

- De regressievergelijking van een subrange van de data kan zeker voldoen wanneer dat deel zich bevindt bij een belangrijke medische beslisgrens.
- Indien bij de regressieberekening een aanmerkelijk verschil bestaat tussen de kleinste kwadraten regressielijn en de Passing-Bablok-regressielijn wees dan op uw hoede.
- Maak altijd een scattergram en een verschilplot (Bland-Altman) zodat u kunt zien wat de werkelijke verschillen zijn tussen de metingen en trends beter zichtbaar worden.
- Een uiteindelijke beslissing over een aanvaardbare methode-overstap kan ook nog gesteund worden door een analyse van de overeenstemming tussen de resultaten in het pathologische gebied, m.a.w. is er concordantie voor de resultaten die verlaagd of verhoogd zijn t.o.v. het referentiegebied.

Literatuur

1. Hendriks HA, Kortlandt W, Verweij WM. Standardised comparison of five new-generatie immunoassay analyzers. *Clin Chem* 2000; 46: 105-111.
2. Fraser CG, Hyltoft Petersen P. Analytical performance characteristics should be judged against objective quality specifications. Editorial. *Clin Chem* 1999; 45: 321-323.
3. Fraser CG, Hyltoft Petersen P, Ricos C, Haeckel R. Proposed quality specifications for the inaccuracy and the imprecision of analytical systems for clinical chemistry. *Eur J Clin Chem Clin Biochem* 1992; 30: 311-317.
4. Ricos C, Alvarez V, Cava F, Garcia-Lario JV, Hernandez A. et al. Current databases on biological variation: pros, cons and progress. *Scand J Clin Lab Invest* 1999; 59: 491-500.
5. Klee G, Schryver PG, Kisabeth RM. Analytical bias specifications based on the analysis of effects on performance of medical guidelines. *Scand J Clin Lab Invest* 1999; 59: 509-512.
6. Lynch M, Thompson A, Gaumont B, et al. Transference of reference values of automated chemiluminescent immunoassays on a new platform. *Clin Chem* 1998; 44: 2061-2062.
7. Westgard JO. Editorial: Points of care using statistics in method comparison studies. *Clin Chem* 1998; 44: 2240-2242.
8. Stöckl D, Dewitte K, Thienpont LM. Validity of linear regression in method comparison studies: is it limited by the statistical model or the quality of the analytical input data. *Clin Chem* 1998; 44: 2340-2346.

Ned Tijdschr Klin Chem 2000; 25: 346-351

Bepalingen van geslachtshormonen in bloed anno 2000: juistheid en standaardisatie

J.H.H. THIJSSSEN

Gebruik makend van gegevens uit het in Nederland door de LWBA georganiseerde externe kwaliteitsborgingsprogramma over 1999 en de eerste helft van 2000 wordt de kwaliteit van bepalingen van oestradiol, progesteron en testosteron besproken. Achtergrond van deze evaluatie is de grote variatie die in de resultaten van metingen van deze steroïden geconstateerd kan worden. De juistheid van de meetresultaten van deze steroïden wordt beoordeeld aan de hand van gegevens uit een Duits extern kwaliteitsborgingsprogramma, dat informatie over concentraties bevat zoals gemeten met referentiemethodes. Aandacht wordt gegeven aan de resultaten afkomstig van automatische analysers. Meting van oestradiol in bloed blijkt met een redelijke kwaliteit en juistheid verricht te kunnen worden bij concentraties van meer dan circa 500 pmol/l, bruikbaar voor cyclusanalyse en voor fertiliteitsdoeleinden. Voor lagere concentraties

worden weinig betrouwbare en juiste concentraties gemeten. Progesteron geeft een acceptabele meting bij concentraties van meer dan circa 10 nmol/l. Voor testosteron zijn metingen in het referentiegebied van normale volwassen mannen (> 10 nmol/l) redelijk acceptabel, maar voor metingen bij vrouwen en kinderen zijn duidelijk minder goede resultaten te verwachten.

Het is alweer zes jaren geleden dat Norbert Tietz in een kritische editorial "Accuracy in Clinical Chemistry - Does Anybody Care?" in het toonaangevende tijdschrift *Clinical Chemistry* (1) zijn zorg verwoordde over met name de duidelijk tekort schietende aandacht voor de juistheid van bepalingen. Bij de gegeven voorbeelden werden ook enkele immunoassays besproken en wel die van thyrotropine (TSH), follitropine (FSH) en trijodothyronine. De conclusie was dat bij deze stoffen onvoldoende aandacht lijkt te hebben bestaan voor een beoordeling van de juistheid van de bepalingen, dus voor het analytisch en klinisch correct zijn van de resultaten. Met name bij de introductie van moderne analyse-automaten is regelmatig sprake geweest van het niet beschikbaar zijn van voldoende informatie om de analytische juistheid van alle mogelijke metingen te kunnen beoordelen. Veelal

