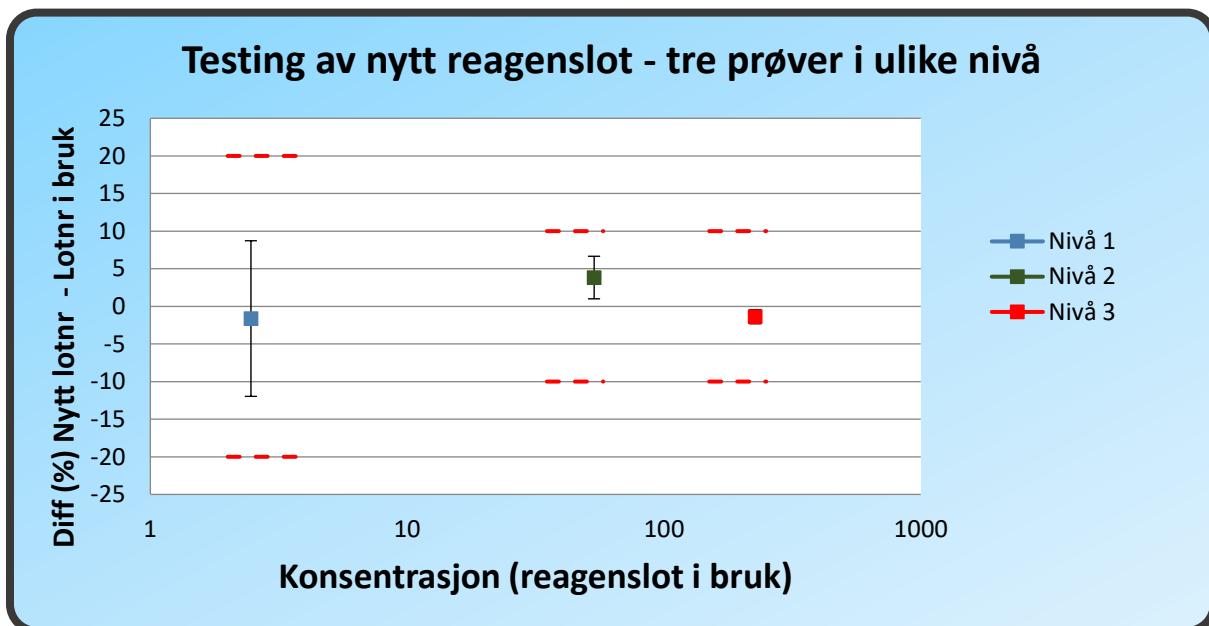


# Veileder for testing av reagens lot-til-lot-variasjon



24.09.2020 12:55:29

## Innhold

<b>1. Definisjoner</b> .....	3
<b>2. Bakgrunn</b> .....	3
<b>2.1. Når bør man teste for lot-til-lot variasjon mellom ulike reagens-lot?</b> .....	4
<b>3. Metodikk</b> .....	5
<b>3.1. Anbefalinger fra CLSI</b> .....	5
<b>3.2. Gruppens valgte metodikk</b> .....	5
<b>3.3. Valg av pasientprøver</b> .....	6
<b>3.4. Antall replikater</b> .....	6
<b>3.5. Valg av akseptgrenser</b> .....	6
<b>4. Praktisk utførelse for laboratoriene</b> .....	11
<b>5. Vurdering av resultater</b> .....	11
<b>5.1. Hva kan man gjøre når akseptgrensene overskrides?</b> .....	13
<b>6. Innsending av resultater til Noklus i prosjektperioden</b> .....	13
<b>7. Referanser</b> .....	15

## 1. Definisjoner

Akseptgrenser: grense for maksimal akseptabel prosent differanse mellom gjennomsnitt av målinger med reagens-lot i bruk og gjennomsnitt av målinger av samme prøvemateriale med ny reagens-lot. I modellen som er valgt i dette prosjektet er det et krav at differansens 90%-konfidensintervall også skal være innenfor akseptgrensene.

Limit of detection (LoD): den laveste konsentrasjonen av en analytt som kan påvises.

Reagens-lot: en spesifikk produksjon av et reagens.

«State of the art»: det beste på området beregnet for rutinedrift i medisinske laboratorier.

