

# Validering av analysekvalitet, Del 2.

Presisjon, måleområde, usikkerhet

PÅL RUSTAD Først Medisinsk Laboratorium, Søren Bulls vei 25, N-1051 Oslo (prustad@furst.no)

Dette er andre del av "Validering av analysekvalitet" som er basert på en "Standard Prosedyre" for metodevalidering som gjelder for Først Medisinsk Laboratorium.

I del I ble tester og riktighet omtalt (KKN 4/2000).

## Presisjon

Definisjon: Overensstemmelse mellom uavhengige måleresultater oppnådd med en måleprosedyre under spesifiserte betingelser [10].

Uttrykkes kvantitativt ved standard avvik ( $s$ ) og/eller variasjonskoeffisient ( $s/M$ ) i % (%CV).

## Repeterbarhet

Definisjon: Overensstemmelse mellom resultatene av påfølgende målinger av samme målestørrelse utført under samme målebetingelser [1].

Samme målebetingelser kan innebære samme målemetode, observatør, måleinstrument, lokaler, bruksbetingelser osv. Målingene bør utføres over kort tid og f.eks. med uttak av prøver fra samme prøvemateriale. Kvantitativt uttrykkes repeterbarhet ved standard avviket  $s_r$ , men målebetingelser bør likevel angis mer spesifikt.

"Innen serie variasjon" er et uttrykk som ofte brukes synonymt med repeterbarhet.

Hensikten med å bestemme repeterbarhet er å lokalisere usikkerhetskildene for målemetoden. Repeterbarheten er uttrykk for variasjon i selve måleprosessen og kan inngå som en del av usikkerhetsbudsjettet (se «Måleusikkerhet» bak).

Innen serie standard avvik er det nødvendig å bestemme for å lage korrekte kontrollgrenser ved bruk av middelvei- og rangeregel hvis mer enn 1 kontroll benyttes for å vurdere serien.

## Reproduserbarhet

Definisjon: Overensstemmelse mellom resultatene av påfølgende målinger av samme størrelse utført under ulike målbetingelser [1].

I motsetning til repeterbarhet skal altså målebetingelsene her varieres mest mulig, dvs forskjellige operatører, bruksbetingelser, målinger over

lang tid, ulike instrumenter. I mange sammenhenger tas også med ulike lokaler og metoder. Kvantitativt angis den med standard avviket  $s_R$ .

For å angi reproduserbarhet innen laboratoriet (samme laboratorium og metode) er foreslått begrepet intern reproduserbarhet (eng.: intermediate precision) som uttrykkes kvantitativt med  $s_I$ .

Variasjonen i måleresultater er selvsagt svært avhengig av hvilke betingelser målingene er utført under. Et standard avvik skal derfor alltid være ledsaget av slike opplysninger. En enkel (men ufullstendig) måte å angi betingelsene er å bruke suffiksene  $r$ ,  $R$  og  $I$ .

SWEDACS retningslinjer (fra 1995) lyder: "Metodens presisjon anges som CV beråknat på minst 30 värden uppmått under en tidsperiod av minst 3 månader. For nya metoder får preliminära CV baserade på kortare inkörningstid, användas" [2a].

Vårt laboratoriums måleusikkerhet er angitt i listen «Analysevariasjon» som %CV og er uttrykk for vår interne reproduserbarhet ved måling.

Eksempel: Mål en lav og en høy prøve 10 ganger med så mange variable før og under målingene som mulig. F.eks. utføres målingene på ulike prøveuttak av prøven i ulike serier med ulike operatører med kalibrering for hver serie.

Spesifiser betingelser. Beregn  $s_I$ .

Krav, repeterbarhet og intern reproduserbarhet: Hvis nødvendig, enten for bruk ved beregning av kontrollgrenser eller for bruk i usikkerhetsbudsjett, beregnes  $s_I$  og  $s_R$  fra paralleller av en prøve i helst 20 serier ( $s_I$  beregnes som standard avvik for alle resultatene og  $s_r$  som rot av gjennomsnittlig kvadratsum av differanser mellom paralleller dividert med  $\sqrt{2}$ ). Man får med dette oppsettet et gjennomsnitt for repeterbarhet over flere serier, og man kan beregne mellom serie variasjon som  $\sqrt{(s_I^2 - s_r^2)}$  (gjelder for store  $n$ , internt regneark

kan brukes for beregning fra mindre antall data).