ASL Endocrinologie, Universitair Medisch Centrum Utrecht

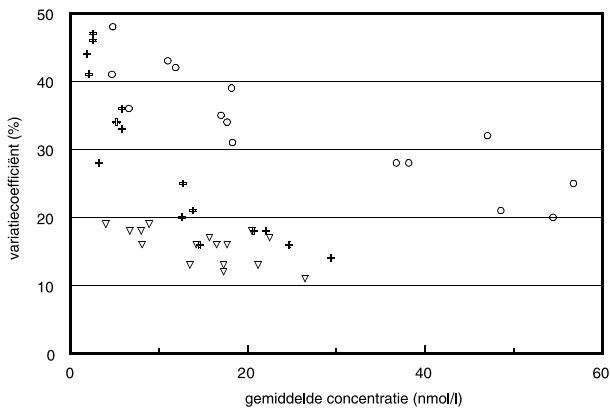
Correspondentie: Prof. Dr. J.J.H. Thijssen, ASL Endocrinologie, Universitair Medisch Centrum Utrecht, KE.03.139.2, Lundlaan 6, 3584 EA Utrecht.
E-mail J.Thijssen@LAB.AZU.NL

wordt volstaan met het geven van correlaties met bestaande bepalingen maar wordt niet getracht een sluitende verklaring te geven over afwijkingen van 1,00 in de berekende hellingshoeken of over duidelijk verschillende resultaten verkregen bij individuele monsters. In het algemeen kan natuurlijk gesteld worden dat het al bijzonder is dat met immunoassays enigszins vergelijkbare getallen als uitkomst te produceren zijn, omdat met vaak matig gedefinieerde reagentia (antilichamen) niet exact te karakteriseren stoffen bepaald worden, immers veel "hormonen" bestaan in werkelijkheid uit heterogene mengsels van meerdere componenten. Echter, steroidhormonen behoren juist niet tot de categorie van verbindingen die in principe uitsluitend op hun biologische eigenschappen te karakteriseren zijn, het zijn nauwkeurig bekende stoffen, die uitstekend in SI-eenheden uit te drukken zijn. Zoals Ekins (2) het verwoord heeft: analyse van een steroidhormoon is gericht op het meten van het aantal moleculen van die éne stof, met een unieke chemische structuur, in een monster. Resultaten kunnen uitgedrukt worden in eenheden (per volume) gebaseerd op het aantal van de betreffende moleculen en metingen dienen identieke resultaten te geven, ongeacht het meetprincipe. In theorie moet het dus zonder meer mogelijk zijn om voor steroiden aan goede analytische principes te voldoen. Bepalingen van steroidhormonen hebben in de afgelopen decennia een behoorlijke evolutie doorgemaakt. Van metingen na extractie en chromatografische zuivering zijn bepalingen alleen na extractie mogelijk geworden en meer recent zijn voor de geslachtshormonen een aantal directe bepalingen beschikbaar gekomen. Mede vanwege de binding aan specifieke transporteiwitten in bloed moeten daarbij maatregelen genomen worden om die totale concentratie van de gewenste steroiden te meten. In het volgende stuk zal ingegaan worden op meting van de concentratie van het totaal van de verschillende geslachtshormonen en niet op pogingen om de biobeschikbare (bioavailable) of vrije concentratie van bijvoorbeeld testosteron te meten. Daarover kan hier volstaan worden met de opmerking dat de thans daarvoor verkrijgbare directe bepalingen zeer slechte informatie blijken te geven (3). De kwaliteit van bepalingen van de voornaamste geslachtshormonen bij de mens (oestradiol, progesteron en testosteron) heeft de afgelopen jaren opvallend weinig publiciteit gehad, in de afgelopen vijf jaren is slechts één publicatie (4) te vinden met enige informatie over de analytische kwaliteit van bepalingen van een van deze drie hormonen. In de nu volgende bijdrage zal getracht worden gegevens verkregen uit externe programma's voor kwaliteitsborging te gebruiken die het mogelijk maken enige uitspraken te doen over de juistheid van metingen van deze steroiden. Een nadeel van het gebruik van gegevens uit kwaliteitsborgingsprogramma's is natuurlijk dat daarmee uitsluitend een retrospectieve beoordeling te geven is, immers alle getallen zijn in het verleden geproduceerd, over de eventueel nieuwste ontwikkeling bestaan nog geen resultaten. Echter, veel fabrikanten trachten reagentia voor bepalingen te leveren die over langere periodes stabiele uitkomsten geven. In onder-

staande beschrijving worden bovendien slechts resultaten van recente bepalingen gebruikt, die gedurende 1999 en de eerste helft van 2000 verkregen zijn.

