld wordt, een breed opgezet onderzoek naar pathologische verschijnselen bij één persoon, dan wel of gesproken wordt over bevolkingsonderzoek ter opsporing van dragers van een bepaalde ziekte. In het algemeen zal echter aan leden van de referentiepopulatie de eis gesteld worden dat ze gezond zijn. Het begrip gezond is echter een zeer moeilijk begrip. Het houdt volgens de definitie van de World Health Organisation (WHO) in :

"A state of complete physical, mental and social well-being and not merely the absence of disease or infirmity."

Bovendien wordt in de medische begripsvorming de dichotomie gezond-ziek steeds meer vervangen door een continuum. Uitgaande van zo'n glijdende overgang van de ene naar de andere kategorie is het denkbaar, dat iemand volgens allerlei klinische en biochemische onderzoeken nog als gezond bestempeld wordt, terwijl sommige testen toch al pathognomonisch zijn voor toekomstige ziekte. Het ondubbelzinnig vaststellen van gezondheid is dus een welhaast onmogelijke opgave. In de literatuur worden dan ook zeer verschillende definities van gezondheid gehanteerd. In de meeste gevallen (3, 15, 76, 92, 96) wordt er van uitgegaan dat, wanneer iemand normaal werkt en zich gezond voelt (subjectief), er sprake is van gezondheid. Dit is natuurlijk een zeer eenvoudige voorstelling van zaken, die echter in de praktijk makkelijk gehanteerd kan worden. Ook wordt er nogal eens vanuit gegaan, dat bloeddonoren gezond zijn (o.a. 8, 39, 65). In landen waar donoren niet betaald worden voor een donatie, zal dit vaak inderdaad het geval zijn.

In landen echter waar gehonoreerde donoren voor komen, is het af te raden deze groep als gezond te beschouwen, wegens het grote percentage alcoholisten en drugsverslaafden en andere ongezonden met geldgebrek (2, 90).
Bovendien, ook al kunnen donoren als gezond beschouwd worden, kan deze groep niet zonder meer voor alle bepalingen als referentiepopulatie gebruikt worden. Een bekend voorbeeld is de ferritine bepaling in serum : door het regelmatig afstaan van bloed zal een gedeelte van de donoren een verlaagd ferritine gehalte hebben.

In enkele publicaties wordt een meer uitgebreide definitie van gezondheid gegeven (82) :

"The normal population consisted of Mayo-clinic employees, whose health was confirmed by routine physical examination. Exclusion criteria included evidence of pulmonary disease, cardiac disease, hepatic disease, gastro-intestinal disease, endocrine disease, neurologic disease, hypertension (defined by systolic blood pressure greater than 90 mm Hg), obesity (greater than 15% of ideal weight), gout, cancer, infection (including upper respiratory infection) hysterectomy or oophorectomy in women less than 50 years old, xanthalasma, hyperlipidemia and medication (except occasional use of aspirin). All electrocardiograms were normal."

"Systolic" in deze definitie is waarschijnlijk een verschrijving en moet zijn "diastolic".
Deze omschrijving lijkt erg volledig maar is in de dagelijkse praktijk natuurlijk vrijwel onhanteerbaar. Nog volledieger zijn echter de aanbevelingen van de Scandinavische vereniging voor klinische chemie (1), die zo uitgebreid zijn (33 pagina's) dat waarschijnlijk niemand ze ooit in zijn geheel zal kunnen gebruiken. De kenmerken waaraan alle leden van de referentie-

populatie moeten voldoen, kunnen ook nog andere objectief vast te stellen gegevens bevatten. Voorbeelden zijn: geslacht, leeftijdscategorie, tijd van bloedafname etc. Op al deze problemen wordt uitgebreid ingegaan door Hoeke (55).

Punt 2 waaraan de referentiepopulatie moet voldoen, nl. een aselecte steekproef uit de totale populatie vormen voor alle kenmerken die niet in punt 1 nader omschreven zijn, is zo mogelijk nog moeilijker te realiseren. Bijzonder illustratief in dit verband is hoofdstuk 9 uit het boek van Kendall en Stuart, dat handelt over het verkrijgen van aselecte getallen (64).
Een groep bloeddonoren, zo ze al voldoet aan de eis van gezondheid, voldoet meestal zeker niet wat dit punt betreft. Meestal komt de leeftijdsopbouw helemaal niet overeen met die van de totale bevolking ; het aantal mannen is over het algemeen veel groter dan het aantal vrouwen etc. Het niet aselect zijn van bijna iedere referentiepopulatie heeft Whitehead gebracht tot zijn veel (o.a. 56) geciteerde definitie van normaal :

".....dat een normaal persoon tot het mannelijk geslacht behoort, tussen 20 en 25 jaar oud is en medicijnen studeert."

Het selecteren van een groep mensen die gebruikt kan worden voor de bepaling van referentiewaarden, is dus geen eenvoudige opgave. Vandaar dat er vele pogingen gedaan zijn (3, 8, 20, 29, 35, 36, 58, 75, 84, 86, 87, 89, 90, 91, 96, 114, 121) om referentiewaarden te berekenen m.b.v. de totale hoeveelheid ongeselecteerde getallen die dagelijks in een klinisch chemisch laboratorium geproduceerd wordt. De gepubliceerde statistische methoden om dit mogelijk te maken (6, 10, 58, 86, 91) gaan alle uit van een aantal vooronderstellingen. De belangrijkste hiervan is, dat het grootste aantal uitslagen van een patiëntenpopulatie niet beïnvloed is door ziekte. Het grootste deel van de uitslagen is dus "normaal". De tweede veronderstelling is, dat de frequentieverdeling van de uitslagen afkomstig van gezonden,

Gaussisch is. Pathologische uitslagen zullen nu, vooral in de staarten van de frequentieverdeling, een vervorming van de Gausscurve veroorzaken. Met statistische methoden is het echter mogelijk (dit is de derde veronderstelling), om de oorspronkelijke Gausscurve te herkennen. Een aantal methoden (6, 91) gaat er bovendien van uit (vierde veronderstelling), dat de pathologische uitkomsten voor het overgrote deel aan één kant van de frequentieverdeling liggen en dat ook bekend is aan wélke kant.

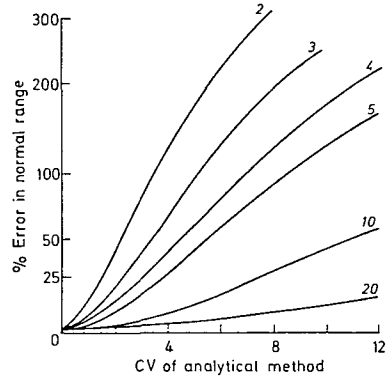
Het gebruik van deze zogenaamde "indirecte methoden" (3) om referentiewaarden te berekenen, heeft kritiek uitgelokt.
Elveback (29) richt haar bezwaren vooral op het feit, dat er bij alle indirecte methoden aangenomen wordt dat de frequentieverdeling Gaussisch is. Amador (3) concludeert uit het feit dat de referentiewaarden voor een aantal bepalingen, berekend uit "gezonden", niet gelijk zijn aan die, berekend uit patiënten, dat de indirecte methoden niet deugen. Hierbij moet aangetekend worden, dat een aantal van de bepalingen waarop de indirecte methoden werden toegepast, beslist niet Gaussisch verdeeld is. (Voorbeelden zijn de enzymbepalingen LDH (EC 1.1.1.27) en ASAT (SGOT; EC 2.6.1.1.), welke als parameters bij de diagnostiek en het verloop van het myocard infarct worden gebruikt.)
Uit andere publicaties over het vergelijken van referentiewaarden, berekend uit gezonden en patiënten, blijkt de overeenkomst zeer acceptabel te zijn (8, 20, 75).
Aangezien het "patiënt zijn" invloed kan hebben op de waarde van bepaalde parameters, is het zeer de vraag of een groep gezonden en een groep patiënten zonder meer vergeleken mogen worden (57). Zo kan door bedrust de concentratie van bepaalde stoffen in serum hoger worden (2). Deze aspecifieke effecten, die niet door ziekte veroorzaakt worden, moeten in de ligging van het referentiewaardengebied tot uitdrukking komen. In meer algemene zin is ook dit een punt, dat bij de vergeljking van referentiewaarden, berekend uit gezonden en patiënten, in het oog moet worden gehouden.

1.4. Factoren die van invloed (kunnen) zijn op de referentiewaarden.

Er zijn vele, deels onbekende, factoren, die de uitslag van klinisch chemische en hematologische bepalingen beïnvloeden. Het is gebruikelijk en ook logisch om bij de bepaling van referentiewaarden vooral rekening te houden met dié factoren, waarvan bekend is, dat de invloed ervan erg groot is. Zo is het gebruikelijk, om bij de beoordeling van een hemoglobine uitslag, rekening te houden met het geslacht van de patiënt, aangezien de waarden voor vrouwen gemiddeld lager liggen dan voor mannen. Wanneer echter bij de bepaling van de referentiewaarden voor hemoglobine geen splitsing naar geslacht gemaakt wordt, wil dat niet zeggen dat de resulterende waarden onbruikbaar zijn. Het referentiewaarden gebied is dan uiteraard wel veel breder. Een verhoging van het hemoglobine gehalte van een vrouw of een verlaging van dit gehalte van een man, zal dan veel groter moeten zijn om een bepalingsuitkomst buiten het referentiewaarden gebied op te leveren.
Het is een kwestie van keuze, met welke factoren bij de bepaling van referentiewaarden rekening wordt gehouden. De gemaakte keuze moet echter wel bekend zijn. In het algemeen zullen de referentiewaarden gebieden smaller worden, naarmate er met meer significante factoren rekening wordt gehouden (9).
Ook bij "ongeselecteerde" getallen kan een splitsing gemaakt worden naar geslacht, leeftijd, e.d. Het ongeselecteerde karakter van de getallen die gebruikt worden bij de zogenaamde indirecte methoden, is dus maar betrekkelijk.
De kritiek van Amador (3), dat de indirecte methoden geen rekening houden met deze factoren, is dan ook onterecht.

Factoren die de uitkomsten van bepalingen, en dus de ligging van het referentiewaardengebied, beïnvloeden, kunnen opgesplitst worden in :

-Niet-instrumentele invloeden. Hieronder vallen bijv.: leeftijd, geslacht, houding tijdens bloedafname, soort anticoagulans, dieet, rookgewoonten, tijd van de dag, inspanning, tijd in de mentruele cyclus, e.d. (21, 63, 71, 108).



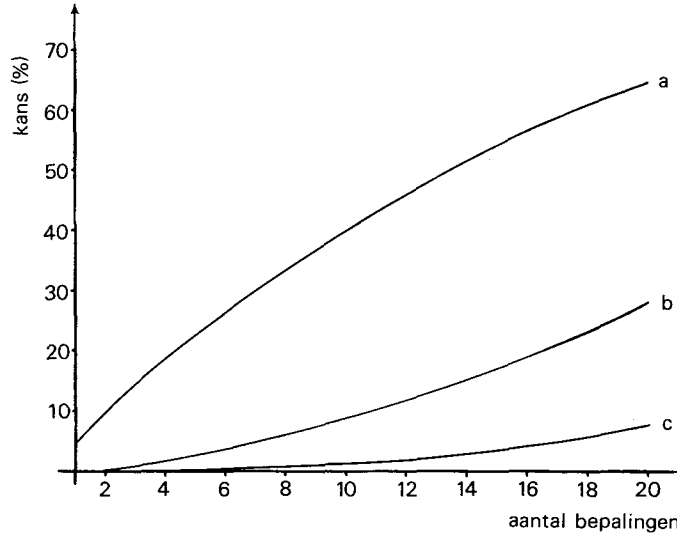
Figuur 1.
"The effect of increasing the C.V. of the analytical method (horizontal axis) on the normal range. Figures at the end of each line correspond to the true C.V. of the normal population. The vertical axis shows the percentage error of the observed C.V. of this normal population (log scale)."
Overgenomen van Gowenlock en Broughton (38), met toestemming.

-Instrumentele invloeden.

Naarmate de analytische variabiliteit stijgt, door welke oorzaak dan ook, wordt het referentiewaardengebied breder (38,68). Illustratief in dit verband is figuur 1, overgenomen van Gowenlock et al. (38). Hierin is aangegeven dat de breedte van het referentiewaarden gebied toeneemt, wanneer de variatiecoëfficiënt van de meting (CV, uitgezet op de X-as) groter wordt. Deze toename zal bij gegeven analytische variatie groter zijn, naarmate de oorspronkelijke breedte van het gebied (bij een meetfout gelijk aan 0) kleiner is. De oorspronkelijke breedte van het referentiewaarden gebied is, uitgedrukt als variatiecoëfficiënt, aangegeven aan het einde van iedere curve.

-Mathematische invloeden.

De ligging van de grenzen van het referentiewaardengebied wordt uiteraard beïnvloed door de definitie van het begrip "referentiewaarde" waarvan wordt uitgegaan en door de wijze van verwerking van de gegevens uit de referentiepopulatie (92,110). Van alle factoren die mogelijk van invloed zijn op de referentiewaarden van een bepaling, moet bekend zijn of er al dan niet rekening mee is gehouden. De eis is dus niet en kan vaak niet zijn, dat er rekening mee gehouden is, maar als er geen rekening mee gehouden is, moet dit feit bekend zijn.



Figuur 2.
De kans om minstens 1 (curve a), minstens 2 (curve b) en minstens 3 (curve c) "pathologische" uitkomsten te vinden bij N uitgevoerde onafhankelijke bepalingen.

1.5. Meerdimensionale referentiewaarden.

Stel, dat referentiewaarden voor 'n bepaling worden gedefinieerd als die waarden, waarbinnen de resultaten van 95% van de gezonde bevolking terecht komen. De kans, dat de uitslag van een

"gezonde" buiten het referentiewaardengebied valt, is dan 0,05. Wanneer nu een tweede bepaling wordt uitgevoerd, die onafhankelijk is van de eerste, dan is de kans om als gezonde buiten de referentiewaarden van deze tweede bepaling terecht te komen ook weer 0,05. De kans dat de uitslagen van beide bepalingen "normaal" zijn, is nu 0,95 x 0,95 = 0,9025 (101).

Hoe groter het aantal onafhankelijke bepalingen is, dat uit een serum- of bloedmonster gedaan wordt, des te groter wordt de kans op het vinden van minstens één pathologische uitslag. Dit zijn dus vals positieve uitslagen. Bij 13 à 14 uitgevoerde bepalingen wordt in ongeveer de helft van alle gevallen minstens één afwijkende uitslag gevonden (figuur 2).

Nog illustratiever is tabel 1, overgenomen van Simonson (107). Hierin is het aantal individuen getabelleerd dat onderzocht moet worden, om één persoon te vinden die voor 4, 8 en 16 onafhankelijke bepalingen uitkomsten heeft die respectievelijk tussen het 40ste en 60ste percentiel of tussen het 25ste en 75 ste percentiel liggen.

Volgens deze berekening moeten dus meer dan 150 miljard mensen (ongeveer 40 maal de wereldbevolking) onderzocht worden, om één

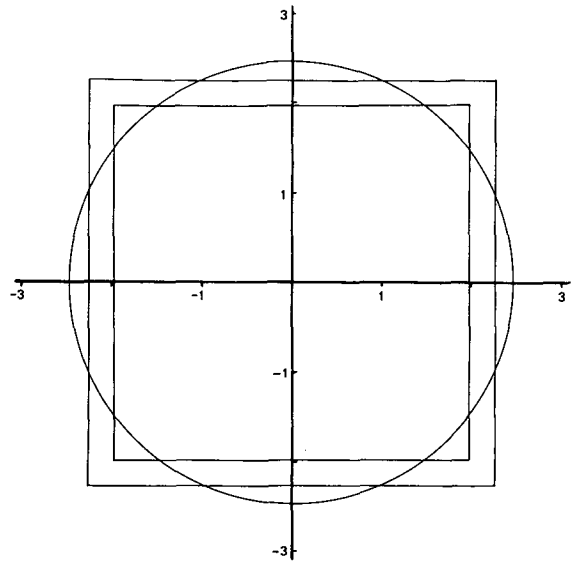
Tabel 1.
Het aantal personen, dat onderzocht moet worden om er één te vinden, die voor alle (onafhankelijke) bepalingen uitkomsten heeft tussen respectievelijk het 40ste en 60ste percentiel (kolom A) of tussen het 25ste en 75ste percentiel (kolom B). Overgenomen van Simonson (107), met toestemming.

aantal onafhankelijke bepalingen	aantal te onderzoeken personen om er één te vinden, die voor alle bepalingen uitkomsten heeft tussen	
	$P_{40} - P_{60}$	$P_{25} - P_{75}$
4	625	16
8	390.625	256
16	152.587.890.625	65.536

persoon te vinden, die voor 16 onafhankelijke bepalingen een uitkomst heeft tussen het 40ste en 60ste percentiel.
Simonson bekijkt het probleem van meerdimensionale referentiewaarden vanuit de cardiologie. Hij stelt, dat het normale ECG met 12 afleidingen (I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6), meer dan 75 onafhankelijke gegevens kan leveren. In het licht van het voorgaande is de kans, dat er iemand gevonden kan worden met een volkomen normaal ECG, dan ook uiterst minimaal.
Het uitvoeren van meer dan één bepaling in een bloed- of serummonster, leidt dus tot een zeer sterke toename van de kans om afwijkende uitslagen te vinden bij overigens gezonde personen. Om verschillende redenen is dit ongewenst. De belangrijkste is wel, dat het vinden van veel afwijkende uitslagen, die niet door ziekte veroorzaakt worden, nonchalance bij de aanvrager van het onderzoek in de hand kan werken, waardoor de kwaliteit van het medisch handelen daalt. Een andere belangrijke factor is, dat onnodig vervolgonderzoek wordt gestimuleerd, met als gevolg sterke stijging van de kosten van laboratoriumonderzoek, zonder dat daar extra rendement, in de vorm van betere patiëntenzorg, tegenover staat.

Healy (47) en Werner & Marsh (120) stellen voor, om het referentiewaardengebied van iedere afzonderlijke bepaling breder te maken, naarmate er meer bepalingen uit een monster worden uitgevoerd, en wel zodanig, dat de kans op het vinden van een pathologische uitslag gelijk blijft aan 5%. Bij één aangevraagde bepaling heeft het referentiewaardengebied dan een breedte van $3{,}92\ \sigma$ ($\mu \pm 1{,}96\ \sigma$) zodat 95% van alle uitslagen afkomstig van normalen, binnen dit gebied valt. Wanneer echter twee bepalingen worden aangevraagd, krijgt het referentiewaardengebied een breedte van $4{,}48\ \sigma$ ($\mu \pm 2{,}24\ \sigma$), zodat voor iedere bepaling afzonderlijk 97,47% van alle uitslagen afkomstig van gezonden binnen dit gebied ligt. De kans nu, dat de uitslagen van beide bepalingen normaal zijn, is echter 95%

$(0,9747^2 \times 100\%)$. Er wordt op deze manier bereikt, dat 95% van alle gezonden een niet afwijkende uitslagencombinatie krijgt toebedeeld. Een deel van de werkelijk pathologische uitkomstencombinaties, zal op deze manier echter ook als "normaal" beoordeeld worden.



Figuur 3.
Het conventionele tweedimensionale referentiewaardengebied (kleinste vierkant), het referentiewaardengebied volgens Healey (47) (grootste vierkant) en het correcte gebied (o.a. Chew (18)) voor twee onafhankelijke bepalingen ($\mu = 0$; $\sigma = 1$) (39).

In figuur 3 stelt het kleinste vierkant het gebied voor, waarbinnen een uitslagencombinatie van twee bepalingen als "normaal" wordt bestempeld : binnen het referentiewaardengebied van iedere bepaling afzonderlijk, liggen de uitslagen van 95% van alle gezonden voor die bepaling. De lengte en breedte van dit vierkant is 3,92 σ. Het grootste vierkant heeft een lengte en breedte van 4,48 σ, en is zo gekozen, dat 95% van alle uitslagencombinaties afkomstig van gezonden, erbinnen liggen. Het correcte tweedimensionale referentiewaardengebied, wanneer

de twee parameters een Gauss verdeling hebben en onafhankelijk van elkaar zijn, is echter het gebied binnen de cirkel, met een straal van $\sqrt{5,99}$ (zie 4.4) (18, 22, 95).
De theorie van meerdimensionale referentiewaarden wordt uitgebreid besproken door Chew (18). Wanneer de twee bepalingen niet onafhankelijk zijn, verandert de cirkel uit figuur 3 in een ellips. Hoe hoger de correlatie van de uitkomsten van de bepalingen onderling is, des te slanker wordt de ellips.
Voor combinaties van meer dan twee bepalingen wordt het zeer moeilijk, zo niet onmogelijk, een grafische voorstelling van het meerdimensionale referentiewaardengebied te verkrijgen. De gepubliceerde toepassingen beperken zich dan ook tot "zinnige" combinaties van bepalingen, waarbij "zinnig" slaat op het algemeen bekende verband dat er tussen de betreffende bepalingen bestaat. De aard van dit verband kan verschillend zijn: zo zal bij een chronische nieraandoening een relatie bestaan tussen de serumgehaltes van ureum en creatinine op basis van een onvoldoende excretiefunctie van de nier, hoewel de stofwisseling van creatinine en ureum onderling niet gekoppeld zijn. De relatie tussen het Ca-gehalte van het bloed en het albumine gehalte is van geheel andere aard: hier speelt een rol de partiële binding van het Ca aan albumine.
Enkele gepubliceerde meerdimensionale referentiewaardengebieden zijn : IgG-albumine in liquor cerebrospinalis (7), thyroxine index-triiodo thyronine index-thyrotropine (62), Ca-albumine in serum (95) en de combinatie hemoglobine-Fe (122).
Ook echter, wanneer er niet direct een verband tussen verschillende parameters bestaat, is het zinvol om meerdimensionale referentiewaardengebieden toe te passen.

1.6. Samenvatting en probleemstelling.

Het begrip "normaal" heeft vele betekenissen. Zo betekent "normaal" in medische zin "niet pathologisch". In het algemeen

spraakgebruik is normaal synoniem met "ideaal" en in de statistiek tenslotte, betekent normaal meestal "Gaussisch verdeeld". Om deze verwarring te voorkomen, wordt niet meer de term "normaalwaarde" gebruikt, maar de term "referentiewaarde".

Een zeer algemene definitie van het begrip referentiewaarde is : "Dié getallen, die aangeven binnen welke grenzen de uitkomsten van een nader te specificeren percentage van de referentiepopulatie voor een nader aangeduide bepaling terecht komen."

De referentiepopulatie uit deze definitie moet voldoen aan twee eisen :

1. Alle leden van de referentiepopulatie voldoen aan bepaalde gezondheidskenmerken en eventueel aan een aantal objectief vast te stellen kenmerken zoals geslacht, leeftijd etc.
2. De samenstelling van de populatie moet zodanig zijn, dat wat betreft de niet in 1 omschreven kenmerken, gesproken kan worden van een aselecte steekproef uit de totale populatie.

In de praktijk blijkt het nu zeer moeilijk, zo niet onmogelijk, om aan deze twee eisen te voldoen. Het stringent selecteren van een groep mensen, die gebruikt kan worden voor de bepaling van referentiewaarden, is dan ook een vrijwel hopeloze opgave. Een pragmatische benadering, die het probleem van de selectie van een referentiepopulatie omzeilt, is, om met statistische methoden te pogen, uit de totale hoeveelheid ongeselecteerde getallen die dagelijks in een laboratorium geproduceerd wordt, dat gedeelte te selecteren, dat niet beïnvloed is door ziekte.
Een eerste doel van het in dit proefschrift beschreven onderzoek is, om de gepubliceerde methoden op dit gebied met elkaar te vergelijken.
Bij alle gepubliceerde methoden voor de berekening van referentiewaarden uit ongeselecteerde getallen, wordt aangenomen, dat de frequentieverdeling van deze getallen Gaussisch is. Een tweede doel van het hier beschreven onderzoek is nu, een methode te ontwikkelen waarbij deze veronderstelling niet gemaakt wordt.

2. Methoden voor de berekening van referentiewaarden.

2.1. Inleiding.

Onder een "homogene" populatie zal verstaan worden, een groep mensen die voldoet aan de criteria gesteld in 1.3.2. In het algemeen bestaat een homogene populatie uit "gezonden". De uitkomsten van een klinisch chemische bepaling, gevonden bij leden van een homogene populatie, kunnen in principe allemaal gebruikt worden voor de berekening van referentiewaarden voor die bepaling en populatie. De tot dusver gangbare methoden voor de berekening van referentiewaarden uit een homogene populatie, kunnen opgesplitst worden in twee groepen :

1. De methode, waarbij aangenomen wordt, dat de frequentieverdeling van de uitkomsten Gaussisch is of op een eenvoudige manier tot een Gaussische frequentieverdeling getransformeerd kan worden.
2. De methoden die niet uitgaan van een bepaalde frequentieverdeling : de zogenaamde verdelingsvrije methoden. Van deze groep zijn de percentielenmethode (30, 53) en de methode die uitgaat van de range van de waarnemingen, de meest bekende.

Een "inhomogene" populatie is een groep mensen, die wat betreft de gezondheidscriteria uit 1.3.2. niet allemaal gelijk zijn. Een inhomogene populatie bestaat meestal uit "gezonden" en "zieken", waarbij het grootste aantal tot de eerste categorie behoort (hoofdpopulatie) en een kleiner deel tot de tweede categorie (contaminerende populatie). De uitkomsten van een klinisch chemische bepaling, gevonden bij leden van een inhomogene populatie, kunnen niet zonder meer allemaal gebruikt worden voor de berekening van referentiewaarden voor de hoofdpopulatie. Op een statistische manier moet getracht worden, de hoofdpopulatie te scheiden van de contaminerende populatie(s). Alle gepubliceerde methoden om dit te bereiken (6, 10, 58, 86, 91) gaan uit van de veronderstelling, dat de resultaten gevonden bij

leden van de hoofdpopulatie een frequentieverdeling vertonen volgens Gauss. Deze veronderstelling beperkt de toepassing van deze methoden zeer (29).
Verdelingsvrije methoden voor de berekening van referentiewaarden kunnen niet toegepast worden bij een inhomogene populatie. Wanneer blijkt dat de bepalingsresultaten van de hoofdpopulatie niet verdeeld zijn volgens een Gauss curve, moet dus een andere veronderstelling over de vorm van de frequentieverdeling van deze deelgroep gemaakt worden.
In hoofdstuk 5 zal blijken dat de niet Gaussische frequentieverdelingen, die in de klinische chemie voor een vrij groot aantal bepalingen gevonden worden (zie tabel 4), goed beschreven kunnen worden met gamma verdelingen. Dit maakt het mogelijk om ook voor bepalingen waarvan de resultaten een "scheve" frequentieverdeling vertonen, referentiewaarden te berekenen, uitgaande van een inhomogene populatie.

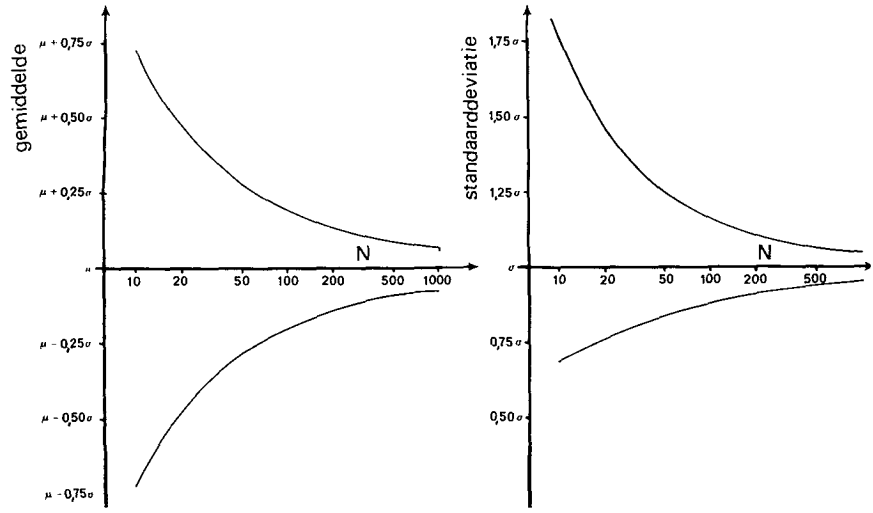
2.2. Methoden die uitgaan van een homogene populatie.

2.2.1. De frequentieverdeling is Gaussisch.

Wanneer aangenomen kan worden, dat de uitkomsten van een bepaling een Gauss curve vormen (voor toetsingscriteria, zie 5.2), is de voor de hand liggende methode om referentiewaarden vast te stellen, het berekenen van het gemiddelde ($\hat{\mu}$) en de standaarddeviatie ($\hat{\sigma}$) van uitkomsten, gevonden bij een aantal leden van de (homogene) populatie. Deze waarden zijn schattingen van het werkelijke gemiddelde (μ) en standaarddeviatie (σ) van de populatie. De referentiewaarden worden nu meestal gedefinieerd als $\hat{\mu} \pm 1{,}96\ \hat{\sigma}$ (waarbij 1,96 meestal wordt afgerond tot 2). Zo gedefinieerd, zal gemiddeld ongeveer 95% van de leden van de (homogene) populatie een bepalingsuitkomst binnen het referentiewaardengebied hebben ; van de resterende 5% heeft de helft een uitkomst boven de bovengrens en de andere helft een uitkomst

lager dan de benedengrens. Hoewel deze methode eenvoudig is, kunnen toch een aantal kanttekeningen gemaakt worden (99).

Uit een, meestal vrij klein, aantal waarnemingen, wordt een schatting $\hat{\mu}$ van het gemiddelde μ en een schatting $\hat{\sigma}$ van de standaarddeviatie σ, gemaakt. Wanneer het aantal waarnemingen zeer groot is, zullen de geschatte en de werkelijke waarden vrijwel gelijk zijn. Bij een kleiner aantal waarnemingen kan het verschil tussen geschatte en werkelijke waarden aanzienlijk zijn (figuur 4). Wanneer uit een vrij klein aantal waarnemingen op deze manier referentiewaarden worden berekend, zal het in de praktijk dan ook vaak voorkomen, dat veel meer of veel minder dan 5% van de resultaten, gevonden bij leden van de populatie, buiten de grenzen van het referentiewaardengebied vallen.
Bezemer (9) en Rümke en Bezemer (99) hebben voorgesteld, om in plaats van $\hat{\mu} \pm 1{,}96\,\hat{\sigma}$, als referentiewaarden dié grenzen te kiezen waarboven en waar beneden (met zekere betrouwbaarheid) ten-



Figuur 4.
A. 95% voorspellingsinterval voor de schatting $\hat{\mu}$ van μ als functie van het aantal waarnemingen (N).
B. Idem voor de schatting $\hat{\sigma}$ van σ.

minste 2,5% van de "normale" uitslagen ligt. Wanneer $\hat{\mu} \pm 1{,}96\,\hat{\sigma}$ als referentiewaarden worden aangehouden, ligt gemiddeld ongeveer 2,5% van de "normale" uitslagen boven en beneden deze waarden. Bij het voorstel van Bezemer en Rümke, ligt in de praktijk meestal méér dan 2,5% van de waarnemingen voorbij een of beide grenzen, maar dit verlies aan specificiteit wordt goedgemaakt door een toename van de gevoeligheid van de test. M.a.w. er wordt op tijd gewaarschuwd voor een mogelijk afwijkende uitslag en er wordt op de koop toe genomen, dat relatief vaak ten onrechte gewaarschuwd wordt.

Wanneer grenzen vastgesteld moeten worden voor descriptieve doeleinden, stellen dezelfde auteurs voor, om in plaats van de boven beschreven binnengrenzen (die grenzen waarbuiten met een zekere betrouwbaarheid minstens 5% van de waarnemingen ligt), buitengrenzen te berekenen: die grenzen waarbuiten met een zekere betrouwbaarheid maximaal 5% van de waarnemingen zich bevindt.
Door Lumsden (77) wordt voorgesteld om referentiewaarden voor bepalingen, uitgevoerd bij dieren, ook te definiëren als dié grenzen waarbuiten maximaal 5% van de uitkomsten van de bepaling ligt.

2.2.2. Er worden geen veronderstellingen gedaan over de vorm van de frequentieverdeling.

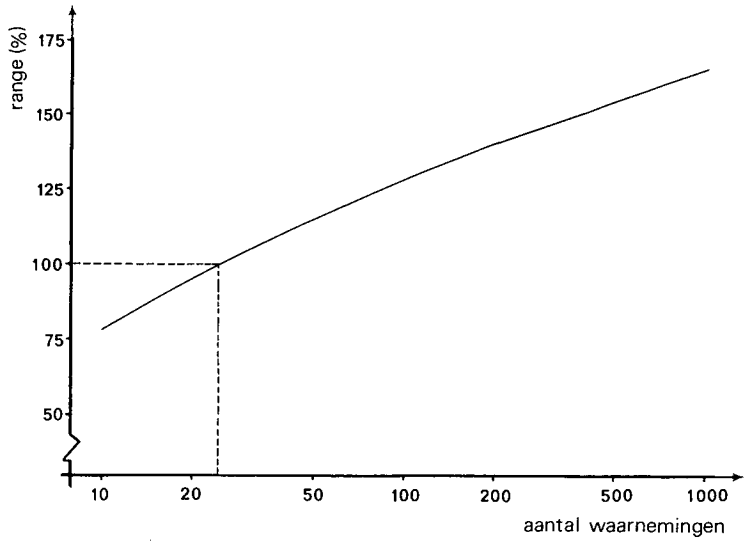
2.2.2.1. De range van de waarnemingen.

Een niet aanbevelenswaardige, hoewel ogenschijnlijk de meest eenvoudige, manier om referentiewaarden vast te stellen, is de volgende :
Men neemt de laagste en de hoogste waarneming uit de referentiepopulatie en definiëert deze range als referentiewaardengebied.
Hoewel deze methode op het eerste gezicht aantrekkelijk lijkt,

is het al spoedig duidelijk, dat de zo bepaalde range afhankelijk is van het aantal waarnemingen. De kans dat een extreme waarde gevonden wordt, neemt nl. toe, naarmate er meer waarnemingen gedaan worden. Bij deze methode wordt geen enkele veronderstelling over de vorm van de frequentieverdeling gedaan. Ze kan dus ook toegepast worden, wanneer de frequentieverdeling Gaussisch is. In dit speciale geval kan vrij eenvoudig worden berekend, hoe breed de range is als functie van het aantal waarnemingen (figuur 5). Het aantal waarnemingen, waarbij verwacht mag worden, dat de range overeenkomt met de gebruikelijke breedte van het referentiewaardengebied (3,92 σ), is dus ongeveer 25. Bij minder waarnemingen is de range te klein, bij meer waarnemingen te groot. Opgave van de range van de uitkomsten, gevonden bij "normalen", zonder het aantal te vermelden, is dus volstrekt zinloos.

2.2.2.2. De percentielenmethode.

Wanneer A% van alle waarnemingen uit een frequentieverdeling kleiner is dan X, dan wordt in de statistiek de waarde X het A^{de} percentiel genoemd. Voordat referentiewaarden kunnen worden bepaald m.b.v. de percentielenmethode, moeten eerst de waarnemingen in opklimmende grootte gerangschikt worden. Wanneer N het totale aantal waarnemingen is, dan is het A^{de} percentiel het (AxN/100)ste getal in de rij gesorteerde getallen. Stel bv. dat N=400, dan is het 2,5de percentiel gelijk aan het 10de getal uit de reeks naar grootte gesorteerde getallen en het 97,5de percentiel is gelijk aan het 390ste getal uit de reeks. Wanneer (AxN/100) geen geheel getal is, dan wordt meestal lineaire interpolatie toegepast om de waarde van het A^{de} percentiel te vinden (30, 53). Stel, er wordt gevonden, dat het 2,5de percentiel gelijk is aan het 8,4de getal uit de reeks waarnemingen. Voor de percentielwaarde wordt nu het 8ste getal vermeerderd met 40% van het verschil tussen het 9de en 8ste getal. Ook andere



Figuur 5.
De range van de waarnemingen (hoogste - laagste waarde) als percentage van de gebruikelijke breedte van het referentiewaarden gebied (3,92 σ), als functie van het aantal waarnemingen (N).

dan lineaire interpolaties zijn uiteraard mogelijk.

De percentielenmethode heeft als grote voordeel, dat er geen veronderstelling over de vorm van de frequentiecurve gedaan hoeft te worden. Het belangrijkste nadeel is, dat geen gebruik gemaakt wordt van alle informatie die in de getallen opgesloten ligt.

Bij een bepaald aantal waarnemingen uit bijv. een Gauss verdeling zijn de referentiewaarden, berekend met de percentielenmethode, dan ook minder betrouwbaar dan de referentiewaarden, berekend als $\bar{y} \pm 1{,}96\,\hat{\sigma}$. Wanneer dus informatie over de vorm van de frequentieverdeling voorhanden is, moet deze ook toegepast worden (30). Zo vinden Rümke en Bezemer (99) dat de zgn. binnengrenzen (die waarden, waarbuiten met een zekere betrouwbaarheid minimaal 5% van alle "normale" uitkomsten van een bepaling liggen) voor een Gauss verdeling, berekend met de percentielenmethode, nauwer zijn, dan binnengrenzen berekend met de methode

die gebaseerd is op een Gauss verdeling. M.a.w. om "zeker" te zijn dat minimaal 2,5% boven de bovengrens en beneden de benedengrens ligt, moet bij de percentielenmethode een extra verlies aan specificiteit op de koop toe worden genomen.