Krav: I valideringsøyemed beregnes  $s_I$  fra minst 5 (helst >20) måleresultater av samme prøve i flere serier.

Man skal i den grad det er nødvendig for bruk i kontrollregler også angi mellom serie variasjonen. For bruk som måleusikkerhet i laboratoriets informasjon til kundene, skal  $s_I$  beregnes over et tidsrom på minst 3 måneder i rutine.

Hvis den intern reproducerbarhet ved måling uttrykt som variasjonskoeffisient er  $CV_a$  og de biologiske variasjonene, hhv. intra-, interindividuell og total variasjon er hhv.  $CV_w$ ,  $CV_b$  og  $CV_t$ , der  $CV_t = \sqrt{(CV_w^2 + CV_b^2)}$  kan man antyde følgende mål for analytisk kvalitet  $CV_a$  [4]:

	Monitorering	Screening
Optimal:	<0.25 $CV_w$	<0.25 $CV_t$
Ønskelig:	<0.5 $CV_w$	<0.5 $CV_t$
Minimum:	<0.75 $CV_w$	<0.75 $CV_t$

### Overdraging (carry over)

Definisjon: I hvilken grad måleresultatet påvirkes av målestørrelsen i seriens forrige prøve (egen definisjon).

For komponenter hvor det kan forekomme store forskjeller i måleresultater (narkotiske stoffer, allergener, autoantistoffer, enzymer osv.), er det meget aktuelt å måle overdraging, spesielt der det ikke er gitt opplysninger om dette fra produsenten av utstyret.

Som regel er problemstillingen at en høy prøve foran en lav prøve gir en falsk forhøyet verdi for den lave prøven.

For å teste ut dette trenger man en høy (H) og en lav (L) prøve. Disse måles i en serie etter flg. mønster: H-LI-L2-H-LI-L2-H-LI-L2... Overdraging (C) beregnes som endring i L (LI-L2) som følge av overdraging i forhold til forskjellen mellom høy og lav prøve:  $C = (LI-L2)/(H-L2)^1$

Test: I dette tilfellet er det bare presisjonen for den lave prøven som er interessant, og denne kan i forsøket beregnes fra alle parene av L2-LI (s). Kriteriet for signifikant overdraging blir:  $[M(L2) - M(LI)]\sqrt{n/s} > t_{n-1,p}$  (obs! - ensidig test, her tester vi om  $M(L2) > M(LI)$ , ikke om de er forskjellige). Her er spørsmålet om analytisk mål

viktig og trenger litt omtanke. Hvis vi maksimum kan tillate at overdraging for en lav prøve er D i komponentens enheter, trenger vi minst dette antallet (n) av L1L2-par i serien for at denne forskjellen skal bli signifikant på p nivå:

$n > [t_{n-1,p} \times s/D]^2$  eller uttrykt ved overdraging

$n > \{t_{n-1,p} \times s / (C \times [M(H) - M(L2)])\}^2$ .

Eksempel: Cannabis: Grensen mellom negativ og positiv cannabis er 20 ng/L. Det maksimale man kan tillate at en lav cannabis kan øke som følge av en høy prøve foran, er 9 ng/L. Hvis den høyeste positive som man kan forvente å få tilsendt til laboratoriet er 300 ng/L, vil dette tilsvare en overdraging på  $9/300 = 3\%$ . Man velger en negativ (f.eks. 5 ng/L) som lav prøve og en høy positiv (f.eks. 205 ng/L) som høy prøve. Innen serie standard avvik for målingen er ca 3 ng/L i det lave området (og dermed  $3\sqrt{2}$  ng/L for differansen mellom to lave). Vi velger et lavt signifikansnivå ( $p=0.01$ ) og prøver med ca 6 par ( $n-1=5$  dvs  $t_{5,0.01}=3.4$ ). Vi trenger derfor  $n > \{3.4 \times (3\sqrt{2}) / [3\%(205-5)]\}^2$  dvs mer enn 6 par for å oppnå signifikans for en overdraging på 3%.

Krav: For komponenter og prøver der det forekommer store konsentrasjonsforskjeller og hvor det er viktig å måle nøyaktig i enten høyt eller lavt område og det foreligger utilstrekkelig dokumentasjon på overdraging fra produsenten av pipetteringsutstyret, skal det gjøres forsøk for å bestemme overdraging.

### Måleområde

Definisjon: Det område for målestørrelsen hvor målinger kan utføres med angitt usikkerhet.

Måleområdet er definert av nedre (L) og øvre (H) målegrense.