MATERIALEN en METHODEN

Gebruikt worden gegevens van het kwaliteitsborgingsprogramma dat door de Sectie Landelijke Werkgroep Bindingsanalyse (LWBA) van de Stichting Kwaliteitsbewaking Ziekenhuis Laboratoria (SKZL) te Nijmegen verzorgd wordt. In dit programma worden, zoals bekend, zesmaal per jaar telkens twee monsters van drooggevroren maar verder onbewerkt humaan serum rondgestuurd waarin de deelnemende laboratoria componenten analyseren volgens de in het betreffende laboratorium gebruikte technieken. Aan telkens één van deze twee monsters worden bekende maar wisselende hoeveelheden van een aantal van de te analyseren stoffen toegevoegd, zodat een indruk verkregen wordt van de terugvinding ("recovery") van deze stoffen met de gebruikte analysetechnieken. De resultaten worden centraal bewerkt en weer aan de deelnemers teruggestuurd, daarbij worden gemiddelden van identieke analysemethoden berekend, met een standaard deviatie, na verwijdering van uitbijters volgens de methode van Healy (5). Per laboratorium worden de eigen resultaten gemeld o.a. in vergelijking met deze gemiddelden, alle andere individuele resultaten worden anoniem gegeven. Voor de te beschrijven resultaten werden de gegevens uit de zes rondes van 1999 en die van de eerste twee van 2000 gebruikt, in totaal dus acht paren van telkens twee monsters, tot en met mei 2000. Het tweede kwaliteitsborgingsprogramma (Deutsche Ringversuche (DR)) waaruit gegevens gebruikt werden, wordt verzorgd door de "Deutsche Gesellschaft für Klinische Chemie", dat vanuit het "Referenzinstitut für Bioanalytik, Abteilung für externe Qualitätssicherung" in Bonn (D) verstuurd wordt. In dit programma worden vier maal per jaar telkens twee monsters van onbewerkt drooggevroren humaan serum rondgestuurd ter analyse van specifieke hormonen, de deelnemende laboratoria analyseren volgens de bij hen gebruikelijke technieken waarvan de resultaten weer centraal bewerkt worden. Ook bij dit programma worden de resultaten van het betreffende laboratorium vergeleken met die van andere gebruikers van dezelfde reagentia en met die van de overige deelnemers. Het bijzondere van het programma, aangeduid met HM, is dat voor o.a. een aantal steroidhormonen de concentraties gemeten zijn met behulp van een referentiemethode, gebaseerd op GC-MS. De gebruikte referentie methode is voor o.a. geslachtshormonen gevalideerd door een grondige vergelijking met een onafhankelijk laboratorium dat analoge technieken hanteert (6). Volgens wettelijke Duitse richtlijnen worden resultaten per laboratorium beoordeeld op hun juistheid (Richtigkeit). In de volgende beschrijving zijn resultaten gebruikt verkregen over het gehele jaar 1999 en de eerste helft van 2000, in totaal dus over zes paren van telkens twee monsters. Voor zover mogelijk zijn resultaten verwerkt met reagentia van leveranciers met dezelfde naam als in Nederland.



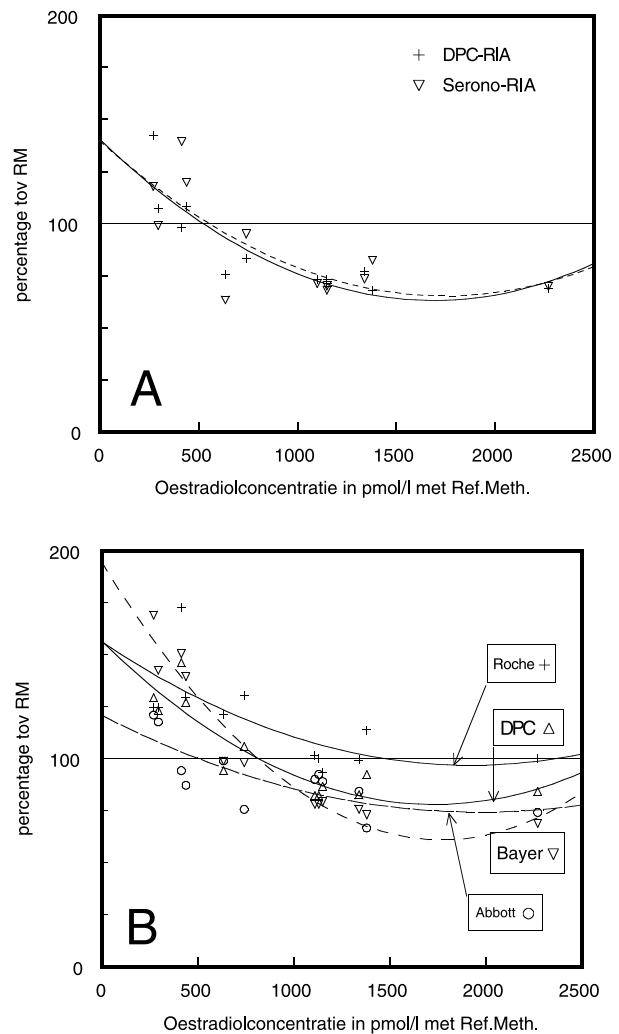
Figuur 1. Variatiecoëfficiënt geslachtshormonen berekend per individueel monster dat door de LWBA in 1999 en de eerste helft van 2000 binnen Nederland gebruikt is. Voor oestradiol (O), progesteron (+) en testosteron (∇) wordt de VC uitgezet tegen de gemiddelde concentratie van het betreffende hormoon in dit monster berekend uit de gerapporteerde resultaten. Voor oestradiol is deze concentratie in pmol/l door 10 gedeeld.