Anderzijds is het uiteraard zeer gevaarlijk om parametrische berekeningsmethoden te hanteren, wanneer goede informatie over de vorm van de verdeling ontbreekt (110, 119). In een dergelijk geval verdient een verdelingsvrije berekening, zoals de percentielenmethode, de voorkeur.

2.3. Methoden, uitgaande van een inhomogene populatie.

2.3.1. De methode van Hoffmann.

Iedere Gauss curve, met gemiddelde μ en standaarddeviatie σ, kan getransformeerd worden tot de standaard Gauss curve met $\mu=0$ en $\sigma=1$, door de x-waarden om te rekenen als volgt: $c = (x-\mu)/\sigma$. Wanneer c uitgezet wordt tegen x, ontstaat een rechte lijn met helling $1/\sigma$. Bij iedere waarde van c kan de oppervlakte onder de Gauss curve van $-\infty$ tot c, uitgedrukt als percentage van de totale oppervlakte onder de curve, berekend worden. Wanneer de c-waarden lineair worden uitgezet, dan vormen de erbij horende percentages een zogenaamde probit- of waarschijnlijkheidsschaal. Deze schaal is uiteraard niet-lineair. Wanneer nu getallen, aselect getrokken uit een Gauss verdeling, in opklimmende grootte gerangschikt worden, zal het percentage getallen tot een bepaalde waarde x, uitgezet op waarschijnlijkheidspapier (papier, waarvan één as verdeeld is volgens een waarschijnlijkheidsschaal) een rechte lijn opleveren. Wanneer geen waarschijnlijkheidspapier beschikbaar is, kunnen eerst de berekende percentages met behulp van een tabel omgezet worden in een probit waarde (o.a. 22). Deze probit waarden kunnen lineair uitgezet worden tegen x.

Hoffmann (57, 58) heeft nu gemeend, dat het gebruik van waarschijnlijkheidspapier voor de bepaling van het gemiddelde en de standaarddeviatie ook geoorloofd was, wanneer een gedeelte van de waarnemingen afkomstig is uit een andere dan de hoofdpopulatie. (De hoofdpopulatie wordt in de praktijk meestal gevormd door "gezonden" ; waarnemingen niet afkomstig uit deze populatie, kunnen daarom voor het gemak aangeduid worden als "pathologisch".) Juist door het relatieve karakter van deze plot, waarbij uit cumulatieve frequenties, met behulp van een totaal aantal, percentages worden berekend, kan het berekende gemiddelde en standaarddeviatie van de hoofdpopulatie niet meer zijn dan een vrij grove schatting. Dit wordt ook door Hoffmann zelf ingezien (58) :

"The slope of the line which would be used to estimate the normals is not quite as steep as it should be. This means that a graphic estimate of the standard deviation, on which the limits of clinical normality are based, will be a little too broad, but if the example is reasonable the bias is not very large."

Niet alleen wordt echter de helling van de Hoffmann plot afgeplat, wanneer een gedeelte van de getallen pathologisch is, ook het gemiddelde verschuift naar de kant (te hoog of te laag) waar het grootste gedeelte van de pathologische getallen ligt. Hoffmann adviseert dan ook, om de 5% grens, die uit de grafiek wordt afgelezen ($\mu - 1{,}645\sigma$), te benoemen als $\mu - 1{,}96\sigma$ en de 95% grens ($\mu + 1{,}645\sigma$) als $\mu + 1{,}96\sigma$. In figuur 6 is een voorbeeld uitgewerkt, waaruit de moeilijkheden van de Hoffmann methode duidelijk naar voren komen.
Samengevat zijn de nadelen van de Hoffmann methode, het feit dat aangenomen wordt dat de frequentieverdeling Gaussisch is en het relatieve karakter van de plot. Door dit relatieve karakter van de methode, heeft een onjuiste waarde voor het totale aantal waarnemingen (doordat een gedeelte van de waarnemingen pathologisch is), invloed op alle berekende percentages. Hierdoor

wordt het verloop van de hele curve verstoord. In de publicaties, waarin van de Hoffmann methode gebruik wordt gemaakt (3, 29, 70, 87, 89, 96), komen de auteurs dan ook vrijwel allen tot de conclusie, dat deze methode nauwelijks geschikt is voor het berekenen van referentiewaarden uit ongeselecteerde bepalingsuitkomsten. De voordelen van de Hoffmann plot zijn, dat de methode eenvoudig is en dat er vrij weinig (ongeveer 100) getallen nodig zijn om een vloeiende curve te produceren. Deze voordelen zijn echter alleen van belang, wanneer het overgrote deel van de waarnemingen afkomstig is van "gezonden". In dat geval levert het eenvoudig berekenen van gemiddelde en standaarddeviatie even betrouwbare referentiewaarden op.

2.3.2. De methode van Neumann.

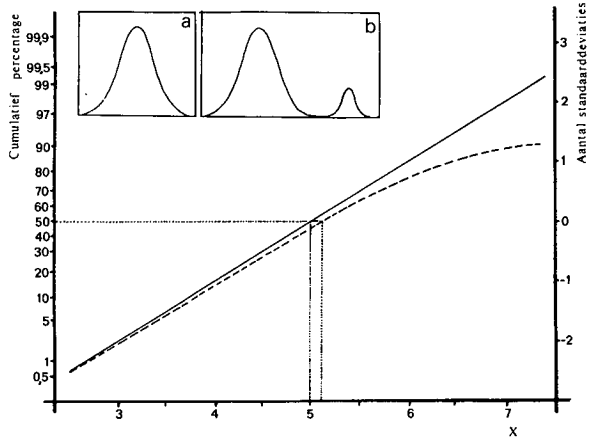
Het totale aantal waarnemingen bij een inhomogene populatie, bestaat uit de som van het aantal "normale" en het aantal "pathologische" getallen. Ieder percentage in de Hoffmann plot zal dus afwijken van de juiste waarde, omdat het totale aantal "normalen" niet precies bekend is. Neumann (86) heeft dit heel duidelijk ingezien en een methode voorgesteld om het aantal normalen in een gemengde populatie te schatten.

Wanneer aangenomen wordt, dat de pathologische uitkomsten aan één kant van de frequentieverdeling liggen, verloopt de berekening in detail als volgt (zie figuur 7 en tabel 2) :
Komende vanaf de ongecontamineerde kant van de frequentieverdeling, worden op de gebruikelijke manier de cumulatieve percentages uitgezet op waarschijnlijkheidspapier. Ten gevolge van de "pathologische" bepalingsuitkomsten in de staart van de frequentieverdeling, zal deze plot op een bepaald punt duidelijk gaan afwijken van een rechte lijn (punt A in figuur 7). Het punt waar de curve afbuigt, wordt als afkappunt genomen (truncation limit): het aantal waarnemingen beneden deze grens wordt op 100%

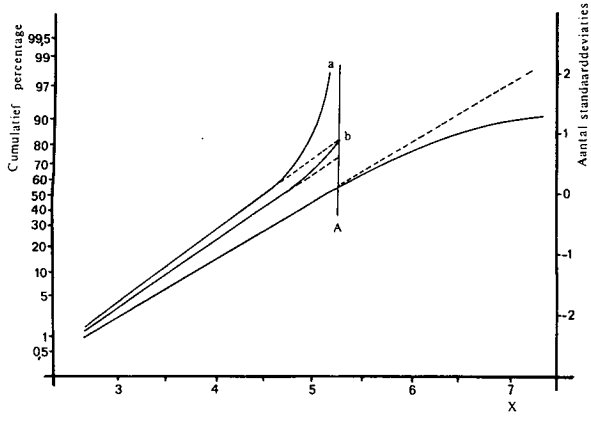
Tabel 2.
Voorbeeld voor de Hoffmann- en Neumann methode.

x	cumulatief percentage bij N = 100	cumulatief percentage bij N = 110	Neumann 1 N = 60	Neumann 2 N = 74	Neumann 3 N = 82
2,67	1	0,91	1,7	1,4	1,2
2,95	2	1,82	3,3	2,7	2,4
3,12	3	2,73	5,0	4,1	3,7
3,35	5	4,55	8,3	6,8	6,1
3,72	10	9,09	16,7	13,6	12,2
4,16	20	18,18	33,3	27,2	24,3
4,48	30	27,27	50,0	40,8	36,5
4,75	40	36,36	66,7	54,4	48,7
5,00	50	45,45	88,3	68,0	60,8
5,25	60	54,55	100	81,6	73,0
5,52	70	63,64			
5,84	80	72,73			
6,28	90	81,82			
6,64	95	86,36			
6,88	97	88,18			
7,05	98	89,09			
7,33	99	90,00			

Het cumulatieve percentage (van het totale aantal N) in kolom 2 is het percentage waarnemingen, dat kleiner of gelijk is aan de overeenkomstige x-waarde in kolom 1 (zie figuur 6A). Wanneer door pathologisch hoge uitkomsten (figuur 6B) het totale aantal waarnemingen 10% hoger wordt, veranderen de cumulatieve percentages in de waarden van kolom 3. Zoals uit figuur 6 blijkt, heeft de resulterende Hoffmann plot een kleinere helling dan de theoretische en wordt de 50% lijn bij een hogere x-waarde doorsneden. Bovendien is de lijn slechts lineair tot een x-waarde van 5,25, overeenkomend met 54% van het totale aantal waarnemingen. Dit punt is het afkappunt in de Neumann methode (zie figuur 7). Het aantal waarnemingen kleiner dan 5,25 wordt nu gelijk gesteld aan 100%, zodat de cumulatieve percentages van kolom 4 ontstaan. De Hoffmann plot (lijn a in figuur 7) buigt in de buurt van het afkappunt steil naar boven (naar 100%) af. Het lineaire stuk van lijn a wordt nu geëxtrapoleerd naar het afkappunt. Bij dit punt wordt een percentage afgelezen van 81,6% Dit percentage wordt nu geïnterpreteerd als : het gebruikte totaal (N=60) is niet 100% van het werkelijke totaal, maar slechts 81,6%. Een betere benadering van het werkelijke totaal is nu 60/0,816 = 73,5. Met dit nieuwe totaal resulteert kolom 5 in bovenstaande tabel en lijn b in figuur 7. Deze procedure wordt herhaald totdat het berekende totaal niet meer verandert.



Figuur 6.
Voorbeeld bij de Hoffmann methode. Voor toelichting zie de tekst en tabel 2.



Figuur 7.
Voorbeeld bij de Neumann methode. Voor toelichting zie de tekst en tabel 2.

gesteld (N1). Dit is een onderschatting van het aantal "normalen", omdat er ook boven de truncation limit normale uitslagen liggen. Er wordt nu met behulp van dit nieuwe totale aantal een tweede Hoffmann plot geconstrueerd. Het resultaat is een lijn, die in de buurt van het afkappunt steil naar boven (naar 100%) afbuigt. Het rechte gedeelte van de lijn wordt geëxtrapoleerd naar het afkappunt, en uit het bij dit punt afgelezen percentage (P) wordt een nieuwe benadering van het totale aantal berekend als: N2 = N1/P x 100. Deze procedure moet een aantal malen worden herhaald, totdat de benadering van het aantal "normalen" in de frequentieverdeling niet meer verandert. De laatste Hoffmann plot geeft dan "correcte" grenzen voor het referentie interval.
De hier beschreven iteratieve berekening van het totale aantal "normalen" convergeert, althans voor de 260 theoretische verdelingen uit hoofdstuk 3. Hoeke (55) heeft gevonden, dat het berekende totale aantal voor sommige frequentieverdelingen over de juiste waarde heen schoot.

Het zal duidelijk zijn, dat de methode Neumann een grote verbetering inhoudt van de door Hoffmann voorgestelde procedure. Met de Neumann methode is het ook mogelijk, twee elkaar gedeeltelijk overlappende Gauss curves te scheiden, door eerst het gemiddelde en de standaarddeviatie van één curve uit te rekenen met de hierboven geschetste iteratieprocedure, vervolgens de zo gevonden frequentieverdeling af te trekken van de totale verdeling, en op het verschil dezelfde procedure nog eens toe te passen. Wanneer de verdelingen twee zuivere Gauss curves zijn, niet vertekend door "pathologische" uitkomsten, zal, na het vinden van de eerste verdeling en het aftrekken van deze van de totale verdeling, de Neumann plot, toegepast op het verschil, meteen een rechte lijn opleveren.

De veronderstelling, dat de pathologische uitkomsten aan één kant van de frequentieverdeling liggen, is onnodig. Dezelfde

iteratieve rekenwijze kan gevolgd worden, wanneer gewerkt moet worden met twee afkappunten.
Samengevat, is het grote voordeel van de Neumann methode, dat het aantal pathologische uitkomsten in de frequentieverdeling het berekende referentiewaardengebied in principe niet beïnvloedt, zoals bij de Hoffmann plot wél het geval is. Ook Neumann doet de veronderstelling, dat de frequentieverdeling Gaussisch is of op een eenvoudige manier tot een Gaussische verdeling getransformeerd kan worden. Pathologische uitkomsten kunnen zowel te hoog als te laag zijn. De tijdrovende iteratieve berekeningswijze kan als een nadeel beschouwd worden. De methode is o.a. toegepast door Richardson et al. (96), Amador et al. (3) en Hoeke (55).

2.3.3. De Pryce conventie.

Bij de methode die door Pryce (91) is voorgesteld, wordt het volgende aangenomen :

1. De frequentieverdeling is Gaussisch, of kan op een eenvoudige manier tot een Gauss verdeling getransformeerd worden.
2. De pathologische uitkomsten komen slechts aan één kant van de frequentieverdeling voor.
3. Het is bekend, aan wélke kant de pathologische uitkomsten voorkomen.

De berekening van de referentiewaarden verloopt dan als volgt :
De uitkomst met de hoogste frequentie (de modus) wordt als schatting van het gemiddelde genomen. Aangezien bij een Gauss verdeling modus, mediaan (de middelste waarneming) en gemiddelde samenvallen, is dit geoorloofd, mits de frequentieverdeling van "pathologische" uitkomsten de frequentieverdeling van "normale" uitkomsten niet te ver overlapt. Aan één kant van het gemiddelde (links of rechts, dit is voor de betreffende bepaling bekend) zijn alle uitkomsten "normaal". Dit aantal, afgetrokken van het aantal waarnemingen aan de andere kant van het gemiddelde, geeft

het aantal "pathologische" uitkomsten in de frequentieverdeling. Het verschil tussen het gemiddelde en het 16de percentiel (wanneer de "pathologische" uitkomsten aan de hoge kant van de frequentieverdeling liggen) of het verschil tussen het 84ste percentiel en het gemiddelde (wanneer de "pathologische" uitkomsten aan de lage kant liggen), wordt als schatting genomen van de standaarddeviatie.

Wanneer de pathologische uitkomsten zowel in het hoge als in het lage gebied voorkomen, stelt Pryce voor om de helft van de afstand tussen de punten, waarbinnen de centrale 68% van de waarnemingen ligt, als standaarddeviatie aan te nemen. Deze schatting van de standaarddeviatie zal in de praktijk echter steeds te hoog zijn, omdat het totale aantal "normale" uitkomsten in dit geval niet bekend is. Als benadering van deze waarde wordt het totale aantal waarnemingen (normale + pathologische) genomen. Voor tweezijdig gecontamineerde frequentieverdelingen zal de Pryce conventie dan ook onjuiste referentiewaarden opleveren. In tabel 3 is een aantal klinisch chemische en hematologische bepalingen onderverdeeld in drie kategorieën: bepalingen, waarvan de pathologische uitkomsten meestal aan één kant voorkomen (te laag of te hoog) en bepalingen, waarvoor dit niet het geval is. Voor deze laatste categorie kan de Pryce conventie dus niet gebruikt worden. De methode is o.a. toegepast door Amador et al. (3), Cook et al. (20) en Payne et al. (90).

2.3.4. De methode van Becktel.

De methode van Becktel (6) is in principe gelijk aan de door Pryce (91) voorgestelde procedure. De berekening van het gemiddelde en de standaarddeviatie verloopt echter iets anders.

De waarnemingen worden naar opklimmende grootte gerangschikt. Uit het totale aantal waarnemingen (N) wordt het grootste oneven

Tabel 3.
Een aantal klinisch chemische en hematologische bepalingen, onderverdeeld in drie kategorieën : bepalingen, waarvan de "pathologische" uitkomsten meestal aan één kant (te laag of te hoog) van de frequentieverdeling liggen en bepalingen, waarvoor dit niet het geval is.

meestal te laag	meestal te hoog	onbeslist
albumine	alkalische fosfatase	APTT
totaal eiwit	ammoniak	bicarbonaat
erytrocyten	bilirubine	chloride
hemoglobine	bezinkingssnelheid erytrocyten	leucocyten
folaat	bloedingstijd	MCV
vitamine B12	CEA	MCH
kreatinine clearance	cholesterol	MCHC
overlevingstijd erytrocyten	CPK	natrium
ijzer	diastase	stollingstijd
totale ijzerbindings capaciteit	galactose	thyroxine (T4)
O_2 verzadiging	glucose	T3 index
pO_2	γGT	trombocyten
kalium	koolmonoxide (HbCO)	TSH
calcium	kreatinine	pH
trombotest (%)	LDH	pCO_2
	lipiden	actuele bicarbonaat
	fosfolipiden	base excess
	prolactine	cortisol
	reticulocyten	fibrinogeen
	ASAT	fosfaat
	ALAT	immuunglobulinen
	ureum	lactaat
	urinezuur	lithium
	zure fosfatase	magnesium
	zweetproef (Cl^-)	
	protrombinetijd (sec.)	

getal (J) berekend, dat kleiner of gelijk is aan 1/2 N. Er wordt nu het minimum bepaald van de volgende verschillen : $x_{(J)}-x_{(1)}$, $x_{(J+1)}-x_{(2)}$ ---------- $x_{(N)}-x_{(N-J+1)}$, waarbij $x_{(N)}$ het N^{de} getal in de rij gerangschikte getallen is. Als een schatting van het gemiddelde wordt nu de x-waarde genomen met een rangnummer gelijk aan het gemiddelde van de rangnummers van de x-waarden, waarvoor het kleinste verschil gevonden wordt. Wanneer echter de mediane waarneming verder af ligt van de "pathologische" staart van de frequentieverdeling, wordt deze waarneming als schatting van het gemiddelde genomen.
Omdat weer bekend is, aan welke kant van het gemiddelde de "pathologische" waarnemingen liggen, kan nu het totale aantal "normalen" in de frequentieverdeling berekend worden.

Nadat het 5de, 10de, 15de, 20ste, 25ste, en 30ste percentiel geschat zijn door lineaire interpolatie (wanneer de pathologische uitkomsten alleen in het hoge gebied voorkomen; wanneer de pathologische uitkomsten alleen in het lage gebied liggen, worden het 95ste, 90ste, 85ste, 80ste, 75ste en 70ste percentiel geschat) wordt de standaarddeviatie berekend als :

$$\hat{\sigma} = \frac{6\hat{\mu} - P_5 - P_{10} - P_{15} - P_{20} - P_{25} - P_{30}}{6}$$

waarbij : P_A = het A^{de} percentiel

Deze formule is eigenlijk een benadering en komt voort uit het feit, dat het 5de percentiel gevonden wordt bij $\mu - 1{,}65\,\sigma$, het 10de bij $\mu - 1{,}28\,\sigma$, het 15de bij $\mu - 1{,}04\,\sigma$, het 20ste bij $\mu - 0{,}84\,\sigma$, het 25ste bij $\mu - 0{,}66\,\sigma$ en het 30ste bij $\mu - 0{,}5\,\sigma$. Het totaal is gelijk aan $6\mu - 5{,}99\,\sigma$. Bij benadering geldt dus, dat $\mu - \sigma$ gelijk is aan de som van de genoemde percentielwaarden gedeeld door 6.

Een veronderstelling, die door Becktel expliciet vermeld wordt, is, dat het grootste deel van de pathologische uitkomsten niet te dicht bij het gemiddelde van de normalen ligt. Wanneer dit

wel zo is, kan de schatting van het gemiddelde fout uitvallen. De methode van Becktel is in de praktijk gebruikt door Best et al. (8).

Samengevat, hebben beide laatstgenoemde methoden het voordeel, dat ze erg eenvoudig zijn. Voor deze eenvoud wordt echter betaald door de extra veronderstelling (t.o.v. de Neumann methode), dat de pathologische uitkomsten slechts aan één kant van de frequentieverdeling voorkomen. Hoewel deze veronderstelling in de praktijk vaak gerechtvaardigd kan worden, zijn er ook vele bepalingen waarvoor deze niet geldt (zie tabel 3).

2.3.5. De methode van Bhattacharya.

Wanneer de resultaten van een bepaling geturfd worden in klassen van gelijke breedte h, en wanneer de frequentieverdeling van de resultaten Gaussisch is, kan aangetoond worden, dat de logaritme van het quotiënt van twee opeenvolgende frequenties ($\log (f_{i+1}/f_i) = \Delta\log$), uitgezet tegen het midden van de klasse i, een rechte lijn oplevert (y=a+bx, waarbij y= Δ log en x=klassemidden). Uit de helling (=b) en de doorsnijding van de X-as (=-a/b) van deze lijn, kunnen waarden voor het gemiddelde en de standaarddeviatie worden afgeleid (zie 4.1). Er geldt :

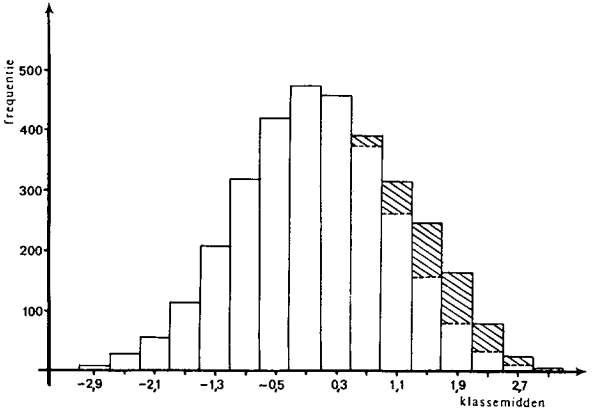
$$\hat{\mu} = -a/b + \tfrac{1}{2}h \qquad \hat{\sigma}^2 = -h/b - h^2/12$$

Voor het toepassen van de vergelijking voor $\hat{\sigma}^2$ moet gelden, dat de breedte van de klassen (h) niet al te groot is. In hoofdstuk 4 zal hier nader op worden ingegaan.

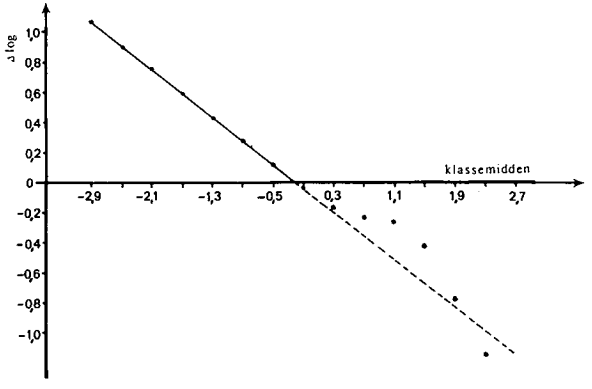
Een groot voordeel van de Bhattacharya plot (10) ten opzichte van de methoden van Hoffmann (58) en Neumann (86), is het feit, dat bij deze plot gerekend wordt met de waargenomen discrete

frequenties, in tegenstelling tot de relatieve en cumulatieve benadering van de twee andere methoden. Dit wil zeggen, dat, wanneer de frequentieverdeling van de uitkomsten van een bepaling opgebouwd gedacht wordt uit een zuivere Gauss curve voor de "normale" populatie, gedeeltelijk overlapt door 'n verdeling van "pathologische" waarden, deze waarden de Bhattacharya plot niet zullen beïnvloeden in dat gedeelte, waar er geen overlap van beide verdelingen bestaat. De berekening van $\hat{\mu}$ en $\hat{\sigma}$ wordt dus in feite teruggebracht tot het selecteren van het rechte gedeelte van de Bhattacharya plot. In het deel van de frequentieverdeling, waar er overlap bestaat tussen de "normale" en de "pathologische" frequentieverdeling, wordt de Bhattacharya plot meestal zeer onregelmatig (zie figuur 8).

De methode is door Bhattacharya in feite ontworpen, om elkaar gedeeltelijk overlappende Gauss curves te scheiden. Wanneer we te maken hebben met P elkaar overlappende Gauss curves, en wanneer er in de totale frequentieverdeling voor iedere Gauss curve een gebied is aan te wijzen, waar de invloed van de andere curves verwaarloosd mag worden, dan kunnen in de Bhattacharya plot ook P rechte stukken aangewezen worden, die rechtstreeks waarden opleveren voor de betreffende $\hat{\mu}$ en $\hat{\sigma}$. Ook in de situatie van twee Gauss curves (A en B), waarbij Gauss curve B volledig door Gauss curve A wordt overlapt, maar er wel een gebied bestaat, waar A niet beïnvloed wordt door B, kan de Bhattacharya plot waarden voor de twee gemiddelden en standaarddeviaties leveren. In eerste instantie zal een recht stuk in de plot gezien worden, dat het gevolg is van het ongecontamineerde deel van curve A. Uit dit recht stuk kan het gemiddelde en de standaarddeviatie van curve A berekend worden. Wanneer nu van de totale frequenties de frequenties van curve A afgetrokken worden, levert nogmaals toepassen van de Bhattacharya plot op het verschil, een tweede rechte op, die het gevolg is van curve B.



Figuur 8.
Voorbeeld bij de Bhattacharya methode.
A. Zuivere Gauss curve gecontamineerd met "pathologische" uitkomsten (geärceerd).



B. De Bhattacharya plot wordt in het ongecontamineerde deel niet beïnvloed door deze "pathologische" uitkomsten. De uit het rechte deel van de plot berekende waarden voor het gemiddelde en de standaarddeviatie zijn gelijk aan de theoretische.

Aangezien de Bhattacharya plot berust op het berekenen van een quotiënt van twee frequenties, en de relatieve fouten in de frequenties additief zijn in het quotiënt, moet het totaal aantal verzamelde getallen in de frequentieverdeling vrij groot zijn, om met succes deze plot te kunnen toepassen. In de praktijk kan als vuistregel worden aangehouden, dat met minder dan 1500 geturfde getallen de berekende referentiewaarden niet betrouwbaar zijn. Wanneer het percentage pathologische uitkomsten aanzienlijk is, moet het aantal verzamelde getallen nog groter zijn.

Samengevat geldt, dat de Bhattacharya plot, door het feit dat gerekend wordt met de waargenomen discrete frequenties, de nadelen van de Hoffmann benadering niet kent. Ten opzichte van de Neumann methode, die het grote bezwaar van de Hoffmann methode ook oplost, kan gelden, dat met de Bhattacharya plot geen iteratieve berekeningswijze nodig is, waardoor deze eenvoudiger kan worden toegepast. Een nadeel, voortkomend uit het specifieke karakter van de plot, is het feit, dat het aantal verzamelde getallen vrij groot moet zijn. Dit maakt de methode minder geschikt voor dié klinisch chemische bepalingen, die niet in grote series worden uitgevoerd. De Bhattacharya methode is onder andere toegepast door White (121), Gindler (35) en Naus et al. (84).

3. Vergelijking van methoden die uitgaan van een inhomogene populatie.

3.1. Keuze van methoden voor de vergelijking.

Bij het beschouwen van de methoden voor het berekenen van referentiewaarden m.b.v. ongeselecteerde patiëntengegevens, valt een aantal verschillen en overeenkomsten meteen op. Zo gaan alle methoden ervan uit, dat de frequentieverdeling van de bepaling voor de hoofdpopulatie Gaussisch is. Voor veel bepalingen in de klinische chemie en hematologie kan deze veronderstelling inderdaad gemaakt worden, voor andere zal ze beslist leiden tot onjuiste resultaten (tabel 4).

Tabel 4.
Een aantal klinisch chemische en hematologische bepalingen, waarvan de resultaten niet verdeeld zijn volgens een Gauss curve.

enzymen	andere bepalingen
ASAT	bilirubine
ALAT	bezinkingssnelheid erytrocyten
CPK	leucocyten
alkalische fosfatase	trombocyten
γGT	vitamine B12
	folaat

De methode van Hoffmann is in feite principiëel ongeschikt voor het berekenen van referentiewaarden uit een frequentieverdeling, waarin het percentage pathologische resultaten aanzienlijk is. Het totale aantal waarnemingen (dat gebruikt wordt voor het berekenen van de cumulatieve percentages) is te groot ; daardoor zullen alle berekende percentages te laag uitvallen, met als mogelijk gevolg een afwijkend gemiddelde (verschoven naar de kant waar de pathologische getallen liggen) en een te grote

standaarddeviatie. De Hoffmann methode zal dan ook om theoretische redenen niet betrokken worden in een vergelijking van methoden. Ook de Pryce conventie geeft (bij tweezijdig gecontamineerde frequentieverdelingen) geen juiste resultaten, omdat de punten waarbinnen 68% van de waarnemingen ligt rond het gemiddelde, niet goed te bepalen zijn. De reden is weer, dat het totale aantal waarnemingen te groot is door de pathologische uitslagen, waardoor alle percentages te laag zijn. De methode van Becktel is zelfs helemaal niet bruikbaar bij tweezijdig gecontamineerde verdelingen. Dat de methode van Neumann slechts bruikbaar zou zijn voor éénzijdig gecontamineerde verdelingen, zoals gesteld door Hoeke (55), is niet juist (86).

Op grond van het bovenstaande lijkt het zinvol, om in een vergelijking van methoden te betrekken: de Bhattacharya plot, de Neumann methode en de Pryce conventie. Deze laatste uitsluitend voor éénzijdig gecontamineerde verdelingen.

3.2. De wijze van vergelijking.

Amador et al. (3) hebben een aantal methoden voor de berekening van referentiewaarden vergeleken door ze toe te passen op : 1. ongeselecteerde patiëntengegevens en 2. gegevens van bloeddonoren. De toegepaste methoden waren : de methode van Hoffmann, de methode van Neumann, de Pryce conventie en nog een drietal andere. Deze werden toegepast op bepalingsuitslagen van : nuchtere bloedsuiker, ureum, ASAT (EC 2.6.1.1), LDH (EC 1.1.1.27) en cholesterol. De toegepaste methoden gaan er alle van uit, dat de frequentieverdelingen Gaussisch zijn. Toch zijn verscheidene bepalingen, waarop deze methoden werden toegepast, beslist niet Gaussisch verdeeld, bijv. ASAT en LDH. Om te bepalen of de berekende referentiewaarden uit ongeselecteerde patiëntengegevens de juiste waren, werden ze vergeleken met referentie-

waarden berekend uit donorbloed. Hier wordt echter een gevaarlijke stap gezet : er wordt verondersteld dat het gezonde gedeelte van de patiëntenpopulatie in alle opzichten vergelijkbaar is met de groep donoren. Hoewel gegevens over leeftijdsopbouw, verdeling over beide geslachten etc., ontbreken, lijkt het toch zeer onwaarschijnlijk, dat dit het geval zal zijn. Wanneer dan ook uit het niet gelijk zijn van de referentiewaarden berekend uit patiënten en donoren, geconcludeerd wordt, dat de rekenmethode niet deugt, is deze conclusie op zijn minst aanvechtbaar. Wanneer bovendien in aanmerking genomen wordt, dat verscheidene van de bepalingen, waarop de methoden zijn toegepast, niet Gaussisch verdeeld zijn, kan zonder meer gesteld worden, dat dit veel geciteerde artikel van Amador et al. aanvechtbaar is.

Bij het vergelijken van methoden voor de berekening van referentiewaarden, moet uitgegaan worden van waarnemingen uit een verdeling, waarvan de parameters (gemiddelde en standaarddeviatie, wanneer de verdeling Gaussisch is) vastliggen. De garantie, dat waarnemingen inderdaad afkomstig zijn uit een bekende verdeling, kan verkregen worden door aselect getallen te laten genereren door een computer, uit een nader te specificeren frequentieverdeling. Dit is o.a. gedaan door Hoeke (55). In zijn vergelijking werden betrokken : de methode van Hoffmann, de methode van Neumann, de Bhattacharya plot en de "best passende parabool" methode. Deze laatste methode berust op de waarneming van Hald, dat het klassemidden, uitgezet tegen de logaritme van de klassefrequentie, een parabool oplevert voor een Gauss curve. Aangezien dit niet wezenlijk verschilt van de benadering van Bhattacharya, waar het verschil van de logaritmen van de frequenties van twee opvolgende klassen, uitgezet tegen het klassemidden, een rechte lijn oplevert, is deze methode niet verder besproken. Hoeke heeft per experiment 1000 getallen laten trekken. Hij werkte met een hoofdpopulatie met gemiddelde 100 en standaarddeviatie 5 (N(100,5)) en met contaminerende populaties:

ofwel N(115, 5) ofwel N(130, 15). Er werden nu verschillende combinaties van hoofdpopulatie en contaminerende populatie gemaakt. Met behulp van de boven vermeldde methoden werd vervolgens bekeken, of inderdaad het gemiddelde (100) en de standaarddeviatie (5) van de hoofdpopulatie konden worden teruggevonden. Toepassing van de methode van Bhattacharya gaf "zeer teleurstellende resultaten". Er werd geen recht gedeelte in de plot gevonden. Dit is te wijten, enerzijds aan het kleine aantal (1000) getrokken getallen, en anderzijds aan het niet "smoothen" van de frequenties of toepassen van een gewogen kleinste kwadraten berekening. Door het verbinden van twee punten uit de plot, liggende rond een Δ log van 0, met een rechte lijn, heeft Hoeke toch schattingen gemaakt voor μ en σ. De schattingen voor μ kloppen heel redelijk, zoals verwacht kan worden, omdat dié twee punten uit de plot genomen worden, die het dichtste bij de doorsnijding van de X-as liggen. De schattingen van de standaarddeviaties kloppen minder goed, ook zoals te verwachten is, omdat er maar twee punten uit de plot genomen worden, die, door statistische fluctuaties in de frequenties, een heel andere helling dan de juiste kunnen opleveren.

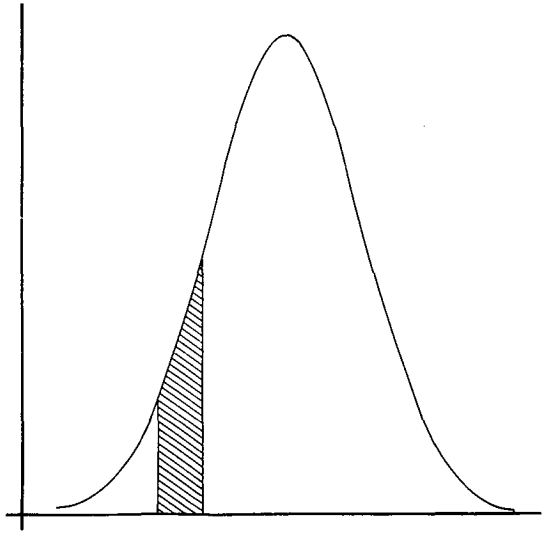
Over de best passende parabool methode wordt gezegd dat :
"de methode weliswaar minder gevoelig is voor 'random' variaties van de klassefrequenties dan de methode van Bhattacharya, maar dat de invloed van contaminatie beduidend groter is."
De methode van Bhattacharya is uiteraard gevoeliger voor random variaties, omdat de mogelijke relatieve fout in een quotiënt gelijk is aan de som van de relatieve fouten in teller en noemer. Bij de best passende parabool methode wordt niet de logaritme van een quotiënt van twee frequenties genomen, maar de logaritme uit de frequentie van iedere klasse afzonderlijk. Het is dus logisch, dat deze waarde minder gevoelig is voor statistische fluctuatie. Dat de invloed van contaminatie aanzienlijk is, komt waarschijnlijk door het feit dat voor de berekening van de parabool, een ongewogen kleinste kwadratenmethode toegepast wordt.

Samengevat kan gesteld worden, dat voor de vergelijking van de verschillende rekenmethoden, uitgegaan moet worden van door de computer gegenereerde getallen uit bekende verdelingen. De zo verkregen frequentieverdelingen kunnen al dan niet gecontamineerd worden met "pathologische" getallen. De gegevens van Hoeke, hoewel verkregen met een dergelijke benadering, geven geen uitsluitsel over de kwaliteit van de rekenmethoden, deels door gebrekkige uitwerking, deels door onjuiste toepassing van de methoden. De vergelijkingsmethode van Amador et al. (3) is volkomen ongeschikt, omdat geconstateerde verschillen het gevolg kunnen zijn van verschillen tussen groepen (patiënten - donoren) of van verschillen in rekenmethode. Het in alle opzichten vergelijkbaar zijn van het gezonde deel van de patiëntengroep en de groep donoren, is principiëel niet aan te tonen.

