

Test av riktighet med NFKK Reference Serum X utført i danske, islandske og norske laboratorier

Pål Rustad, Først Medisinsk Laboratorium, Oslo. E-post: prustad@furst.no



Innledning

I forbindelse med innføring av NORIP referanseintervaller i de nordiske land, er det sertifiserte referansematerialet "NFKK Reference Serum X" (senere kalt X) tilbudt nordiske medisinsk biokjemiske laboratorier for testing av riktighet for de aktuelle analysene.

Uavhengig av om targetverdiene for X er riktige eller ikke, er sporbarheten for NORIP referanseintervaller knyttet til disse targetverdiene. X er beskrevet tidligere i dette tidsskriftet [1] og i andre fora [2] hvor det er lagt vekt på å dokumentere tre fundamentale egenskaper ved materialet som er avgjørende for tilliten til det:

- Kommutabilitet (karakteriserer matriks) – X er en fersk-frosset serumpool fra menn og er kun blandet, sterilfiltrert og fordelt til 5 mL polypropylenrør før frysing ved -80°C .
- Sporbarhet (riktighet av targetverdier) – X har targetverdier sporbare til referansemetoder for 18 av komponentene, og konsensusverdier fra NORIP for 7 komponenter av hvilke 5 er enzymer [3].
- Holdbarhet – materialet er oppbevart frosset ved -80°C og sendt til laboratoriene på tørris.

For å lette arbeidet i laboratoriet og gjøre vurderingen av avvik i riktighet sikrere og mer sammenlignbar, ble det i samarbeid med Norsk Klinisk-kjemisk Kvalitetssikring (NKK) foreslått en protokoll for måling og utarbeidet et regneark for beregning [4].

Det ble også laget en rettleiding publisert på hjemmesiden til NKK hvor ulike problemstillinger ved kontroll av riktighet ved bruk av X er diskutert [5] (tilsvarende rettleiding finnes også på EQUALIS' hjemmeside). Her er spesielle komponenter og målesystemer/metoder diskutert, samt for eksempel hvor realistisk det foreslåtte kvalitetsmål er for enkelte komponenter.

For bedre å kunne vurdere årsaken til eventuelle avvik var det nødvendig å kunne skille mellom feil som er generelle for målesystemet og feil som ikke er det. Det var derfor naturlig å vurdere data fra flere laboratorier med samme målesystem.

Metode

Regnearket

I regnearket er targetverdier inkl. standard usikkerhet for X registrert, samt kvalitetsmål for bias basert på 37.5% av total biologisk variasjon.

X ble tilsendt laboratoriene på tørris og protokollen for analyse av X gikk ut på:

- Mål i en serie ti replikater annen hver en av X og av bruksstandard (lokal "kalibrator" kalt C senere).
- Registrer måleverdiene for X og C i regnearket.
- Registrer targetverdier (inkl. standard usikkerhet hvis tilgjengelig) for C i regnearket.

Flg. beregninger blir utført i regnearket:

- Middelerverdi og standard avvik for X og C (kalt M_X og M_C)
- "Korrigert" middelerverdi for X beregnet som $M_X^* = M_X T_C / M_C$ der T_C er targetverdi for bruksstandard
- Ukorrigerte og korrigerte relative avvik (hhv $M_X / T_X - 1$ og $M_X^* / T_X - 1$) fra targetverdi blir så vurdert

Produsent	Type	Ant.	Sum	
Abbott	Aeroset	3	7	
	Architect	4		
Bayer Advia	1650	10	10	
Beckman	Synchron	1	1	
DB Dimension	RxL	1	1	
Konelab	30i	1	1	
Olympus	AU 640	1	1	
	Roche Cobas	Ukjent		1
	I400	3		
	I700	5		
	I800	9		
Roche Hitachi	Mira	1	19	
	911	2		
	912	6		
	917	5		
Vitros	Modular	11	24	
	250	9		
Vitros	950	9	18	

Tabell 1: Målesystemer benyttet. I plottene over avvik (Figur 1 og 2) er laboratoriene samlet i hovedgrupper som angitt til venstre i tabellen.

for statistisk signifikans ($p=0.05$) på bakgrunn av usikkerhetsangivelsene, og om de var viktige i forhold til kvalitetsmål for bias.

Resultater

I samarbeid med de respektive nordiske EQA-organisasjoner ble 34 regneark fra Danmark, 3 fra Island og 39 fra Norge (utført vår-høst 2004) samlet inn.