RESULTATEN

In figuur 1 wordt geïllustreerd dat enige onrust over de vergelijkbaarheid van metingen van geslachtshormonen terecht is, de variatiecoëfficiënt (VC) voor alle metingen in Nederland bedraagt 13 - 19% voor testosteron (gemiddelde van alle gerapporteerde concentraties tussen 4 en 27 nmol/l), 14 - 48% voor progesteron (idem tussen 2 en 30 nmol/l) en 20 - 49% voor oestradiol (idem tussen 50 en 600 pmol/l). Voor progesteron en oestradiol is er een onmiskenbare sterke toename van de VC bij de lage concentraties, voor testosteron is deze toename minder uitgesproken.

Oestradiol

Uit het jaaroverzicht 1999 van de LWBA voor oestradiol blijkt dat er zeer duidelijke systematische verschillen bestaan tussen metingen die gebruik maken van reagentia van verschillende leveranciers. Bij gemiddelde concentraties van 150 en 430 pmol/l respectievelijk, zijn de resultaten van gebruikers van reagentia van Abbott, Bayer en Roche ongeveer 20-40% hoger dan die van gebruikers van reagentia van Chiron, Clinical Assays, DPC en Wallac. Uit deze gegevens is uiteraard niet aan te geven welke van deze resultaten analytisch de beste juistheid hebben. In eerste instantie kan hierover iets gezegd worden op grond van de terugvinding die per set reagentia berekend wordt in het LWBA programma. Het groeperen van reagentia in twee klassen, handmatige versus volledig gemechaniseerde bepalingen, levert dan als eerste conclusie op dat de terugvinding met meer handmatige bepalingen (DPC-RIA, Orion, Wallac en BioMérieux) varieert tussen 51 en 75% ($n = 31$). Drie van de vier automaten (Abbott, Bayer en Roche) geven terugvindingen van 76-101% ($n = 24$), terwijl de DPC-LEIA 61-78% ($n = 8$) laat zien. Hierbij dient wel het volgende bedacht te worden: de toevoegingen bedroegen tussen de 108 en 540 pmol/l en de uiteindelijk bereikte concentraties lagen alle onder de 700 pmol/l, dus alle waren relatief laag. In de DR werden



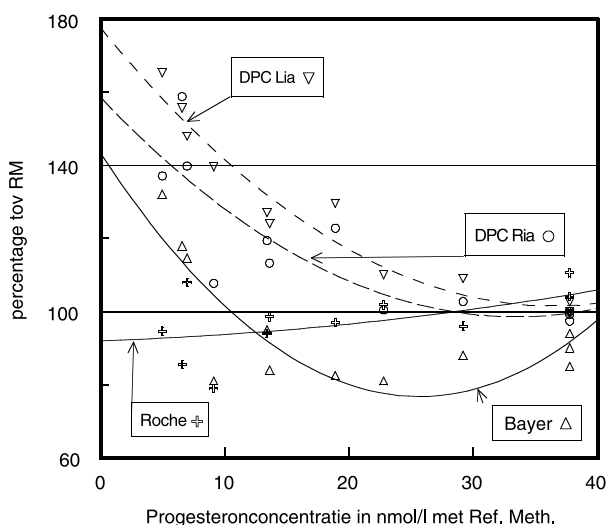
Figuur 2. Meetresultaten voor oestradiol over 1999 en de eerste helft van 2000, gerapporteerd in de DR na gebruik van reagentia die in Duitsland verkocht worden. Onder A worden de resultaten van twee handmatig uitgevoerde RIA's getoond terwijl onder B de resultaten verkregen met automaten te zien zijn. De gemiddelde waarde per methode is uitgezet tegen de concentratie gemeten m.b.v. een referentiemethode (RM). Per bepalingmethode is een polynoom berekend die zo goed mogelijk door de verschillende punten getrokken kan worden.