3.3. Het genereren van verdelingen.

Alle beschreven methoden zullen, wanneer ze worden toegepast op zuivere Gauss curves, ongetwijfeld de juiste waarden voor het gemiddelde en de standaarddeviatie opleveren. De vraag moet echter gesteld worden, hoe goed de resultaten zijn, wanneer de frequentieverdeling vervormd wordt. Deze vervorming kan in de praktijk het resultaat zijn van twee dingen : 1. Door statistische toevalligheden zal het aantal waarnemingen in iedere klasse een bepaalde spreiding vertonen. 2. De staarten van de, in principe zuivere, Gauss curve kunnen gecontamineerd worden door andere verdelingen (pathologische uitkomsten). Deze pathologische verdelingen kunnen nog zeer belangrijk zijn binnen de $\mu \pm 1{,}96\sigma$ grenzen van de oorspronkelijke verdeling. Door Hoeke (55) is m.b.v. een computer een groot aantal getallen getrokken uit een Gauss verdeling met bekende parameters. Wanneer deze getallen in klassen worden gesorteerd, zal het aantal per klasse onderhevig zijn aan de hierboven vermeldde statistische fluctuaties. Aan deze benadering, hoewel principiëel juist, kleven

twee nadelen. In de eerste plaats kost het trekken van aselecte getallen veel rekentijd. Vooral omdat ieder sampling experiment een aantal keren herhaald moet worden, kan dit een belangrijk nadeel zijn. In de tweede plaats moet aangetoond worden, dat de getrokken getallen inderdaad aselect zijn. Bij grote computers (zoals de Burroughs B7700) zijn de getrokken getallen inderdaad aselect. De "random number" generatoren van tafelcomputers leveren echter vaak getallen die zeer duidelijk niet aselect zijn. Een waarschuwing voor dit verschijnsel is niet overbodig.



Figuur 9.
Illustratie bij de afleiding, dat de standaarddeviatie van het aantal waarnemingen in iedere klasse, gelijk is aan de wortel uit dit aantal. Voor verdere toelichting zie de tekst.

Er is echter een eenvoudiger benadering mogelijk om Gauss verdelingen te construeren die door statistische fluctuaties vertekend zijn. Wanneer nl. een zuivere Gauss curve in klassen verdeeld wordt, dan zal de verwachting van het aantal in iedere klasse gevonden getallen, evenredig zijn met de oppervlakte van die klasse onder de Gauss curve. De standaarddeviatie van dit

aantal is bij benadering gelijk aan de wortel uit dit aantal. Dit laatste kan als volgt beredeneerd worden :

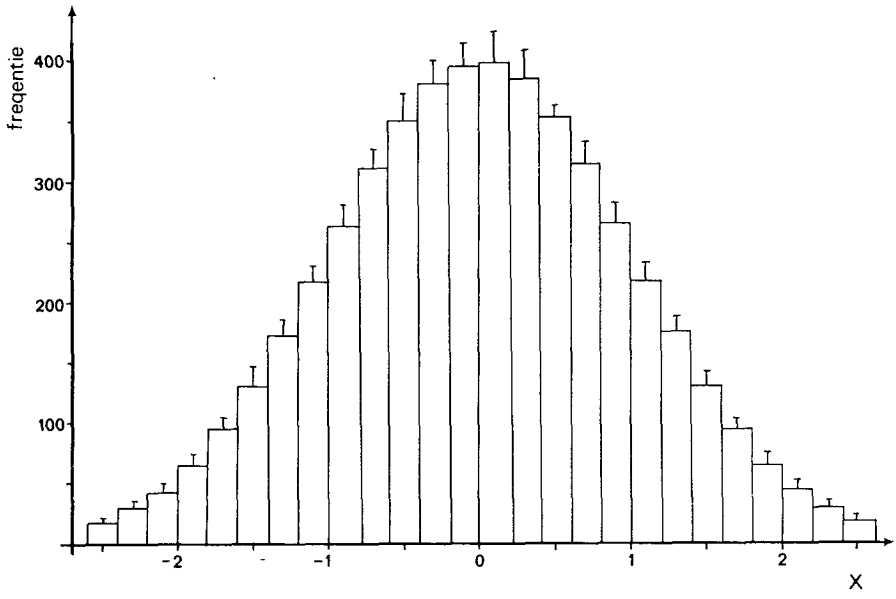
Wanneer de situatie van figuur 9 beschouwd wordt, is de kans om in het gearceerde gebied terecht te komen gelijk aan de oppervlakte van dit gebied, gedeeld door de totale oppervlakte onder de Gauss curve. Deze kans wordt aangeduid met p. De kans om buiten het gearceerde gebied onder de Gauss curve terecht te komen, is nu uiteraard gelijk aan (1-p). Wanneer nu n waarnemingen gedaan worden, is de kans dat er precies x in het gearceerde gebied liggen :

$$\binom{n}{x} p^x (1 - p)^{n-x} = P(n, x)$$

waarbij : $P(n, x)$ = kans, dat van n waarnemingen er precies x in het gearceerde gebied liggen.

Dit is de vergelijking van de binomiale verdeling. De verwachting van x is gelijk aan np en de standaarddeviatie is gelijk aan $\sqrt{np\ (1 - p)}$. Wanneer p klein is en n groot, kan de binomiale verdeling goed benaderd worden door de Poisson verdeling met gemiddelde np en standaarddeviatie gelijk aan $\sqrt{np}$.

Om nu door statistische invloeden vertekende Gauss curves te verkrijgen, is als volgt te werk gegaan :
Er werd uitgegaan van een theoretische Gauss curve met gemiddelde 0 en standaarddeviatie 1. Tussen -2,6 en 2,6 werden 26 klassen met een breedte van 0,2 geconstrueerd. De oppervlakte van iedere klasse onder de Gauss curve werd berekend m.b.v. de Chebyshev-Hermite polynomen, beschreven in hoofdstuk 4. De benadering werd doorgevoerd tot en met h^{10}. In de klasse met de kleinste oppervlakte betekent dit een nauwkeurigheid tot in meer dan 13 decimalen. De oppervlakten zijn berekend en niet opgezocht in een tabel, omdat de getabelleerde oppervlakten, vooral in de staarten van de frequentieverdeling, onvoldoende nauwkeu-



Figuur 10.
De gemiddelde frequenties in iedere klasse van de 20 theoretische, alleen door statistische invloeden vertekende, Gauss curves. Bovendien is van iedere klassefrequentie de standaarddeviatie aangegeven.

rige frequenties opleveren.

De berekende oppervlakten werden vermenigvuldigd met 5000 om frequenties te krijgen. Het totale aantal getallen in alle klassen samen is dus gelijk aan 4953.

Tot hier is de frequentieverdeling nog steeds zuiver. Om nu statistische fluctuaties een rol te laten spelen, werd bij de frequentie in iedere klasse, x $\sqrt{\text{frequentie}}$ opgeteld. De getallen x werden genomen uit een tabel van aselecte getallen, getrokken uit een standaard normale verdeling. Op deze manier werden 20 frequentieverdelingen berekend. Hiervoor waren dus 20 x 26 = 520 aselecte getallen nodig. Wanneer deze 20 frequentieverdelingen zouden zijn gegenereerd door een computer, zouden 20 x 4953 = 99060 aselecte trekkingen nodig zijn geweest. In figuur 10 is de range voor de frequenties in iedere klasse

aangegeven. Om nu frequentieverdelingen te verkrijgen die één- of tweezijdig gecontamineerd zijn met "pathologische" verdelingen, werd de volgende procedure gevolgd :
Van 2 van de 20 frequentieverdelingen, die alleen vervormd zijn door statistische fluctuaties, wordt de ene beschouwd als de hoofdpopulatie. Het gemiddelde van de andere verdeling wordt verplaatst van 0 naar 3,8. Bij de frequenties van de hoofdpopulatie wordt nu een bepaald percentage van de contaminerende (verschoven) populatie opgeteld. Aangezien er 20 oorspronkelijke verdelingen beschikbaar zijn, werden 20 links-éénzijdig gecontamineerde verdelingen en 20 rechts-éénzijdig gecontamineerde verdelingen berekend. Voor de percentages werd gekozen voor 10, 20, 30, en 40%, zodat er in totaal 160 éénzijdig gecontamineerde frequentieverdelingen ontstonden.
Voor tweezijdig gecontamineerde verdelingen was de procedure analoog. Hier werden 80 tweezijdig gecontamineerde frequentieverdelingen berekend. De details van de berekeningen zijn samengevat in tabel 5 en 6.

3.4. Het verwerken van de verdelingen.

Er zijn nu 260 berekende verdelingen :
20 "zuivere" Gauss curves met gemiddelde 0 en standaarddeviatie 1; 80 links-éénzijdig gecontamineerde Gauss curves; 80 rechts-éénzijdig gecontamineerde Gauss curves en 80 tweezijdig gecontamineerde Gauss curves. Met behulp van de methoden van Bhattacharya (10), Neumann (86) en Pryce (91) werden voor deze 260 verdelingen schattingen van het gemiddelde en de standaarddeviatie berekend.

Het principe van de 3 methoden is uiteengezet in hoofdstuk 2. Omdat het aantal te verwerken curves nogal groot is, en ook omdat de berekeningen nogal gecompliceerd zijn, werd besloten computerprogramma's voor de verschillende methoden te schrijven.

Tabel 5.
Berekening van rechts-éénzijdig gecontamineerde frequentieverdelingen.
i = nummer van de klasse.
j = volgnummer van de frequentieverdeling.

10 %	$f_{i,j} = f_{i,j} + 0{,}1\ f_{i-19,j+1}$	$20 \leq i \leq 26$	$1 \leq j \leq 19$
	$f_{i,20} = f_{i,20} + 0{,}1\ f_{i-19,1}$	$20 \leq i \leq 26$	$j = 20$
20 %	$f_{i,j} = f_{i,j} + 0{,}2\ f_{46-i,j+1}$	$20 \leq i \leq 26$	$1 \leq j \leq 19$
	$f_{i,20} = f_{i,20} + 0{,}2\ f_{46-i,1}$	$20 \leq i \leq 26$	$j = 20$
30 %	$f_{i,j} = f_{i,j} + 0{,}3\ f_{46-i,j+2}$	$20 \leq i \leq 26$	$1 \leq j \leq 18$
	$f_{i,19} = f_{i,19} + 0{,}3\ f_{46-i,1}$	$20 \leq i \leq 26$	$j = 19$
	$f_{i,20} = f_{i,20} + 0{,}3\ f_{46-i,2}$	$20 \leq i \leq 26$	$j = 20$
40 %	$f_{i,j} = f_{i,j} + 0{,}4\ f_{i-19,j+2}$	$20 \leq i \leq 26$	$1 \leq j \leq 18$
	$f_{i,19} = f_{i,19} + 0{,}4\ f_{i-19,1}$	$20 \leq i \leq 26$	$j = 19$
	$f_{i,20} = f_{i,20} + 0{,}4\ f_{i-19,2}$	$20 \leq i \leq 26$	$j = 20$

Tabel 6.
Berekening van tweezijdig gecontamineerde frequentieverdelingen. Bij de contaminatie percentages 10 en 20 moet frequentieverdeling 20 gecombineerd worden met 1. Bij de contaminatie percentages 30 en 40 moet frequentieverdeling 19 gecombineerd worden met 1 en frequentieverdeling 20 met 2.
Voor verdere toelichting zie tabel 5.

10 %	$f_{i,j} = f_{i,j} + 0{,}1 \begin{cases} f_{i+19,j+1} \\ f_{i-19,j+1} \end{cases}$	$1 \leq i \leq 7$	$1 \leq j \leq 19$
		$20 \leq i \leq 26$	$1 \leq j \leq 19$
20 %	$f_{i,j} = f_{i,j} + 0{,}2 \begin{cases} f_{8-i,j+1} \\ f_{46-i,j+1} \end{cases}$	$1 \leq i \leq 7$	$1 \leq j \leq 19$
		$20 \leq i \leq 26$	$1 \leq j \leq 19$
30 %	$f_{i,j} = f_{i,j} + 0{,}3 \begin{cases} f_{8-i,j+2} \\ f_{46-i,j+2} \end{cases}$	$1 \leq i \leq 7$	$1 \leq j \leq 18$
		$20 \leq i \leq 26$	$1 \leq j \leq 18$
40 %	$f_{i,j} = f_{i,j} + 0{,}4 \begin{cases} f_{i+19,j+2} \\ f_{i-19,j+2} \end{cases}$	$1 \leq i \leq 7$	$1 \leq j \leq 18$
		$20 \leq i \leq 26$	$1 \leq j \leq 18$

Vooral de Neumann methode levert bij het programmeren een aantal specifieke problemen op. Deze zijn: het vertalen van cumulatieve percentages in probits en het selecteren van de lineaire gedeelten in de plots.

Cumulatieve percentages kunnen op waarschijnlijkheidspapier direct uitgezet worden. Wanneer gebruik gemaakt moet worden van gewoon millimeterpapier, moeten deze percentages eerst omgezet worden in probit waarden met behulp van een tabel. In het computerprogramma voor de Neumann plot, werden deze probit waarden op de volgende manier benaderd :
'n Gauss curve wordt tussen -3 en +3 sd opgedeeld gedacht in 600 klassen met een breedte van 0,01 sd. De oppervlakte van iedere klasse wordt benaderd door :

$$\text{oppervlakte van klasse } i = 0{,}01 \left(\frac{f_i + f_{i+1}}{2} \right)$$

waarbij : i = nummer van de klasse $(1 < i < 600)$

$$f_i = \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2} (0{,}01\, i - 3{,}01)^2 \right)$$

De op deze wijze benaderde oppervlakte van een klasse met een breedte van 0,01 sd, heeft een maximale afwijking van de werkelijke oppervlakte van 0,003%. Deze benadering kan als voldoende nauwkeurig worden beschouwd, zodat toepassing van de Chebyshev-Hermite polynomen hier niet noodzakelijk is.
Om een cumulatief percentage nu te vertalen in een probit waarde, worden zoveel van de 600 klasse oppervlakten bij elkaar opgeteld, totdat de som van deze oppervlakten het cumulatief percentage juist overschrijdt. Het volgnummer van de klasse, waarbij dit gebeurt, is recht evenredig met de probitwaarde.

Om de grens van het lineaire gedeelte van de plot, en dus de afkappunten, te bepalen, werd als volgt te werk gegaan :
In 10 punten van de plot, rond het 50% punt (die volgens Hoff-

mann (58) het grootste gewicht moeten krijgen), wordt de helling bepaald door het verschil in Y-waarde van twee opvolgende punten. Van deze 10 hellingen wordt het gemiddelde en de standaarddeviatie berekend. 'n Afkappunt van de Neumann plot wordt nu gedefinieerd als dat punt, waarbij het verschil in Y-waarde tussen dit punt en het volgende, meer afwijkt van het gemiddelde verschil dan 2,5 sd.

Voor de Bhattacharya methode werd gebruik gemaakt van de gewogen kleinste kwadratenmethode, die in 4.2.2 wordt beschreven.

3.5. Resultaten en discussie.

De resultaten van de berekening van het gemiddelde en de standaarddeviatie van de 260 theoretische frequentieverdelingen zijn samengevat in tabel 7 (gemiddelde) en 8 (standaarddeviatie). Uit tabel 7 kan geconcludeerd worden, dat het gemiddelde, berekend met de methode van Bhattacharya, zeer weinig afwijkt van de theoretische waarde (maximaal 0,016). Het gemiddelde wordt enigszins verschoven naar de zijde waar de "pathologische" uitkomsten liggen. Deze verschuiving is bijna recht evenredig met het contaminatiepercentage. Het gemiddelde van de tweezijdig gecontamineerde verdelingen is vrijwel gelijk aan dat van de ongecontamineerde Gauss curves.
De methode van Neumann levert gemiddelden die maximaal 0,076 afwijken van de werkelijke waarde. Dit is meer dan 4 maal zoveel dan bij de methode van Bhattacharya. Verder valt op, dat de spreiding in de gemiddelde waarden bij de methode van Neumann groter is.
De Pryce conventie geeft voor iedere vorm van contaminatie, zoals gehanteerd in dit hoofdstuk, uiteraard hetzelfde gemiddelde. Voor het gemiddelde volgens Pryce werd immers het midden van de klasse met de hoogste frequentie gekozen.

Tabel 7.
Het gemiddelde van de 260 theoretische frequentieverdelingen berekend met de methoden van Bhattacharya (10), Neumann (86) en Pryce (91). Ieder gemiddelde in de tabel is het gemiddelde van 20 berekende waarden. Tussen haakjes is aangegeven de standaarddeviatie van deze 20 berekende waarden.

contaminatie	Bhattacharya	Neumann	Pryce
geen	0,009 (0,023)	-0,032 (0,036)	0,000 (0,189)
rechts 10%	0,010 (0,022)	-0,027 (0,054)	"
rechts 20%	0,012 (0,023)	-0,018 (0,043)	"
rechts 30%	0,013 (0,023)	-0,017 (0,052)	"
rechts 40%	0,016 (0,023)	-0,021 (0,058)	"
links 10%	0,008 (0,023)	-0,047 (0,037)	"
links 20%	0,007 (0,022)	-0,064 (0,044)	"
links 30%	0,005 (0,022)	-0,070 (0,042)	"
links 40%	0,003 (0,022)	-0,076 (0,049)	"
2 zijdig 10%	0,010 (0,022)	-0,019 (0,037)	"
2 zijdig 20%	0,009 (0,023)	-0,024 (0,039)	"
2 zijdig 30%	0,009 (0,022)	-0,043 (0,044)	"
2 zijdig 40%	0,010 (0,022)	-0,036 (0,053)	"

Tabel 8.
De standaarddeviatie van de 260 theoretische frequentieverdelingen, berekend met de methoden van Bhattacharya (10), Neumann (86) en Pryce (91). Iedere standaarddeviatie in de tabel is het gemiddelde van 20 berekende waarden. Tussen haakjes is aangegeven de standaarddeviatie van deze 20 berekende waarden.

contaminatie	Bhattacharya	Neumann	Pryce
geen	0,992 (0,032)	0,992 (0,037)	1,044 (0,072)
rechts 10%	0,995 (0,032)	1,026 (0,046)	0,991 (0,089)
rechts 20%	0,999 (0,032)	1,019 (0,038)	"
rechts 30%	1,004 (0,032)	1,027 (0,044)	"
rechts 40%	1,012 (0,033)	1,039 (0,062)	"
links 10%	0,995 (0,032)	0,997 (0,037)	0,991 (0,092)
links 20%	0,999 (0,032)	1,024 (0,040)	"
links 30%	1,004 (0,032)	1,008 (0,032)	"
links 40%	1,013 (0,033)	1,037 (0,062)	"
2 zijdig 10%	0,998 (0,032)	1,024 (0,033)	1,079 (0,076)
2 zijdig 20%	1,006 (0,032)	1,033 (0,037)	1,115 (0,079)
2 zijdig 30%	1,017 (0,033)	1,062 (0,055)	1,152 (0,082)
2 zijdig 40%	1,033 (0,034)	1,075 (0,049)	1,190 (0,085)

De standaarddeviatie (tabel 8), berekend met de Bhattacharya plot, wordt groter, naarmate de contaminatie toeneemt, ongeacht of deze links, rechts of beiderzijds bestaat. De oorzaak hiervan moet gezocht worden in de gewichten, die aan ieder punt toegekend worden. De onregelmatigheden in de Bhattacharya plot, die veroorzaakt worden door de "pathologische" getallen in de staarten van de frequentieverdeling, hebben toch, zij het een kleine, invloed op het verloop van de berekende rechte lijn.
De Neumann methode geeft waarden voor de standaarddeviatie, die gemiddeld ongeveer 3,5 maal zoveel afwijken dan de waarden van de Bhattacharya plot. Uit de spreiding van de standaarddeviatie blijkt weer, evenals bij het gemiddelde, dat de Neumann plot vaker uitschieters oplevert dan de methode Bhattacharya.
De standaarddeviaties berekend volgens de Pryce conventie, zijn uiteraard, voor wat betreft de éénzijdig gecontamineerde verdelingen, onafhankelijk van het contaminatie percentage. De standaarddeviatie wordt immers berekend uit de niet gecontamineerde helft van de verdeling. Bij de tweezijdig gecontamineerde verdelingen, neemt de berekende standaarddeviatie vrij sterk toe met het contaminatiepercentage.
Dit kan ook verwacht worden, omdat het gebied, waarbinnen zich 68% van alle waarnemingen bevindt, te breed wordt, door de te hoge schatting van het totale aantal waarnemingen. Dit totaal is te hoog, omdat alle "pathologische" uitslagen meegeteld worden.

De rekentijden voor de verschillende methoden verhouden zich ongeveer als 1 (Pryce) : 10 (Bhattacharya) : 50 (Neumann). Voor éénzijdig gecontamineerde verdelingen is de "prijs-kwaliteitsverhouding" van de methode van Pryce dan ook bijzonder hoog. De zeer lange rekentijden, nodig voor de Neumann methode, leveren zeker geen verbetering t.o.v. de Bhattacharya plot. Het integrale karakter maakt bovendien de methode weinig doorzichtig. Gemiddeld zijn de berekende gemiddelden en standaarddeviaties, zowel voor de Neumann- als voor de Bhattacharya methode, zeer

goed. De spreiding in de waarden van de Neumann plot is echter beduidend groter. In de praktijk wordt dan ook de voorkeur gegeven aan de Bhattacharya plot. Daar komt nog bij, dat voor niet-Gaussische frequentieverdelingen, de Bhattacharya plot makkelijker is aan te passen dan de Neumann methode. Op deze problematiek wordt verder ingegaan in hoofdstuk 5.

4. Meer gedetailleerde bespreking van de Bhattacharya plot.

4.1. Wiskundige achtergrond van de plot.

Bij een zuivere Gauss curve (gemiddelde = μ ; standaarddeviatie = σ) is de verwachting van het aantal waarnemingen in een klasse met breedte h en middelste waarde x, evenredig met de oppervlakte onder de curve tussen x - 0,5h en x + 0,5h. Mathematisch kan dit als volgt worden uitgedrukt :

$$f(x) = \int_{x-h/2}^{x+h/2} \frac{N}{\sigma\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}(v-\mu)^2/\sigma^2\right) dv \qquad \text{I}$$

Wanneer de substitutie $v = x + \sigma u$ uitgevoerd wordt, verandert vergelijking I in :

$$f(x) = \int_{-h/2\sigma}^{h/2\sigma} \frac{N}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}(x+\sigma u-\mu)^2/\sigma^2\right) du \qquad \text{II}$$

Wanneer $(x - \mu)/\sigma$ bovendien vervangen wordt door t en de kansdichtheidsfunctie van de standaard normale verdeling voorgesteld wordt door Z, kan vergelijking II vereenvoudigd worden tot :

$$f(x) = N \int_{-h/2\sigma}^{h/2\sigma} Z(t+u)\, du \qquad \text{III}$$

De integraal van Z kan niet exact opgelost worden, maar wanneer Z(t+u) vervangen wordt door de Taylor-reeks :

$$Z(t+u) = Z(t) + \frac{u}{1!} Z^{(1)}(t) + \frac{u^2}{2!} Z^{(2)}(t) + \ldots + \frac{u^n}{n!} Z^{(n)}(t) + \ldots \qquad \text{IV}$$

kan een benadering van de integraal gegeven worden. Wanneer vergelijking IV wordt ingevuld in III, ontstaat de volgende uitdrukking :

$$f(x) = N \sum_{n=0}^{\infty} \left[\frac{Z^{(n)}(t)}{n!} \int_{-h/2\sigma}^{h/2\sigma} u^n \, du \right] \qquad V$$

$Z^{(n)}$ is de n^{de} afgeleide van Z.

Er geldt :

$$Z^{(1)}(t) = -t\, Z(t)$$
$$Z^{(2)}(t) = (t^2 - 1)\, Z(t)$$
$$Z^{(3)}(t) = (-t^3 + 3t)\, Z(t)$$
$$Z^{(4)}(t) = (t^4 - 6t^2 + 3)\, Z(t)$$

De polynomen in bovenstaande uitdrukkingen zijn de zogenaamde Chebyshev-Hermite polynomen. Een meer compleet overzicht van de afgeleiden van Z wordt gegeven in (64).

Wanneer bovenstaande afgeleiden ingevuld worden in vergelijking V en wanneer termen met h^5 en hogere machten van h verwaarloosd worden, ontstaat een benadering van f(x), de verwachting van het aantal waarnemingen in een klasse met breedte h en middelste waarde x :

$$f(x) = \frac{hN}{\sigma\sqrt{2\pi}}\, e^{-t^2/2} \left(1 + \frac{t^2 - 1}{6} \left(\frac{h}{2\sigma}\right)^2\right) \qquad VI$$

$$\log\left(f(x)\right) = \log\left(\frac{hN}{\sigma\sqrt{2\pi}}\right) - \frac{t^2}{2} + \log\left(1 + \frac{t^2 - 1}{6} \left(\frac{h}{2\sigma}\right)^2\right) \qquad VII$$

Wanneer $|x| < 1$ geldt :

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \qquad VIII$$

Toepassen van vergelijking VIII op vergelijking VII en verwaarlozen van termen met h^4 en hogere machten van h, resulteert in :

$$\log (f(x)) = \log \left(\frac{hN}{\sigma\sqrt{2\pi}} \right) - \frac{t^2}{2} + \frac{t^2 - 1}{6} \left(\frac{h}{2\sigma} \right)^2 \qquad \text{IX}$$

Vergelijking IX kan anders gerangschikt worden tot :

$$\log (f(x)) = \log \left(\frac{hN}{\sigma\sqrt{2\pi}} \right) - \frac{h^2}{24\sigma^2} - \left(\tfrac{1}{2} - \frac{h^2}{24\sigma^2} \right) t^2 \qquad \text{X}$$

Voor de volgende klasse, met middelste waarde x + h, geldt, dat de eerste twee termen van log (f(x+h)) gelijk zijn aan de eerste twee termen van log (f(x)) in vergelijking X. In de laatste term van vergelijking X, moet t vervangen worden door t + h/ σ . De logaritme van het quotiënt van de frequenties in twee opvolgende klassen met middelste waarden resp. x en (x + h), kan dus uitgedrukt worden als :

$$\log \frac{f(x+h)}{f(x)} = \log (f(x+h)) - \log (f(x)) =$$

$$- \left(\tfrac{1}{2} - \frac{h^2}{24\sigma^2} \right) \left(\left(t + \frac{h}{\sigma} \right)^2 - t^2 \right) \qquad \text{XI}$$

Wanneer t weer vervangen wordt door (x - μ)/ σ , kan vergelijking XI herschreven worden als :

$$\log \frac{f(x+h)}{f(x)} = - \frac{h}{\sigma^4} \left(\sigma^2 - \frac{h^2}{12} \right) \left(x - \mu + \tfrac{1}{2}h \right) \qquad \text{XII}$$

Het grafisch uitzetten van log (f(x+h)/f(x)) tegen x, resulteert dus in een rechte lijn : y = ax + b. De helling (=a) en de doorsnijding van de X-as (=-b/a) worden gegeven door :

$$a = - \frac{h}{\sigma^4} \left(\sigma^2 - \frac{h^2}{12} \right) \qquad -b/a = \mu - \tfrac{1}{2}h \qquad \text{XIII}$$

Er is echter een uitdrukking nodig voor σ^2 als functie van de klassebreedte h en de helling van de rechte lijn a. Wanneer σ^2 vervangen wordt door p, kan vergelijking XIII herschreven worden tot :

$$ap^2 + hp - \frac{h^3}{12} = 0 \qquad \text{XIV}$$

De oplossing van deze eenvoudige kwadratische vergelijking wordt gegeven door :

$$p_{1,2} = \frac{-h \pm \sqrt{h^2 + 4a\frac{h^3}{12}}}{2a} \qquad \text{XV}$$

Wanneer $|ha/3|$ klein is, kan de wortel uit vergeljking XV benaderd worden door : $h\,(1 + ha/6)$. De twee oplossingen van vergelijking XIV zijn dus bij benadering :

$$p_1 = \frac{h^2}{12} \qquad p_2 = -\frac{h}{a} - \frac{h^2}{12} \qquad \text{XVI}$$

De eerste oplossing is onafhankelijk van de helling (a) en moet daarom verworpen worden.

Samenvattend :
Wanneer de resultaten van een bepaling geturfd worden in klassen van gelijke breedte h, en wanneer de frequentieverdeling van de resultaten Gaussisch is, geldt, dat de logaritme van het quotiënt van twee opeenvolgende frequenties, grafisch uitgezet tegen het midden van de eerste klasse, bij benadering een rechte lijn geeft (vergelijking XII). Het gemiddelde van de bepaling kan berekend worden uit de doorsnijding van deze lijn met de X-as (vergelijking XIII) en de standaarddeviatie uit de helling.

(vergelijking XVI). De laatste term in de vergelijking voor σ^2 is Sheppards correctie en komt voort uit het feit dat de klassebreedte niet oneindig klein is.

4.1.1. De breedte van de klassen.

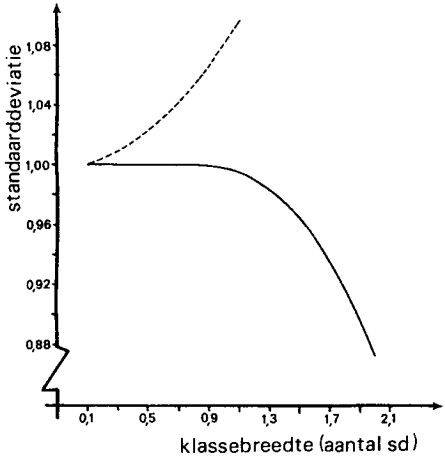
Aangezien in bovenstaand mathematisch exposé een aantal vereenvoudigingen is aangebracht, zijn de vergelijkingen, die een verband leggen tussen μ en σ^2 enerzijds en de doorsnijding van de X-as en de helling van de rechte lijn anderzijds, slechts benaderingen. De drie vereenvoudigingen zijn :

1. Het vervangen van Z(t+u) in vergelijking III door een Taylor reeks en het verwaarlozen van termen met h^5 en hogere machten van h.
2. Het vervangen van de logaritme in vergelijking VII door een machtreeks en het verwaarlozen van termen met h^4 en hogere machten van h.
3. De benadering van de wortel in vergelijking XV.

Deze drie vereenvoudigingen kunnen alleen maar worden aangebracht, wanneer h voldoende klein is. Om te bepalen hoe breed de klassen maximaal mogen zijn, is het volgende experiment opgezet:

Een Gauss curve met gemiddelde 0 en standaarddeviatie 1, werd verdeeld in 7 klassen met een breedte h. De middens van de klassen waren resp. -3h, -2h, -h, 0, h, 2h en 3h. De oppervlakten van de klassen werden berekend met behulp van de Taylor reeksontwikkeling uit vergelijking IV. Slechts termen met h^{11} en hogere machten van h werden verwaarloosd. Deze ver doorgevoerde benadering van de integraal is voldoende exact.
De zo berekende klasse-oppervlakten kunnen als theoretische frequenties worden beschouwd. Voor verschillende waarden van h tussen 0,1 en 2,0 werd nu een Bhattacharya plot geconstrueerd

met behulp van deze "exacte" frequenties. De helling en de doorsnijding van de X-as van iedere Bhattacharya plot resulteerde in waarden voor resp. $\hat{\mu}$ en $\hat{\sigma}$. Deze waarden werden vergeleken met de theoretische $\mu = 0$ en $\sigma = 1$. Het resultaat is getekend in figuur 11.
Uit deze figuur kan de conclusie getrokken worden, dat de berekende waarde voor σ voldoende nauwkeurig is, wanneer h

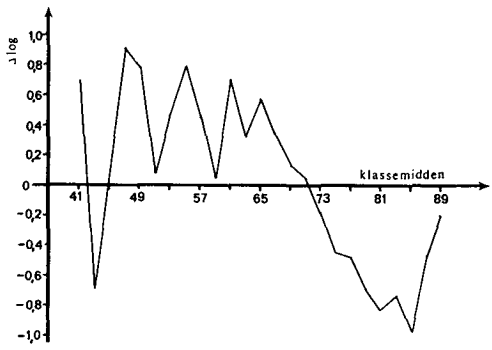


Figuur 11.
Berekende waarde voor σ mét (getrokken lijn) en zonder (stippellijn) Sheppard's correctie als functie van de breedte van de klassen.

kleiner is dan 1 σ. Voor grotere waarden van h is de standaarddeviatie, berekend uit de Bhattacharya plot, te klein. De berekende waarde van μ blijkt in dit experiment onafhankelijk te zijn van de breedte van de klassen.

4.2. Berekening van de Bhattacharya plot.

Zoals blijkt uit figuur 12, kan de Bhattacharya plot verdeeld worden in drie stukken, die duidelijk van elkaar kunnen worden onderscheiden : een onregelmatig deel, resulterend uit patholo-



Figuur 12.
Bhattacharya plot berekend uit 3804 bepalingsuitkomsten van totaal eiwit in serum. Duidelijk zijn de drie delen, waarin de plot uiteenvalt, te onderscheiden.

gisch lage uitkomsten, een recht stuk, dat het zuiver Gaussische deel van de frequentiecurve representeert, en een onregelmatig deel, resulterend uit pathologisch hoge uitkomsten.
De berekening van het gemiddelde en de standaarddeviatie van een bepaling met behulp van de Bhattacharya plot, wordt over het algemeen uitgevoerd als volgt :
-Het rechte deel van de plot wordt "op het oog" geselecteerd.
-Door de punten van dit rechte stuk wordt, met de methode der kleinste kwadraten een lijn berekend.

Tegen deze procedure zijn echter enkele bezwaren aan te voeren. In de eerste plaats kan het selecteren van het rechte stuk van de plot bijzonder moeilijk zijn, vooral wanneer het aantal resultaten in de frequentiecurve aan de lage kant is. Om ook in een dergelijk geval een recht stuk in de Bhattacharya plot te kunnen onderscheiden, kan op de waargenomen frequenties een smoothing-procedure worden toegepast (zie 4.3). Smoothing houdt echter ook in, dat in het hoge en lage gebied punten verloren gaan. Het aantal keren, dat een smoothing-procedure kan worden uitgevoerd, is, mede om deze reden, dan ook (zeer) beperkt. In de praktijk betekent dit meestal, dat, wanneer na één keer

smoothen van de frequenties nog steeds geen duidelijk recht gedeelte in de Bhattacharya plot kan worden onderscheiden, de conclusie getrokken moet worden, dat het aantal geturfde resultaten in de frequentieverdeling blijkbaar te klein is.
In de tweede plaats, wanneer het wél mogelijk is een recht stuk in de plot te onderscheiden, is het niet correct, om een rechte lijn door de punten van dit rechte stuk te berekenen met de (ongewogen) kleinste kwadraten methode. Ieder punt van de Bhattacharya plot is namelijk het resultaat van een quotiënt van twee frequenties en is dus betrouwbaarder naarmate deze frequenties hoger zijn. Bovendien speelt de afhankelijkheid van de punten een rol. Daarom is het beter een gewogen kleinste kwadratenmethode toe te passen.

Voordat het ontwikkelen van een weegprocedure voor de automatische berekening van de Bhattacharya plot besproken wordt, zal nu eerst een korte samenvatting worden gegeven van de kleinste kwadratenmethode (23).

4.2.1. Berekening van een rechte lijn met behulp van de kleinste kwadratenmethode.

Stel, dat een rechte lijn berekend moet worden door een aantal punten (x_i, y_i, waarbij x_i de onafhankelijke en y_i de afhankelijke variabele is. Er wordt aangenomen dat fouten alleen in y_i kunnen voorkomen.
In het algemeen wordt dié lijn als de beste beschouwd, waarbij de som van de kwadraten van de afstanden in verticale richting van de punten tot de lijn, minimaal is. Wanneer de te berekenen lijn voorgesteld wordt door y=ax+b, moeten bij de kleinste kwadratenmethode de waarden van a en b zodanig bepaald worden, dat:

$$\sum_{i=1}^{n} (y_i - (ax_i + b))^2 \qquad \text{minimaal is.} \qquad \text{XVII}$$

Dit minimum kan bepaald worden door vergelijking XVII partiëel te differentiëren naar a en b en deze twee afgeleiden gelijk te stellen aan 0.

$$-2 \sum_{i=1}^{n} (y_i - (ax_i + b)) x_i = 0 \qquad \text{XVIII}$$

$$-2 \sum_{i=1}^{n} (y_i - (ax_i + b)) = 0$$

In matrix notatie wordt dit :

$$X^T (Y - XB) = 0 \qquad \text{of :} \qquad B = (X^T X)^{-1} X^T Y \qquad \text{XIX}$$

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad B = \begin{bmatrix} a \\ b \end{bmatrix}$$

Vergelijking XIX kan algemeen worden toegepast en is niet beperkt tot het hier gegeven tweedimensionale voorbeeld. De enige voorwaarden zijn, dat de vergelijking waarin y wordt uitgedrukt als functie van x lineair is in de parameters en dat het aantal verschillende punten waarin waarnemingen zijn gedaan groter of gelijk is aan het aantal parameters (dan is $X^T X$ regulier).

4.2.2. De gewogen kleinste kwadratenberekening.

Bij de ongewogen kleinste kwadratenmethode heeft ieder punt in principe evenveel invloed op de positie van de berekende curve. Wanneer de invloed van ieder punt op het verloop van de curve

anders moet zijn, bijv. omdat de meetnauwkeurigheid niet voor alle punten gelijk is, moet een gewogen kleinste kwadratenmethode worden toegepast.
De berekening volgens vergelijking XIX blijft ook in dit geval in principe gelijk. Het enige dat verandert is, dat de elementen van de X en Y matrices, die behoren bij een bepaald punt i, vermenigvuldigd moeten worden met het gewicht van dat punt : w_i.
Dit gewicht is groter, naarmate de invloed van het betreffende punt op de positie van de curve groter moet zijn. Hoe het gewicht van een meetpunt gekozen moet worden, hangt sterk af van de nauwkeurigheid, waarmee de positie van dat punt vastgelegd is. Hoe onnauwkeuriger een meting is, des te minder invloed zal het meetpunt mogen hebben op het verloop van een te berekenen curve. In het geval van de Bhattacharya plot, is de positie van een punt $(i, \log (f_{i+1}/f_i))$ nauwkeuriger bepaald, naarmate de waargenomen frequenties in klasse i en i + 1 groter zijn. De standaarddeviatie van een waargenomen frequentie is namelijk gelijk aan $\sqrt{f}$ en de relatieve fouten in f_{i+1} en f_i zijn additief in het quotiënt. Het is dus voor de hand liggend om voor een punt van de Bhattacharya plot $(f_{i+1} + f_i)^x$ als gewicht te kiezen, waarbij x nog nader bepaald moet worden.
Door Bhattacharya wordt in dit verband opgemerkt :

"While matching the straight line it is better to fit closely to the points where the frequency is large even if the apparent discrepancy becomes somewhat large where the frequency is small."