Målesystemene representert i undersøkelsen er vist i Tabell 1. Dataanalysen er presentert på NORIP's hjemmeside [6] i form av plott av enkeltlaboratoriers avvik sortert etter målesystem (se Figur 1 og 2). I plottene er resultatene samlet i hovedgruppene vist lengst til ven-

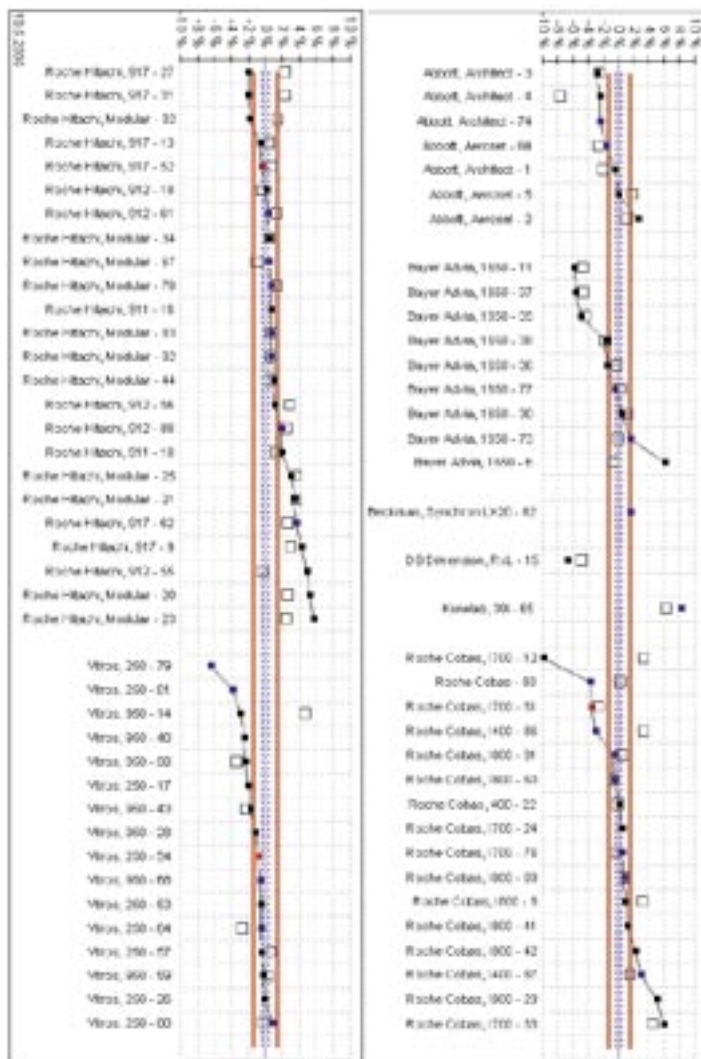
stre i tabell 1. Hvert laboratorium har fått et konfidensielt nummer slik at man kan vurdere eget avvik mot andre som har samme målesystem. Relative avvik er presentert som "ukorrigerte" (fylte kvadrater) og "korrigerte" (åpne kvadrater) avvik. I plottene er også vist ekspandert usikkerhet ($k=2$) for targetverdi (stiplede blå linjer) og kvalitetsmål (hele røde linjer). Ikke alle laboratoriene har målt det anbefalte antall replikater på 10, og noen har heller ikke målt bruksstandard parallelt, det siste vil fremgå av Figur 1 og 2 ved at det ikke finnes et åpent kvadrat for disse laboratoriene.

Selv om antallet laboratorier i hver gruppe er lite og analysen kan ha skjedd med opptil 6 måneders mellomrom, kan man likevel for enkelte målesystemer og komponenter trekke noen klare konklusjoner.

For alle målesystemene har man rettlinjede standardkurver unntatt for Vitros (og for TIBC dersom denne er målt som transferrin). Man kan derfor ikke vente at en proporsjonal korrigering skal fungere like godt for Vitros som for de andre målesystemene. Det er da heller ikke mange av Vitros-brukerne som har målt bruksstandard sammen med X.