in de genoemde periode sera gedistribueerd die hogere oestradiol concentraties bevatten dan de LWBA-sera in Nederland, de met de referentiemethode gemeten concentraties varieerden tussen 235 en 2270 pmol/l. In Figuur 2 worden de resultaten getoond van de gemiddelden van metingen met behulp van enkele specifieke sets reagentia, onder A voor twee handmatige en onder B voor vier volledig geautomatiseerde technieken. Per set reagentia worden gemiddelden van 7 - 71 laboratoria gebruikt. Per leverancier is een curve getekend die zo goed mogelijk door de getallen getrokken kan worden. De getoonde twee handmatige methoden (fig. 2A) illustreren dat de, bij de LWBA geconstateerde, niet optimale terugvinding, ook te zien is in de vergelijking van resultaten met deze technieken met de getallen verkregen m.b.v. een referentiemethode. Zeker boven concentraties van circa 500 pmol/l liggen de meetresultaten aanzienlijk lager terwijl bij de lagere spiegels enige overwaardering

van de juiste concentraties te zien is. De vier directe methodes (fig. 2B) laten zien dat één van deze moderne technieken redelijk juiste getallen geeft boven concentraties van zo'n 1000 pmol/l, terwijl de drie andere weer tot te lage waarden tenderen. Bij lagere concentraties treedt bij alle technieken een zeer duidelijke afwijking op, de gemeten spiegels worden nadrukkelijk te hoog. Uit de getoonde resultaten kan als conclusie getrokken worden dat veel van de commercieel verkrijgbare reagentia te lage waarden geven voor oestradiol bij concentratie van meer dan 500 - 1000 pmol/l. Onder deze concentratie is eigenlijk altijd sprake van niet juiste resultaten, er is dan sprake van een wisselende, matig tot sterk positieve afwijking op.

Progesteron

Het jaaroverzicht 1999 van de LWBA toont voor progesteron eveneens een soort tweedeling in de verkregen resultaten, Abbott en DPC leveren getallen die in het algemeen zo'n 25% hoger liggen dan die van Biomérieux, Orion, Roche en Wallac, bij gemiddelde concentraties van ongeveer 4 en 19 nmol/l. De kwaliteit van de terugvinding is aanzienlijk beter en minder gevarieerd dan bij oestradiol, over de gehele periode wordt gemiddeld $91 \pm 4\%$ terug gevonden van de toegevoegde 8 tot 26 nmol/l. Tussen de vier automatische technieken bestaan wel significante verschillen in recovery maar de onderlinge verschillen zijn veel kleiner, met als uitersten DPC met gemiddeld 101% en Bayer/Chiron met 82%. De op grotere aantallen deelnemende laboratoria gebaseerde bevindingen uit de DR worden in figuur 3 geïllustreerd. Bij deze deelnemers zijn er meer die handmatig uitgevoerde RIA's gebruiken waardoor de resultaten met deze reagentia apart aan te geven zijn. In grote lijnen wordt bovenstaande bevestigd d.w.z. dat de gemiddelde resultaten met de gebruikte reagentia minder spreiding rond de 100% lijn laten zien, zij het dat dat voor drie van de



Figuur 3. Meetresultaten over 1999 en de eerste helft van 2000 voor progesteron uit Duitsland voor één handmatige RIA en drie automatische methoden. Gemiddelde resultaten per methode zijn uitgezet tegen de concentratie gemeten met een referentiemethode. De curves geven polynomen berekend op grond van de meetresultaten per bepalingstechniek.

vier vermelde leveranciers slechts geldt voor absolute concentraties hoger dan zo'n 10 nmol/l. Op grond van de getoonde gegevens kan geconcludeerd worden dat de juistheid van de bepalingen van progesteron in bloed boven concentraties van 10 nmol/l in het algemeen redelijk is, d.w.z. dat de verkregen spiegels minder dan 20% afwijken van de werkelijke concentraties.