Een tweede benadering voor de keuze van gewichten voor de punten van de Bhattacharya plot is de volgende :
De onregelmatige delen van de Bhattacharya plot, veroorzaakt door pathologisch hoge en lage uitkomsten, mogen in feite helemaal geen invloed uitoefenen op het verloop van de rechte lijn. Deze punten moeten eigenlijk beschouwd worden als uitbijters. De

ongewogen kleinste kwadratenmethode heeft als belangrijkste nadeel, dat een uitbijter grote invloed heeft op de positie van de berekende curve. In het algemeen geldt, dat hoe verder een punt verwijderd is van de theoretische curve, hoe groter de invloed van dat punt is op de ligging van de berekende curve. Deze eigenaardigheid van de kleinste kwadratenmethode, die in feite in strijd is met de logica, kan tegengegaan worden door voor ieder punt i als gewicht te kiezen $(\overline{\delta}/(\delta_i + \overline{\delta}))^y$, waarbij δ_i de afstand van het punt i tot de lijn is, $\overline{\delta}$ de gemiddelde afstand van alle punten tot de lijn en y nog nader bepaald moet worden (118).
Wanneer de afstand van de punten tot de curve gebruikt wordt voor de berekening van de gewichten van die punten, houdt dat in, dat de berekening iteratief uitgevoerd moet worden. Als eerste benadering kan de ongewogen kleinste kwadratenmethode gebruikt worden. Hierna kan aan ieder punt een gewicht worden toegekend volgens bovenstaande formule en een gewogen kleinste kwadraten berekening worden uitgevoerd. Na deze tweede berekening kunnen de gewichten van ieder punt weer worden aangepast, waarna een derde curve berekend kan worden. Deze procedure wordt net zo lang herhaald, totdat de berekende curve niet meer verandert.

De twee beschreven benaderingen van de weegfactor benadrukken elk een ander aspect van de Bhattacharya plot. De eerste benadering legt vooral de nadruk op het feit, dat een punt van de plot meer betrouwbaar is naarmate het aantal waarnemingen, waaruit het berekend is, toeneemt. De tweede benadering is meer pragmatisch en gaat er van uit, dat naarmate de afstand van een punt tot de lijn toeneemt, de invloed ervan op het verloop van de lijn afneemt.
In de praktijk zal meestal een combinatie van deze twee aspecten een rol spelen. Naarmate de frequentie in een klasse kleiner wordt, wordt de onnauwkeurigheid van een punt, berekend uit deze frequentie, groter. Dit houdt tegelijkertijd in, dat de waarschijnlijkheid dat het punt ver van de lijn ligt, groter wordt.

Een combinatie van de twee boven beschreven weegfactoren levert het volgende :

$$w_i = \frac{(f_{i+1} + f_i)^x}{(1 + \delta_i/\bar{\delta})^y} \qquad XX$$

waarbij :

w_i = gewicht van punt i.
f_i = frequentie in klasse i.
δ_i = afstand van punt i tot de lijn.
$\bar{\delta}$ = gemiddelde afstand van alle punten tot de lijn.
x, y = nader te bepalen exponenten.

Door x en y te variëren (van 0 oplopend met stappen van 0,5 tot 3) kunnen verschillende weegfactoren opgesteld worden. Deze werden toegepast op de 40% links éénzijdig gecontamineerde frequentieverdelingen uit hoofdstuk 3. Er worden dus op deze manier voor iedere weegfactor 20 gemiddelden en 20 standaarddeviaties berekend. Voor iedere weegfactor worden nu vier parameters bepaald : het gemiddelde en de standaarddeviatie van de 20 berekende gemiddelden en het gemiddelde en de standaarddeviatie van de 20 berekende standaarddeviaties. Door middel van het toekennen van scores aan iedere weegfactor voor elk van de vier parameters (waarbij de weegfactor die het beste resultaat voor een parameter oplevert de laagste score krijgt toebedeelt en de weegfactor die het slechtste resultaat oplevert de hoogste score), kan vastgesteld worden, dat de beste weegfactor gevonden wordt voor x=2 en y=0 :

$$w_i = (f_{i+1} + f_i)^2 \qquad XXI$$

De afstanden van de punten tot de lijn blijken dus niet in de weegfactor tot uitdrukking te komen. Rekentechnisch gezien is dit resultaat interessant, omdat geen iteratieve rekenwijze nodig blijkt te zijn.

Het opstellen van een weegprocedure voor een kleinste kwadraten berekening kan ook meer mathematisch aangepakt worden. Wanneer vergelijking XIX de oplossing van een ongewogen kleinste kwadraten berekening voorstelt, is de overeenkomstige vergelijking voor een gewogen kleinste kwadraten berekening de volgende :

$$B = (X^T V^{-1} X)^{-1} X^T V^{-1} Y \qquad \text{XXII}$$

waarbij voor V meestal de variantie-covariantie matrix gekozen wordt. Voor de Bhattacharya plot kan afgeleid worden, dat V bij benadering (wanneer de frequenties niet te klein zijn) gelijk is aan (125) :

$$V = \begin{bmatrix} \frac{1}{f_1} + \frac{1}{f_2} & -\frac{1}{f_2} & & & \\ -\frac{1}{f_2} & \frac{1}{f_2} + \frac{1}{f_3} & & \emptyset & \\ & & \ddots & & -\frac{1}{f_{k-1}} \\ & \emptyset & & -\frac{1}{f_{k-1}} & \frac{1}{f_{k-1}} + \frac{1}{f_k} \end{bmatrix}$$

Met beide procedures (de berekening volgens vergelijking XXII en de berekening met de weegfactor uit vergelijking XXI) werden het gemiddelde en de standaarddeviatie berekend van de 260 theoretische frequentieverdelingen uit hoofdstuk 3. De resultaten zijn samengevat in tabel 9. Uit deze tabel blijkt het volgende :
De theoretische weegprocedure (volgens vergelijking XXII) levert betere resultaten op voor de niet gecontamineerde, alleen door statistische invloeden vertekende, Gauss verdelingen. Wanneer echter de contaminatie toeneemt, levert deze procedure duidelijk minder goede resultaten. Bij de afleiding van de variantie-covariantie matrix is ook alleen maar rekening gehouden met statistische fouten in de waargenomen frequenties, en niet met mogelijke contaminatie. Op grond van deze vergelijking wordt de voorkeur gegeven aan de (mathematisch minder goed gefundeerde) weegprocedure volgens vergelijking XXI.

Tabel 9A.
Gemiddelden van de 260 theoretische frequentieverdelingen uit hoofdstuk 3, berekend met de twee verschillende gewogen kleinste kwadraten berekeningen. Ieder gemiddelde in de tabel is het gemiddelde van 20 berekende waarden. Tussen haakjes is vermeld de standaarddeviatie.

contaminatie	weging volgens vergelijking XXII	weging volgens vergelijking XXI
geen	0,004 (0,017)	0,009 (0,023)
rechts 10%	0,031 (0,016)	0,010 (0,022)
rechts 20%	0,062 (0,017)	0,012 (0,023)
rechts 30%	0,094 (0,016)	0,013 (0,023)
rechts 40%	0,122 (0,017)	0,016 (0,023)
links 10%	-0,024 (0,017)	0,008 (0,023)
links 20%	-0,055 (0,016)	0,007 (0,022)
links 30%	-0,086 (0,017)	0,005 (0,022)
links 40%	-0,115 (0,018)	0,003 (0,022)
2 zijdig 10%	0,003 (0,016)	0,010 (0,022)
2 zijdig 20%	0,004 (0,018)	0,009 (0,023)
2 zijdig 30%	0,005 (0,018)	0,009 (0,022)
2 zijdig 40%	0,000 (0,018)	0,010 (0,022)

Tabel 9B.
Standaarddeviaties van de 260 theoretische frequentieverdelingen uit hoofdstuk 3 berekend met de twee verschillende gewogen kleinste kwadraten berekeningen. Iedere standaarddeviatie in de tabel is het gemiddelde van 20 berekende waarden. Tussen haakjes is de standaarddeviatie vermeld.

contaminatie	weging volgens vergelijking XXII	weging volgens vergelijking XXI
geen	1,003 (0,011)	0,992 (0,032)
rechts 10%	1,034 (0,010)	0,995 (0,032)
rechts 20%	1,069 (0,012)	0,999 (0,032)
rechts 30%	1,103 (0,014)	1,004 (0,032)
rechts 40%	1,133 (0,014)	1,012 (0,033)
links 10%	1,035 (0,009)	0,995 (0,032)
links 20%	1,071 (0,011)	0,999 (0,032)
links 30%	1,106 (0,013)	1,004 (0,032)
links 40%	1,137 (0,012)	1,013 (0,033)
2 zijdig 10%	1,066 (0,010)	0,998 (0,032)
2 zijdig 20%	1,132 (0,011)	1,006 (0,032)
2 zijdig 30%	1,199 (0,014)	1,017 (0,033)
2 zijdig 40%	1,268 (0,017)	1,033 (0,034)

4.3. Manuele berekening na smoothen van de frequenties.

Wanneer het gemiddelde en de standaarddeviatie manueel bepaald moeten worden uit een Bhattacharya plot, moet het rechte gedeelte uit deze plot duidelijk te herkennen zijn. Bij de Bhattacharya plot wordt echter gerekend met de waargenomen discrete frequenties, dit in tegenstelling tot de cumulatieve, relatieve benadering van de Hoffmann- en Neumann methode. Dit houdt in, dat statistische fluctuaties in het aantal waarnemingen per klasse, een grote invloed zullen hebben op log (f_{i+1}/f_i), de y-waarde in de Bhattacharya plot, vooral wanneer het totale aantal waarnemingen klein is (kleiner dan ongeveer 1500).
Door deze statistische fluctuaties kan het soms zeer moeilijk zijn, het rechte deel van de Bhattacharya plot te herkennen. Hoeke (55) heeft daaruit geconcludeerd, dat in de praktijk de Bhattacharya plot nauwelijks waarde heeft.

Het toepassen van een smoothing procedure heeft tot gevolg, dat statistische fluctuaties worden glad gestreken, waardoor uiteindelijk de selectie van het rechte deel van de Bhattacharya plot vergemakkelijkt wordt. Dit is belangrijk, omdat "pathologische" uitslagen, die het niet-lineaire deel van de plot veroorzaken, geen invloed mogen uitoefenen op de ligging van de "op het oog" getrokken rechte lijn.
Smoothing kan beter uitgevoerd worden op de waargenomen klassefrequenties dan op de y-waarden van de Bhattacharya plot. De waargenomen klassefrequenties zijn nl. minder afhankelijk van elkaar dan de y-waarden van de Bhattacharya plot. De noemer in de uitdrukking voor y_i is nl. gelijk aan de teller in y_{i-1} en de teller in y_i is de noemer in y_{i+1}.

Aangezien de klassen allemaal even breed zijn, ligt het toepassen van de smoothing, zoals beschreven door Savitzky en Golay (104), voor de hand. Bij deze smoothing wordt door een oneven

aantal datapunten (x_i, y_i) met de methode van de kleinste kwadraten een polynoom berekend. De bij de middelste x-waarde afgelezen waarde voor y, is de gecorrigeerde y-waarde. Vervolgens wordt de berekening herhaald, met links één datapunt minder en rechts één meer, totdat alle punten gebruikt zijn.
De berekening van de gecorrigeerde y-waarden (in het geval van een frequentieverdeling zijn deze y-waarden de waargenomen aantallen per klasse) kan gereduceerd worden tot het toepassen van een zeer eenvoudige vergelijking, wanneer de x-waarden onderling gelijke afstanden hebben. Als voorbeeld moge dienen de berekening van een gecorrigeerde waarde voor y_3, wanneer met de kleinste kwadratenmethode een parabool wordt berekend door vijf opeenvolgende datapunten : $(-2, y_1)$, $(-1, y_2)$, $(0, y_3)$, $(1, y_4)$ en $(2, y_5)$.

Een parabool wordt gegeven door de vergelijking :
$y = ax^2 + bx + c$. Wanneer uit een aantal meetpunten (x, y), waarden voor a, b en c berekend moeten worden met behulp van de methode van de kleinste kwadraten, geldt de volgende matrix vergelijking (zie 4.2) :

$$B = (X^T X)^{-1} X^T Y \qquad \text{XXIII}$$

In dit voorbeeld geldt :

$$X = \begin{bmatrix} 4 & -2 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} \qquad B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Zodat :

$$X^TX = \begin{bmatrix} 34 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 5 \end{bmatrix} \qquad (X^TX)^{-1} = \frac{1}{70} \begin{bmatrix} 5 & 0 & -10 \\ 0 & 7 & 0 \\ -10 & 0 & 34 \end{bmatrix}$$

$$X^T Y = \begin{bmatrix} 4y_1 + y_2 + y_4 + 4y_5 \\ -2y_1 - y_2 + y_4 + 2y_5 \\ y_1 + y_2 + y_3 + y_4 + y_5 \end{bmatrix}$$

De gecorrigeerde waarde voor y_3 is gelijk aan de constante c uit de vergelijking van de parabool, aangezien de bij y_3 behorende x-waarde gelijk is aan 0. Uitwerking van de vergelijking $B = (X^T X)^{-1} X^T Y$ levert voor c de volgende waarde op :

$$c = \hat{y}_3 = \frac{-3y_1 + 12y_2 + 17y_3 + 12y_4 - 3y_5}{35} \qquad \text{XXIV}$$

Een vergelijking analoog aan de bovenstaande kan afgeleid worden voor een willekeurig polynoom en een willekeurig aantal punten. Door Savitzky en Golay (104) worden uitgebreide tabellen gegeven voor de constanten die in die gevallen moeten worden toegepast. In deze tabellen is echter een aantal fouten geslopen. Deze zijn gecorrigeerd door Steinier et al. (109).

Om het effect van smoothen op een zuivere Gauss curve vast te stellen, werd als volgt te werk gegaan :

-Een zuivere Gauss curve ($\mu = 0$, $\sigma = 1$) werd verdeeld in klassen met een breedte van 0,2.

-De oppervlakten van de 26 klassen tussen -2,6 en +2,6, berekend zoals beschreven in 3.3, werden als waargenomen frequenties beschouwd.

-Op de zo resulterende frequentieverdeling werd smoothing toegepast met behulp van verschillende polynomen (lineair, kwadratisch en 4e graads). Het aantal punten voor de smoothing bedroeg minimaal 3 en maximaal 19.

-Met de gecorrigeerde frequenties werd de Bhattacharya plot geconstrueerd en hieruit werden het gemiddelde en de standaarddeviatie berekend.

De resultaten zijn vermeld in tabel 10.

Tabel 10.
De invloed van verschillende typen smoothing (9 p 4e betekent bijv. dat een 9 punts smoothing is toegepast m.b.v. een 4e graads polynoom) op de berekende standaarddeviatie van een zuivere Gauss verdeling. Het berekende gemiddelde blijkt in dit experiment niet te veranderen. Ook de lineariteit blijft bij ieder type smoothing perfect (r = 1).

smoothing	$\hat{\sigma}$	afwijking t.o.v. 1
geen	1,000	0,0 %
3 p 1e	1,026	2,6 %
5 p 1e	1,079	7,9 %
7 p 1e	1,159	15,9 %
9 p 1e	1,273	27,3 %
11 p 1e	1,432	43,2 %
5 p 2e	0,999	-0,1 %
7 p 2e	0,999	-0,1 %
9 p 2e	1,002	0,2 %
11 p 2e	1,014	1,4 %
13 p 2e	1,039	3,9 %
15 p 2e	1,085	8,5 %
17 p 2e	1,153	15,3 %
19 p 2e	1,249	24,9 %
7 p 4e	0,999	-0,1 %
9 p 4e	1,000	0,0 %
11 p 4e	1,000	0,0 %
13 p 4e	1,001	0,1 %
15 p 4e	1,004	0,4 %
17 p 4e	1,014	1,4 %
19 p 4e	1,032	3,2 %

Het gemiddelde blijkt in dit experiment niet te veranderen, ongeacht het type smoothing. Ook de lineariteit van de Bhattacharya plot blijft perfect, zoals blijkt uit de correlatiecoëfficient van de berekende rechte lijn (r = 1). De berekende standaarddeviatie neemt echter toe naarmate het gebruikte polynoom eenvoudiger wordt, en naarmate het aantal punten, dat gebruikt wordt voor de smoothing, toeneemt.
Dit laatste is ook begrijpelijk, aangezien de beschrijving van een deel van de Gauss curve door een polynoom slechter wordt, naarmate het polynoom eenvoudiger en het deel van de Gauss curve groter wordt.

In de praktijk wordt de voorkeur gegeven aan de meest eenvoudige smoothing, die een zuivere Gauss curve niet vertekent : 5 punts smoothing met behulp van een parabool.

4.4. Berekening van meerdimensionale referentiewaardengebieden met behulp van de Bhattacharya plot.

Zoals in 1.5 uiteengezet is, wordt het tweedimensionale referentiewaardengebied voor twee onafhankelijke bepalingen, die beide een Gaussische frequentieverdeling hebben, gevormd door een ellips. Deze ellips is slanker, naarmate de correlatie tussen de twee bepalingen groter wordt en vereenvoudigt tot een cirkel, wanneer de correlatiecoëfficiënt gelijk is aan 0.
Zo wordt in de k-dimensionale ruimte het referentiewaardengebied voor een combinatie van k onafhankelijke bepalingen, die alle een Gaussische frequentieverdeling hebben, gevormd door een ellipsoïde. De vergelijking van deze ellipsoïde wordt gegeven door (18) :

$$S = \left\{ Y \mid (Y - \mu)^T V^{-1} (Y - \mu) \leqslant \chi^2(\alpha, k) \right\} \qquad \text{XXV}$$

waarbij :

S = referentie ellipsoide.
Y = vector van k resultaten.
μ = vector van de verwachtingen van de k bepalingen.
V = variantie-covariantie matrix.
k = de dimensie van de ellipsoide
= het aantal onafhankelijke bepalingen.
$\chi^2(\alpha, k)$ = α-fractiel van een χ^2 verdeling met k vrijheidsgraden (α is meestal gelijk aan 0,95).

Wanneer k = 1, met andere woorden wanneer een referentiewaardengebied berekend wordt voor één enkele variabele, kan vergelijking XXV vereenvoudigd worden tot :

$$S = \left\{ y \mid (y - \mu) \frac{1}{\sigma^2} (y - \mu) \leqslant 3{,}84 \right\} \qquad \text{XXVI}$$

waarbij : μ = gemiddelde van de bepaling.
σ = standaarddeviatie.
$\chi^2(0,95,\ 1) = 3,84$ (o.a. 22).

Vergelijking XXVI kan ook geschreven worden als :

$$\mu - 1,96\,\sigma \leqslant y \leqslant \mu + 1,96\,\sigma \qquad \text{XXVII}$$

Dit is waarschijnlijk de meest gebruikte vergelijking voor de berekening van ééndimensionale referentiewaarden (zie 2.2.1).

Wanneer k = 2 en wanneer het gemiddelde en de standaarddeviatie van beide bepalingen respectievelijk gelijk zijn aan 0 en 1, wordt de variantie-covariantie matrix gegeven door :

$$V = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \qquad \text{en hieruit volgt : } V^{-1} = \frac{1}{1 - r^2} \begin{bmatrix} 1 & -r \\ -r & 1 \end{bmatrix} \qquad \text{XXVIII}$$

Aangezien $\chi^2(0,95,\ 2)$ gelijk is aan 5,99 (22), kan vergelijking XXV voor dit tweedimensionale voorbeeld geschreven worden als :

$$S = \left\{ Y \mid (y_1 \quad y_2) \ \frac{1}{1 - r^2} \begin{bmatrix} 1 & -r \\ -r & 1 \end{bmatrix} \ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leqslant 5,99 \right\}$$

$$\text{of : } y_1^2 - 2ry_1y_2 + y_2^2 \leqslant 5,99\ (1 - r^2) \qquad \text{XXIX}$$

waarbij : y_1 = uitkomst van bepaling 1.
y_2 = uitkomst van bepaling 2.
r = correlatiecoëfficiënt, die het verband tussen de twee bepalingen aangeeft.

Deze laatste vergelijking stelt in het algemeen een ellips voor. Wanneer r = 0 wordt dit een cirkel met straal $\sqrt{5,99}$.

Uit het bovenstaande blijkt, dat voor de berekening van een k-dimensionale referentiewaarden ellipsoïde, de volgende gegevens over de k bepalingen bekend moeten zijn :

1. De gemiddelden.
2. De variantie-covariantie matrix, die bestaat uit :
   a. de varianties van de k bepalingen en
   b. de covarianties van iedere combinatie van twee van de k bepalingen.

Van een inhomogene populatie, die bestaat uit zowel "normale" als "pathologische" getallen, kan het gemiddelde en de standaarddeviatie van de normale populatie eenvoudig bepaald worden met de Bhattacharya plot, aangenomen dat de frequentieverdeling Gaussisch is. Het enige probleem dat dus nog opgelost moet worden, alvorens uit een inhomogene populatie meerdimensionale referentiewaardengebieden te kunnen berekenen, is het bepalen van de covarianties.

Aangezien voor de covariantie van de bepaling Y_1 en Y_2 geldt dat cov $(y_1, y_2) = \sigma_1 \sigma_2 r$, zou de gedachte kunnen opkomen de covariantie te bepalen via het berekenen van de correlatiecoëfficiënt, wanneer de resultaten van bepaling Y_1 en Y_2 grafisch tegen elkaar worden uitgezet. Deze gedachte is echter onjuist, omdat een (groot) gedeelte van de bepalingsuitkomsten "pathologisch" is en dus niet gebruikt mag worden voor de berekening van de correlatiecoëfficiënt.

Een andere benadering, die wél juist is, is gebaseerd op het volgende :

Wanneer zowel de resultaten van bepaling Y_1 als die van bepaling Y_2 een Gaussische frequentieverdeling hebben, heeft een lineaire combinatie van de resultaten van deze twee bepalingen ook een Gaussische frequentieverdeling (64). Dit betekent, dat, wanneer de Bhattacharya plot toegepast kan worden op de resul-

taten van bepaling Y_1 en Y_2, de plot ook toegepast kan worden op de som van de individuele uitkomsten voor bepaling Y_1 en Y_2, en ook op het verschil van deze uitkomsten.
De som van de individuele uitkomsten voor bepaling Y_1 en Y_2 heeft als gemiddelde $(\mu_1 + \mu_2)$ en als variantie :
$\sigma_1^2 + \sigma_2^2 + 2 \text{ cov } (y_1, y_2)$.
Wanneer het verschil van de individuele uitkomsten gebruikt wordt, is het gemiddelde $(\mu_1 - \mu_2)$ en de variantie :
$\sigma_1^2 + \sigma_2^2 - 2 \text{ cov } (y_1, y_2)$.
De covariantie van twee bepalingen kan dus berekend worden als :

$$\text{cov } (y_1, y_2) = \frac{\sigma^2_{\text{som}} - \sigma^2_{\text{verschil}}}{4} \qquad \text{XXX}$$

In de praktijk zal het toepassen van meerdimensionale referentiewaardengebieden voor combinaties van meer dan twee bepalingen, om praktische redenen alleen kunnen geschieden met behulp van een computer.
In 6.5 wordt een publicatie gereproduceerd, waarin als voorbeeld de berekening van een referentiewaardenellips voor de combinatie albumine-totaal eiwit in serum wordt uitgewerkt.

5. Niet Gaussische frequentieverdelingen.

5.1. Inleiding.

In de tot nu toe gegeven beschouwingen, is steeds de veronderstelling gemaakt, dat de resultaten van een klinisch chemische bepaling verdeeld zijn volgens een Gauss curve. Hoewel deze veronderstelling voor veel bepalingen terecht gemaakt wordt, is er ook een vrij groot aantal klinisch chemische onderzoeken, waarbij de veronderstelling van een Gauss verdeling leidt tot ernstige fouten in de berekening van het referentiewaarden gebied. De niet Gaussisch verdeelde bepalingen zijn vooral enzym activiteitsmetingen, maar ook een aantal andere (bijv. bilirubine) heeft duidelijk een asymmetrische frequentieverdeling (15, 70).
In dit hoofdstuk zullen nu de volgende zaken aan de orde komen :

1. De criteria, op grond waarvan besloten wordt, dat de veronderstelling dat de frequentieverdeling Gaussisch is, al dan niet correct is.
2. Het bespreken van een procedure, die met goed succes toegepast kan worden, wanneer blijkt dat de frequentieverdeling niet Gaussisch is.

5.2. Toetsen of een frequentieverdeling Gaussisch is.

Wanneer alle waarnemingen beschouwd kunnen worden als aselecte trekkingen uit één verdeling, is het toetsen op het al dan niet Gaussisch zijn van deze verdeling vrij eenvoudig. Drie gebruikelijke methoden in de statistiek zijn de χ^2-toets voor aanpassing, de Kolmogorov-Smirnov toets (17, 74) en de toets van Shapiro en Wilk (105). Deze laatste is vooral zeer geschikt, wanneer het aantal waarnemingen klein is (<20).
Wanneer niet alle waarnemingen uit één populatie afkomstig zijn, is het veel moeilijker, om te toetsen of de frequentieverdeling van de hoofdpopulatie Gaussisch is. Dit lijkt een ernstig pro-

bleem, maar blijkt in de praktijk erg mee te vallen. In 5.4.4 wordt hier nader op in gegaan.
Opgemerkt moet verder worden, dat het toepassen van bovenstaande toetsen op een gecontamineerde Gauss verdeling, tot onjuiste conclusies kan leiden.

5.3 Methoden om de frequentieverdeling mathematisch te beschrijven, wanneer de vorm ervan niet Gaussisch is.

5.3.1. Inleiding.

Wanneer blijkt, dat de frequentieverdeling van de uitkomsten van een bepaling niet beschreven kan worden met een Gauss curve, wordt het berekenen van referentiewaarden gecompliceerder. Het gebruik van verdelingsvrije methoden, waarbij geen veronderstellingen over de vorm van de frequentieverdeling gedaan hoeven te worden, is alleen mogelijk, wanneer de bepalingsuitkomsten afkomstig zijn van een qua gezondheidskenmerken homogene groep mensen (zie 2.2). Wanneer de groep mensen niet homogeen is, zoals in dit proefschrift wordt aangenomen, met andere woorden, wanneer een gedeelte van de groep bestaat uit "gezonden" en een ander deel uit "zieken", kunnen alleen referentiewaarden worden berekend voor de hoofdpopulatie, wanneer de frequentieverdeling van uitkomsten verkregen uit deze deelgroep, in de totale frequentieverdeling herkend kan worden. Dit is slechts mogelijk, wanneer de mathematische vorm van de frequentieverdeling van de hoofdpopulatie bekend is.
Er zijn verschillende benaderingen beschreven, om een willekeurige waargenomen frequentieverdeling mathematisch aan te pakken (64) :

1. Het opstellen van een familie van frequentieverdelingen, waaruit dan dié verdeling geselecteerd wordt, die de waargenomen verdeling het beste beschrijft (de benadering van Pearson).

2. Het beschrijven van een waargenomen frequentieverdeling door een reeksontwikkeling in de afgeleiden van de Gauss verdeling.
3. Het transformeren van de waarnemingen, zodanig, dat de getransformeerde variabele verdeeld is volgens een bekende functie, bij voorkeur de Gauss functie.

5.3.2. De benadering van Pearson.

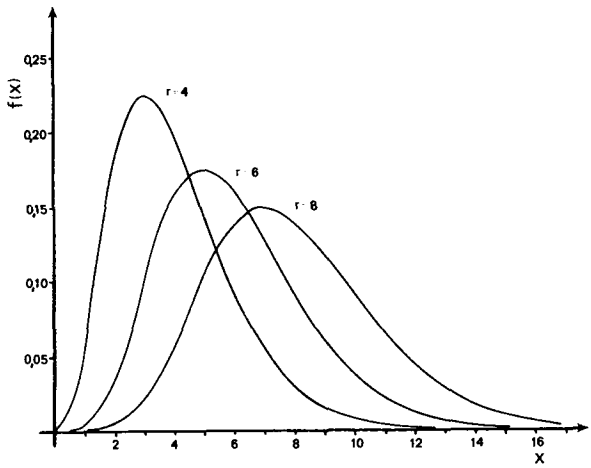
Pearson heeft een familie vergelijkingen opgesteld, waarvan blijkt, dat de meeste in de praktijk waargenomen frequentieverdelingen erdoor beschreven kunnen worden. Al deze vergelijkingen voldoen aan de differentiaalvergelijking :

$$\frac{df}{dx} = \frac{(x - a)\, f}{b_0 + b_1 x + b_2 x^2}$$

Afhankelijk van de waarden van de constanten b_0, b_1 en b_2 ontstaan er verschillende typen verdelingen, waarvan de voornaamste

Tabel 11.
Belangrijkste Pearson verdelingen.
Alle parameters (m_1 en m_2 voor de Beta verdeling; λ en r voor de Gamma verdeling en ν voor de Student verdeling) zijn positief.

kansdichtheidsfunctie	definitiegebied	naam (type)
$f(x) = \frac{\Gamma(m_1+m_2)}{\Gamma(m_1)\,\Gamma(m_2)}\, x^{m_1-1}\,(1-x)^{m_2-1}$	$0 \leqslant x \leqslant 1$	Beta (I)
$f(x) = \frac{\lambda^r}{\Gamma(r)}\, x^{r-1}\, e^{-\lambda x}$	$0 \leqslant x < \infty$	Gamma (III)
$f(x) = \frac{\Gamma(\frac{1}{2}+\frac{1}{2}\nu)}{\Gamma(\frac{1}{2}\nu)\sqrt{\pi\nu}}\,(1+x^2/\nu)^{-\frac{1}{2}(\nu+1)}$	$-\infty < x < \infty$	Student (VII)



Figuur 13.
Gamma verdelingen voor $\lambda = 1$ en verschillende waarden voor r.

zijn samengevat in tabel 11. Voor de klinisch chemische praktijk lijken het meest in aanmerking te komen (uiteraard) type VII, wanneer de verdeling symmetrisch is en type III, wanneer dit niet het geval is. De vorm van de type III verdelingen doet sterk denken aan de frequentieverdelingen van de resultaten van enzym activiteits metingen (fig 13). Opgemerkt moet verder worden, dat χ^2-verdelingen een ondergroep vormen van de Gamma verdelingen, nl. die, waarbij $\lambda = 2$ en $r = 0{,}5\nu$.

5.3.3. Het beschrijven van een waargenomen frequentieverdeling door een reeksontwikkeling.

Functies kunnen vaak in een reeks worden ontwikkeld, waarbij de termen van de reeks bijv. machten van de variabele zijn (Taylor reeksen) of trigonometrische functies (Fourier reeksen).
Voor kansdichtheidsfuncties kan ook een reeksontwikkeling worden ontworpen, die gebaseerd is op de Chebyshev-Hermite polynomen en

de eigenschappen daarvan, die ook gebruikt zijn in 4.1. Wanneer een frequentieverdeling ontwikkeld kan worden in de afgeleiden van de Gauss verdeling, ontstaat de volgende uitdrukking :

$$f(x) = Z(x)\left[1 + 1/2\,(\mu_2 - 1)\,H_2 + 1/6\,\mu_3\,H_3 + 1/24\,(\mu_4 - 6\mu_2 + 3)\,H_4 + \ldots\right]$$

waarbij : $Z(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}x^2}$

μ_i = i-de moment van f

$H_0 = 1$

$H_1 = x$ en in het algemeen :

$$H_n = x^n - \frac{n^{(2)}}{2.1!}\,x^{n-2} + \frac{n^{(4)}}{2^2.2!}\,x^{n-4} + \ldots$$

$$n^{(i)} = n\,(n-1)\,(n-2)\ldots(n-i+1)$$

Bovenstaande vergelijking stelt de zogenaamde Gram-Charlier serie type A voor. Het aantal verschillende vormen van frequentieverdelingen dat met de Gram-Charlier serie beschreven kan worden is vrij beperkt, althans veel beperkter dan het aantal vormen dat met de Pearson functies beschreven kan worden.
De onbekende μ's in bovenstaande vergelijking moeten geschat worden uit een waargenomen frequentieverdeling. Dit kan alleen maar eenvoudig, wanneer de frequentieverdeling niet gecontamineerd is. Wanneer dit wél het geval is, dus wanneer de waargenomen frequentieverdeling bestaat uit 2 (of meer) elkaar gedeeltelijk overlappende verdelingen ("normaal" en "pathologisch"), is het berekenen van de momenten niet direct mogelijk. Bovenstaande vergelijking is dan veel te gecompliceerd om in de praktijk bruikbaar te zijn.

5.3.4. Het transformeren van de waarnemingen.

Wanneer blijkt, dat de frequentieverdeling van een variabele niet Gaussisch is, kan geprobeerd worden om door middel van een (liefst zo eenvoudig mogelijke) transformatie, de frequentieverdeling wél Gaussisch te maken. Waarschijnlijk de meest toegepaste transformatie in de klinische chemie is het nemen van de logaritme uit de uitkomsten van de niet Gaussisch verdeelde bepaling. Vaak blijkt dan, dat de logaritme van de uitkomsten wél een nagenoeg Gaussische vorm heeft (33, 37).
Johnson (61) heeft een familie van kansdichtheidsfuncties opgesteld die de eigenschap hebben, dat ze op een eenvoudige manier getransformeerd kunnen worden tot een Gauss vorm. Deze groep functies is vrijwel net zo algemeen als het Pearson systeem (64). Johnson beschrijft drie typen kansdichtheidsfuncties :

-Het S_L of lognormale systeem, waarbij de transformatiefunctie g(y) gelijk is aan log (y).
-Het S_B systeem, waarbij g(y) = log (y/(1-y)). Deze transformatie kan alleen toegepast worden, wanneer de waarnemingen fracties zijn ($0 < y < 1$).
-Het S_U systeem, waarbij $g(y) = \sinh^{-1}(y)$.

Harris (45) heeft aangetoond, dat het mogelijk is om door middel van transformaties nagenoeg Gaussisch verdeelde variabelen te vinden voor een aantal klinisch chemische bepalingen. De transformaties die in eerste instantie worden toegepast zijn :
$y = \sqrt{x + C}$ en $y = \log (x + C)$.
Als toetsingsgrootheid voor het al dan niet Gaussisch zijn, worden γ_1 en γ_2 gebruikt, maten voor respectievelijk "skewness" en "kurtosis". Deze beide indices zijn voor een Gauss verdeling gelijk aan 0. Wanneer na de eerste transformatie γ_1 en/of γ_2 nog significant van 0 afwijken, wordt een tweede transformatie toegepast, nl. $z = \sinh^{-1}(y)$. Voor alle beschouwde bepalingen

kunnen zo waarden voor γ_1 en γ_2 bereikt worden, die vrijwel gelijk zijn aan 0. Hoewel de tweede transformatie voor een aantal bepalingen verbetering van de waarden van γ_1 en/of γ_2 betekent, wordt de vorm van de uiteindelijke verdeling niet wezenlijk beïnvloed door deze tweede transformatie (45).

Reed (93) heeft als reactie op het artikel van Harris (45) de door laatstgenoemde voorgestelde transformaties toegepast op een aantal theoretische frequentieverdelingen. Deze verdelingen zijn overgenomen uit het artikel van Johnson (61) en behoren alle tot het S_B systeem.
Reed vergelijkt de referentiewaarden berekend met verdelingsvrije methoden met de referentiewaarden, die resulteren na transformatie van de variabele volgens Harris. De conclusie is, dat de transformaties die door Harris zijn voorgesteld, slechts te prefereren zijn boven verdelingsvrije methoden, wanneer de oorspronkelijke frequentieverdeling al nagenoeg symmetrisch is.
Er zijn echter twee punten van kritiek op de publikatie van Reed. In de eerste plaats zijn de meeste frequentieverdelingen, waarop de transformaties toegepast worden, "vreemd van vorm". In de tweede plaats is het natuurlijk verwonderlijk, dat frequentieverdelingen die juist zo ontworpen zijn, dat een transformatie naar de Gaussische vorm op een eenvoudige manier mogelijk is, niet goed getransformeerd kunnen worden.

Een transformerende functie moet toegepast worden op de oorspronkelijke gegevens. Deze zijn echter veelal niet meer voorhanden, doordat meetresultaten omgerekend en afgerond worden. De fouten die zo ontstaan in de getransformeerde variabele kunnen aanzienlijk zijn. Bovendien zijn deze fouten afhankelijk van de absolute grootte van de oorspronkelijke variabele. Het is met andere woorden alleen maar aan te raden transformaties toe te passen, wanneer de oorspronkelijke gegevens, niet omgerekend en niet afgerond, met voldoende nauwkeurigheid bekend zijn.

5.4. Keuze van de rekenprocedure, wanneer de frequentieverdeling niet Gaussisch is.

5.4.1. Inleiding.

Uit het bovenstaande blijken twee methoden het meest in aanmerking te komen voor het berekenen van referentiewaarden uit een niet Gaussisch verdeelde populatie. Deze methoden zijn :

- Aannemen, dat de verdeling beschreven kan worden met een Gamma-functie. De vorm van Gamma-functies lijkt sterk op de vorm van frequentieverdelingen die in de praktijk worden waargenomen (30).
- Het transformeren van de oorspronkelijke gegevens naar de Gaussische vorm m.b.v. de formule : $y = \log (x + C)$. Transformeren is echter alleen mogelijk, wanneer de onbewerkte, oorspronkelijke gegevens beschikbaar zijn.