Man ville forvente at spredningen av korrigerte avvik ville være mindre enn for ukorrigerte avvik fordi

(Fortsætter side 20)



Figur 1: Kalsium: Ukorrigerte (fylte kvadrater, svarte: Island og Norge, blå: Danmark, røde: Roche og Vitros referanselaboratorium) og korrigerte (åpne kvadrater) avvik fra targetverdi for NFKK Reference Serum X (3-10 replikater). Blå linjer representerer targetverdi for X med ekspandert usikkerhet (95% - stiplede linjer). Røde linjer representerer kvalitetsmål som $0.375 \cdot CV_b$ hvor CV_b er total biologisk variasjon angitt som variasjonskoeffisient. Plottet viser betydelige avvik i riktighet for flere målesystemer, men også stor variasjon mellom laboratorier innen samme målesystem. Generelt er nok presisjonen for analysen et problem til tross for at utmerkede metoder finnes tilgjengelige på markedet.

(Fortsat fra side 19)

mellom-serie variasjonen i prinsippet er tilnærmet eliminert.

Resultatet av en slik vurdering er vist i Figur 3 hvor spredningen av middelverdier er angitt før og etter korrigering for de ulike hovedgrupper av instrumenter. Figuren viser tydelig at korrigeringen har den forventede virkning til tross for at man introduserer en ekstra usikkerhetskomponent ved korrigering.

Konklusjon

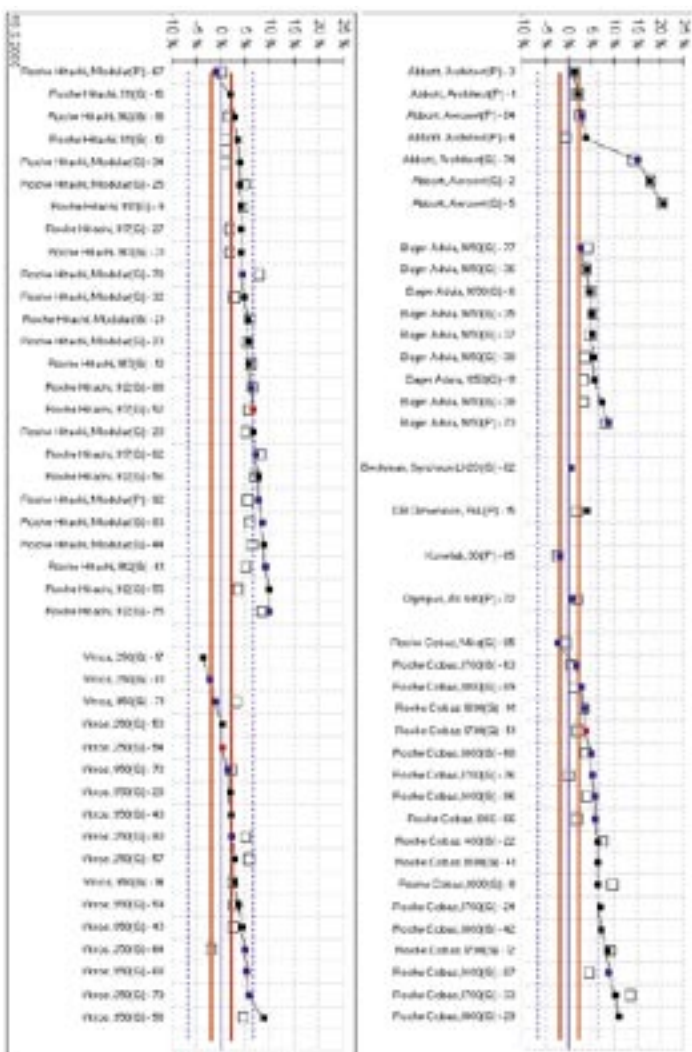
Fordi materialets kommutabilitet, sporbarhet, for-

sendelse, oppsett for måling og beregninger for vurdering er så optimale som vel mulig i denne undersøkelsen, er det grunn til å legge stor vekt på de resultater som fremkommer. Den eneste ulempen er at konsentrasjonene for enkelte komponenter i X er for lav til at man kan trekke konklusjoner om avvik på andre og viktigere nivåer, som for eksempel ved referansegrenser og beslutningsgrenser (gjelder spesielt flere enzymer og bilirubin).