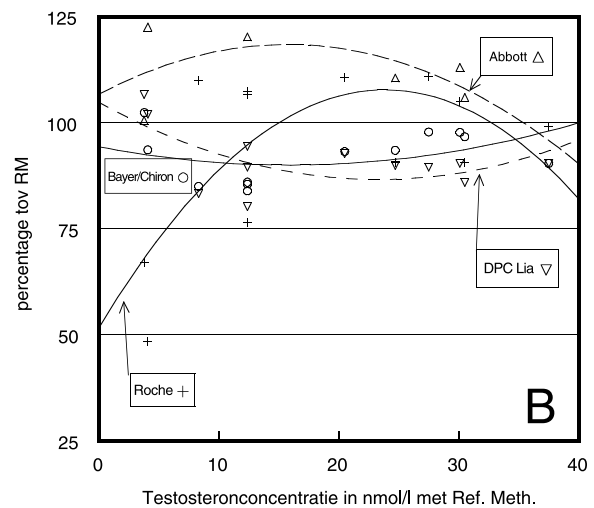
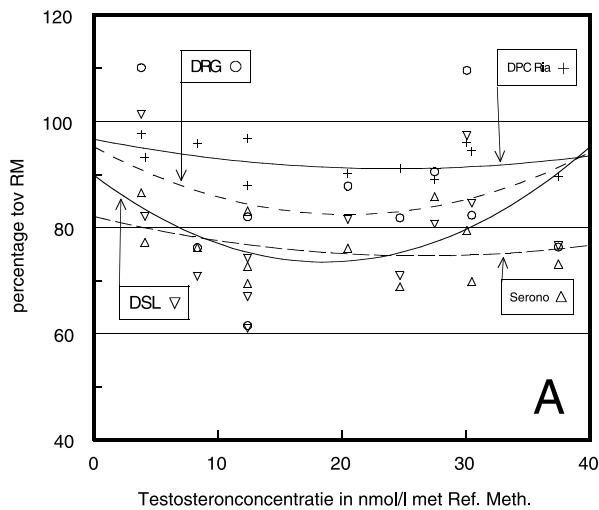
Testosteron

Het LWBA jaaroverzicht 1999 laat een minder uitgesproken tweedeling zien in resultaten voor testosteron, bij gemiddelde concentraties van 9,5 tot 20 nmol/l. Wel lijkt er meer variatie te bestaan binnen gebruikers van reagentia van DPC met luminescentie als detectie. In Nederland blijken overwegend reagentia van slechts één leverancier in grotere aantallen laboratoria gebruikt te worden en vandaar dat slechts iets te zeggen is over recoveries bij gebruik van de RIA en bij de geautomatiseerde LIA van DPC. Toevoegingen van 3,3 tot 10,8 nmol/l werden met de RIA voor gemiddeld 81% en met de LIA voor 72% teruggevonden maar gezien de spreiding in de getallen is het verschil niet significant. In de DR zijn grotere aantallen gebruikers voor verschillende reagentia geregistreerd. Daardoor is het mogelijk om weer onderscheid te maken tussen gebruikers van handmatige bepalingen (zie figuur 4A) en de automatische systemen (zie figuur 4B). Uit fig. 4A is te zien dat alle handmatige methoden tenderen naar het leveren van resultaten die gemiddeld 15 - 20% lager liggen dan de werkelijke concentraties over het geteste gebied tussen de 4 en 38 nmol/l. Opvallend anders zijn de resultaten die met automaten verkregen zijn: bij concentraties van meer dan 10 nmol/l liggen de gemiddeld gevonden waarden veelal binnen de 10% van de referentiewaarden. Slechts met één van de getoonde technieken worden afwijkende waarden gevonden bij concentraties van rond de 5 nmol/l, de overige drie blijven redelijk in de buurt van de echte concentraties.

Op grond van met name de uit Duitsland verkregen gegevens ziet het er naar uit dat testosteron bij mannen redelijk goed bepaalbaar lijkt m.b.v. moderne automatische technieken, bij concentraties van meer dan ongeveer 5 nmol/l.

DISCUSSIE

Uit de gepresenteerde gegevens lijkt de conclusie gerechtvaardigd te zijn dat bepalingen van de drie besproken geslachtshormonen een aanzienlijke verbetering hebben laten zien in de afgelopen jaren. Voor een aantal klinische toepassingen van de besproken bepalingen is dit een gunstige ontwikkeling. Deze verbetering vertaalt zich naar de constatering dat oestradiol mits aanwezig in concentraties van meer dan 500 pmol/l redelijk juist te meten is met moderne reagentia, ook met directe automatische technieken. Dit betekent dat diagnostiek voor de kliniek acceptabel verricht lijkt te kunnen worden binnen het gebied van de gynaecologie, t.b.v. monitoring van oestradiol gedurende de menstruele cyclus en tijdens stimulatie van



Figuur 4. Meetresultaten voor testosteron verkregen binnen de DR voor monsters die in 1999 en de eerste helft van 2000 rondgestuurd zijn. In figuur 4A worden de resultaten van handmatige bepalingen uitgezet tegen de concentraties die met een referentiemethode gemeten zijn. In figuur 4B wordt op analoge wijze het resultaat voor een aantal analyse-automaten getoond. De curves geven polynomen berekend op grond van de meetresultaten per bepalingstechniek.