5.4.2. De Gamma verdeling.

Stel, dat de uitkomsten van een bepaling voor een homogene groep mensen verdeeld zijn volgens een Gamma verdeling :

$$f(x) = \frac{\lambda^r}{\Gamma(r)} \; x^{r-1} \; e^{-\lambda x}$$

waarbij : $r > 0 \quad \lambda > 0 \quad x \geqslant 0$

$$\Gamma(r) = \int_0^\infty x^{r-1} \; e^{-x} \, dx$$

$$= (r - 1)! \text{ als r geheel is.}$$

Wanneer de bepalingsuitkomsten verkregen zijn uit een inhomogene populatie, kan het probleem omschreven worden als : vind dat gedeelte van de frequentieverdeling, dat beschreven kan worden met een Gamma functie en bereken uit die functie de referentiewaarden voor de hoofdpopulatie. Wanneer de benadering van

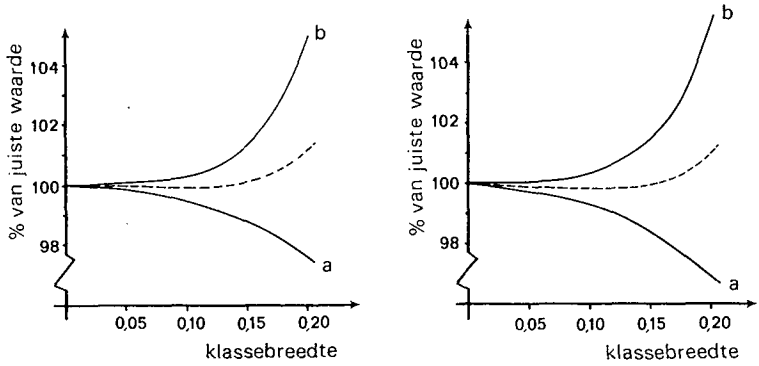
Bhattacharya weer gevolgd wordt, en dus de logaritme van het quotiënt van twee opeenvolgende frequenties (f_{i+1} en f_i) uitgezet wordt tegen het midden van klasse i, ontstaat uiteraard geen rechte lijn, zoals het geval is bij een Gauss verdeling.
Wanneer aangenomen wordt, dat de oppervlakte van een klasse met breedte h evenredig is met de functiewaarde voor de middelste x-waarde in die klasse, geldt :

$$\log \frac{f(x+h)}{f(x)} = (r - 1) \log \left(1 + \frac{h}{x}\right) - \lambda h$$

Deze benadering is slechts exact, wanneer de klassebreedte oneindig klein is. Voor eindige klassebreedte kan afgeleid worden, dat de logaritme van het quotiënt van twee opeenvolgende frequenties bij benadering gelijk is aan (22a) :

$$\log \frac{f(x+h)}{f(x)} = (r - 1) \log \left(1 + \frac{h}{x}\right) - \lambda h + \frac{h^3}{12} \left[\frac{(r - 1)\lambda}{x^2} - \frac{(r - 1)(r - 2)}{x^3}\right]$$

Bovenstaande vergelijking is niet lineair in de parameters. Om waarden voor r en λ te berekenen uit een waargenomen frequentieverdeling zal dan ook een iteratieve rekenprocedure gevolgd moeten worden.
Om na te gaan in hoeverre bovenstaande vergelijkingen inderdaad gebruikt kunnen worden om gamma verdelingen te beschrijven, waarvan de klassen een eindige breedte h hebben, zijn berekeningen uitgevoerd met theoretische gamma verdelingen.
Voor verschillende waarden van r en λ werden theoretische frequentieverdelingen berekend met toenemende klassebreedte. Wanneer uit deze verdelingen Bhattacharya plots berekend worden en hieruit, met behulp van een gewogen kleinste kwadratenbereke-



Figuur 14.
De uit theoretische gamma verdelingen berekende waarde van r (linker figuur) en λ (rechter figuur), uitgedrukt als percentage van de werkelijke waarde. Voor curve a is de vergelijking met 2 termen gebruikt en voor curve b de vergelijking met 3 termen. De gestippelde curve is het gemiddelde van deze twee benaderingen. De klassebreedte is aangegeven in eenheden r/λ.

ning, waarden voor r en λ, blijkt het volgende (zie fig. 14) : Bij toenemende klassebreedte worden de berekende waarden van r en λ kleiner dan de theoretische, wanneer de eenvoudige vergelijking voor de Bhattacharya plot gebruikt wordt. Wanneer de meer gecompliceerde vergelijking met drie termen toegepast wordt, worden de (op een iteratieve manier) berekende waarden van r en λ juist groter dan de theoretische. Het gemiddelde van deze twee benaderingen levert echter, tot een klassebreedte van 0,15 r/λ, "exact" juiste waarden voor r en λ.
In tabel 12 is als voorbeeld uitgewerkt, de berekening van schattingen voor r en λ uit een theoretische gamma verdeling met r=4 en λ =1.

De meest eenvoudige manier om een gamma functie aan te passen door een waargenomen frequentieverdeling, nl. het berekenen van r en λ uit het gemiddelde μ (=r/λ) en de variantie σ^2 (=r/λ^2), kan slechts toegepast worden, wanneer alle waarnemingen afkomstig zijn uit één populatie.

Tabel 12.
De Bhattacharya benadering toegepast op een gamma verdeling (r=4 λ =1 en klassebreedte =0,4).

klasse nummer =i	klasse midden =j	theoretische frequentie $=f_i$	$x=\log\left(1+\frac{h}{j}\right)$	$y=\log\frac{f_{i+1}}{f_i}$	berekende frequentie $=\hat{f}_i$
1	0,4	0,33	0,693	1,555	0,28
2	0,8	1,56	0,405	0,799	1,53
3	1,2	3,47	0,288	0,459	3,45
4	1,6	5,50	0,223	0,269	5,49
5	2,0	7,19	0,182	0,147	7,19
6	2,4	8,34	0,154	0,063	8,34
7	2,8	8,88	0,134	0,001	8,88
8	3,2	8,89	0,118	-0,046	8,89
9	3,6	8,49	0,105	-0,083	8,49
10	4,0	7,81	0,095	-0,114	7,81
11	4,4	6,97	0,087	-0,139	6,97
12	4,8	6,07	0,080	-0,159	6,07
13	5,2	5,17	0,074	-0,177	5,17
14	5,6	4,33	0,069	-0,193	4,33
15	6,0	3,57	0,065	-0,206	3,57

De theoretische frequenties (kolom 3) zijn berekend als :

$$f_i = \left[\sum_{k=0}^{r-1} \frac{e^{-\lambda p}(\lambda p)^{r-1-k} - e^{-\lambda(p+h)}(\lambda(p+h))^{r-1-k}}{(r-1-k)!}\right] * 100$$

waarbij : p = ondergrens van de klasse.

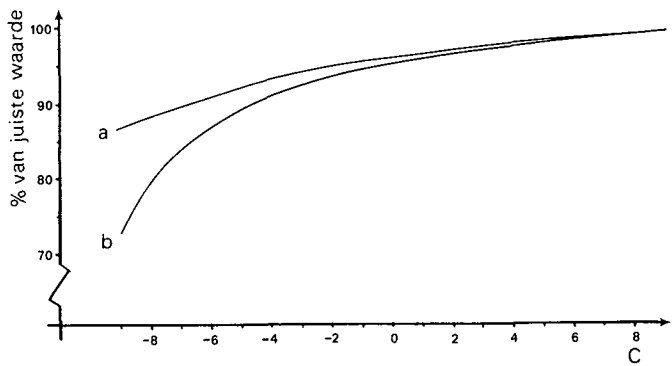
De berekening van x (kolom 4) uit kolom 2 en de berekening van y (kolom 5) uit kolom 3 behoeft geen toelichting. Met een gewogen kleinste kwadratenmethode (zie 4.2.2) kunnen, op de wijze als aangegeven in de tekst, waarden voor r en λ worden geschat : $\hat{r} = 4{,}001$ $\hat{\lambda} = 0{,}999$. Met behulp van deze waarden kan de laatste kolom van bovenstaande tabel worden berekend.

In dit proefschrift wordt er van uitgegaan, dat alle resultaten van een klinisch chemische bepaling in een frequentieverdeling verwerkt worden, zodat ervan uitgegaan moet worden dat deze niet homogeen is. De berekende $\hat{\mu}$ en $\hat{\sigma}^2$ zijn dan niet gelijk aan respectievelijk het eerste en tweede moment van de hoofdpopulatie. Er kan echter in dit geval met een benadering, analoog aan die van Bhattacharya voor een Gauss verdeling, een gamma functie worden aangepast aan de waargenomen frequentieverdeling. Er wordt van uitgegaan, dat voor de kleinste kwadratenberekening een zelfde weegprocedure kan worden toegepast als beschreven in 4.2.2.
Aangezien de klassebreedte niet oneindig klein is, zijn de berekende waarden van r en λ slechts benaderingen. Deze benaderingen zijn echter vrijwel exact voor een zuivere gamma verdeling, wanneer de klassebreedte niet groter is dan $0,15\ r/\lambda$.

5.4.3. Log-Gauss verdelingen, benaderd met een gamma verdeling.

Wanneer in de praktijk blijkt, dat de frequentieverdeling van de resultaten van een bepaling niet Gaussisch is, wordt vaak een logaritmische transformatie toegepast, in de hoop dat de logaritme van de resultaten wél een Gauss verdeling vertoont.
Het is minder gebruikelijk, hoewel juister, om de transformatie $\log (x + C)$ toe te passen, waarbij C zodanig gekozen wordt, dat de frequentieverdeling van de getransformeerde variabele zo goed mogelijk een Gauss verdeling benadert. De scheefheid van de oorspronkelijke frequentieverdeling bepaalt in feite de waarde van de constante C : hoe schever de frequentieverdeling, des te kleiner moet C gekozen worden.
In het voorgaande is al ingegaan op de problemen verbonden aan het transformeren. Ook is gewezen op de gelijkenis tussen gamma verdelingen en log-Gauss verdelingen. Om de overeenkomst tussen deze twee verdelingen te concretiseren, is door een

20-tal log-Gauss verdelingen met verschillende scheefheid, een gamma verdeling aangepast. De (theoretische bekende) 2,5% en 97,5% percentielen zijn vergeleken met de waarden, die berekend werden uit de aangepaste gamma verdeling. De resultaten zijn grafisch weergegeven in figuur 15.



Figuur 15.
Log-Gauss verdelingen, aangepast met een gamma verdeling. De uit de aangepaste gamma verdeling berekende waarde voor het 2,5de en het 97,5de percentiel zijn uitgezet als functie van de "scheefheid" van de log-Gauss verdeling. De gevonden percentielwaarden zijn uitgedrukt als percentage van de werkelijke waarde. De toegepaste vergelijking voor de klassefrequenties van de log-Gauss verdelingen is :

$$f(x) = \int_{x-\frac{1}{2}h}^{x+\frac{1}{2}h} \left[\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}\left(\frac{\log(x+C) - \mu}{\sigma}\right)^2\right) \right] dx$$

waarbij : $\mu = (\log((35+C)(10+C)))/2$

$\sigma = (\log((35+C)/(10+C)))/5{,}5$

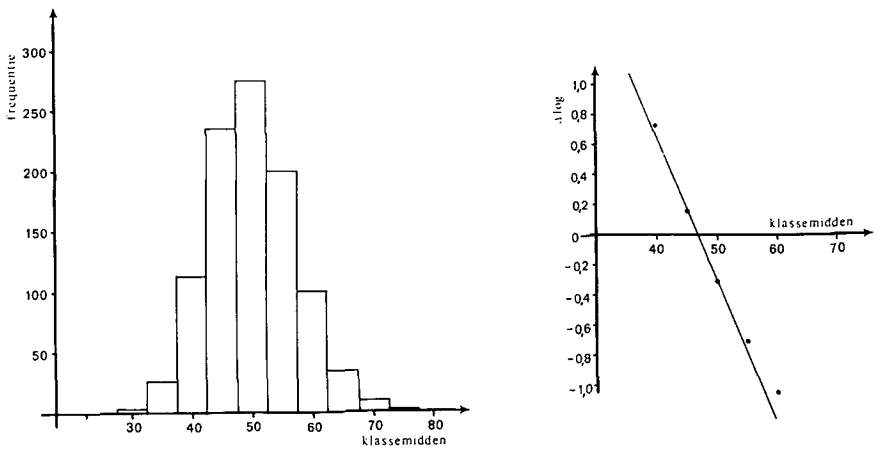
$h = 1$

$10 \leqslant x \leqslant 35$

$-9 \leqslant C \leqslant 10$

5.4.4. Keuze van de benadering, wanneer er twijfel is over de vorm van de frequentieverdeling.

Zoals reeds opgemerkt is in 5.2, zijn er geen eenvoudige toetsen beschikbaar, die toegepast kunnen worden op gecontamineerde frequentieverdelingen, om te beslissen of de hoofdpopulatie verdeeld is volgens een Gauss curve. In de meeste gevallen zal echter in de praktijk geen twijfel bestaan over het al dan niet Gaussisch verdeeld zijn van de uitkomsten van een klinisch chemische bepaling. In de eerste plaats is van de meeste bepalingen bekend, hoe de vorm van de frequentieverdeling van de uitkomsten is. In de tweede plaats kan in de Bhattacharya plot, die berekend wordt uit een waargenomen frequentieverdeling, al dan niet een recht gedeelte onderscheiden worden.
In twijfelgevallen, d.w.z. wanneer het niet duidelijk is of het vloeiend verlopende deel van de Bhattacharya plot lineair is of licht gebogen, is de keuze tussen het berekenen van een Gauss- of een gamma curve door de waargenomen frequentieverdeling moei-



Figuur 16.
Theoretische gamma verdeling (r=50, λ =1) ten onrechte (op grond van de nagenoeg rechte Bhattacharya plot) benaderd met een Gauss verdeling. De berekende referentiewaarden (2,5de en 97,5de percentiel) zijn niet erg verschillend van de theoretische. Voor verdere toelichting zie de tekst.

lijk. In deze gevallen zijn de resultaten dus verdeeld volgens een Gauss curve of volgens een gamma verdeling met vrij grote r, d.w.z. een vrijwel symmetrische verdeling, die aanleiding geeft tot een Bhattacharya plot die slechts zeer licht gebogen is. Beide benaderingen zullen in deze situatie vrijwel gelijke referentiewaarden opleveren.

Als voorbeeld is in figuur 16 een gamma verdeling ($r=50$ $\lambda=1$) benaderd met een Gauss curve. Uit deze figuur blijkt, dat de Bhattacharya plot inderdaad slechts zeer licht gebogen is. Wanneer door de punten van deze plot een rechte lijn berekend wordt (en dus in feite op grond van de figuur ten onrechte aangenomen wordt, dat de frequentieverdeling Gaussisch is), resulteren als referentiewaarden (berekend als $\hat{\mu} \pm 1{,}96\hat{\sigma}$): 35,6 - 63,1. Deze waarden wijken slechts weinig af van de theoretische 2,5 en 97,5 percentielwaarden, die resp. gelijk zijn aan 37,1 en 64,8.

6. Toepassingen.

6.1. Inleiding.

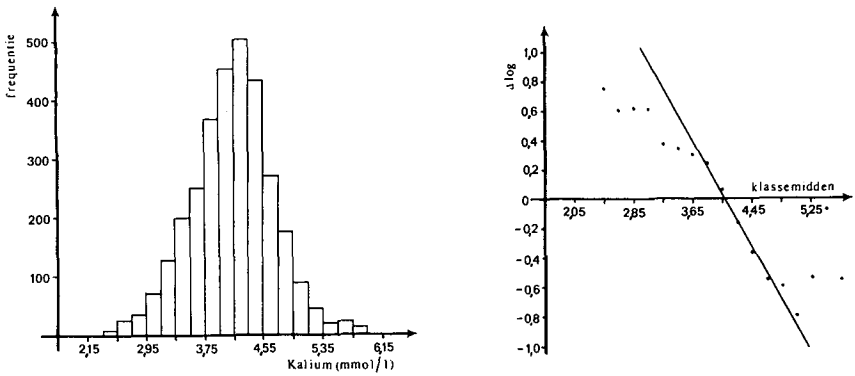
In dit hoofdstuk zal het berekenen van referentiewaarden uit patiëntengegevens toegelicht worden aan de hand van enkele min of meer willekeurig gekozen voorbeelden van bepalingen die in ons laboratorium worden uitgevoerd.
Er is een onderverdeling gemaakt tussen bepalingen, waarvan de frequentieverdeling Gaussisch is (als voorbeelden worden behandeld de bepaling van kalium, calcium en thyroxine) en bepalingen waarvoor dit niet het geval is (als voorbeelden ASAT, ALAT en urinezuur). Tot slot worden twee publicaties gereproduceerd, de eerste handelend over de toepassing van de Bhattacharya plot voor de berekening van referentiewaarden voor een aantal hematologische bepalingen en de tweede over de berekening van meerdimensionale referentiewaarden met behulp van patiëntengegevens.

Het verzamelen van de gegevens gebeurt gedeeltelijk handmatig, gedeeltelijk ook automatisch. Voor twee grote analyse automaten in ons laboratorium (SMA 12/60 en Hemalog 8/90) wordt nl. per analyseserie een ponsband geproduceerd, waarop naast de gegevens van de patiënt (naam, geboortedatum, geslacht en de aanduiding poliklinisch, klinisch of donor) de analyse uitkomsten staan. In het kader van de kwaliteitsbewaking wordt uit deze gegevens per bepaling een daggemiddelde berekend. Hiervoor worden alleen dié patiëntengegevens gebruikt, die niet meer dan 2,5 sd afwijken van het gemiddelde van het betreffende referentiewaardengebied. Met behulp van deze ponsband kunnen, naast het berekenen van een daggemiddelde, ook alle analyseresultaten eenvoudig verzameld worden, opgesplitst naar geslacht, leeftijdskategorieën en dergelijke. Dit maakt het mogelijk om ook voor "moeilijke" groepen, zoals bejaarden en kinderen, referentiewaarden vast te stellen.

6.2. Enkele voorbeelden van bepalingen waarvan de frequentieverdeling Gaussisch is.

6.2.1. Kalium.

Kalium is het belangrijkste intracellulaire kation. De concentratie binnen de cellen is ongeveer een factor 23 groter dan de concentratie in de extracellulaire vloeistof. Dit grote concentratieverschil wordt in stand gehouden door een actief transportmechanisme enerzijds en diffusie (ten gevolge van actief transport van natrium uit de cel) anderzijds. Wanneer serum na bloedafname niet vrij snel van de bloedcellen wordt gescheiden, kan door diffusie van kalium uit de erythrocyten een aspecifieke verhoging van de serumconcentratie optreden.
Een verhoging van de extracellulaire kaliumconcentratie (in vivo) heeft tot gevolg dat de membraanpotentiaal minder negatief wordt, met onder andere als gevolg dat de hartslag langzamer en minder krachtig wordt. Een plasmaconcentratie boven 10 mmol/l is meestal dodeljk. De belangrijkste oorzaken van hyperkaliaemie zijn : een onvermogen van het lichaam om kalium uit te scheiden, bijv. ten gevolge van nierafwijkingen, of celbeschadiging, waardoor kalium vanuit de cellen in de extracellulaire vloeistof komt (bijv. bij hemolyse). Een verlaging van de kaliumconcentratie in plasma kan het gevolg zijn van verminderde opname of verhoogde uitscheiding, zoals bij ernstige diarree of braken. Ook een hoge aldosteronconcentratie (zoals bijv. kan optreden bij de ziekte van Conn) heeft een verhoogde uitscheiding van kalium en vaak een verlaagde plasmaconcentratie tot gevolg.
De bepaling van kalium geschiedt met emissie vlamfotometrie met lithium als interne standaard. Het door kalium geëmitteerde licht wordt gemeten bij 768 nm. De in ons laboratorium gebruikte vlamfotometer is van de firma Technicon.
In de periode waarin de patiëntenuitkomsten van kalium verzameld zijn (januari en februari 1981), was de variatiecoëfficiënt van



Figuur 17.
A. Frequentieverdeling van 3148 kalium uitkomsten en
B. de hieruit berekende Bhattacharya plot.

Tabel 13.
Details van de Bhattacharya plot voor kalium. Eenheden van de eerste kolom : mmol/l.

klassemidden	frequentie	frequentie na smoothen	Δ log
2,05	1	–	–
2,25	2	–	–
2,45	8	9,9	0,748
2,65	25	21,0	0,601
2,85	36	38,4	0,615
3,05	70	71,0	0,604
3,25	128	129,8	0,374
3,45	199	188,8	0,342
3,65	250	265,6	0,301
3,85	367	359,1	0,247
4,05	453	459,6	0,071
4,25	503	493,7	-0,156
4,45	434	422,2	-0,376
4,65	271	289,9	-0,545
4,85	176	168,2	-0,578
5,05	90	94,3	-0,783
5,25	45	43,1	-0,528
5,45	22	25,4	–
5,65	24	–	–
5,85	14	–	–
>	30		
	3148		

de dagelijks geanalyseerde autonorm contrôlesera zowel in het hoge (6 mmol/l) als in het lage (3,5 mmol/l) gebied kleiner dan 1,5%. Dit getal geeft een indruk over de reproduceerbaarheid van de bepaling. De Wellcome kwaliteitscontrôle uitslagen in deze periode waren ten opzichte van laboratoria met dezelfde analysemethode gemiddeld 0,32 sd te hoog (laagste uitkomst -1,7 sd en hoogste uitkomst +1,2 sd ; in de verzamelperiode zijn 5 analyserondes geweest). Dit geeft een, zij het beperkte, indruk over de juistheid van de verkregen uitkomsten.

De frequentieverdeling van de verzamelde analyseresultaten is weergegeven in figuur 17A. Uit de hieruit berekende Bhattacharya plot (figuur 17B en tabel 13) blijkt, dat er een duidelijk recht gedeelte onderscheiden kan worden. Er kan dus aangenomen worden, dat de frequentieverdeling van de kalium uitkomsten Gaussisch is. De met de gewogen kleinste kwadratenmethode berekende rechte lijn heeft als vergelijking :

$$\Delta \log = -0{,}900 \text{ (klassemidden)} + 3{,}68$$

De helling is dus gelijk aan -0,900. De doorsnijding met de X-as wordt gevonden bij 4,09. Met een klassebreedte van 0,2 volgt hieruit voor het geschatte gemiddelde en standaarddeviatie :

$$\hat{\mu} = 4{,}09 + \tfrac{1}{2}\,0{,}2 = 4{,}19 \qquad \hat{\sigma}^2 = -\frac{0{,}2}{-0{,}900} - \frac{0{,}2^2}{12} = 0{,}219$$

zodat de berekende referentiewaarden zijn: 3,27 - 5,11 mmol/l ($\hat{\mu} \pm 1{,}96\,\hat{\sigma}$). Door Tietz (113) wordt een referentiewaardengebied opgegeven van 3,5 - 5,3 mmol/l. Hoewel deze waarden 0,2 mmol/l hoger zijn, is de breedte van het gebied gelijk. Het valt op (zie figuur 17A), dat het grootste deel van de "pathologische" uitkomsten verlaagd is. Verhoogde kalium uitkomsten zijn veel zeldzamer.

6.2.2. Calcium.

Meer dan 99% van de totale hoeveelheid calcium in het menselijk lichaam bevindt zich in de botten. De resterende hoeveelheid heeft een aantal zeer belangrijke biochemische functies. Zo spelen Ca^{++} ionen een wezenlijke rol in de bloedstolling en als cofactoren voor een aantal enzymen.

Het overgrote deel van het calcium in het bloed komt extracellulair voor en wel in twee vormen : voor circa 45% eiwitgebonden (voornamelijk aan albumine) en voor circa 55% vrij. Deze laatste kategorie kan nog onderverdeeld worden in calcium in de vorm van complexen (citraten, fosfaten e.d., enkele procenten van het totaal) en calcium in geïoniseerde vorm. Alleen deze laatste kategorie is fysiologisch actief.

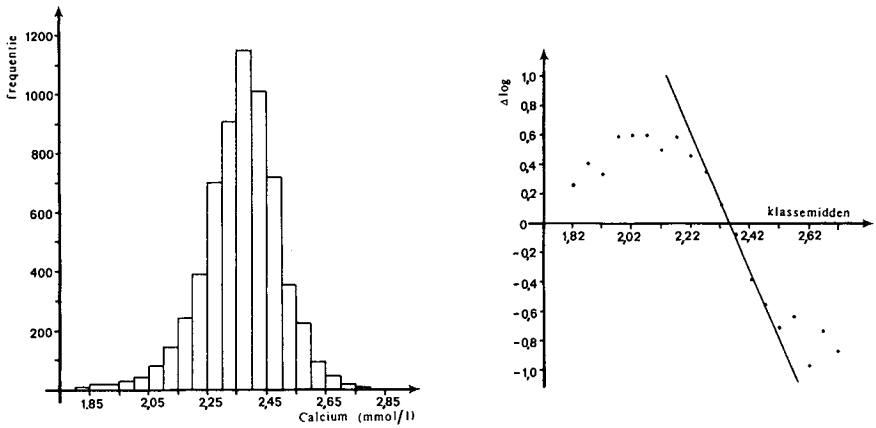
De calcium concentratie in plasma wordt gereguleerd door een aantal hormonen, waarvan de belangrijkste zijn :

-Parathyroid hormoon : een verlaging van de Ca^{++} ionen concentratie in het bloed stimuleert de bijschildklieren tot afgifte van dit hormoon. De effecten ervan zijn: mobiliseren van calcium uit de botten, stimuleren van de calcium opname uit de darm en verminderen van de calcium uitscheiding via de nieren. Alle drie effecten dragen bij aan een verhoging van de calcium concentratie in plasma.

-Calcitonine : een hoge Ca^{++} ionenconcentratie stimuleert de schildklier tot afgifte van dit hormoon, waarvan het belangrijkste effect is het afremmen van de botresorptie.

-Vitamine D. De actieve vorm van vitamine D, 1,25-dihydrocholecalciferol, verhoogt de opname van calcium in het bloed vanuit de darm en heeft ook een stimulerend effect op de botresorptie.

De totale hoeveelheid calcium in serum (vrij en gebonden aan eiwit) wordt in ons laboratorium bepaald op de SMA 12 m.b.v. de o-cresolphtaleïne-complexon methode. Hierbij wordt calcium van eiwit gedissociëerd door toevoeging van HCl. Hierna wordt het calcium gedialyseerd in een reagensstroom van o-cresolphta-



Figuur 18.
A. Frequentieverdeling van 6220 calcium uitkomsten en
B. de hieruit berekende Bhattacharya plot.

Tabel 14.
Details van de Bhattacharya plot voor calcium. Eenheden van de eerste kolom : mmol/l.

klassemidden	frequentie	frequentie na smoothen	Δlog
1,62	1	–	–
1,67	3	–	–
1,72	2	2,7	0,098
1,77	4	3,0	1,031
1,82	6	8,5	0,267
1,87	16	11,2	0,411
1,92	12	16,8	0,340
1,97	28	23,7	0,592
2,02	41	42,8	0,601
2,07	78	78,2	0,598
2,12	144	142,2	0,496
2,17	241	233,6	0,593
2,22	391	422,8	0,457
2,27	702	667,6	0,360
2,32	910	957,3	0,136
2,37	1152	1096,9	-0,067
2,42	1013	1025,5	-0,379
2,47	721	702,0	-0,552
2,52	358	404,2	-0,710
2,57	227	198,8	-0,629
2,62	92	105,9	-0,968
2,67	46	40,2	-0,729
2,72	19	19,4	–
2,77	8	–	–
2,82	5	–	–

leïne-complexon, waaraan toegevoegd 8-hydroxyquinoline.(Mg wordt door 8-hydroxyquinoline gebonden, waardoor geen storing van dit ion kan optreden.) Na toevoegen van een alkalische diëthylamine buffer ontstaat bij een pH van ongeveer 11 een gekleurd complex, waarvan de extinctie gemeten wordt bij 570 nm.
De calibratie van de bepaling gebeurt met Lock-In serum van de firma General Diagnostics. (In de periode waarin de hier vermeldde gegevens verzameld zijn lotnr. OA861, waarvan de calcium concentratie opgegeven wordt als 2,45 mmol/l.)
In de periode waarin de patiëntenuitkomsten van calcium verzameld zijn (14-7-1981 t/m 30-11-1981), was de variatiecoëfficiënt van de dagelijks geanalyseerde autonorm sera in het hoge gebied (2,96 mmol/l) 1,14% en in het lage gebied (1,88 mmol/l) 1,53%. De Wellcome kwaliteitscontrôle uitslagen in dezelfde periode waren ten opzichte van laboratoria met dezelfde analysemethode gemiddeld 0,30 sd te laag (range -0,84 sd tot 0,37 sd; 9 analyse rondes).

De frequentieverdeling van de 6220 verzamelde analyseresultaten is weergegeven in figuur 18A. De Bhattacharya plot, die hieruit berekend is, vertoont een duidelijk lineair gedeelte, zodat aangenomen kan worden dat de Ca uitkomsten verdeeld zijn volgens een Gauss curve (zie figuur 18B en tabel 14).
De berekende rechte lijn heeft als vergelijking :

$$\Delta \log = -4{,}743 \text{ (klassemidden)} + 11{,}143.$$

De helling is dus gelijk aan -4,743. De doorsnijding van de X-as wordt gevonden bij 2,349. Met een klassebreedte van 0,05 volgt hieruit :

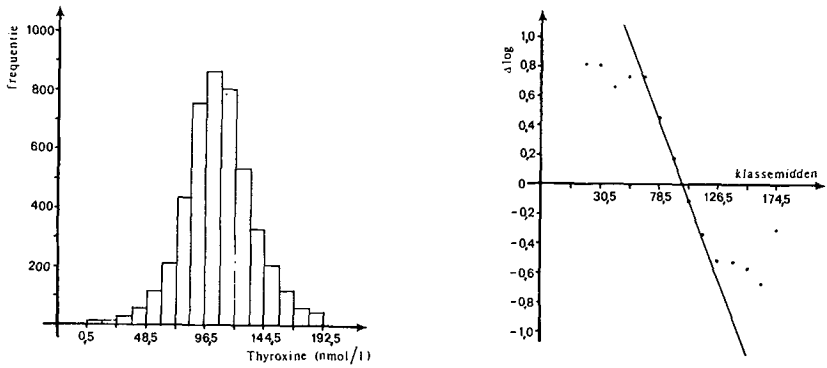
$$\hat{\mu} = 2{,}349 + \tfrac{1}{2}\,0{,}05 = 2{,}374 \qquad \hat{\sigma}^2 = -\frac{0{,}05}{-4{,}743} - \frac{0{,}05^2}{12} = 0{,}01034$$

zodat de berekende referentiewaarden zijn ($\hat{\mu} \pm 1{,}96\,\hat{\sigma}$) :

2,17-2,57 mmol/l. Door Tietz (113) wordt een referentiewaardengebied opgegeven van 2,12-2,60 mmol/l. Er werd geen verschil geconstateerd in de referentiewaarden voor mannen en vrouwen. Ook werd geen afhankelijkheid van de leeftijd vastgesteld. In dezelfde periode waarin de patiëntengegevens verzameld zijn, is ook serum van 547 (gezonde) bloeddonoren geanalyseerd. De hieruit berekende referentiewaarden zijn : 2,17-2,58 mmol/l. Dit gebied is gelijk aan dat berekend uit patiëntengegevens. Uit figuur 18A blijkt dat het grootste deel van de pathologische Ca uitslagen verlaagd is.

6.2.3. Thyroxine.

Ongeveer 98% van de hoeveelheid schildklierhormoon is thyroxine (T4). De resterende 2% is triiodothyronine (T3). Schildklierhormoon wordt in het bloed grotendeels getransporteerd, gebonden aan thyroxine bindend globuline (TBG), thyroxine bindend prealbumine en albumine. Slechts een zeer klein gedeelte (< 0,1% van het totaal) van het schildklierhormoon komt ongebonden in het bloed voor. T4 wordt in de circulatie grotendeels omgezet in T3, dat een ongeveer 4x zo sterke werking heeft.
De T4 concentratie in het bloed wordt geregeld via een negatief terugkoppelingsmechanisme : een hoge T4 concentratie remt de afgifte van TRF (thyrotropine releasing factor) door de hypothalamus ; hierdoor ontstaat een vermindering van de TSH (thyroid stimulating hormone) afgifte door de hypofyse, met als gevolg een verminderde uitscheiding van T4 en T3 door de schildklier.
Het belangrijkste biologisch effect van schildklierhormonen is in de meeste organen een toename van het zuurstofverbruik, waarbij de extra geleverde energie niet opgeslagen wordt in energierijke fosfaten, maar omgezet wordt in warmte. De meest opvallende verschijnselen van hyperthyreoidie (ziekte van Graves) zijn dan ook : het niet kunnen verdragen van warmte, sterk transpireren, gewichtsverlies, diarree, spierzwakte, extreme



Figuur 19.
A. Frequentieverdeling van 4535 thyroxine uitkomsten en
B. de hieruit berekende Bhattacharya plot.

Tabel 15.
Details van de Bhattacharya plot voor thyroxine. Eenheden van de eerste kolom : nmol/l.

klassemidden	frequentie	frequentie na smoothen	Δ log
6,5	11	–	–
18,5	11	–	–
30,5	25	23,9	0,812
42,5	54	54,0	0,663
54,5	111	104,8	0,736
66,5	209	218,9	0,735
78,5	433	456,5	0,458
90,5	752	722,0	0,183
102,5	860	867,0	-0,113
114,5	800	774,0	-0,336
126,5	533	552,9	-0,515
138,5	323	330,2	-0,519
150,5	201	196,5	-0,562
162,5	113	112,0	–
174,5	57	–	–
186,5	42	–	–

vermoeidheid maar niet kunnen slapen, beven van de handen. De verschijnselen van hypothyreoidie (myxoedeem) zijn over het algemeen tegengesteld aan die bij hyperthyreoidie.

De bepaling van thyroxine wordt in ons laboratorium uitgevoerd met de radioimmunoassay kit van de firma Clinical Assays. Per bepalingsserie worden 3 contrôlesera geanalyseerd naast ongeveer 30 patiëntenmonsters. De variatiecoëfficiënt van de resultaten van deze 3 monsters waren in de periode waarin de patiëntengegevens verzameld werden resp. : in het lage gebied (42 nmol/l) 7,8%, in het middengebied (79 nmol/l) 5,6% en in het hoge gebied (185 nmol/l) 6,0%. De afwijking ten opzichte van laboratoria die dezelfde radioimmunoassay kit gebruiken, bedroeg in de periode waarin de thyroxine uitkomsten verzameld zijn gemiddeld -0,44 sd (range -0,89 sd tot +0,14 sd ; 6 analyse rondes).

De frequentieverdeling van de 4535 verzamelde thyroxine uitkomsten is weergegeven in figuur 19A. Uit het duidelijk lineaire gedeelte in de Bhattacharya plot (figuur 19B en tabel 15), die hieruit is berekend, blijkt, dat de frequentieverdeling van de thyroxine uitkomsten Gaussisch is. De berekende rechte lijn heeft als vergelijking :

$$\Delta \log = -0{,}02237 \text{ (klassemidden)} + 2{,}2027.$$

Met behulp van de helling (= -0,0223), de doorsnijding van de X-as (=98,5) en de klassebreedte (=12) wordt als schatting van het gemiddelde en de standaarddeviatie gevonden :

$$\hat{\mu} = 98{,}5 + \tfrac{1}{2}\,12 = 104{,}5 \qquad \hat{\sigma}^2 = -\frac{12}{-0{,}02237} - \frac{12^2}{12} = 524{,}4$$

De referentiewaarden voor thyroxine zijn dus: 59,6-149,4 nmol/l. Van de pathologische thyroxine uitkomsten is het grootste gedeelte verhoogd, hoewel ook een vrij groot aantal uitkomsten

verlaagd is.
Aangezien de uitkomsten van de bepaling van thyroxine sterk afhankelijk zijn van de gebruikte analysemethode, is een exacte literatuurvergelijking moeilijk. In de bijsluiter van de door ons gebruikte radioimmunoassay kit, wordt als referentiewaarden aanbevolen 58 - 148 nmol/l.

In het algemeen wijzen afwijkende thyroxine uitkomsten op het bestaan van schildklierpathologie. Vaak kan medicamenteuze therapie de schildklierfunctie weer min of meer normaliseren. Er kan dus gesteld worden, dat de kans groot is, dat mensen waarbij een afwijkende thyroxine uitkomst gevonden is, medicamenteuze therapie zullen krijgen. Vervolgens zal bijv. elke 3 maanden ter contrôle van de therapie een thyroxine bepaling worden uitgevoerd. Deze mensen, waarbij bij juiste therapie normale thyroxine uitkomsten gevonden worden, mogen uiteraard niet als gezond worden beschouwd. Het is echter denkbaar dat de analyseuitkomsten van deze personen in het middengebied van de frequentieverdeling tot een verstoring aanleiding kunnen geven.
Er blijkt dat ongeveer 10% van de geanalyseerde monsters afkomstig is van mensen die schildkliertherapie ontvangen. Van deze uitkomsten blijkt ongeveer 25% in het gebied te liggen waar de Bhattacharya plot recht is. Dit aantal is dermate klein, dat niet verwacht kan worden, dat de berekende referentiewaarden hierdoor ernstig zullen worden beïnvloed.