Det er funnet flere avvik som bør bekymre i første rekke laboratoriene og brukerne av laboratoriesvarene, men også produsentene av målesystemene:

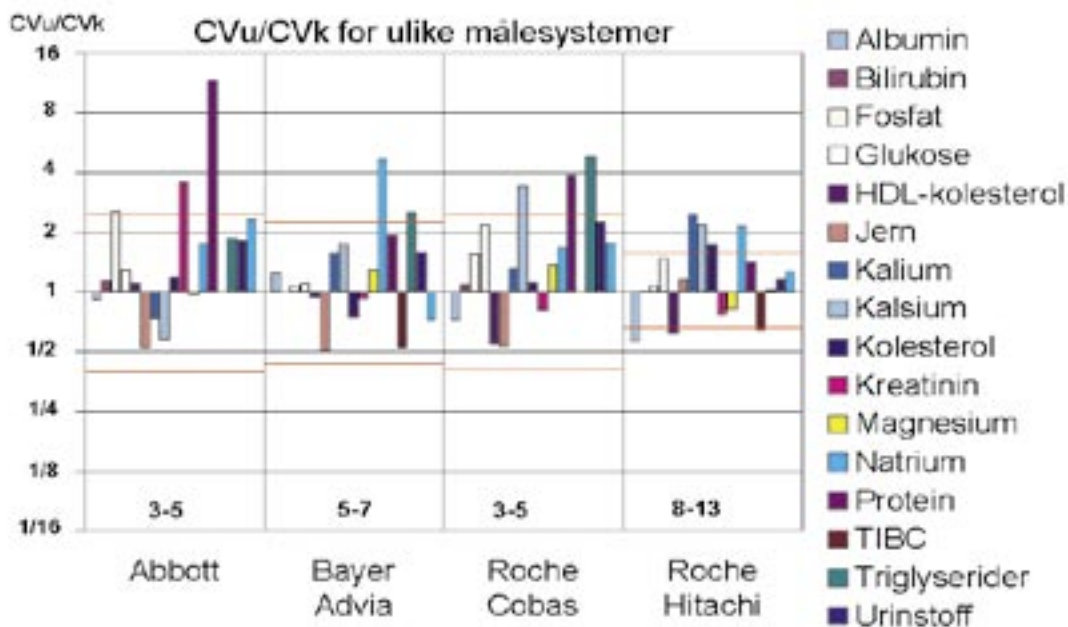
- Mest bemerkelsesverdig er avviket som ble funnet for Abbott Architect / Aeroset for magnesium, ca +20%! Resultatet er rapportert sentralt til Abbott, og midlertidig korreksjonsfaktor anbefalt av Abbott Norge AS benyttes nå i Norge.
- Det er også verdt å legge merke til Roche Hitachi/Modular's avvik for kalsium på gjennomsnittlig +2-3%, som er betydningsfullt for denne komponenten. Roche benytter atomabsorpsjon som referansemetode mens X er sporbar til IDMS. Roche angir nå, som

(Fortsætter side 24)



Figur 2: Albumin: Ukorrigerte (fylte kvadrater, svarte: Island og Norge, blå: Danmark, røde Roche og Vitros referanselaboratorium) og korrigerte (åpne kvadrater) avvik fra targetverdi for NFKK Reference Serum X (3-10 replikater). Blå linjer representerer targetverdi for X med ekspandert usikkerhet (95% - stiplede linjer). Røde linjer representerer kvalitetsmål som 0.375-CVb hvor CVb er total biologisk variasjon angitt som variasjonskoeffisient. Generelt ligger alle målesystemene i gjennomsnitt for høyt. Selv om måleusikkerheten er stor for targetverdien for X, er denne sporbar til referansemetode (RID med referansematerialet CRM 470) gjennom to uavhengige undersøkelser (DGKC 1997 og DGKC 2002) med verdien 41.5 g/L i begge tilfeller.

(Fortsat fra side 20)



Figur 3: Forholdet mellom variasjonen av ukorrigerede (CVu) og korrigerede (CVk) middelværdier av X for hvert målesystem (bare norske resultater). De røde horisontale linjer angir et 95% konfidensintervall for forholdet. Ordinatskalaen er logaritmisk for at verdier over og under 1 visuelt skal være sammenlignbare. Det er generelt en tydelig lavere variasjon for korrigerede verdier enn for ukorrigerede.

et supplement, også "kalibrator"-verdier målt med IDMS.