de ovaria t.b.v. ovulatie-inductie. Zeer problematisch blijft het gebruik van alle besproken technieken ter bepaling van oestradiol voor andere klinische toepassingsgebieden. Het blijft onmogelijk om met commerciële reagentia bij groepen als kinderen vóór de puberteit, vrouwen na de menopauze of bij mannen betrouwbare en juiste informatie te krijgen over oestradiol concentraties. Voor dergelijke toepassingen moeten nog steeds de technieken met extractie en chromatografische zuivering gebruikt worden. In de voorbeelden die onder resultaten besproken zijn, is geen van deze methoden apart vermeld. Opgemerkt dient echter te worden dat deze arbeidsintensieve technieken de enige zijn die ook bij zeer lage concentraties gegevens opleveren die zeer kort bij de echte oestradiolconcentraties blijken te liggen. Dat blijkt duidelijk uit de resultaten van de monsters die in Duitsland gebruikt worden. Analooft geldt voor de bepaling van progesteron dat concentraties boven de 10 nmol/l met redelijke juistheid te meten zijn d.w.z. dat concentraties van dit steroïd in de luteale fase van de normale menstruele cyclus en tijdens een zwangerschap veelal betrouwbaar te meten zijn. Ook hier de duidelijke beperking dat bij metingen van progesteron voor andere doeleinden een selectieve keuze van reagentia noodzakelijk is. Tenslotte testosteron: ook voor het mannelijke geslachtshormoon, in het gebruikelijke referentiegebied voor mannen, boven de 10 nmol/l, kan gesteld worden dat meting ervan acceptabel begint te worden. Voor dit steroïd lijken de moderne automatische technieken gemiddeld waarden te leveren met een meer juiste concentratie dan met de handmatige immunoassays. De spreiding van meetresultaten van automaten t.o.v. de waarden verkregen met een referentie methode is duidelijk kleiner. Ook hier de nadrukkelijke beperking voor metingen van testosteron bij mannen en wel bij concentraties hoger dan circa 4 nmol/l. Dit betekent dat op grond van de beschikbare getallen geen harde

uitspraak te doen is over de juistheid van metingen van dit hormoon bij lagere concentraties, zoals die voorkomen bij vrouwen, kinderen en hypogonade mannen. De beschikbare data bevatten geen getallen van monsters met voldoende lage testosteron concentraties. Uit andere gegevens blijkt echter dat er nog steeds aanzienlijke verschillen (>50% verschil) bestaan tussen testosteron concentraties gemeten bij gezonde vrouwen met reagentia van verschillende leveranciers. Ook incidentele metingen bij kinderen met problemen rond de geslachtelijke ontwikkeling geven aanleiding tot grote bezorgdheid over de juistheid van de meetresultaten bij het gebruik van commerciële reagentia bij de genoemde groepen. Samenvattend kan gesteld worden dat voor zeer belangrijke gebieden van diagnostiek redelijk correcte metingen mogelijk zijn voor geslachtshormonen. Aangegeven is voor welke, eveneens belangrijke diagnostische gebieden, goede resultaten nog ontoegankelijk zijn bij gebruik van commerciële reagentia.

Literatuur

1. Tietz NW. Accuracy in clinical chemistry - does anybody care? *Clin Chem* 1994; 40: 859-861.
2. Ekins R. Immunoassay standardisation. *Scand J Clin Lab Invest* 1991; 51(Suppl 205): 33-46.
3. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. *J Clin Endocrinol Metab* 1999; 84: 3666-3672.
4. Middle J. Standardisation of steroid hormone assays. *Ann Clin Biochem* 1998; 35: 354-363.
5. Healy MJR. A mean difference estimator of standard deviation in symmetrically censored normal samples. *Biometrika* 1978; 65: 643-646.
6. Thienpont LM, Nieuwenhove B van, Stöckl D, Reinauer H, Leenheer AP de. Determination of reference method values by isotope dilution-gas chromatography/mass spectrometry: a five years' experience of two European Reference Laboratories. *Eur J Clin Chem Clin Biochem* 1996; 34: 853-860.

Summary

Thijssen JJH. Sex steroids measurements in blood: accuracy and standardisation. Ned Tijdschr Klin Chem 2000; 25: 346-351.

Using data obtained during 1999 and the first part of 2000 from the LWBA, the National Dutch External Quality Assessment Scheme (EQAS), an attempt has been made to judge the quality of determinations in blood of the major sex steroids, oestradiol, progesterone and testosterone. The background of the evaluation is the unacceptably high variation in the results of sex steroid measurements observed in the Dutch EQAS. Therefore an additional evaluation of the accuracy was made, based on comparison of results by using specific sets of

reagents with those obtained using reference methods, as available from the German EQAS. Special emphasis was given to the performance of automated analysers.