## 2. Bakgrunn

Tema for workshop på Fagmøte for sykehus- og private laboratorier våren 2018 var lot-til-lot variasjon. Under og etter møtet ble det tydelig at mange laboratorieansatte ønsket en anbefaling for når man tester for lot-til-lot reagensvariasjon og et forslag til hvordan dette kan gjøres.

Samtidig tok ansatte ved Oslo universitetssykehus (OUS) initiativ til et prosjekt med lot-til-lot variasjon som tema. OUS erfarer tidvis at det foreligger relativt store nivåforskjeller mellom lot og foreslår at brukere ved norske laboratorier begynner å dele informasjon og resultater fra testing av nye reagens-lot for å øke kunnskapen innen dette feltet.

Basert på disse innspillene opprettet Noklus et prosjekt og etablerte en prosjektgruppe.

Prosjektgruppen ble etablert september 2018 og består av:

- Erik Amundsen (overlege, Oslo universitetssykehus)
- Joakim Eikeland (overlege, Oslo universitetssykehus)
- Kristine Solem (kvalitetsleder, St.Olavs hospital)
- Ann Helen Kristoffersen (overlege, Haukeland universitetssjukehus)
- Gro Gidske (produksjonsleder for kontrollmateriale / rekrutteringsstilling, Noklus)
- Eva Rønneseth (avdelingsingeniør /rekrutteringsstilling, Noklus)
- Anne Vegard Stavelin (forsker, Noklus)
- Sverre Sandberg (leder, Noklus)
- Anne Elisabeth Solsvik (kvalitetskoordinator/ rekrutteringsstilling, Noklus, (prosjektleder))

Prosjektgruppens arbeid er basert på eksisterende retningslinjer/ guidelines, litteratursøk og egne erfaringer.

Mandat er som følger:

- Utarbeide en protokoll som er praktisk gjennomførbar både for små og store laboratorier med tilhørende verktøy for å undersøke lot-til-lot variasjon mellom reagens-lot i bruk og ny reagens-lot for utvalgte analyser
- Invitere norske medisinske laboratorier til å delta i et prosjekt der denne protokollen følges. I en prosjektperiode på to år sendes alle resultater fra testing av nye reagens-lot til prosjektgruppen. I prosjektperioden kan deltakerne kontakte prosjektgruppen ved prosjektleder for å få informasjon om resultater delt fra andre deltakere.
- Utrede om det er hensiktsmessig at Noklus etablerer et nasjonalt overvåkingssystem for lot-forskjeller basert på erfaring fra prosjektet
- Publisere resultater og erfaringer fra prosjektet

### 2.1. Når bør man teste for lot-til-lot variasjon mellom ulike reagens-lot?

En ideell analysemetode har ingen lot-til-lot variasjon mellom hverken ulike reagens-lot, kalibrator-lot eller andre forbruksvarer som påvirker analysekvaliteten. Om ett nytt reagens-lot skal testes for lot-til-lot variasjon avhenger dels av erfaringer med lot variasjoner for aktuell metode, hvilken betydning slike endringer kan ha for tolkning av analysesvar og tilgjengelige ressurser i laboratoriet (personell og reagenskostnader). Kjennskap til lot-til-lot variasjon er spesielt viktig når pasienter følges over lang tid, som for eksempel for HbA1c og enkelte tumormarkører, og for analyser der man har viktige kliniske beslutningsgrenser. For hver enkelt analyse bør det gjøres en vurdering av hvilke konsekvenser et eventuelt nivåskifte har for enkeltpasienter og for viktige pasientgrupper.

Det er vanligvis ikke nødvendig å teste nye reagens-lot med tanke på lot-til-lot variasjon på flere instrument i samme laboratorium (1). Laboratorier med sammenlignbare transportkjeder, oppbevaringsforhold og kalibreringsrutiner kan også vurdere om et samarbeid er hensiktsmessig.

Selv om man vurderer at det ikke er sannsynlig med klinisk viktige lot-forskjeller og at man derfor velger å ikke teste for slike forskjeller, må alltid intern analytisk kvalitetskontroll analyseres og vurderes. Dersom resultater fra intern analytisk kvalitetskontroll indikerer en endring i nivå bør man vurdere å gjøre en videre testing for å finne viktige nivåforskjeller. Vær klar over at det kan være en signifikant viktig lot-forskjell som påvirker pasientprøver selv om dette ikke fanges opp med intern kvalitetskontroll (matrix-effekter) og omvendt (2). For mange analyser er det derfor nyttig å overvåke analysens nivå ved å følge pasientmedian over tid.