- Albuminresultatene ligger generelt høyt. X er sporbar til referansemotoden RID med målestANDARD CRM 470. Usikkerheten i targetverdien er relativt stor, men likevel er det grunn til å anta at målesystemene gir for høye verdier (unntatt Vitros). Det er verdt å merke seg at overført verdi for X i to uavhengige undersøkelser (overført verdi i NORIP fra CAL med referansemotodeverdi fra DGKC fra 1997 og overført verdi i den nordiske riktighetsstudien fra IMEP 17 Material 1 med referansemotodeverdi fra DGKC i 2002) gir identiske targetverdier for X, nemlig 41.5 g/L.
- For jern ligger både Abbott- og Bayer-systemene ca 15% for lavt, hvilket er unødvendig selv om kvalitetsmålet for denne komponenten ikke er strengt pga høy biologisk variasjon.
- For natrium er det påfallende hvor stor effekt

korrigeringen har på variasjonen innen en gruppe – standard avvik for laboratorienes middelværdier blir halvert etter korrigering for både Roche Hitachi og Bayer-instrumentene.

- For kreatinin er det naturlig at Jaffé-metodene på grunn av interferenser gir for høye målresultater, spesielt gjelder dette der kalibratoren har en høy kreatininverdi, for eksempel Bayer og Roche-instrumentene. Abbotts kalibrator er en normalverdi hvilket gir lite avvik for X, men desto større avvik ved høyere verdier. Forsøk viser at den kompenserte Jaffé-metoden til Roche gir riktige resultater. For andre Jaffé-metoder kan viktig informasjon hentes fra EQUALIS' hjemmeside, om EQUALIS's kreatininfrie serum. Dette kan benyttes for å finne hvilken korreksjon som er nødvendig for å oppnå referansemotodenivå [7]. De enzymatiske metodene og korrigeret Vitros-metode (iflg. Ortho: ny = gammel*1.02 - 8.1 [umol/L])

viser god riktighet for kreatinin. Det er tydelig at de fleste Vitros-brukere enda ikke har tatt i bruk korreksjonen for å gi overensstemmelse med referansemåtenivå.

- For urea viser Bayer 1650 ca 7 % for høye svar mens Roche Hitachi ligger ca 5% for høyt.
- For enzymene er som nevnt targetverdiene for X lave, og kvalitetsmålene er vide, slik at det er lite å kommentere for disse.
- For de to komponentene (av ikke-enzymene) HDL kolesterol og TIBC hvor targetverdiene for X ikke er sporbare til referansemåtenivå, er det ubetydelige avvik for alle målesystemer.

Det er heldigvis vist stor interesse for resultatene av denne undersøkelsen fra flere av de store produsentene, for eksempel Roche, Ortho og Abbott, og viktige tiltak er allerede satt i verk for å verifisere funnene og eventuelt rette opp påviste feil.

Selv om tydelige (statistisk signifikante) og betydelige (viktige) gjennomsnittlige avvik er funnet, kan man spørre hva som er årsaken til at spredningen innen samme målesystem for mange komponenters vedkommende er så stor som den er. Det er ikke sikkert at alle laboratoriene har fulgt de prosedyrer som er anbefalt – for eksempel ble det oppfordret til at man skulle sende inn originaldata fra instrumentene mens det er tydelig at dette for eksempel ikke er gjort for noen enzymer for Vitros. Man vet heller ikke om materialene er behandlet som foreskrevet, for eksempel om X og bruksstandard er målt i samme serie (og ikke to ulike serier). Noen har kanskje målt alle 25 komponenter i samme serie hvilket kan ha betydning for resultatet (f.eks natrium). Variasjonen kan også være et uttrykk for målesystemets robusthet, dvs hvor godt produsent av instrument og reagens klarer å holde

en jevn kvalitet.

Litteratur

1. Mortensen PM, Rustad P, Simonsson P. NFKK Reference Serum X. Klinisk Biokemi i Norden, no 4, 2003.
2. Henriksen GM, Pedersen MM, Nørgaard I, Blom M, Blou L, Blaabjerg O, Uldall A. Minimally processed fresh frozen human reference sera: preparation, testing, and application to international external quality assurance. Scand J Clin Lab Invest 2004; 64; p293-308
3. NORIP traceability: <http://www.furst.no/norip/cal.htm>
4. Regneark: http://www.furst.no/norip/X/x_spreadsheet.xls
5. Vurdering av analysekvalitet i forbindelse med innføring av NORIP referanseintervall: <http://www.legeforeningen.no/index.gan?id=39967&subid=0>.
6. <http://www.furst.no/norip/X/nor.htm>
7. Sertifikat for kreatininritt serum: <http://www.equalis.se/NORIP/Certifikat%20Kreatininritt%20serum%201.pdf>