6.3. Enkele voorbeelden van bepalingen waarvan de frequentieverdeling niet Gaussisch is.

6.3.1. De transaminases ASAT (SGOT) en ALAT (SGPT).

De transaminases vormen een groep enzymen die de interconversie van aminozuren en α ketozuren katalyseren. Deze interconversie

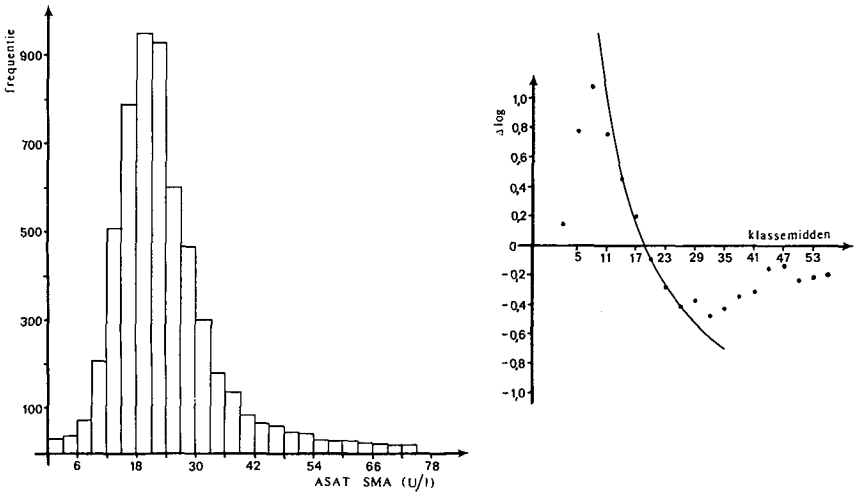
gebeurt door overdracht van een aminogroep. Transaminases zijn intracellulaire enzymen en een verhoging van de serumconcentratie wijst dan ook op celbeschadiging. De intracellulaire concentraties van ASAT en ALAT zijn echter niet gelijk in verschillende weefsels. In hartspier is de ASAT concentratie bijv. een factor 17 hoger dan de concentratie van ALAT. Na een myocard infarct is de ASAT activiteit in serum dan ook veel sterker verhoogd dan die van ALAT. De stijging van de serumactiviteit begint 4 tot 6 uur na het begin van het infarct; de piekactiviteit wordt bereikt na 24 tot 36 uur, terwijl na 4 à 5 dagen de activiteit in serum weer normaal is. Er bestaat geen absolute correlatie tussen de maximale stijging van de ASAT activiteit en de prognose, hoewel zeer hoge enzymactiviteiten vaak samengaan met dodelijk verlopende infarcten.
Bij virale hepatitis en andere leverziektes waarbij celbeschadiging optreedt, is de activiteit van beide enzymen in serum zeer sterk verhoogd. De ALAT activiteit over het algemeen echter meer dan die van ASAT. De stijging van de enzymactiviteiten is minder bij intrahepatische cholestase.
Andere ziektes waarbij celbeschadiging optreedt, zullen ook een stijging van de transaminase activiteit in serum te zien geven, waarbij de hoogte van de stijging grofweg samenhangt met de intracellulaire concentratie van het enzym en de uitgebreidheid van de beschadiging. Zo zal hemolyse over het algemeen slechts een lichte stijging van de transaminase activiteit in serum tot gevolg hebben door de vrij lage concentratie van zowel ASAT als ALAT in erythrocyten.

6.3.1.1. ASAT (SGOT ; EC 2.6.1.1).

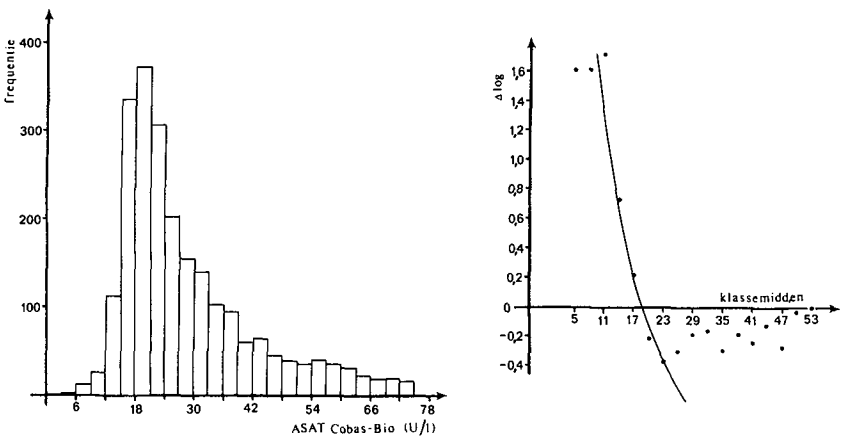
ASAT katalyseert de volgende reactie :
asparaginezuur + α-ketoglutaarzuur ⇆
oxaalazijnzuur + glutaminezuur.

Het gevormde oxaalazijnzuur wordt onder verbruik van NADH omgezet in malonzuur door malaatdehydrogenase. De afname van de absorptie bij 340 nm, veroorzaakt door het verbruik van NADH, kan ofwel na een vaste tijd gemeten worden (eindpuntsmeting) ofwel kinetisch worden vervolgd, waarbij de snelheid van verdwijnen van NADH een maat is voor de ASAT activiteit. In ons laboratorium worden beide meetprincipes gebruikt : op de SMA (12 kanaals continuous flow analyse automaat van de firma Technicon) wordt de bepaling van de ASAT activiteit als eindpuntsmeting uitgevoerd en op de Cobas-Bio (centrifugaal analyser van de firma Hoffmann La Roche) als kinetische bepaling, beide bij 37 °C. De calibratie van de ASAT meting op de SMA geschiedt met Lock-In serum (lotnr. OA861, waarvan de waarde wordt opgegeven als 90 U/l) van General Diagnostics. Op de Cobas-Bio wordt de helling van de lijn, die de afname van de extinctie bij 340 nm weergeeft, omgerekend naar een ASAT activiteit m.b.v. de factor 1323. Deze factor wordt berekend m.b.v. de micromolaire extinctie coëfficiënt van NADH bij 340 nm (0,0063 $cm^2/\mu mol$), die gemeten is met NADH van de firma Boehringer (nicotinamid adenin dinucleotid reduziert, Reinheitsgrad 1) op een Zeiss spectrofotometer type DM4.
De (kinetische) meting van de ASAT activiteit op de Cobas-Bio kan in feite als een absolute meting worden beschouwd, aangezien op dit apparaat de werkelijke snelheid waarmee het substraat verbruikt wordt, gemeten kan worden. Een eindpuntsmeting (zoals op de SMA wordt toegepast) levert slechts een indirecte waarde voor de enzymactiviteit.

Er werden 5988 ASAT uitkomsten van de SMA en 2702 ASAT uitkomsten van de Cobas-Bio verzameld (figuur 20A en 21A). De frequentieverdelingen zijn duidelijk asymmetrisch, zoals ook blijkt uit het feit dat de Bhattacharya plots (figuur 20B en 21B) geen recht gedeelte vertonen. Toepassing van de in hoofdstuk 5 beschreven methode levert de volgende schattingen op



Figuur 20.
A. Frequentieverdeling van 5988 ASAT uitkomsten geänalyseerd op de SMA en
B. de hieruit berekende Bhattacharya plot.



Figuur 21.
A. Frequentieverdeling van 2702 ASAT uitkomsten geänalyseerd op de Cobas-Bio en
B. de hieruit berekende Bhattacharya plot.

Tabel 16.
Details van de Bhattacharya plots voor ASAT. Eenheden van de eerste kolom : U/l.

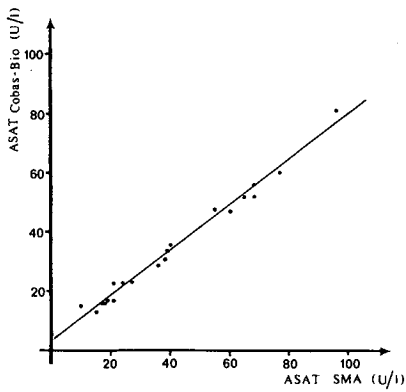
	SMA 12/60			Cobas-Bio		
klassemidden	frequentie	frequentie na smoothen	Δlog	frequentie	frequentie na smoothen	Δlog
2	31	-	-	0	-	-
5	36	-	-	1	-	-
8	87	78,9	1,080	12	5,1	1,625
11	205	232,3	0,761	25	25,9	1,720
14	506	497,2	0,457	112	144,6	0,729
17	788	785,4	0,197	334	299,8	0,219
20	952	956,2	-0,087	372	373,2	-0,206
23	929	876,3	-0,280	306	303,6	-0,360
26	601	662,5	-0,409	202	211,9	-0,300
29	465	440,0	-0,368	154	157,0	-0,183
32	301	304,5	-0,468	140	130,7	-0,154
35	181	190,7	-0,418	103	112,0	-0,291
38	136	125,5	-0,339	94	83,7	-0,177
41	84	89,4	-0,307	59	70,1	-0,230
44	66	65,8	-0,151	64	55,7	-0,116
47	60	56,6	-0,126	46	49,6	-0,272
50	46	49,9	-0,231	39	37,8	-0,019
53	44	39,6	-0,207	35	37,1	0,008
56	29	32,2	-0,195	40	37,4	-0,027
59	27	26,5	-0,004	35	36,4	-0,190
62	27	26,4	-0,091	31	30,1	-0,269
65	24	24,1	-0,202	23	23,0	-0,212
68	20	19,7	-	17	18,6	-
71	16	-	-	19	-	-
74	16	-	-	16	-	-
>	311			423		
	5988			2702		

voor r en λ van de aangepaste gamma verdeling :

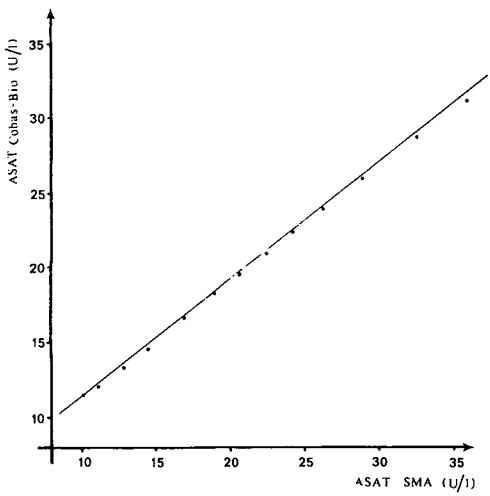
SMA	:	$\hat{r} = 11,93$	$\hat{\lambda} = 0,536$
Cobas-Bio	:	$\hat{r} = 15,60$	$\hat{\lambda} = 0,727$

Het 97,5de percentiel wordt gevonden bij een ASAT waarde van 36,7 (SMA) respectievelijk 33,4 (Cobas-Bio). Het valt op dat het percentage "pathologische" ASAT uitkomsten hoger is voor de Cobas-Bio dan voor de SMA. De oorzaak hiervan is waarschijnlijk, dat de analyses die uitgevoerd worden op de Cobas-Bio "gericht" worden aangevraagd, d.w.z. de aanvrager van het onderzoek heeft in veel gevallen, bijv. op grond van het klinische beeld, reden om een afwijkende uitkomst te vermoeden. Wanneer echter een SMA-12 profiel wordt aangevraagd (een normale laboratorium aanvrage bij bijv. iedere ziekenhuis opname), heeft de aanvrager van het onderzoek in de meeste gevallen geen enkele reden om een afwijkende ASAT uitkomst te verwachten. Op grond van het bovenstaande is het dus logisch, dat het percentage "pathologische" ASAT uitkomsten groter is op de Cobas-Bio dan op de SMA.

Het verschil in referentiewaarde, berekend uit patiëntenresultaten van twee verschillende analyse apparaten, is erg groot. Zoekend naar een verklaring, werden ongeveer 35 patiënten monsters geanalyseerd op beide apparaten. Uit figuur 22, waarin de resultaten lineair tegen elkaar zijn uitgezet, blijkt dat de SMA uitkomsten een factor 1,33 hoger zijn dan de uitkomsten van de Cobas-Bio. Aangezien de Cobas-Bio uitkomsten als juist beschouwd mogen worden (zie boven), is het vanzelfsprekend om de oorzaak van dit grote verschil in de SMA te zoeken. De calibratie van de SMA gebeurt met Lock-In serum van de firma General Diagnostics. De ASAT uitkomst van dit serum wordt door de firma opgegeven als 90 U/l (voor de SMA methode). Wanneer deze opgegeven waarde een bepaalde factor afwijkt van de werkelijke waarde (zoals hier waarschijnlijk het geval is),



Figuur 22.
Verband tussen de ASAT uitkomsten van de SMA en de Cobas-Bio. De berekende rechte heeft als vergelijking : y = 4,2 + 0,75x.



Figuur 23.
De getrokken lijn stelt het berekende verband voor tussen de ASAT uitkomsten van de SMA en de Cobas-Bio (figuur 22). De punten zijn percentielen, berekend uit de frequentieverdelingen van patiëntenresultaten voor de beide apparaten (figuur 20 en 21). Voor verdere toelichting zie de tekst.

worden alle patiëntenuitkomsten een zelfde factor te hoog of te laag gevonden. Wanneer aangenomen wordt dat de opgegeven waarde van het Lock-In serum voor ASAT onjuist is en in plaats van 90 U/l, 90/1,33 = 68 U/l moet zijn, dan wordt de SMA voor deze bepaling dus fout gecalibreerd. Deze calibratiefout kan het verschil in berekende referentiewaarde nagenoeg verklaren.
In figuur 23 zijn de percentielwaarden, berekend uit de frequentieverdeling van de SMA resultaten, uitgezet tegen die van de Cobas-Bio. De resulterende rechte valt nagenoeg samen met de lijn die het verband aangeeft tussen de analyse uitkomsten die voor 35 patiëntenmonsters op beide apparaten werden gevonden.
Het (na correctie) toch nog iets ruimere gebied, berekend uit de SMA resultaten, komt waarschijnlijk voort uit het feit, dat de dag-tot-dag variatie van de SMA groter is dan die van de Cobas Bio, zoals blijkt uit de resultaten van de dagelijkse Autonorm contrôles : de variatiecoëfficiënt in het lage gebied (50 U/l) is voor de SMA in de verzamelperiode 4,70% en voor de Cobas-Bio 2,16% en in het hoge gebied (110 U/l) respectievelijk 3,64% en 2,08%.

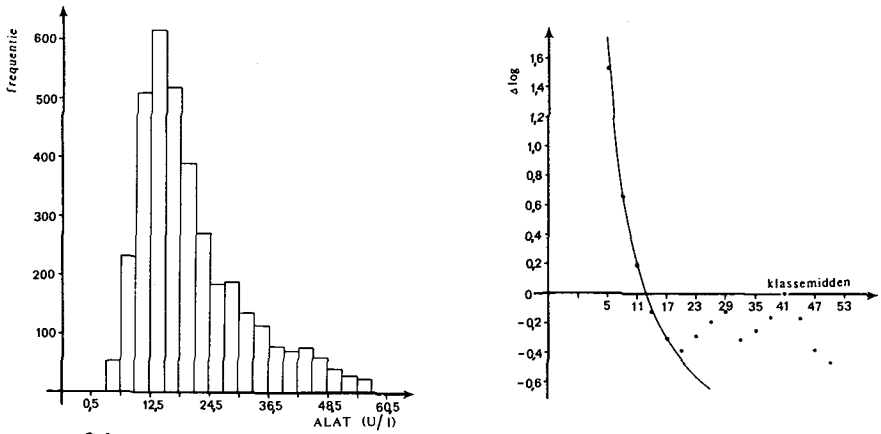
De gemiddelde afwijking t.o.v. laboratoria met dezelfde analyse methode bedroeg in de verzamelperiode voor de Cobas-Bio -0,17 sd (range -0,7 sd tot +0,6 sd ; aantal enquêtes 10).

Er werd een gering verschil in referentiewaarde gevonden voor mannen en vrouwen. Dit verschil (ongeveer 2 U/l voor het 97,5de percentiel) is te klein om in de praktijk van belang te zijn. Ook door Siest et al. (106) wordt een verschil van een dergelijke grootte gerapporteerd, althans voor de modale waarde. Uit 413 ASAT uitkomsten, gevonden in serum van mannelijke bloeddonoren, werd een waarde van 40 U/l gevonden voor het 97,5de percentiel. Dit is goed in overeenstemming met de waarde van 39 U/l, berekend uit de resultaten van 2722 mannelijke patiënten in dezelfde periode. Deze uitkomsten zijn geproduceerd op de SMA, voordat de onjuiste calibratie gecorrigeerd werd.

6.3.1.2. ALAT (SGPT ; EC 2.6.1.2).

ALAT katalyseert de reactie :
alanine + α-ketoglutaarzuur ⇌ pyrodruivezuur + glutaminezuur.
Het gevormde pyrodruivezuur wordt in een indicatorreactie door lactaatdehydrogenase (LDH ; EC 1.1.1.27) omgezet in melkzuur onder verbruik van NADH. De afname van de concentratie van NADH kan bij 340 nm vervolgd worden. In ons laboratorium wordt de ALAT activiteit kinetisch gemeten bij 37 °C op de Cobas-Bio met de reagentiaset van de firma Boehringer. Met de factor 1323 (berekend m.b.v. de micromolaire extinctiecoëfficiënt van NADH van 0,0063 $cm^2/\mu mol$) wordt de helling van de rechte lijn, die de afname van de NADH concentratie als functie van de tijd weergeeft, omgerekend naar een ALAT activiteit in U/l.
In de periode waarin de ALAT uitkomsten verzameld zijn, was de variatiecoëfficiënt van de dagelijks geanalyseerde autonorm contrôlesera in het lage gebied (50 U/l) 3,26% en in het hoge gebied (135 U/l) 1,89%. De afwijking t.o.v. laboratoria met dezelfde analysemethode, zoals berekend uit de Wellcome kwaliteitscontrôle resultaten, bedroeg in deze periode gemiddeld -0,94 sd (range -1,4 sd tot -0,4 sd ; aantal enquêtes 8).

Er werden 4071 ALAT uitkomsten verzameld, waarvan de frequentieverdeling is weergegeven in figuur 24A. Uit de Bhattacharya plot (figuur 24B en tabel 17) blijkt duidelijk, dat er geen lineair gedeelte gevonden kan worden, zodat geconcludeerd moet worden, dat de frequentie verdeling van ALAT uitkomsten niet Gaussisch is. Met de benadering uit hoofdstuk 5 worden voor de parameters van de aangepaste gamma verdeling de volgende schattingen gevonden : $\hat{r} = 7,55$ en $\hat{\lambda} = 0,462$.
Het 97,5de percentiel wordt gevonden bij 30,1 U/l. Deze bovengrens van het referentiewaarden gebied is lager dan die gevonden is voor ASAT (33,4 U/l). Door Bauer et al. (5a) wordt als bovengrens van het referentiewaardengebied voor ASAT en ALAT opgegeven resp. 30 en 25 U/l (bij 37°C). Ook door Tietz et al. (113)



Figuur 24.
A. Frequentieverdeling van 4071 ALAT uitkomsten en
B. de hieruit berekende Bhattacharya plot.

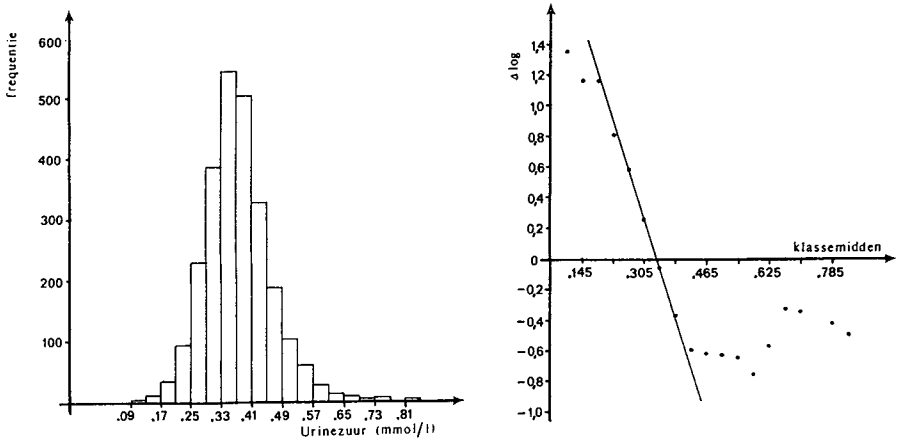
Tabel 17.
Details van de Bhattacharya plot voor ALAT. Eenheden van de eerste kolom : U/l.

klassemidden	frequentie	frequentie na smoothen	Δ log
2	3	–	–
5	54	–	–
8	233	252,8	0,660
11	508	488,8	0,201
14	616	597,4	-0,123
17	517	528,3	-0,305
20	387	389,4	-0,381
23	270	265,9	-0,285
26	182	200,0	-0,189
29	186	165,6	-0,123
32	134	146,4	-0,316
35	116	106,7	-0,248
38	77	83,2	-0,159
41	70	71,0	0,000
44	75	71,0	-0,172
47	60	59,8	-0,383
50	39	40,8	-0,473
53	28	25,4	–
56	22	–	–
59	46	–	–
>	448		
	4071		

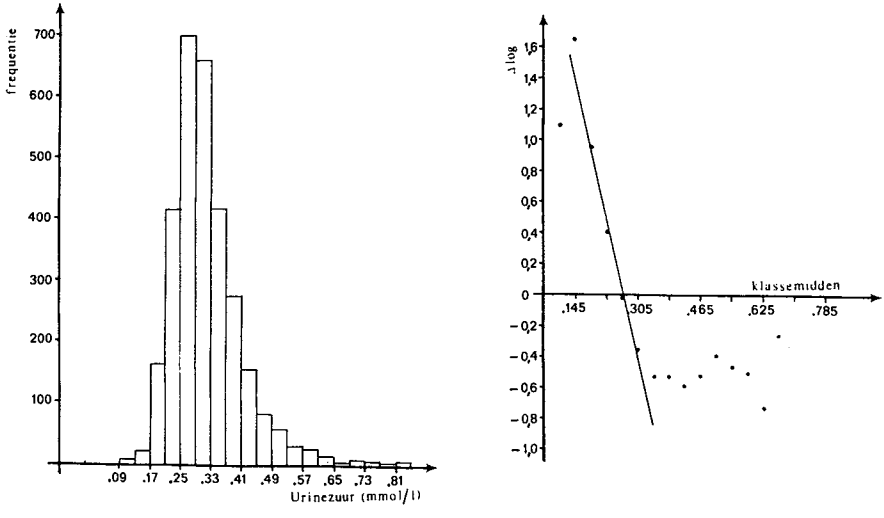
wordt een lagere bovengrens voor ALAT dan voor ASAT opgegeven (bij 30 °C). Siest et al. (106) vinden, vooral voor mannen vanaf 20 jaar, zeer hoge 97,5 percentielwaarden voor ASAT en ALAT. Vooral voor ALAT zijn de gevonden grenzen opvallend. De grenzen zijn echter berekend met de percentielenmethode uit de analyse resultaten van mensen die een keuring moesten ondergaan. Het is goed denkbaar dat deze hoge grenzen het gevolg zijn van het vrij hoge alcoholgebruik in Frankrijk, waardoor dus veel licht verhoogde waarden gevonden worden voor de transaminases bij overigens gezonde mensen.

6.3.2. Urinezuur.

Bij de mens, de mensapen en de Dalmatische hond is urinezuur het eindproduct van het purine metabolisme. Bij de andere zoogdieren kan urinezuur geoxideerd worden tot allantoïne, dat beter in water oplosbaar is dan urinezuur. Een verhoogd urinezuurgehalte in serum komt vooral voor bij jicht, waarvan de verschijnselen veroorzaakt worden door microkristallen van urinezuur in gewrichtsholten. Ook na het eten van voedsel dat rijk is aan nucleoproteïnen (zoals lever en nier), kan een verhoogd urinezuurgehalte in serum gevonden worden. Familiaire idiopathische hyperuricaemie is een zeldzame oorzaak van een verhoogd urinezuurgehalte en kan berusten op een overproductie bij een normale excretie of op een normale productie bij een verlaagde excretie.
Urinezuur wordt na filtratie door de glomeruli voor ongeveer 90% in de tubuli weer teruggeresorbeerd. Aangezien een verhoogd urinezuurgehalte door veel extrarenale factoren kan worden veroorzaakt, is de bepaling eigenlijk niet geschikt voor het controleren van de nierfunctie.
De bepaling van urinezuur is in ons laboratorium een onderdeel van het SMA-12 pakket. De bepalingsmethode berust op de oxidatie



Figuur 25.
A. Frequentieverdeling van 2535 urinezuur uitkomsten afkomstig van mannen en
B. de hieruit berekende Bhattacharya plot.



Figuur 26.
A. Frequentieverdeling van 3042 urinezuur uitkomsten afkomstig van vrouwen en
B. de hieruit berekende Bhattacharya plot.

Tabel 18.
Details van de Bhattacharya plots voor urinezuur. Eenheden van de eerste kolom : mmol/l.

klassemidden	MANNEN			VROUWEN		
	frequentie	frequentie na smoothen	Δ log	frequentie	frequentie na smoothen	Δ log
0,025	1	-	-	0	-	-
0,065	0	-	-	2	-	-
0,105	4	2	1,37	6	0	-
0,145	10	10	1,17	20	32	1,66
0,185	33	31	1,17	163	168	0,96
0,225	92	100	0,81	416	440	0,40
0,265	227	224	0,58	700	659	-0,02
0,305	384	399	0,26	658	643	-0,36
0,345	541	519	-0,05	416	448	-0,53
0,385	502	492	-0,37	272	264	-0,53
0,425	326	339	-0,59	155	156	-0,58
0,465	186	189	-0,61	83	87	-0,52
0,505	101	103	-0,62	58	52	-0,39
0,545	59	55	-0,64	31	35	-0,46
0,585	27	29	-0,75	25	22	-0,50
0,625	15	14	-0,56	14	13	-0,72
0,665	8	8	-0,32	4	7	-0,25
0,705	5	6	-0,34	7	5	-0,08
0,745	6	4	-	5	5	-
0,785	2	-	-	2	-	-
0,825	6	-	-	5	-	-

van urinezuur tot allantoïne in alkalisch milieu door fosforwolfraamzuur. Hierbij ontstaat een blauw gekleurd product, waarvan de concentratie evenredig is met de urinezuur concentratie en waarvan de extinctie gemeten wordt bij 660 nm.
De calibratie van de bepaling gebeurt met Lock-In serum van de firma General Diagnostics (concentratie urinezuur = 0,43 mmol/l; lotnr. OA861).
De variatiecoëfficiënt van de dagelijks geanalyseerde autonorm contrôlesera was in de periode, waarin de urinezuur uitkomsten werden verzameld, in het lage gebied (0,26 mmol/l) 1,72% en in het hoge gebied (0,59 mmol/l) 1,02%. De afwijking t.o.v. laboratoria met dezelfde analyse methode was in deze periode gemiddeld +0,69 sd (range 0,4 sd tot 1,5 sd ; aantal enquêtes 8).
In de periode 14-07-1981 t/m 30-11-1981 werden alle urinezuur uitkomsten van patiënten verzameld, opgesplitst naar geslacht en leeftijd. Bovendien werd een onderverdeling gemaakt tussen poliklinische en klinische bloedafname. In deze periode werden 2535 uitkomsten van mannen verzameld (figuur 25A), 1477 poliklinisch en 1058 klinisch, en 3042 uitkomsten van vrouwen (figuur 26A), 1811 poliklinisch en 1231 klinisch. Gemiddeld blijken de uitkomsten van mannen duidelijk hoger te zijn dan die van vrouwen. Er werd geen leeftijdsafhankelijkheid vastgesteld en ook geen verschil tussen de poliklinische en klinische uitkomsten.
Zoals uit het rechte stuk in de Bhattacharya plots van figuur 25B en 26B blijkt, kan de veronderstelling dat de frequentieverdeling van de urinezuuruitkomsten Gaussisch is, zowel voor mannen als voor vrouwen, veilig gedaan worden. De berekende vergelijkingen zijn :

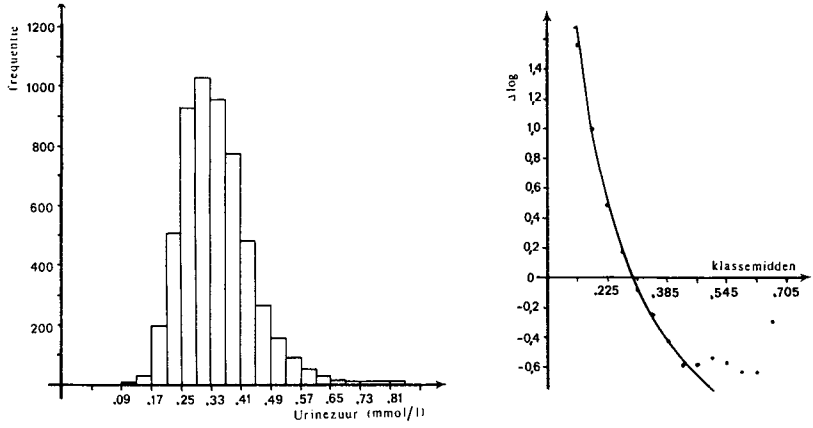
mannen : $\Delta \log = - 8{,}422$ (klassemidden) + 2,855.
vrouwen: $\Delta \log = -10{,}274$ (klassemidden) + 2,723.

Voor schattingen voor het gemiddelde en de standaarddeviatie volgt hieruit (klassebreedte = 0,04 mmol/l) :

mannen : $\hat{\mu} = 0,339 + \frac{1}{2}\,0,04 = 0,359$ $\hat{\sigma}^2 = -\frac{0,04}{-8,422} - \frac{0,04^2}{12} = 0,00462$

vrouwen : $\hat{\mu} = 0,265 + \frac{1}{2}\,0,04 = 0,285$ $\hat{\sigma}^2 = -\frac{0,04}{-10,274} - \frac{0,04^2}{12} = 0,00376$

zodat de berekende referentiewaarden ($\hat{\mu} \pm 1,96\,\hat{\sigma}$) zijn : voor mannen 0,23-0,49 mmol/l en voor vrouwen 0,17-0,41 mmol/l. Door Tietz et al (113) wordt als referentiewaarden opgegeven: mannen 0,21-0,43 mmol/l en vrouwen 0,16-0,36 mmol/l. In dezelfde periode, waarin de patiënten uitkomsten verzameld zijn, werd ook serum geanalyseerd van 351 mannelijke en 116 vrouwelijke bloeddonoren. De referentiewaarden die uit de resultaten hiervan berekend werden, zijn : mannen 0,25-0,48 mmol/l en vrouwen 0,17-0,40 mmol/l. Dit is fraai in overeenstemmming met de waarden die uit patiëntengegevens berekend zijn.



Figuur 27.
A. Frequentieverdeling van 5577 urinezuur uitkomsten (2535 afkomstig van mannen en 3042 afkomstig van vrouwen) en
B. de hieruit berekende Bhattacharya plot.

Wanneer geen splitsing gemaakt wordt naar geslacht, is de frequentieverdeling duidelijk asymmetrisch (figuur 27A), zoals ook blijkt uit de Bhattacharya plot waarin geen lineair gedeelte onderscheiden kan worden. Toepassing van de in hoofdstuk 5 beschreven rekenmethode levert voor de schattingen van de parameters van de aangepaste gamma verdeling op : $\hat{r}$=15,51 $\hat{\lambda}$=46,7. Met behulp van deze waarden kunnen de referentiewaarden berekend worden als : 0,18-0,51 mmol/l.

J. Clin. Chem. Clin. Biochem.
Vol. 18, 1980, pp. 621–625

The Use of Patient Data for the Calculation of Reference Values for Some Haematological Parameters

By *A. J. Naus, A. Borst* and *P. S. Kuppens*

Department of Hematology and Clinical Chemistry, St. Laurentius Hospital, Mgr. Driessenstraat 6, 6043 CV Roermond, The Netherlands

(Received January 28/June 18, 1980)

Summary: We have investigated the use of patient data for the calculation of reference values for the parameters which are determined by the Hemalog. For this purpose we used the *Bhattacharya* plot. All the parameters, with the exception of leukocytes, appear to meet the main underlying assumption of this plot, namely that the fequency distribution is *Gauss*ian. In the case of leukocytes, however, the frequency distribution could be resolved into two overlapping *Gauss*ian curves, thus making it possible to calculate reference values for this parameter also.

The reference values as calculated from 14,500 unselected data (excluding children) are in general agreement with the literature. Significant differences were however detected between a group of patients and a group of blood donors.

When a *Bhattacharya* plot has to be constructed with relatively few data, smoothing of the observed frequencies is very helpful in deciding which part of the plot is linear. Smoothing was carried out using the least squares method with a quadratic equation. Since the classes are equally spaced, this involves only a simple numerical transformation of the frequencies.

Die Verwendung von Patienten-Daten für die Ermittlung von Referenzwerten für einige hämatologische Kenngrößen

Zusammenfassung: Wir haben die Verwendung von Patienten-Daten für die Ermittlung von Referenzwerten für Kenngrößen, die mit dem Hemalog bestimmt wurden, untersucht und dafür die *Bhattacharya*-Darstellung benutzt. Alle Kenngrößen außer den Leukocyten scheinen die dieser Darstellung hauptsächlich zugrundeliegende Annahme, daß eine *Gauss*'-Verteilung vorliegt, zu erfüllen. Für die Leukocyten konnte die Häufigkeitsverteilung in zwei sich überlappende *Gauss*-Kurven aufgelöst werden, so daß auch für diese Kenngröße Referenzwerte ermittelt werden konnten.

Die aus 14.500 unausgewählten Daten (Kinder ausgenommen) ermittelten Referenzwerte stimmen mit Literaturangaben überein. Signifikante Unterschiede wurden jedoch zwischen einer Gruppe von Patienten und einer Gruppe von Blutspendern gefunden.

Wenn die *Bhattacharya*-Darstellung aus relativ wenig Daten konstruiert wird, ist die Glättung der beobachteten Häufigkeiten für die Entscheidung, welcher Teil der Darstellung linear ist, sehr hilfreich. Die Glättung wurde unter Verwendung der Methode der kleinsten Quadrate mit einer quadratischen Gleichung durchgeführt. Da die Klassen gleichen Raum einnehmen, beinhaltet dies nur eine einfache numerische Transformation der Häufigkeiten.

Introduction

The Hemalog (Technicon, Tarrytown, New York), determines simultaneously in a blood sample platelets (PLTS), leukocytes (WBC), erythrocytes (RBC), haemoglobin (Hb) and packed cell volume (PCV) and calculates the mean corpuscular volume MCV = PCV/RBC), mean corpuscular haemoglobin (MCH = Hb/RBC) and mean corpuscular haemoglobin concentration (MCHC = Hb/PCV = MCH/MCV) (1).

During the development of a quality control program for the Hemalog, the results of which will be reported in a subsequent paper, the need was felt for reference values of greater accuracy than those currently available. The problems, however, encountered in finding a group of persons that can be used for the determination of these values are numerous (2). It certainly is not acceptable in our view to use the laboratory staff or a group of blood donors for this purpose, because they do not form a true representation of the whole population. In case of

0340-076X/80/0018-0621$2.00
© by Walter de Gruyter & Co. · Berlin · New York

the laboratory staff for instance, the majority of people are female and below the age of 30. Most of our blood donors are male and between 30 and 40. So automatically a selection is made when choosing one of these groups for the calculation of reference ranges. The danger that these ranges are biased when a selection is made beforehand is very great. So in our view it is better to make no selection at all. Simply take all the results produced during a certain time in your laboratory for a certain test and use these. Of course a number of these results are "abnormal" and should not be used for the calculation of mean and standard deviation.

Applying the *Bhattacharya* plot automatically means that this is achieved (3). The only assumption that has to be made when using this plot is that the frequency distribution is *Gaussian*. If not, the plot cannot be applied.

However, the number of abnormal results in the population used for the calculation of reference ranges can be very great. In these cases the method of *Hoffmann* (4) gives an S-shaped curve as was discussed by *White* (5).

Because no selection is made a true representation of the population can be expected. This is even more so, because a very large number of test results (of which the majority are "normal") can be accumulated in a relatively short time.

Materials and Methods

The essence of the *Bhattacharya* plot is the following: The results for a certain assay are accumulated in classes, which must be equally spaced. The logarithm of the quotient of the frequencies in class $i + 1$ and class i is then plotted against the midpoint of class i, partly resulting in a straight line. This straight line represents the part of the distribution that is truly *Gaussian*, excluding all "abnormal" results. The slope of this line and the intercept with the X-axis result in the standard deviation and the mean of the distribution respectively. Mathematically it can be expressed as follows.

$$f_i = \frac{N}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2\right), \qquad \text{(eq 1)}$$

where:

x_i = the midpoint of class i.
f_i = the number of results in class i.
N = the total number of results in the *Gaussian* distribution.
μ = the mean of the distribution
σ = the sd of the distribution.

Equation 1 is the well known frequency density function for a *Gaussian* distribution from which the following can be easily derived.

$$\log \frac{f_{i+1}}{f_i} = -\frac{1}{2}\frac{h^2}{\sigma^2} - \frac{h(x_i - \mu)}{\sigma^2}, \qquad \text{(eq 2)}$$

where:

h = the width of the classes.

Equation 2 represents a straight line with $x = x_i$ and $y = \log \frac{f_{i+1}}{f_i}$.