Måleresultat under nedre målegrense (L) utgis som <L og over øvre målegrense (H) som >H. NB! - dette betyr ikke at den sanne verdien er <L eller >H, men at den målte verdien er det.

Vær oppmerksom på at man ved å tillate at prøver kan oppkonsentreres eller fortynnes, kan området hvor kvantitative resultater kan utgis, utvides betydelig. Nedenfor omtales måleområdet som om slike oppkonsentreringer/fortynninger ikke utføres.

## Kvantifiseringsgrenser

Nedre og øvre kvantifiseringsgrense brukes gjerne for å angi et område der målingene har en usikkerhet mindre enn en angitt verdi. Hvis den spesifiserte måleusikkerhet er  $\%CV=20$ , kalles nedre kvantifiseringsgrense også ofte "funksjonell sensitivitet". Hvis måleresultater nær 0 ikke opptrer for en komponent, skal man angi nedre målegrense som en kvantifiseringsgrense.

## Påvisningsgrense

Definisjon: Det laveste måleresultat som med angitt sikkerhet representerer en sann verdi  $>0$ .

Hvis sikkerheten (konfidensnivået) skal være minst 99%, vil påvisningsgrensen være  $s_0 \times z_{0,99} \times \sqrt{2} = s_0 \times 3.3$  der  $s_0$  er standard avvik for en 0-prøve (eller en prøve nær 0) [2b].

### Eksempel

Mål en prøve nær eller lik nedre måle grense (L), helst 0-prøve hvis det er mulig, og en nær øvre målegrense (H) i minst 10 serier. Beregn  $s_L$  og  $s_H$  (hhv. standard avvik ved målestørrelse L og H). Beregn påvisningsgrense som  $3.3s_L$ .

Velg CV for kvantifiseringsgrenser (CV<sub>k</sub>). Beregn nedre kvantifiseringsgrense som  $s_L/CV$  (CV<sub>k</sub>=0.2 gir nedre kvantifiseringsgrense på  $5 \times s_L$ ). Beregn øvre kvantifiseringsgrense som det høyeste måleresultat som kan måles med  $CV < CV_k$ . Det vil si at alle prøver med høyere måleresultater, må fortynnes og måles på nytt hvis høyere konsentrasjoner enn øvre kvantifiseringsgrense skal utgis.

## Deteksjonsgrense

I motsetning til påvisningsgrense er dette en nedre grense for måleresultater som med angitt sikkerhet ikke kan representere måleresultater fra en 0-prøve. Med samme signifikansnivå er deteksjonsgrensen det dobbelte av påvisningsgrensen.

Oftest brukes deteksjonsgrense som et mål for nedre målegrense. Om man bruker det ene eller andre er ikke så viktig, men det skal tydelig være angitt hvordan grensen er beregnet, inklusive betingelsene for målingene som er utgangspunktet for det benyttede standard avvik.

For vårt laboratorium gjelder:

Påvisningsgrense beregnes med formelen  $s_0 \times 3.3$  ( $p=0.01$ ) der  $s_0$  skal være basert på målinger av 0-prøve (eller i nærheten av 0) under betingelser for intern reproduserbarhet.

## Måleusikkerhet

Definisjon: Parameter tilknyttet et måleresultat som karakteriserer spredningen av verdier som kan tilskrives målestørrelsen [1].

## Riktighet

Man forutsetter at alle kjente systematiske feil korrigeres slik at de ikke influerer på usikkerhetsbudsjettet.

## Presisjon

Intern reproduserbarhet gir et mål for standard usikkerhet (u) for resultatet (M). Hvis det er andre bidrag til usikkerhet enn selve måleprosessen (viktige bidrag kan være prøvetaking, oppbevaring av prøve, usikkerhet i oppgitt verdi på bruksstandard, interferenser osv.), bør disse usikkerhetene adderes til usikkerheten som kan måles ved bruk av en kontrollprøve. Slike usikkerheter kan enten måles i form av standard avvik (type A usikkerheter) eller hentes fra andre kilder f.eks. litteratur, kvalifiserte gjetninger osv. (type B usikkerheter). Man kan i mange tilfeller beregne den kombinerte standard usikkerhet ved å bruke formelen:  $u^2 = u_1^2 + u_2^2 + \dots$  der  $u_i$  er de ulike usikkerhetskomponentene.