### 3. Metodikk

#### 3.1. Anbefalinger fra CLSI

Clinical and Laboratory Standard Institute (CLSI) har utgitt en veileder som gir en anbefaling for når og hvordan laboratorier bør evaluere nye reagens-lot (1). Basert på en valgt kritisk differanse og analysens upresisitet bestemmes en avvisningsgrense. Avvisningsgrense kan leses av i tabeller avhengig av ønsket statistisk styrke som et multiplum av kritisk differanse. Flere pasientprøver i et eller flere valgte nivå analyseres med reagens-lot i bruk og ny reagens-lot, og gjennomsnittlig differanse må være under avvisningsgrensen. Antall paralleller som må analyseres bestemmes ut fra metodens upresisitet og valgt statistisk styrke.

Ulempen med denne metoden er at den er for arbeidskrevende og at det kreves mange analyseringer. Metoden tar heller ikke hensyn til en kumulert differanse mellom flere reagens-lot. Det kan også være krevende for enkelte laboratorier å bestemme kritisk differanse. Gruppen har derfor valgt en annen metodikk.

#### 3.2. Gruppens valgte metodikk

Det er ulike måter å teste for lot-til-lot variasjon på og mange hensyn å ta (3). En av de største utfordringene er å oppnå akseptabel statistisk styrke. Andre viktige momenter er valg av akseptgrenser og å kunne påvise kumulative endringer.

Prosjektgruppen har vurdert to alternative metoder som vi vet brukes til å teste for lot-forskjeller:

- 1) Flere (10-20) pasientprøver med resultater fordelt i måleområdet analyseres med reagens-lot i bruk og med ny reagens-lot med påfølgende metodesammenligning der man undersøker for systematisk avvik mellom ny reagens-lot og reagens-lot i bruk.
- 2) En til tre pasientprøver i valgte nivå analyseres et visst antall ganger. Middelerdi til hver enkelt pasientprøve beregnes. Middelerdi til en pasientprøve analysert med reagens-lot i bruk sammenlignes med middelerdi til den samme pasientprøven analysert med ny reagens-lot. Differanse med tilhørende konfidensintervall vurderes mot en valgt akseptgrense.

Ved å analysere flere replikater av samme pasientprøve reduseres usikkerheten til middelerdien. Prosjektgruppen har valgt fremgangsmåten der en til tre pasientprøver analyseres flere ganger som den mest hensiktsmessige ettersom denne metoden antas å være gjennomførbart for de fleste norske medisinske laboratorier.

Vårt forslag til fremgangsmåte gjelder primært for reagens, men kan også brukes ved skifte av kalibrator-lot og andre komponenter som kan gi et nivåskifte i analyseresultatene. Denne veilederen gjelder bare for kvantitative analyser.

### 3.3. Valg av pasientprøver

Pasientprøver fra rutinedrift benyttes til testing av nye reagens-lot. Interne analytiske kvalitetskontroller kan ha en annen matriks enn pasientprøver og kan gi feil konklusjon (4). Det velges en pasientprøve i hvert nivå (se Tabell 1). Dersom en pasientprøve ikke gir tilstrekkelig prøvevolum kan man blande to eller flere pasientprøver til en pool. Velges denne løsningen må det kommenteres i protokollen.

I modellen velges pasientprøver med analyseresultat nær kliniske beslutningsgrenser og / eller nær øvre og nedre referansegrense. I tillegg må intervall for hvert nivå ikke være så snevert at man ikke finner egnede pasientprøver i egen rutinedrift.

En mulig feilkilde er at enkeltprøver kan oppføre seg annerledes enn majoriteten av pasientprøvene. Dersom man mistenker at valg av prøve ikke er representativ bør det undersøkes med en annen pasientprøve(-r) i samme nivå.

D-dimer har en viktig beslutningsgrense, og det er derfor anbefalt testing i kun ett nivå, men med pooling av flere pasientprøver for å få nok volum til 12 replikater (se Tabell 1).