Determinations of oestradiol do show a reasonable accuracy when concentrations of more than 500 pmol/l were analysed. At lower concentrations none of the commercial methods gave accurate results, all values were too high. For progesterone acceptable measurements are possible at concentrations above 10 nmol/l. In the usually accepted male reference range of more than 10 nmol/l, testosterone determinations do perform in an acceptable way. However, no reliable results are to be expected in the low concentration ranges, in women and children.

Ned Tijdschr Klin Chem 2000; 25: 351-353

KC + automaten: automatisch kwaliteit?

E. LENTJES

De titel kan gelezen worden als: kwaliteitscontrole en automaten of klinisch chemici en automaten. Met de kwaliteitscontrole bedoel ik de externe controle voor de bindingsanalyse, m.a.w. wat kunnen we uit deze gegevens halen m.b.t. de prestaties van de automaten. In het geval van de klinisch chemici wil ik nadruk leggen op de keuze voor en het omgaan met een automaat en als er geautomatiseerd wordt leidt dit dan automatisch tot een betere kwaliteit?

De gegevens die kunnen worden gehaald uit de rondzendingen van de SKZL sectie Landelijke Werkgroep Bindingsanalyse (LWBA) hebben betrekking op de interlaboratoriumspreiding, maar er kan ook gekeken worden naar de prestaties van een individueel laboratorium (denk aan het scoringssysteem waar ik vorig jaar over heb verteld). In 1996 heb ik voor de tumormarkers een precision profile gemaakt van de methoden. Over een periode van vijf jaar bleek de interlab variatie coëfficiënt (VC) van jaar tot jaar lager te worden. Dit betekent dat de methoden naar elkaar toe kruipen, mogelijk door activiteiten van de fabrikanten (kalibratie) en/of door betere prestaties van de individuele laboratoria. Geldt deze verbetering ook voor de hormonen? Om deze vraag te beantwoorden heb ik bepalingen van twee kleinere (cortisol en oestradiol) en van twee grotere moleculen (TSH en LH) nader bekeken. De reden is dat automaten in het algemeen redelijk overweg kunnen met de immunometrische assays (TSH en LH), maar problemen hebben met de competitieve assays (steroiden). Voor alle vier de be-

palingen zijn in figuur 1 de gemiddelde interlab-CV's over alle methoden van 1985, 1990, 1995 en 1999/2000 weergegeven. Hieruit blijkt voor TSH en voor een deel voor LH dat de CV van jaar tot jaar kleiner is geworden, voor TSH met name tussen 1985 en 1990. Voor LH is er iets vreemd aan de hand: na een aanvankelijke daling van de CV blijkt in 1999 de CV weer te stijgen. Kijken we daarentegen naar cortisol en oestradiol dan blijkt het omgekeerde het geval: de interlab-CV neemt van 1985 tot heden toe. Voor cortisol is dit maar een paar procent, maar voor oestradiol is dit bij de lagere concentraties tientallen procenten. Hieronder ga ik nader in op de vier bepalingen.

TSH

We weten dat er in door de industrie hard is gewerkt aan de verbetering van de TSH-bepaling, wat heeft geresulteerd in nieuwere generaties assays, m.a.w. assays met een betere gevoeligheid en betere reproduceerbaarheid. In de afgelopen 15 jaar zien we dat de radioactieve bepalingen fors werden teruggedrongen ten gunste van de methoden met fluorescentie- en luminescentielabels, die meer en meer geautomatiseerd werden uitgevoerd. Begin 1999 werden nog maar 5% van de methoden uitgevoerd met behulp van een radioactief label, terwijl dat in 1985 meer dan 95% was. In de Youden plots van de LWBA-ronde-overzichten van 2000 zijn de verschillen tussen de diverse methoden nagenoeg weggewerkt. Toch is er onderscheid te maken tussen de methoden zoals te zien is in figuur 2. Hierin zijn de scores per laboratorium, ingedeeld naar methode, weergegeven. De score wordt berekend uit de spreiding rond de regressielijn van de resultaten van een laboratorium afgezet tegen de rondegemiddelden. De scores zijn vervolgens gedeeld door het gemiddelde van alle rondegemiddelden waardoor een soort CV wordt verkregen. Zoals te zien is onderscheiden Bayer (Advia), Roche (Elec-

Afdeling Klinische Chemie, Leids Universitair Medisch Centrum, Leiden

Correspondentie: Dr. E.G.W.M. Lentjes, Leids Universitair Medisch Centrum, Afdeling Klinische Chemie, Postbus 9600, 2300 RC Leiden.