Forholdet mellom intern reproduserbarhet ( $s_I$ ) og repeterbarhet ( $s_r$ ) gir et mål på hvor feil oppstår i oppbevaring, prøvepreparering og måling. Hvis  $s_I \gg s_r$  (f.eks.  $s_I > 2s_r$ ) og disse er de viktigste bidrag til den kombinerte usikkerheten, bør man kreve en drøftelse, evt. videre validering for å finne årsaken til disse bidragene til den totale måleusikkerheten.

Komponenter som inngår i usikkerhetsbudsjettet kan være repeterbarhet ( $s_r$ ) og mellom serie variasjon:  $\sqrt{(s_I^2 - s_r^2)}$ . For en uttømmende fremstilling av bruk av usikkerhetsbudsjett henvises til [7, 8].

Området  $M \pm u$  angir et område som (i gjennomsnitt ved mange forsøk) med ca 68% sikkerhet inneholder den sanne verdi. Det vanligste er imidlertid å angi et område med større sikker-

het, oftest 95%. Dette gjøres ved å bruke såkalt ekspandert usikkerhet  $U=ku$  der  $k$  kalles dekningsfaktoren ( $k=2$  når sikkerheten er 95%).

### Biologisk variasjon

Dette er en viktig parameter innen medisinsk laboratorievirksomhet. Intraindividuell biologisk variasjon ( $S_w$ ) kan benyttes for å sette analytiske mål for en målemetode. Setter man standard usikkerhet ( $u$ ) inn i stedet for analytisk standard avvik ( $s_a$ ), kan man gruppere kvaliteten som under «Intern reproducerbarhet» for «Monitorering»: Optimalt:  $u < 0.25S_w$ , ønskelig:  $u < 0.50S_w$ , minimum:  $u < 0.75S_w$ .

Fordi kombinert usikkerhet beregnes ved rot av kvadratsum av usikkerhetskomentene, vil det bare være noen få som virkelig er viktige (men det kan være svært forskjellige komponenter som er viktige i ulike områder av måleområdet).

Krav: Måleusikkerheten skal angis ved en beregning som antydnet ovenfor der antatt viktige usikkerhetskomentene inngår. Alle usikkerhetskomentene  $< S_w/3$  kan betraktes som uvesentlige (men bør tas med hvis de er de største). Kommenter usikkerhet i hele måleområdet, spesielt ved viktige nivåer.

Takk til Heidi Steensland som har kommet med nyttige innspill i forbindelse med utarbeidelse av manuskriptet.

### Litteratur

1. VIM: «International vocabulary of basic and general terms in metrology», 1993, (BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML)
- 2a. "Ackreditering av laboratorier inom värksamhetsområdet laboratoriemedicin, klinisk kemi. Vägledning." utgitt av SWEDAC sept. 95.
- 2b. Metodeevaluering og kvalitetsstyring av kemiske måleprosesser. Poulsen et al., Arbejdsmiljøinstituttet, København 1993.

3. Mal for utprøving av instrumenter til PTH. Utkast mars 1994.
4. General strategies to set quality specifications for reliability performance characteristics. C.G. Fraser. Scand J Clin Lab Invest, 1999; 59; 487-490.
5. Current databases on biological variation: pros, cons and progress. C. Ricos et. al. Scand J Clin Lab Invest, 1999; 59; 491-500.
6. Kvalitetsledelse og kvalitetssikring. Terminologi. NS ISO 8402 2. utgave oktober 1994.
7. Rapport om måleusikkerhet for kjemisk analyse. Norsk Akkreditering . 2000.
8. Quantifying Uncertainty in Analytical Measurement. Second Edition. Draft EURACHEM/CITAC Guide: June 1999.
9. Result, error and uncertainty. R. Dybkaer, Scand J Clin Lab Invest 1995; 55, 97-118
10. ISO 3534-1 Statistics - Vocabulary and symbols. Part 1: Probability and general statistical terms. 1993.

1. *Utledning av formelen: Ved overdraging kan man tenke seg at brøkdelen  $D$  av volumet fra en prøve ( $L$ ) erstattes med samme volum fra forrige prøve ( $H$ ). Konsentrasjonen for prøven  $L1$  i serien over blir da:  $L1 = DH + (1-D)L2$  som innsatt i formelen for overdraging gir  $C = (L1 - L2)/(H - L2) = [DH + (1 - D)L2 - L2]/(H - L2) = D$ , dvs overdraging er den brøkdelen av komponenten i forrige prøve som man finner igjen i denne prøven.*