The intercept with the X-axis is found for $x = \mu - 1/2\ h$. The slope of this line equals $-h/\sigma^2$, hence $\sigma^2 = -h/\text{slope}$. However, due to the grouping of data a correction must be applied. This results in.

$$\sigma^2 = -h/\text{slope} - h^2/12, \qquad \text{(eq 3)}$$

where:

$-h^2/12$ is *Sheppards* correction.

The *Bhattacharya* plot is only linear for the part of the frequency distribution that is truly *Gaussian*. At the high and low end the deviations from linearity due to "abnormal" results are in most cases very obvious. To calculate reference ranges using this plot, the linear part is selected manually and through these points a straight line is calculated using the method of least squares. When the number of data is small, the selection of the straight part in the *Bhattacharya* plot can sometimes be difficult due to statistical variations. In these cases a smoothing procedure is indicated. This facilitates the selection of the points of the plot that lie on the straight part and ensures that "abnormal" results (giving rise to the non-linear part of the plot) are not taken into account.

The smoothing procedure must be carried out on the observed frequencies and not on the points of the *Bhattacharya* plot itself, because the y-values of this plot are not independent. Suppose for instance that a point in the *Bhattacharya* plot has a y_i value that is too high, as a result of statistical variations. Since $y_i = \log \frac{f_{i+1}}{f_i}$ this means that f_{i+1} is too high, f_i too low, or both. It further means that correcting y_i automatically changes y_{i+1}, y_{i-1} or both, since the denominator of y_i is the nominator in y_{i-1} and the nominator of y_i is the denominator in y_{i+1}.

Smoothing of the frequency values was done according to *Savitzky* et al. (6, 7). Five consecutive frequencies (f_{i-2}, f_{i-1}, f_i, f_{i+1}, f_{i+2}) are fitted to a parabola using the method of least squares. The frequency read from the parabola at x_i is the smoothed frequency. Next a parabola is calculated through the frequencies at x_{i-1}, x_i, x_{i+1}, x_{i+2} and x_{i+3} and the value read from the parabola at the midpoint (x_{i+1}) again is the smoothed frequency and so on. Since the classes are equally spaced, the smoothed frequencies can simply be calculated using the following expression.

$$f_{i,\,\text{smoothed}} = \frac{-3f_{i-2} + 12f_{i-1} + 17f_i + 12f_{i+1} - 3f_{i+2}}{35} \qquad \text{(eq 4)}$$

An example of the difference between a smoothed and unsmoothed *Bhattacharya* plot is given in figure 1. The corresponding data are summarized in table 1.

Since the frequency distribution for leukocytes appeared to be very skewed, the possibility was investigated that this distribution could be described by the sum of two overlapping *Gaussian* curves.

The procedure that has to be used when overlap occurs is as follows. First, using the observed frequencies, a *Bhattacharya* plot is constructed. The linear part of this plot results in values for the mean and standard deviation of a *Gaussian* distribution. The calculated frequencies of this distribution are now subtracted from the observed ones. With the resulting differences a new *Bhattacharya* plot is constructed. If a second *Gaussian* distribution is hidden under the first one, a straight part in this plot can be seen. From this straight part the mean and standard deviation of a second *Gaussian* distribution can be calculated.

When this procedure is carried out for leukocytes, it appears that the frequency distribution can indeed be described by the sum of two overlapping *Gaussian* curves ($\mu_1 = 6.55$, $sd_1 = 1.56$, $\mu_2 = 9.86$, $sd_2 = 1.75$). The details of the calculation are given in table 2 and the resulting frequency distributions are depicted in figure 2.

Tab. 1. Data for figure 1.
Frequency distribution for 1593 mean corpuscular volume (MCV) values. The unsmoothed and smoothed *Bhattacharya* plot are shown in figure 1.
The calculation of the smoothed line gives the following results (figures indicated by an asterisk) r = - 0.9997, slope = - 0.1018 and intercept with X axis = 91.847 and from this μ = 92.85 and sd = 4.40. Units of X:fl/cell.

X	F	dLN	Fs-moothed	dLN
81	13	0.932	–	–
83	33	0.288	–	–
85	44	1.012	57	0.702*
87	121	0.435	115	0.481*
89	187	0.245	186	0.288*
91	239	0.190	248	0.100*
93	289	- 0.182	274	- 0.116*
95	241	- 0.337	244	- 0.327*
97	172	- 0.368	176	- 0.408
99	119	- 0.531	117	- 0.486
101	70	- 0.585	72	- 0.693
103	39	- 1.099	36	–
105	13	0.000	–	–
107	13	–	–	–

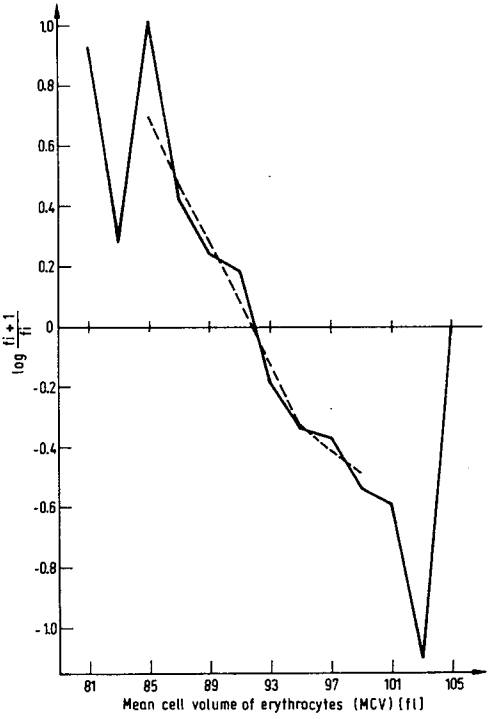


Fig. 1. *Bhattacharya* plot for mean corpuscular volume (MCV) with (dotted line) and without (full line) smoothing.

When two populations are compared, μ_1 is said to be statistically different from μ_2 when the 95% confidence limits of μ_1 and μ_2 do not overlap. For the standard deviations of two populations an analogous statement can be made.

The 95% confidence limits for the standard deviation can be calculated from the variance of the slope of the linear part of the *Bhattacharya* plot.

$$\sqrt{\left(\frac{-h}{\text{slope} - 2\text{sd}_{\text{slope}}} - \frac{h^2}{12}\right)} < \text{sd} < \sqrt{\left(\frac{-h}{\text{slope} + 2\text{sd}_{\text{slope}}} - \frac{h^2}{12}\right)} \qquad \text{(eq 5)}$$

The variance of the slope was calculated using the following equation (10).

$$\text{sd}^2_{\text{slope}} = \frac{\text{sd}_y^2}{\text{sd}_x^2} \frac{(1 - r^2)}{(n - 2)}, \qquad \text{(eq 6)}$$

Tab. 2. Resolution of the frequency distribution of leukocytes (WBC) into two *Gauss*ian distributions.
F (column 2) are the observed frequencies. F_{corr} are the observed frequencies smoothed according to eq. 4. $F_{calc\ 1}$ are the calculated frequencies of the first *Gauss*ian curve (μ_1 = 6.55, sd_1 = 1.56). Differences are the differences between the frequencies in columns 2 and 5 respectively. $F_{calc\ 2}$ are the calculated frequencies of the second *Gauss*ian curve (μ_2 = 9.86, sd_2 = 1.75). $F_{tot\ calc}$ are the sums of the frequencies in columns 5 and 8.
The points indicated by an asterisk (columns 4 and 7) have been taken for the calculation of the straight line.
For further explanation see the text.

Interval mean	F	$F_{corr.}$	dLN	$F_{calc\ 1}$	Difference	dLN	$F_{calc\ 2}$	$F_{tot\ calc}$
1.5	53	–	–	13.9			0.0	13.9
2.5	135	–	–	87.5			0.1	87.6
3.5	343	393.6	1.027*	375.8			1.3	377.1
4.5	1053	1099.1	0.603*	1071.7			9.1	1080.8
5.5	2051	2009.4	0.229*	2027.2			44.7	2071.9
6.5	2585	2525.3	- 0.051	2541.2			157.5	2698.7
7.5	2388	2398.5	- 0.230	2111.7	286.8	0.948	400.7	2512.4
8.5	1889	1904.1	- 0.292	1164.2	739.9	0.296*	735.6	1899.8
9.5	1395	1420.6	- 0.342	425.5	995.1	- 0.095*	974.0	1399.5
10.5	1037	1008.2	- 0.421	103.4	904.8	- 0.339*	930.5	1033.9
11.5	647	661.4	- 0.502	16.9	644.5	- 0.476	641.3	658.2
12.5	393	400.4	–	0.0	400.4	–	318.8	318.8
13.5	275	–	–	0.0	–	–	114.4	114.4
14.5	211	–	–	0.0	–	–	29.6	29.6

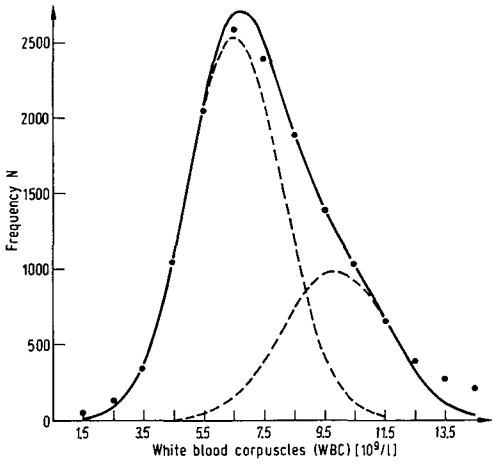


Fig. 2. Resolution of the frequency distribution of leukocytes (WBC) into two overlapping *Gaussian* curves (dotted lines). The full line represents the sum of the frequencies. • are the observed frequencies.

where:

y = interval mean.

$x = \log \frac{f_{i+1}}{f_i}$.

r = coefficient of correlation.

n = number of points on the linear part.

The 95% confidence limits for the mean have been calculated as follows.

$$\text{intcpt} + \frac{1}{2}h - 2sd_{\text{intcpt}} < \mu < \text{intcpt} + \frac{1}{2}h + 2sd_{\text{intcpt}} \quad \text{(eq 7)}$$

The variance of the intercept can be calculated with the following expression.

$$sd^2_{\text{intcpt}} = sd^2_{\text{slope}}\left(sd^2_x \frac{(n-1)}{n} + \mu^2_x\right) \quad \text{(eq 8)}$$

For the collection of data and all calculations, use was made of a Wang PCS II with punched tape reader.

Results and Discussion

By comparing the reference values calculated from a group of hospitalized patients (N = 3500) and a group of blood donors (N = 650), the following conclusions can be drawn.

1. The mean is the same for the two groups for all parameters. For female blood donors, however, the mean is slightly but significantly higher than for patients (tab. 3).

2. The standard deviation for all parameters is smaller for blood donors than for patients (tab. 4).

Because the *Bhattacharya* plot selects in fact the haematologically "normal" people, it is not surprising in

Tab. 3. Comparison between patients and blood donors – mean values. For explanation see the text.

		Patients		Blood donors	
		μ	Range μ	μ	Range μ
PLTS	(10^9/l)	167.5	166.2 – 168.8	169.3	168.7 – 169.8
RBC ♂	(10^{12}/l)	5.07	5.05 – 5.09	5.09	5.07 – 5.11
RBC ♀	(10^{12}/l)	4.56	4.51 – 4.61	4.64	4.62 – 4.66
Hb ♂	(Fe, mmol/l)	9.93	9.85 – 10.01	9.90	9.86 – 9.90
Hb ♀	(Fe, mmol/l)	8.81	8.79 – 8.83	8.91	8.86 – 8.96
PCV ♂	(l/l)	0.471	0.4686 – 0.4726	0.472	0.4704 – 0.4730
PCV ♀	(l/l)	0.432	0.4307 – 0.4333	0.436	0.4334 – 0.4380
MCV	(fl)	93.5	93.19 – 93.77	93.6	93.51 – 93.73
MCH	(Fe, fmol)	1.94	1.940 – 1.942	1.94	1.938 – 1.938
MCHC	(Fe, mmol/l)	21.0	20.92 – 21.00	21.0	20.91 – 20.99

Tab. 4. Comparison between patients and blood donors – standard deviations. For explanation see the text.

		Patients		Blood donors	
		SD	Range SD	SD	Range SD
PLTS	(10^9/l)	43.7	42.7 – 44.8	42.3	42.0 – 42.6
RBC ♂	(10^{12}/l)	0.48	0.46 – 0.51	0.39	0.38 – 0.40
RBC ♀	(10^{12}/l)	0.44	0.42 – 0.47	0.35	0.35 – 0.36
Hb ♂	(Fe, mmol/l)	0.66	0.62 – 0.72	0.48	0.45 – 0.52
Hb ♀	(Fe, mmol/l)	0.70	0.69 – 0.71	0.55	0.53 – 0.57
PCV ♂	(l/l)	0.0302	0.0281 – 0.0329	0.0229	0.0223 – 0.0235
PCV ♀	(l/l)	0.0311	0.0300 – 0.0323	0.0269	0.0257 – 0.0284
MCV	(fl)	4.77	4.56 – 5.01	4.11	4.02 – 4.19
MCH	(Fe, fmol)	0.099	0.098 – 0.101	0.091	0.090 – 0.091
MCHC	(Fe, mmol/l)	0.70	0.67 – 0.74	0.58	0.54 – 0.62

Tab. 5. Reference values.
The reference values have been calculated as $\mu \pm 2SD$,
the reference range for WBC has been calculated as $\mu_1 - 1.8\ SD_1$ $\mu_2 + 1.4\ SD_2$.

		Our method	Ref. (8)	Ref. (9)
PLTS	(10^9/l)	130 – 320	140 – 440	140 – 340
WBC	(10^9/l)	3.8 – 12.3	4.3 – 10.0	4.0 – 10.0
RBC ♂	(10^{12}/l)	4.1 – 6.0	4.5 – 6.3	4.5 – 6.5
RBC ♀	(10^{12}/l)	3.7 – 5.4	4.2 – 5.5	3.9 – 5.6
Hb ♂	(Fe, mmol/l)	8.5 – 11.2	8.7 – 11.2	8.4 – 11.2
Hb ♀	(Fe, mmol/l)	7.4 – 10.1	7.5 – 9.9	7.1 – 10.2
PCV ♂	(l/l)	0.41 – 0.53	0.41 – 0.51	0.40 – 0.54
PCV ♀	(l/l)	0.36 – 0.49	0.37 – 0.47	0.36 – 0.47
MCV	(fl)	84 – 102	82 – 101	76 – 96
MCH	(Fe, fmol)	1.74 – 2.14	1.68 – 2.11	1.68 – 1.99
MCHC	(Fe, mmol/l)	19.6 – 22.1	19.6 – 22.4	18.6 – 21.7

our view that the mean is the same for the two groups. Also the immobilisation period in our hospital is too short to have any influence on the haematopoiesis. It is rather surprising, however, that female blood donors have higher values for haemoglobin, erythrocytes and packed cell volume than female patients. We don't have an explanation for this. The fact that the standard deviation for all parameters is higher for patients than for blood donors can possibly be explained by the fact that the age distribution is much wider for the former group. This fact argues strongly against the use of blood donors for the determination of reference ranges.

The reference values as calculated from a mixed population (14,500 persons, excluding children below the age of 15) are in reasonable agreement with the literature (tab. 5) (8, 9). This population consists of all the patients from whom a sample has been submitted to the laboratory during a 3 month period and comprises about 45% hospitalized-, 45% out patients and 10% blood donors.

An important point indicating that the *Bhattacharya* plot gives correct reference values is the fact that these values are consistent i.e. $\mu_{PCV}/\mu_{RBC} = \mu_{MCV}$, $\mu_{Hb}/\mu_{RBC} = \mu_{MCH}$ and $\mu_{Hb}/\mu_{PCV} = \mu_{MCHC}$. Furthermore the reference range is equally broad for women and men.

For leukocytes it appeared that the frequency distribution could be very well described by the sum of two overlapping *Gaussian* curves. It is not quite clear whether this second curve represents a true subpopulation. The fact, however, that about 30% of all data is part of the second distribution, indicates that they cannot be classified as abnormals. Possibly this group represents persons who suffered from mild infections in the recent past.

Generally it can be concluded that the *Bhattacharya* plot applied to an unselected population is the method of choice for the calculation of reference values for the above mentioned parameters. Smoothing of the frequencies is helpful in deciding which part of the plot is linear when the number of data is small.

References

1. Rutten, W., Scholtis, R., Schmidt, N. & v. Oers, R. (1975), Z. Klin. Chem. Klin. Biochem. *13*, 387–393.
2. Hoeke, J. (1979), Het medisch jaar 1979, 420–429, Bohn, Scheltema & Holkema, Utrecht.
3. Bhattacharya, C. (1967), Biometrics *23*, 115–135.
4. Hoffmann, R. (1963), J. Am. Med. Assoc. *185*, 864–873.
5. White, J. (1978), Clin. Chim. Acta *84*, 353–360.
6. Savitzky, A. & Golay, M. (1964), Anal. Chem. *36*, 1627–1638.
7. Steinier, J., Termonia, Y. & Deltour, J. (1972), Anal. Chem. *44*, 1906–1909.
8. Wintrobe, M. (1974), Clinical Hematology, 7th ed. Lea & Febiger, Philadelphia.
9. Eastham, R. (1974), Clinical Hematology, 4th ed. John Wright & Sons Ltd., Bristol.
10. Diem, K. & Lentner, C. (1969), Wissenschaftliche Tabellen, 7. Auflage, J. R. Geigy AG, Basel.

Ir. A. J. Naus
Dept. of Hematology
and Clinical Chemistry
St. Laurentius Hospital
Mgr. Driessenstraat 6
6043 CV Roermond
The Netherlands

Noot.

De opmerking op pagina 121 :
"When two populations are compared, μ_1 is said to be statistically different from μ_2 when the 95% confidence limits of μ_1 and μ_2 do not overlap. For the standard deviations of two populations an analogous statement can be made."
is niet correct. Wanneer de 95% betrouwbaarheidsintervallen voor $\hat{\mu}_1$ en $\hat{\mu}_2$ (resp. $\hat{\sigma}_1$ en $\hat{\sigma}_2$) elkaar niet overlappen, is het verschil tussen μ_1 en μ_2 statistisch significant ($p<0,05$). Het omgekeerde hoeft echter niet correct te zijn : wanneer de 95% betrouwbaarheidsintervallen elkaar wél overlappen, betekent dit niet dat het verschil niet statistisch significant is. De toetsingsgrootheid voor het verschil van twee gemiddelden kan berekend worden als :

$$\left| \frac{\sqrt{sd^2_{\hat{\mu}_1} + sd^2_{\hat{\mu}_2}}}{\hat{\mu}_2 - \hat{\mu}_1} \right|$$

Wanneer dit groter is dan 2 is het verschil significant ($p<0,05$).
Het verschil in gemiddelde van patiënten en donoren voor PLTS (eerste regel van tabel 3 op pagina 122) blijkt nu ook statistisch significant te zijn. De overige conclusies gebaseerd op tabel 3 en 4 veranderen echter niet.

J. Clin. Chem. Clin. Biochem.
Vol. 20, 1982, 75–80

Determination of n-Dimensional Reference Ellipsoids Using Patient Data

By *A. J. Naus, A. Borst* and *P. S. Kuppens*

Department of Clinical Chemistry, St. Laurentius Hospital, Roermond, The Netherlands

(Received July 27/November 11, 1981)

Summary: A method is described for the calculation of n-dimensional reference ellipsoids, using patient data. The advantages and drawbacks of the use of reference ellipsoids for a set of different parameters, in contrast with the use of a reference range for every single parameter, are discussed. The use of reference ellipsoids in practice is illustrated with an example.

Bestimmung n-dimensionaler Referenzellipsoide mit Patienten-Daten

Zusammenfassung: Eine Methode für die Berechnung von n-dimensionalen Referenzellipsoiden aus Patienten-Daten wird beschrieben. Vorzüge und Nachteile der Verwendung von Referenzellipsoiden für einen Satz verschiedener Kenngrößen im Gegensatz zur Verwendung eines Referenzbereichs für jede einzelne Kenngröße werden erörtert. Die Verwendung von Referenzellipsoiden in der Praxis wird an einem Beispiel illustriert.

Introduction

In clinical chemistry it is customary to compare the result of an analysis with a reference range. In most cases this range is determined in such a way that 95% of a normal population lies within it. A result below the lower or above the upper limit is probably pathological, although the chance that it is normal is still 5% (1–4).

When in the same sample a second (independent) analyte is determined, the chance that the result of this second analysis is within its reference range, again is 95% (assuming that the person is healthy). The chance that both results are "normal" is $0.95^2 = 0.903$. In other words: the more independent analyses are performed in a sample of a healthy person, the greater the chance that one or more of the results are pathological. When 14 independent analyses are performed this chance is about 50% (5). Therefore, in order to be able to better differentiate between normal and pathological, it is advisable to use 2 or more dimensional reference ellipsoids rather than a reference range for every separate determination. A reference ellipsoid can be defined as the area in the n-dimensional space, where the chance that a set of n results of a healthy person lies, is 95%.

In theory it is possible to calculate a reference ellipsoid for every combination of n determinations. In practice, however, it is very difficult to use reference ellipsoids with a dimension greater than 2 without the use of a computer. Even when a computer can be used, it should be emphasized that the danger exists that a set of n results is indeed classified as pathological, but the reason for this classification is no longer apparent. Therefore in our view the use of reference ellipsoids should be limited to dimensions 2 and 3.

Materials and Methods

The general equation for a reference ellipsoid in the k-dimensional space for k variables that all have a *Gaussian* frequency distribution, is given by:

$$S = \{Y | (Y - X)^T V^{-1} (Y - X) \leqslant \chi^2(\alpha, k)\} \qquad \text{eq. 1}$$

where:

S = reference ellipsoid,
Y = vector of k results,
X = vector of the means of the k determinations,
V = variance-covariance matrix,
k = dimension,
$\chi^2(\alpha, k)$ = α-fractile of a χ^2 distribution with k degrees of freedom.

When k = 1, in other words, when a reference range is calculated for a single variable, equation 1 simplifies to:

$$S = \{Y | (y - \mu)\ (1/\sigma^2)\ (y - \mu) \leqslant 3.84\} \qquad \text{eq. 2}$$

where:

μ = mean of the variable,
σ = standard deviation of the variable,
$\chi^2(0.95, 1) = 3.84$.

Furthermore, when $\mu = 0$ and $\sigma = 1$ (standard normal distribution), equation 2 becomes:

$$-1.96 <= Y <= 1.96 \qquad \text{eq. 3}$$

0340-076X/82/0020-0075$02.00
© by Walter de Gruyter & Co. · Berlin · New York

Equation 3 is of course very often used, when the reference range for a single parameter has to be calculated.

When equation 1 is used for a combination of 2 variables, the result is a circle when the coefficient of correlation is 0, or an ellipse when $|r| > 0$. As $|r|$ increases, the ellipse becomes slimmer (fig. 1). When the means of both variables are 0 and the variances 1, the variance-covariance matrix is given by:

$$V = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix} \quad \text{or} \quad V^{-1} = \frac{1}{1-r^2} \cdot \begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix} \qquad \text{eq. 4}$$

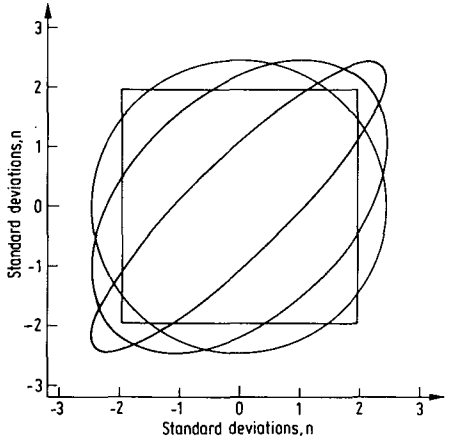


Fig. 1. Reference ellipsoids for a combination of two determinations ($\mu_1 = \mu_2 = 0$; $\sigma_1 = \sigma_2 = 1$) for r = 0, 0.45 and 0.90. The indicated square is the reference area used in the conventional way.

The tabulated value for $\chi^2(0.95, 2)$ is 5.99. The construction of figure 1 is now straightforward. In this figure, the square indicates the reference area used in the conventional way for a combination of these two determinations. It is clear that certain results lie within the square but outside the reference ellipse, in other words they are normal in the conventional way, but pathological with respect to the reference ellipsoid; the reverse can also occur.

In haematology it is customary to calculate from a set of values[1]) for Hb, RBC and PCV, the so called *Wintrobe* indices[1]) MCV, MCH and MCHC (7). These indices have various uses: for example, the combination of a high normal value for Hb and a low normal value for RBC is pathological, although the two results are each within their respective reference ranges. In this example, however, the calculated value for MCH would result in a pathological value, as should be the case. When, in the 3-dimensional space, the area is drawn within which Hb, RBC, PCV and MCV, MCH, MCHC are "normal", the result is a kind of prisma, which closely resembles an ellipsoid.

From the above equations it is clear that the most important thing to determine when a reference ellipsoid has to be calculated, is the variance-covariance matrix. This question can be resolved in two separate problems:

1. the determination of the variances of the separate variables and

2. the determination of the covariance of every combination of two variables.

The mean and variance of an assay can be determined by analysing a group of normal persons. The problems however in finding such a group are numerous. It certainly is not acceptable in our view to use the laboratory staff or a group of blood donors for this purpose, simply because they do not form a true representation of the whole population, although they may all be "normal". Automatically a selection is made when choosing one of these groups for the calculation of reference ranges (8, 9). The danger that these ranges are biased when a selection is made beforehand is very great, so in our view it is better to make no selection at all.

Simply take all the results produced during a certain time in your laboratory and use these. Of course a number of these results is "abnormal" and should not be used for the calculation of mean and standard deviation. In practice, however, most results for a routine test are completely normal. The *Bhattacharya* plot (10–13) is a statistical method that insures that the abnormal results in a frequency distribution do not influence the calculation of mean and standard deviation.

The *Bhattacharya* plot is based on the following:

the results of a determination are accumulated in equally spaced classes. If the frequency distribution is *Gaussian*, the logarithm of the quotient of the frequencies in class (i + 1) and class i, plotted against the midpoint of class i, results in a straight line. The mean and standard deviation of the distribution can be calculated from the x-intercept and the slope of the straight line respectively.

$$\mu = x_{\text{intercept}} - \frac{1}{2}h$$
$$\sigma^2 = -h/\text{slope} - h^2/12 \qquad \text{eq. 5}$$

where:

h = width of the classes.
$h^2/12$ = *Sheppard*'s correction for the grouping of data in classes.

When the number of test results, used to construct the *Bhattacharya* plot is small (less than 1500) the observed frequencies in every class should be smoothed. The method of choice for doing this, is the method of *Savitzky* et al. (14, 15).

So the *Bhattacharya* plot can result in values of μ and σ for a certain assay using unselected patient data. This means that the diagonal elements of the variance-covariance matrix can be determined quite easily. The remaining problem is to calculate values for the covariances or, since $\text{cov}(x, y) = r * \sigma(x) * \sigma(y)$, values for the coefficient of correlation when results for assay X are plotted against those of assay Y. When the frequency distribution of both determination X and determination Y is *Gaussian*, then the frequency distribution of a linear combination of X and Y is also *Gaussian*. This means that, when the *Bhattacharya* plot can be applied to determination $X(\mu_x, \sigma_x)$ and determination $Y(\mu_y, \sigma_y)$, this plot can also be applied to the sum of the results of the two determinations or the difference. When the sum of the results is used the mean is equal to $\mu_x + \mu_y$ and the variance is equal to $\sigma_x^2 + \sigma_y^2 + 2\,\text{cov}_{x,y}$: when the difference is used, the mean is $\mu_x - \mu_y$ and the variance is equal to $\sigma_x^2 + \sigma_y^2 - 2\text{cov}_{x,y}$.

So the covariance for a certain combination of two determinations (both having a *Gaussian* frequency distribution) can be calculated by the following equation:

$$\text{cov}_{x,y} = \frac{\sigma^2_{\text{sum}} - \sigma^2_{\text{diff}}}{4} \qquad \text{eq. 6}$$

Another method that can be used to determine the covariance is the following:

– Starting with a value for r equal 0, all data points are selected that lie within the 99% circle.

– Using the method of least squares a straight line is calculated through these points, resulting in a new value for r.

[1]) Hb = haemoglobin
RBC = red blood corpuscles, erythrocytes
PCV = packed corpuscular volume, haematocrit
MCV = mean corpuscular volume
MCH = mean corpuscular haemoglobin
MCHC = mean corpuscular haemoglobin concentration

– With this new r, again a selection of data points is made; the selection criterion is now that all points lie within the 99% ellipse.

– A straight line is again calculated through these points, resulting in a value for r that is slightly different from the old one.

– This process is repeated until the value of r no longer changes. With this r the covariance between determination X and Y is calculated.

This procedure is depicted in figure 2.

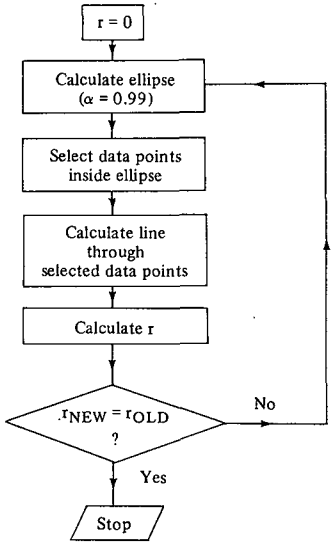


Fig. 2. Flow diagram for the calculation of the coefficient of correlation r between 2 determinations, using patient data.

Tab. 1. Calculation of the *Bhattacharya* plot for total protein.

i	Class mean	Observed frequency	Smoothed frequency	$\log(\frac{f_{i+1}}{f_i})$
1	41	2	–	–
2	43	4	–	–
3	45	2	2.4	- 0.1521
4	47	2	2.0	1.0028
5	49	5	5.6	0.4997
6	51	11	9.3	0.2543
7	53	12	12.0	0.5592
8	55	19	21.1	0.6985
9	57	42	42.5	0.2871
10	59	66	56.6	0.4026
11	61	70	84.7	0.3875
12	63	142	124.8	0.5536
13	65	198	217.2	0.4789
14	67	354	350.6	0.3410
15	69	502	493.1	0.1676
16	71	573	583.2	0.0109
17	73	601	589.6	- 0.1783
18	75	501	493.2	- 0.3844
19	77	321	335.8	- 0.5472
20	79	199	194.2	- 0.6387
21	81	100	102.5	- 0.8213
22	83	44	45.1	- 0.8632
23	85	21	19.0	- 0.7550
24	87	8	8.9	–
25	89	5	–	–
26	91	1	–	–

Tab. 2. Determination of covariance for total protein – albumin.

Determination	Mean	Variance
Total protein	71.9	23.72
Albumin	44.0	9.85
Total protein + albumin	117.5	45.40
Total protein - albumin	28.2	15.44

Results and Discussion

As an example we want to present the calculation of a reference ellipse for the combination total protein and albumin. The results for both determinations are produced by the SMA 12/60. This analyser is coupled to a teletype which punches all results in paper tape. This punched tape is then fed into a Wang PCS II desk top computer.

The frequency distributions for total protein, albumin, total protein + albumin and total protein - albumin are given in figure 3a–d (3840 results for every determination). After applying a 5 point quadratic smooth according to *Savitzky* et al. (14, 15) the *Bhattacharya* plots were constructed (fig. 4a–d). As an example, the details of the calculation of the *Bhattacharya* plot for albumin are given in table 1.

The resulting means and variances are summarized in table 2. From this table it follows that the coefficient of correlation (r) between total protein and albumin is 0.49.

Using the method depicted in figure 2, gives a value for r of 0.52, which is in close agreement.

With these results the 95% reference ellipse can easily be determined as:

$$S = \{Y | (x - 71.9 \;\; y - 44.0) \begin{pmatrix} 4.87^2 & 7.49 \\ 7.49 & 3.14^2 \end{pmatrix} \begin{pmatrix} x - 71.9 \\ y - 44.0 \end{pmatrix} \leqslant 5.99\} \qquad \text{eq. 7}$$

From figure 5 it can be seen that the number of data points that lie within the ellipse, but are abnormal in the conventional way, for total protein, albumin or both, is considerable (n = 153, i.e. 4.0%). More interesting is the rather great number of data points that are normal in the conventional way for total protein and albumin, but lie outside the reference ellipse (n = 43, i.e. 1.1%).

These data points are combinations of low normal total protein and high normal albumin or high normal total protein and low normal albumin, the combination of which is abnormal.

The method, described in this paper, for the calculation of n-dimensional reference ellipsoids, can only be applied when the frequency distribution of every determination is *Gaussian*. When, in the *Bhattacharya* plot, a sufficiently

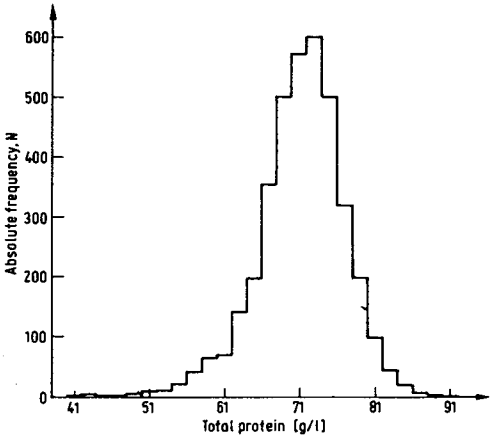


Fig. 3a

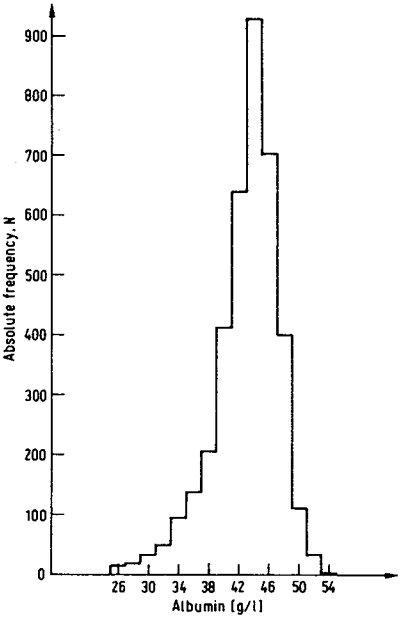


Fig. 3b

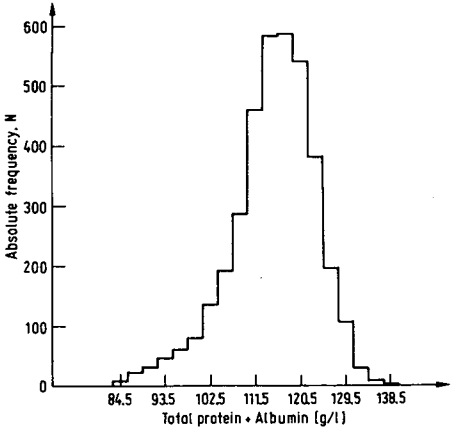


Fig. 3c

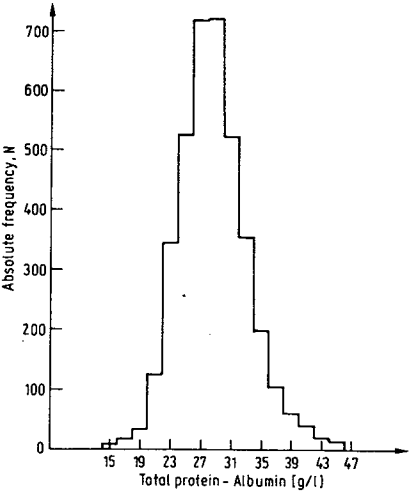


Fig. 3d

Fig. 3. Frequency distribution of 3840 values for total protein, albumin, total protein + albumin and total protein – albumin.

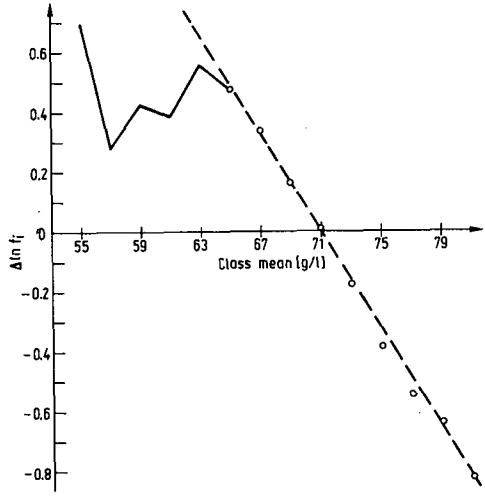


Fig. 4a

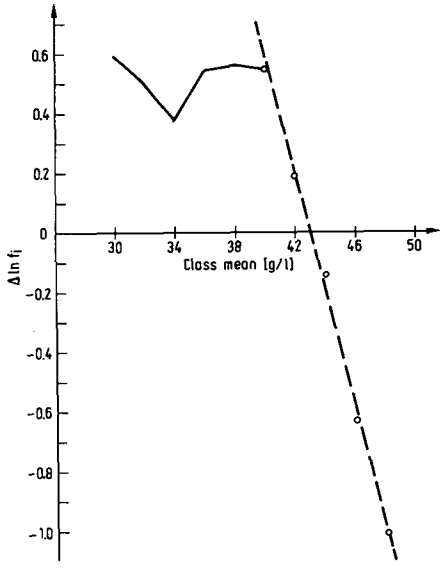


Fig. 4b

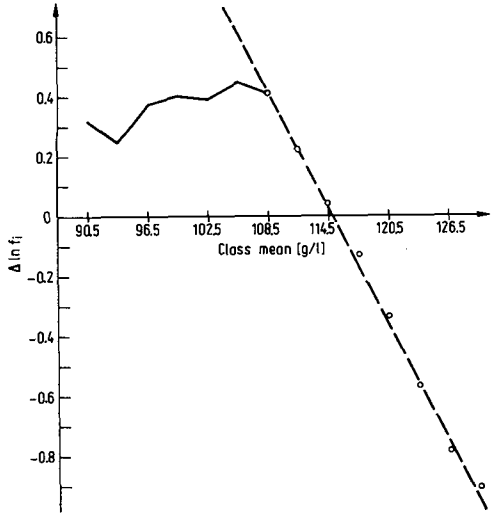


Fig. 4c

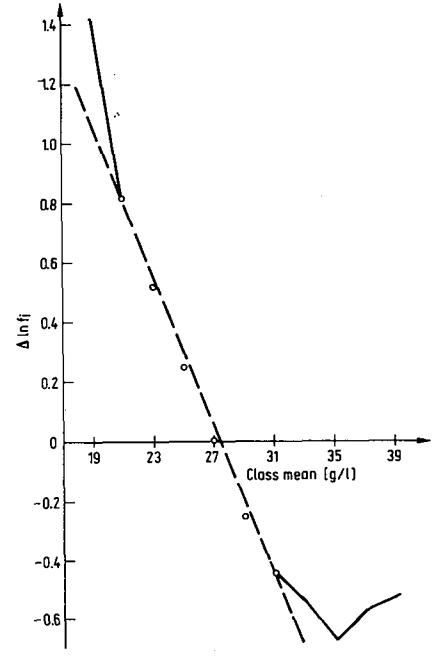
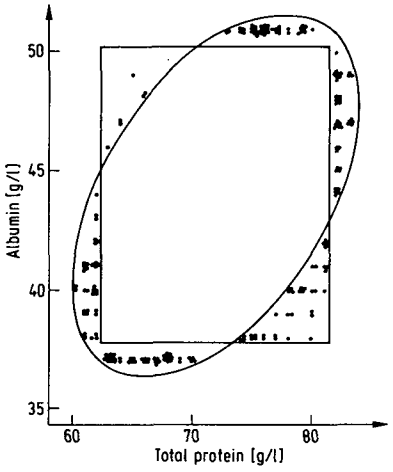


Fig. 4d

Fig. 4. *Bhattacharya* plots constructed with the data of fig. 3 after applying a 5 point quadratic smooth.



long straight part can be detected, the assumption that the frequency distribution is *Gaussian* can safely be made. The question, how to proceed, when the results of a determination are not distributed according to a *Gauss* function, is subject of further research, the results of which will be reported in a subsequent paper.

In our view reference ellipsoids can only be applied efficiently when a computer is used, especially when the dimension of the ellipsoid is greater than 2. Reference ellipsoids can hardly be printed in reference booklets for use by the clinician. This fact is an important drawback for the practical application of this undoubtedly very useful procedure.

Fig. 5. Calculated reference ellipse for the combination total protein albumin, using the data of fig. 3. By the conventional procedure the percentage of normal results is 2.9% less than in the ellipse method.

References

1. Rümke, C. & Bezemer, P. (1972) Ned. Tijdschr. Geneeskd. *116*, 1224–1230.
2. Rümke, C. & Bezemer, P. (1972) Ned. Tijdschr. Geneeskd. *116*, 1559–1568.
3. Dybkaer, R. & Gräsbeck, R. (1973) Scand. J. Clin. Lab. Invest. *32*, 1–7.
4. Dybkaer, R., Jörgensen, K. & Nyboe, J. (1975) Scand. J. Clin. Lab. Invest. *35*, Suppl. 144, 45–74.
5. Gross, R. & Oette, K. (1980) Die Problematik der ungezielten Mehrfachanalyse aus der Sicht des Klinikers. Technilab nov. 1980.
6. Guttman, I. (1970) Statistical tolerance regions. In: Griffin's statistical monographs & courses nr. 26. Griffin, London.
7. Wintrobe, M. (1974) Clinical hematology, 7th ed., Lea & Febiger, Philadelphia.
8. Hoeke, J. (1979) Normale waarden omstreden. In: Het medisch jaar 1979, 420–429, Bohn, Scheltema & Holkema, Utrecht.
9. Alström, T., Gräsbeck, R., Hjelm, M. & Skandsen, S. (1975) Scand. J. Clin. Lab. Invest. *35*, suppl. 144, 1–44.
10. Bhattacharya, C. (1967) Biometrics *23*, 115–135.
11. Naus, A., Borst, A. & Kuppens, P. (1980) J. Clin. Chem. Clin. Biochem. *18*, 621–625.
12. Gindler, E. (1970) Clin. Chem. *16* 124–128.
13. White, J. (1978) Clin. Chim. Acta *84*, 353–360.
14. Savitzky, A. & Golay, M. (1964) Anal. Chem. *36*, 1627–1638.
15. Steinier, J., Termonia, Y. & Deltour, J. (1972) Anal. Chem. *44*, 1906–1909.

Ir. A. J. Naus
Dept. of Clin. Chemistry
St. Laurentiushospital
Mgr. Driessenstraat 6
NL-6043 CV Roermond

Samenvatting.

Doel van het in dit proefschrift beschreven onderzoek is na te gaan in hoeverre het mogelijk is, referentiewaarden te berekenen voor klinisch chemische bepalingen, uitgaande van analyse resultaten verkregen uit een qua gezondheidskenmerken ongeselecteerde populatie.
De redenen waarom het aantrekkelijk is over de mogelijkheid te beschikken, zinnige referentiewaarden te berekenen uit de dagelijkse stroom analyseresultaten die een klinisch chemisch laboratorium verlaten, zijn :

1. De zeer grote moeilijkheden, die het selecteren van een qua gezondheidskenmerken homogene referentiepopulatie oplevert.
2. De niet onaanzienlijke kosten, die het analyseren van vrij grote aantallen monsters afkomstig van leden van de referentiepopulatie met zich meebrengt. Dit speelt een des te belangrijker rol, aangezien allerlei factoren, die van invloed kunnen zijn op de referentiewaarden (o.a. apparatuurkeuze), het noodzakelijk maken referentiewaarden frequent te controleren.
3. De mogelijkheid om kwaliteitscontrôle op de langere termijn te bedrijven.

Na een algemeen inleidend hoofdstuk, worden in hoofdstuk 2 een vijftal uit de literatuur bekende methoden voor de berekening van referentiewaarden, uit analyseresultaten van een inhomogene populatie, besproken. In hoofdstuk 3 worden deze methoden met elkaar vergeleken m.b.v. 260 theoretische frequentieverdelingen. Op grond van deze vergelijking blijkt, dat de door Bhattacharya beschreven methode de beste resultaten oplevert. Bovendien is deze methode relatief eenvoudig :
De resultaten van een bepaling worden in klassen van gelijke breedte geäccumuleerd. Wanneer de frequentieverdeling van de resultaten Gaussisch is, kan aangetoond worden, dat het verschil

van de logaritme van de frequenties in twee elkaar opvolgende klassen, uitgezet tegen het klassemidden, een rechte lijn oplevert.

In hoofdstuk 4 wordt de methode van Bhattacharya meer in detail besproken en wordt afgeleid hoe breed de klassen maximaal mogen zijn. Er wordt een gewogen kleinste kwadraten methode ontworpen die de automatische berekening van referentiewaarden met de methode Bhattacharya mogelijk maakt.

De uitkomsten van veel klinisch chemische bepalingen zijn verdeeld volgens een Gauss curve. Voor een vrij groot aantal bepalingen kan deze veronderstelling echter niet gedaan worden.

Het toepassen van verdelingsvrije methoden voor de berekening van referentiewaarden is niet mogelijk, wanneer de uitkomsten afkomstig zijn van een qua gezondheidskenmerken inhomogene populatie.

Er zal in deze gevallen een andere veronderstelling over de mathematische vorm van de frequentieverdeling van de resultaten van de hoofdpopulatie gedaan moeten worden.

In hoofdstuk 5 worden verschillende methoden besproken om een waargenomen frequentieverdeling mathematisch te beschrijven. Hieruit blijken (althans voor de klinische chemie) twee mogelijkheden naar voren te komen die toegepast kunnen worden, wanneer de waargenomen frequentieverdeling niet Gaussisch blijkt te zijn :

1. Het toepassen van gamma verdelingen en
2. Het transformeren van de bepalingsuitkomsten m.b.v. de vergelijking $y=\log(x+C)$, in de hoop dat de getransformeerde grootheid wél een Gaussische frequentieverdeling vertoont.

Het transformeren van waarnemingen heeft echter als nadeel, dat de oorspronkelijke gegevens gebruikt moeten worden. Deze zijn in de praktijk meestal niet voorhanden. Om deze reden en omdat log Gauss verdelingen vrijwel equivalent zijn met gamma verdelingen, wordt gekozen voor de eerste mogelijkheid.

Analoog aan de benadering van Bhattacharya voor een Gauss

verdeling, wordt een methode uitgewerkt, om uit een waargenomen frequentieverdeling van uitkomsten verkregen uit een inhomogene populatie, referentiewaarden te berekenen voor de hoofdpopulatie, wanneer aangenomen wordt dat de frequentieverdeling van uitkomsten van de hoofdpopulatie beschreven kan worden met een gamma verdeling.

In hoofdstuk 6 tenslotte worden enkele voorbeelden besproken zowel van bepalingen waarvan de resultaten verdeeld zijn volgens een Gauss curve als van bepalingen waarvoor deze veronderstelling niet geldt. Bovendien wordt een voorbeeld uitgewerkt van de berekening van een tweedimensionaal referentiewaardengebied.

Summary.

The aim of the research described in this thesis is to investigate whether it is possible to calculate reference values for determinations, routinely performed in a clinical chemistry laboratory, using the results of an unselected (with respect to health characteristics) patient population.
It is attractive to have the possibility of using the bulk of analytical results produced every day for the calculation of meaningful reference values for the following reasons :

1. The great difficulty of selecting a reference population which is homogeneous with respect to certain health characteristics.
2. Analysing a great number of samples taken from members of a reference population is costly. This point is even more important, because reference values have to be checked frequently in order to establish the effect of certain parameters which may influence these values.
3. The possibility of checking continually the analytical performance of the laboratory.

After a general introduction (chapter 1), five methods for the calculation of reference values using unselected patient data are described in chapter 2. In chapter 3 these methods are compared, using 260 theoretical frequency distributions. From this comparison the method described by Bhattacharya appears to give the best results and it is relatively simple.
The results of a determination are accumulated in equally spaced classes. Assuming the results of the determination have a Gaussian frequency distribution, it can be shown that the difference between the logarithms of the frequencies in two succesive classes, plotted against the midpoint of the first class, produces a straight line.
In chapter 4, the method of Bhattacharya is described in more

detail. It is shown that the width of the classes should not exceed 1 sd. A weighted least squares procedure is described making it possible to calculate reference values automatically, using the Bhattacharya plot.

The results of many determinations performed in a clinical chemistry laboratory have a Gaussian frequency distribution. For many other determinations, however, this assumption cannot be made. The use of non-parametric methods for the calculation of reference values in those cases is only possible, when the population from which the results are derived is homogeneous with respect to certain health characteristics. When this is not the case, another assumption about the mathematical form of the frequency distribution of results from the major ("healthy") part of the population under study, must be made.

In chapter 5, several methods are reviewed for the mathematical description of an observed frequency distribution. Two possibilities appear to be applicable in the field of clinical chemistry when the observed frequency distribution is not Gaussian :

1. The application of Gamma distributions.
2. The transformation of the observed results with the equation $y = \log (x+C)$, hoping that the transformed variable (y) has a Gaussian frequency distribution.

Transformation has the disadvantage that the original (not rounded off) observations must be used and these are usually not available. For this reason, and because log-Gauss distributions resemble Gamma distributions very much, the first possibility has been chosen.

Assuming that the results from the healthy population have a Gamma distribution, a method analogous to the Bhattacharya method for a Gaussian distribution has been devised to calculate reference values for the major ("healthy") part of the population under study, using unselected (with respect to health characteristics) patient data.

In chapter 6, the theory is applied to the results of a number of determinations which have a Gaussian frequency distribution as well as to some others for which this assumption does not hold.
Furthermore the calculation of a twodimensional reference ellips using patient data is illustrated with an example.

Zusammenfassung.

Das Ziel der in dieser Dissertation beschriebenen Forschung ist, zu untersuchen in wiefern es möglich ist, Referenzwerte für klinisch chemische Bestimmungen, aus Ergebnissen einer unselektierten Patientenpopulation zu berechnen.
Es ist einladend über die Möglichkeit zu verfügen Referenzwerte aus dem täglichen Strom Analysenergebnissen, die ein klinisch chemisches Labor verlassen, zu berechnen und zwar aus folgenden Gründen :

1. Die sehr grossen Schwierigkeiten, die es mit sich bringt eine homogene (bezüglich Gesundheit) Referenzpopulation zusammenzusetzen.
2. Die nicht unbedeutenden Kosten, die die Analyse vieler Proben von Gliedern der Referenzpopulation mit sich bringt. Dieser Punkt ist um so wichtiger, weil Referenzwerte frequent kontrolliert werden müssen, um der eventuelle Einfluss von Faktoren, die diese Werte möglicherweise beeinflussen, feststellen zu können.
3. Die Möglichkeit langfristige Qualitätskontrolle auszuüben.

Nach einem allgemeinen Kapitel, werden im 2. Kapitel fünf Methoden beschrieben für die Berechnung von Referenzwerten aus Ergebnissen einer unselektierten Population. Im 3. Kapitel werden diese Methoden mit Hilfe von 260 theoretischen Häufigkeitsverteilungen mit einander verglichen. Aus dieser Vergleichung geht hervor, dasz die durch Bhattacharya beschriebene Methode die besten Resultate ergibt. Ausserdem ist diese Methode relativ einfach :
Die Analysenergebnisse einer Bestimmung werden in Klassen gleicher Breite akkumuliert. Wenn die Frequenzverteilung der Ergebnisse Gaussisch ist, ist nachweislich, dasz die Differenz der Logaritmen von zwei aufeinander folgenden Klassenfrequenzen, ausgesetzt gegen die Mitte der ersten Klasse, eine Gerade ergibt.

Im 4. Kapitel wird die Bhattacharya Methode im grösseren Detail besprochen. Es wird abgeleitet, dasz die Breite der Klassen nicht grösser als 1 sd sein darf. Ausserdem wird eine gewogene kleinste Kwadratenmethode entworfen, die die automatische Berechnung von Referenzwerten mit Hilfe der Bhattacharya plot ermöglicht.

Die Analysenergebnisse vieler klinisch chemischen Bestimmungen haben eine Gaussische Frequenzverteilung. Für ziemlich viele andere Bestimmungen gilt diese Annahme aber nicht. Die Anwendung verteilungsfreier Methoden für die Berechnung von Referenzwerten ist nur gestattet, wenn alle Analysenergebnisse erzielt wurden bei Glieder einer homogenen Population. Wenn dies nicht der Fall ist, können verteilungsfreie Rechenmethoden nicht angewandt werden und muss eine andere Annahme über die mathematische Form der Frequenzverteilung der Analysenresultate des (gesunden) Hauptteils der Population, gemacht werden.

Im 5. Kapitel werden verschiedene Methoden besprochen, die es ermöglichen müssen, eine beliebige Frequenzverteilung mathematisch zu beschreiben. Hieraus ergibt sich, dasz für klinisch chemische Zwecke zwei Möglichkeiten in Frage kommen um eine nicht Gaussische Frequenzverteilung mathematisch zu beschreiben, und zwar :

1. Die Anwendung von gamma Verteilungen.
2. Die Transformation der Analysenergebnisse mit Hilfe der Gleichung : $y = \log (x+C)$, in der Hoffnung, dasz der transformierte Variabele eine Gaussische Häufigkeitsverteilung zeigt.

Die Transformation von Analysenergebnissen hat als Nachteil, dasz die ursprünglichen Daten in die Transformationsformel eingesetzt werden müssen. Diese Daten sind aber selten verfügbar. Darum, und auch weil gamma Verteilungen ungefähr äquivalent sind mit log-Gauss Verteilungen, wird die erste Möglichkeit ausgewählt. Es wird in diesem Kapitel eine Methode beschrieben, analog der Bhattacharya Methode für eine Gaussische Frequenzverteilung, um aus Analysenergebnissen einer gemischten Population

("Gesunde" und "Kranke") Referenzwerte für Gesunde zu berechnen, wenn angenommen wird, dasz die Frequenzverteilung der Analysenresultate dieser Teil der Population mit einer gamma Verteilung beschrieben werden kann.

Im 6. Kapitel schlieszlich, wird die Berechnung von Referenzwerten aus unselektierten Patientendaten durch einige Beispiele erläutert. Ausserdem wird ein Beispiel besprochen für die Berechnung eines zweidimensionalen Ellips aus Patientendaten.

Literatuur.

1. Alström, T., Gräsbeck, R., Hjelm, M., & Skandsen, S.
Scand.J.Clin.Lab.Invest. 35, suppl. 144 (1975).
Recommendations concerning the collection of reference values in clinical chemistry and activity report.

2. Amador, E.
Prog.Clin.Pathol. 5 (1974), 59-83.
Normal ranges.

3. Amador, E., & Hsi, B.
Amer.J.Clin.Pathol. 52 (1969), 538-546.
Indirect methods for estimating the normal range.

4. Annino, J., & Williams, L.
Clin.Chem. 18 (1972), 488-489.
On normal values (letter).

5. Ash, K.
Am.J.Med.Techn. 46 (1980), 504-511.
Reference intervals (normal ranges) : a challenge to laboratorians.

5a. Bauer, J., Ackermann, P. & Toro, G.
Clinical laboratory methods, 8th ed., The C.V. Mosby Company, Saint Louis, 1974.

6. Becktel, J.
Clin.Chim.Acta 28 (1970), 119-125.
Simplified estimation of normal ranges for routine laboratory data.

7. Bernhardt, W., & Weisner, B.
J.Clin.Chem.Clin.Biochem. 14 (1976), 443-447.
Bivariate Wertung von Befunden : Immunoglobulin G und Albumin des Liquor cerebrospinalis.

8. Best, W., Mason, C., Barron, S., & Shepherd, H.
Clin.Chim.Acta 28 (1970), 127-132.
Validation of procedure for setting normal limits on basis of total laboratory experience.

9. Bezemer, P.
Proefschrift Amsterdam (1981).
Referentiewaarden. Een verkenning van methoden voor het bepalen van 'normale waarden'.

10. Bhattacharya, C.
Biometrics 23 (1967), 115-135.
A simple method of resolution of a distribution into Gaussian components.

11. Boyle, R.
Memoirs for the natural history of humane blood, especially the spirit of that liquor. (London 1684).

12. Burnett, R.
Clin.Chem. 21 (1975), 1935-1938.
Accurate estimation of standard deviations for quantitative methods used in clinical chemistry.

13. Büttner, J.
J.Clin.Chem.Clin.Biochem. 15 (1977), 1-12.
Die Beurteilung des diagnostischen Wertes klinisch chemischer Untersuchungen.

14. Campos, J.
Br.J.Radiol. 48 (1975), 913-917.
Application of the Weibull distribution to some instances of cell survival.

15. Cegla, U.
J.Clin.Chem.Clin.Biochem. 12 (1974), 207-209.
Zur statistischen Verteilung von IgA, IgG und IgM bei 20-30 jährigen weiblichen und männlichen Normalpersonen.

16. Cejka, V., Dipert, M., Tyler, S., & Klein, P.
US AEC Argonne Nat.Lab. (1968), 283-285.
The detection and analysis of compound Gaussian distributions.

17. Cembrowski, G., Westgard, J., Conover, W., & Toren, E.
Am.J.Clin.Pathol. 72 (1979), 21-26.
Statistical analysis of method comparison data. Testing normality.

18. Chew, V.
J.Amer.Stat.Assn. 61 (1966), 605-617.
Confidence-, prediction- and tolerance-regions for the multivariate normal distribution.

19. Collen, M., Siegelaub, A., Cutler, J., & Goldberg, R.
Ann. NY Acad.Sci. 161 (1969), 572-580.
Aspects of normal values in medicine.

20. Cook, M., Levell, M., & Payne, R.
J.Clin.Pathol. 23 (1970), 778-780.
A method for deriving normal ranges from laboratory specimens applied to uric acid in males.

21. Cowell, D., & Taylor, W.
Ann.Clin.Biochem. 18 (1981), 76-83.
Ionic fluoride : A study of its physiological variation in man.

22. Diem, K., & Lentner, C.
Wissenschaftliche Tabellen, 7. Auflage (1969), J.R. Geigy AG Basel.

22a. Doornbos, R.
Persoonlijke mededeling.

23. Draper, N., & Smith, H.
Applied regression analysis (1966), Wiley ed., New York.

24. Dybkaer, R., Jörgensen, K., & Nyboe, J.
Scand.J.Clin.Lab.Invest. 35, suppl 144 (1975), 45-74.
Statistical terminology in clinical chemistry reference values.

25. Dybkaer, R., & Gräsbeck, R.
Scand.J.Clin.Lab.Invest. 32 (1973), 1-7.
Editorial : theory of reference values.

26. Dybkaer, R.
Scand.J.Clin.Lab.Invest. 29, suppl 126 (1972), 19.1.
VIIIth international congress on clinical chemistry, Copenhagen 1972.

27. Elion-Gerritzen, W.
Proefschrift Rotterdam (1978).
Requirements for analytical performance in clinical chemistry.

28. Elion-Gerritzen, W.
Am.J.Clin.Pathol. 73 (1980), 183-195.
Analytic precision in clinical chemistry and medical decisions.

29. Elveback, L., Guillier, C., & Keating, F.
JAMA 211 (1970), 69-75.
Health, normality and the ghost of Gauss.

30. Elveback, L., & Taylor, W.
Ann. NY Acad.Sci. 161 (1969), 538-548.
Statistical methods of estimating percentiles.

31. Feinstein, A.
Clin.Pharmacol.Therap. 15 (1974), 528-540.
Clinical biostatistics XXVII. The derangements of the range of normal.

32. Fischer, I.
Clin.Chem. 25 (1979), 1150.
A critical study of normal ranges in individuals.

33. Frommer, D.
Med.J.Aust. 1 (1974), 195-196.
Normal ranges in non-Gaussian distributions.

34. Galen, R.
Arch.Pathol.Lab.Med. 101 (1977), 561-565.
The normal range : a concept in transition.

35. Gindler, E.
Clin.Chem. 16 (1970), 124-128.
Calculation of normal ranges by methods used for resolution of overlapping Gaussian distributions.

36. Glick, J.
Clin.Chem. 18 (1972), 1504-1513.
Statistics of patient test values : application to indirect normal range and to quality control.

37. Gottlieb, C.
J.Immunol. 113 (1974), 51-57.
Application of transformations to normalize the distribution of plaque forming cells.

38. Gowenlock, A., & Broughton, P.
Z.Anal.Chem. 243 (1968), 774-780.
The influence of accuracy and precision on the normal range and on acceptable limits for an analytical result.

39. Grams, R., Johnson, E., & Benson, E.
Amer.J.Clin.Pathol. 58 (1972), 177-219.
Laboratory data analysis system section I-VI.

40. Grams, R.
Med.Instrum. 8 (1974), 3-8.
A statistical systems approach to health and "normal" values

41. Grannis, G., & Lott, J.
Clin.Chem. 24 (1978), 640-651.
A technique for determining the probability of abnormality.

42. Gräsbeck, R. et al.
Clin.Chem. 25 (1979), 1506-1508.
Provisional recommendation on the theory of reference values (1978).

43. Gräsbeck, R., & Fellman, J.
Scand.J.Clin.Lab.Invest. 21 (1968), 193-195.
Normal values and statistics.

44. Harris, E. and DeMets, D.
Clin.Chem. 18 (1972), 244-249.
Effects of intra- and inter-individual variation on distributions of single measurements.

45. Harris, E., & DeMets, D.
Clin.Chem. 18 (1972), 605-612.
Estimation of normal ranges and cumulative proportions by transforming observed distributions to Gaussian form.

46. Harris, E.
Clin.Chem. 21 (1975), 1457-1464.
Some theory of reference values. I.Stratified (categorized) normal ranges and a method for following an individuals clinical laboratory values.

47. Healy, M.
Bull.Acad.R.Med.Belg. 9 (1969), 703-718.
Normal values from a statistical viewpoint.

48. Helm, H.vd, & Hische, E.
Clin.Chem. 25 (1979), 985-988.
Application of Bayes's theorem to results of quantitative clinical determinations.

49. Helm, H.vd, & Hische, E.
NTVG 123 (1979), 1944-1951.
Gevoeligheid, specificiteit en diagnostische waarde van laboratoriumonderzoekingen.

50. Helm, H.vd, & Hische, E.
NTVG 123 (1979), 1983-1987.
Beoordeling van de diagnostische waarde van klinisch chemische onderzoekingen met behulp van het theorema van Bayes.

51. Helm, H.vd, & Hische, E.
Med.NVKC 4 (1979), 129-132.
Discriminatiewaarden.

52. Hermans, J.
Med.NVKC 4 (1979), 123-128.
Discriminatiewaarden.

53. Herrera, L.
J.Lab.Clin.Med. 52 (1958), 34-42.
The precision of percentiles in establishing normal limits in medicine.

54. Hill, P., & Chan, M.
Am.J.Clin.Pathol. 74 (1980), 715-716.
Normal ranges : expected and guaranteed coverage.

55. Hoeke, J.
Proefschrift Rotterdam (1976).
Aspecten van de "gemiddelden van normalen" methode in de klinische chemie.

56. Hoeke, J.
In : Het medisch jaar 1979, 420-429.
"Normale waarden" omstreden.

57. Hoffmann, R.
Clin.Chem. 17 (1971), 456-457.
Establishing normal ranges (letter).

58. Hoffmann, R.
JAMA 185 (1963), 864-873.
Statistics in the practice of medicine.

59. Homburger, H., & Hewan-Lowe, K.
Clin.Chem. 25 (1979), 669-674.
Predictive values of thyroxine, thyrotropin and triiodothyronin concentrations in serum.

60. IFCC Newsletter 16 (1977), 4-11.

61. Johnson, N.
Biometrika 36 (1949), 149-176.
Systems of frequency curves generated by methods of translation.

62. Kågedal, B., Sandström, A., & Tibbling, G.
Clin.Chem. 24 (1978), 1744-1750.
Determination of a trivariate reference region for free thyroxine index, free triiodothyronine index and thyrotropin from results obtained in a health survey of middle aged women.

63. Keating, F., Jones, J., Elveback, L., & Randall, R.
J.Lab.Clin.Med. 73 (1969), 825-834.
The relation of age and sex to distribution of values in healthy adults of serum Ca, inorganic P, Mg, alkaline phosfatase, total proteins, albumin and blood urea.

64. Kendall, M., & Stuart, A.
The advanced theory of statistics. Volume I : distribution theory. 4th ed. Charles Griffin & Co London 1977.

65. Klumpp, T. and Bieth, J.
Clin.Chem. 25 (1979), 969-972.
Automated measurement of the elastase-inhibitory capacity of plasma with a centrifugal analyser.

66. Koller, S.
Verh.Dtsch.Ges.Inn.Med. 81 (1975), 515-528.
Statistische Probleme bei der Ermittlung von Normbereichen und Befundmustern.

67. Koller, S.
Bibl.Haemat. 21 (1965), 125-128.
Problems in defining normal values.

68. Krause, R., Anand, V., Gruemer, H., & Willke, T.
Clin.Chem. 21 (1975), 321-324.
The impact of laboratory error on the normal range : a Bayesian model.

69. Landaas, S., & Skrede, S.
J.Clin.Chem.Clin.Biochem. 19 (1981), 1075-1080.
The levels of serum enzymes, plasma proteins and lipids in normal infants and small children.

70. Landahl, S., Jagenburg, R., & Svanborg, A.
Clin.Chim.Acta 112 (1981), 301-304.
Blood components in a 70-year old population.

71. Leonard, J., & Westlake, A.
Clin.Chim.Acta 82 (1978), 271-279.
The derivation of reference ranges adjusted for age, sex and other variables.

72. Leynse, H., & Leynse, B.
Pharm.Weekblad 107 (1972), 337-347.
De "normale waarden" in de klinische chemie.

73. Leynse, B.
Med.NVKC 4 (1979), 117-122.
Referentie- of normale waarden.

74. Lilliefors, H.
J.Amer.Statist.Ass. 62 (1967), 399-402.
On the Kolmogorov-Smirnov test for normality with mean and variance unknown.

75. Little, A., Williams, R., Parker, S., & Payne, R.
Clin.Chim.Acta 57 (1974), 91-95.
The derivation of biochemical normal ranges from a hospital outpatient population.

76. Lorentz, K., Flatter, B., & Augustin, E.
Clin.Chem. 25 (1979), 1714-1720.
Arylesterase in serum : elaboration and clinical application of a fixed incubation method.

77. Lumsden, J., & Mullen, K.
Can.J.Comp.Med. 42 (1978), 293-301.
On establishing reference values.

78. Mainland, D.
Clin.Chem. 17 (1971), 267-274.
Remarks on clinical "norms".

79. Mainland, D.
Ann. NY Acad.Sci. 161 (1969), 527-537.
Normal values in medicine.

80. McCammon, R.
Ann. NY Acad.Sci. 134 (1966), 559-562.
The concept of normality.

81. Morton, R.
New Eng.J.Med. 290 (1974), 695.
Normal values : definition and cost (letter).

82. Moyer, T., Jiang, N., Tyce, G., & Sheps, S.
Clin.Chem. 25 (1979), 256-263.
Analysis for urinary catecholamines by liquid chromatography with amperometric detection : methodology and clinical interpretation of results.

83. Munan, L., Kelly, A., & Petitclerc, C.
Clin.Biochem. 11 (1978), 75-76.
Statistical and clinical norms : variations on a familiar theme (editorial).

84. Naus, A., Borst, A., & Kuppens, P.
J.Clin.Chem.Clin.Biochem. 18 (1980), 621-625.
The use of patient data for the calculation of reference values for some haematological parameters.

85. Naus, A., Borst, A., & Kuppens, P.
J.Clin.Chem.Clin.Biochem. 20 (1982), 75-80.
Determination of n-dimensional reference ellipsoids using patient data.

86. Neumann, G.
Clin.Chem. 14 (1968), 979-988.
The determination of normal ranges from routine laboratory data.

87. O'Halloran, M., Studley-Ruxton, J., & Wellby, M.
Clin.Chim.Acta 27 (1970), 35-46.
A comparison of conventionally derived normal ranges with those obtained from patients results.

88. Orrell, D.
Clin.Chim.Acta 35 (1971), 483-489.
Albumin as an aid to the interpretation of serum calcium.

89. Owen, J., & Campbell, D.
Clin.Chim.Acta 22 (1968), 611-618.
A comparison of plasma electrolyte and urea values in healthy persons and in hospital patients.

90. Payne, R., & Levell, M.
Clin.Chem. 14 (1968), 172-178.
Redefinition of the normal range for serum sodium.

91. Pryce, J.
Lancet 2 (1960), 333-336.
Level of hemoglobin in whole blood and red blood cells, and proposed convention for defining normality.

92. Reed, A., Henry, R., & Mason, W.
Clin.Chem. 17 (1971), 275-284.
Influence of statistical method used on the resulting estimate of normal range.

93. Reed, A., & Wu, G.
Clin.Chem. 20 (1974), 576-581.
Evaluation of a transformation method for estimation of normal range.

94. Reed, A.
Clin.Chem. 16 (1970), 129-134.
Use of patient data for quality control of clinical laboratory tests.

95. Rehpenning, W., Harm, K., Domesle, A., & Voigt, K.
J.Clin.Chem.Clin.Biochem. 17 (1979), 565-572.
Falsch positive Werte bei der Vielfachanalyse. Die Abschätzung ihrer Häufigkeit mit der Sylvesterschen Formel und ihre Reduktion durch eine multivariate Testgrösse.

96. Richardson, R., Setchell, K., & Woodman, D.
Clin.Chim.Acta 37 (1972), 305-308.
Estimation of the "normal range" for serum cholesterol by different methods.

97. Roberts, L.
Clin.Chim.Acta 16 (1967), 69-78.
The normal ranges, with statistical analysis for seventeen blood constituents.

98. Robertson, E., Van Steirteghem, A., Byrkit, J., & Young, D.
Clin.Chem. 26 (1980), 30-36.
Biochemical individuality and the recognition of personal profiles with a computer.

99. Rümke, C., & Bezemer, P.
NTVG 116 (1972), 1224-1230 en 1559-1568.
Methoden voor de bepaling van normale waarden.
I algemene beschouwingen en beschrijving van enkele gangbare procedures.
II nieuwe methoden.

100. Rümke, C., & Bezemer, P.
NTVG 116 (1972), 1085-1090.
Welke zijn de normale waarden van het differentiële witte bloedbeeld?

101. Rümke, C.
NTVG 119 (1975), 1411-1413.
Ongebruikelijke laboratoriumuitslagen bij normale mensen.

102. Ryan, G., & Masarei, J.
Clin.Chim.Acta 91 (1979), 329-335.
Validity of "corrected" calcium values.

103. Saracci, R.
Am.J.Clin.Pathol. 56 (1971), 252-253.
Multitest screening and normal range.

104. Savitzky, A., & Golay, M.
Anal.Chem. 36 (1964), 1627-1638.
Smoothing and differentiation of data by simplified least squares procedures.

105. Shapiro, S., Wilk, M., & Chen, H.
J.Amer.Statist.Assoc. 63 (1968), 1343-1372.
A comparative study of various tests for normality.

106. Siest, G. et al.
Clin.Chem. 21 (1975), 1077-1087.
ASAT and ALAT activities in plasma : statistical distributions, individual variations and reference values.

107. Simonson, E.
Ann. NY Acad.Sci. 134 (1966), 541-558.
The concept and definition of normality.

108. Statland, B., & Winkel, P.
CRC Crit.Rev.Clin.Lab.Sci 8 (1977), 105-144.
Effects of preanalytical factors on the intraindividual variation of analytes in the blood of healthy subjects : consideration of preparation of the subject and time of venipuncture.

109. Steinier, J., Termonia, Y., & Deltour, J.
Anal.Chem. 44 (1972), 1906-1909.
Comments on smoothing and differentiation of data by simplified least squares procedures.

110. Streeter, A., Shum, H., & O'Neill, B.
Med.J.Aust. 2 (1973), 1019-1020.
Normal ranges in non-Gaussian distributions.

111. Sunderman, F.
Clin.Chem. 21 (1975), 1873-1877.
Current concepts of "normal values" and "discrimination values" in clinical chemistry.

112. Tadano, J., Niwa, M., Saito, H., & Sato, T.
Tokai J.Exp.Clin.Med. 5 (1980), 251-262.
The current concepts of "normal values" and clinical reference values in the clinical laboratory.

113. Tietz, N. *et al.*
Fundamentals of clinical chemistry, 2nd ed., W.B. Saunders Co., Philadelphia, London, Toronto.

114. Tsay, J-Y., Chen, I-W., Maxon, H., & Heminger, L.
Clin.Chem. 25 (1979), 2011-2014.
A statistical method for determining normal ranges from laboratory data including values below the minimum detectable range.

115. Underwood, E.
Brit.Med.Bull. 7 (1951), 265-274.
The history of the quantitative approach in medicine.

116. Vandenberg, S.
Ann. NY Acad.Sci. 134 (1966), 526-537.
Some advances in the statistical analysis of human variation.

117. Vries, S. de.
NTVG 116 (1972), 1102-1103.
Wat is de waarde van de differentiële telling van het witte bloedbeeld?

118. Watts, P., & Raty, D.
Beckton & Dickinson.
Curve fitting with data that is not "normally" distributed.

119. Werner, M., Young, D., Heilbron, D., & Dixon, W.
Clin.Chem. 16 (1970), 809.
Normal ranges and Gaussian distributions (letter).

120. Werner, M., & Marsh, W.
CRC Critical Reviews in Clinical Lab.Sci. (1975), 81-100.
Normal values : theoretical and practical aspects.

121. White, J.
Clin.Chim.Acta 84 (1978), 353-360.
Use of patient data in the control of urea, creatinine and electrolyte estimations.

122. Winkel, P.
Clin.Chem. 19 (1973), 1329-1338.
Patterns and clusters-multivariate approach for interpreting clinical chemistry results.

123. Winkel, P., & Statland, B.
Ann.Biol.Clin. 33 (1975), 174-182.
The biologic application of multivariate techniques. The subject as his own reference.

124. Winkel, P., Lyngbye, J., & Jörgensen, K.
Scand.J.Clin.Lab.Invest. 30 (1972), 339-344.
The normal region: a multivariate problem.

125. Wijnen, J.Th.M.
Persoonlijke mededeling.

126. Zieve, L.
Ann. NY Acad.Sci. 134 (1966), 563-572.
Misinterpretation and abuse of laboratory tests by clinicians.

Levensbeschrijving.

Promovendus werd geboren op 20 november 1951 te Grathem. Van 1964 tot 1970 werd de opleiding gymnasium B gevolgd aan het Bisschoppelijk College te Roermond. In 1970 werd een aanvang gemaakt met de studie scheikundige technologie aan de Technische Hogeschool Eindhoven. Het doctoraalexamen in de studierichting analytische chemie werd in 1975 met lof afgelegd. Hierna werd een dienstverband aangegaan met het St. Laurentius ziekenhuis in Roermond, waar ook van 1978 tot 1982 de opleiding tot klinisch chemicus gevolgd werd (opleider dr. ir. P.S.H. Kuppens).