I lave nivå der laboratoriet ikke gir ut svar på pasientprøver med kvantitative resultater er det ikke nødvendig å teste for lot-variasjon i det lave området, f.eks. dersom man rapporterer troponiner under 10 ng/L som < 10 ng/L er det tilstrekkelig å teste for lot-forskjeller i de to andre nivåene, for troponin T i to nivå fra 10,5 til 17,5 mg/L og fra 150 til 250 ng/L.

### 3.4. Antall replikater

Antall replikater må være tilstrekkelig til at klinisk viktige differanser oppdages med høy sannsynlighet. Det er vektlagt at modellen skal være gjennomførbar i alle norske medisinske laboratorier og prosjektgruppen foreslår derfor et minimum antall replikater (se Tabell 1). Vi har valgt mellom 6 og 12 replikater i hvert nivå, valget ses i sammenheng med valgt akseptgrense, sannsynlighet for alarm når akseptgrensene overskrides og sannsynlighet for falske feilmeldinger (5).

### 3.5. Valg av akseptgrenser

En viktig del av prosjektgruppens arbeid har vært å bestemme hva som skal anbefales brukt som akseptabel differanse mellom to reagens-lot. Prosjektgruppen har spurt norske medisinske laboratorier hva fagmiljøet mener er hensiktsmessige kvalitetskrav ved lot-skifter. Spørsmålet ble sendt til i overkant av 100 avdelingsledere og kontaktpersoner i norske sykehus og ved større private laboratorier som dekker fagområdet medisinsk biokjemi. Det ble sendt til flere personer ved hvert sykehus, og det ble oppfordret til at aktuelle fagpersoner i eget laboratorium skulle svare. Laboratoriene ble bedt om å svare for de analysene laboratoriet utfører.

Til sammen fikk vi 18 svar. Laboratoriene ble bedt om å svare på følgende for hver av de valgte analysene (HbA1c, PSA, D-dimer og troponin T eller I):

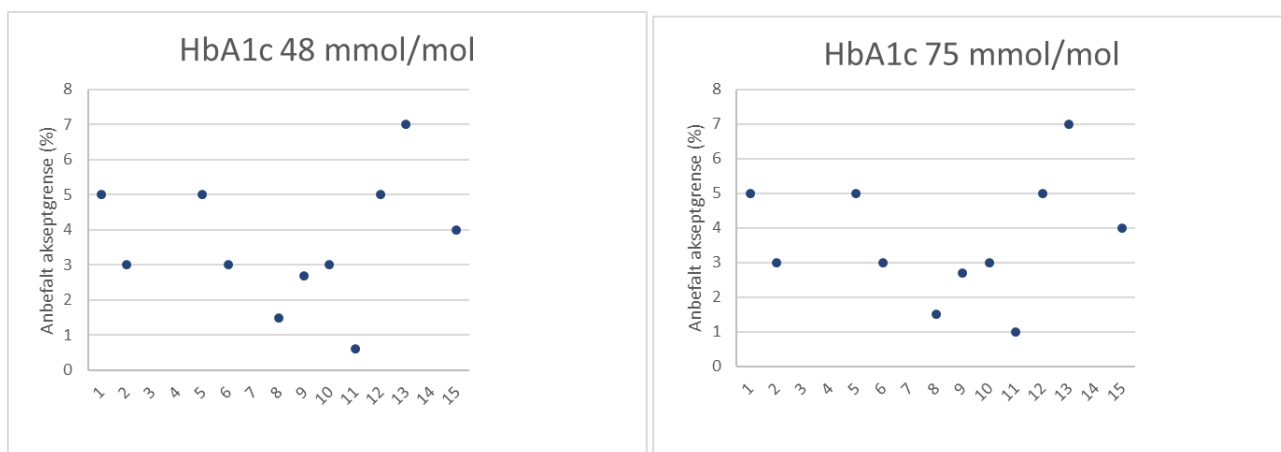
*Forutsatt at reagens-lot i bruk måler pasientresultater i riktig nivå; Hva mener du akseptgrensen bør være for de ulike analysene, dvs. hvor stort maksimalt systematisk avvik (i prosent) vil du akseptere mellom reagens-lot i bruk og ny reagens-lot for at ny lot kan tas i bruk (uten vesentlige tiltak som faktorisering eller endring av referansegrenser)?*

Flere laboratorier oppga at de ikke tester for reagens lot-forskjeller og har følgelig heller ikke svart på hvilke grenser man bør bruke.

### Resultater fra spørsmål til laboratoriene om akseptgrenser:

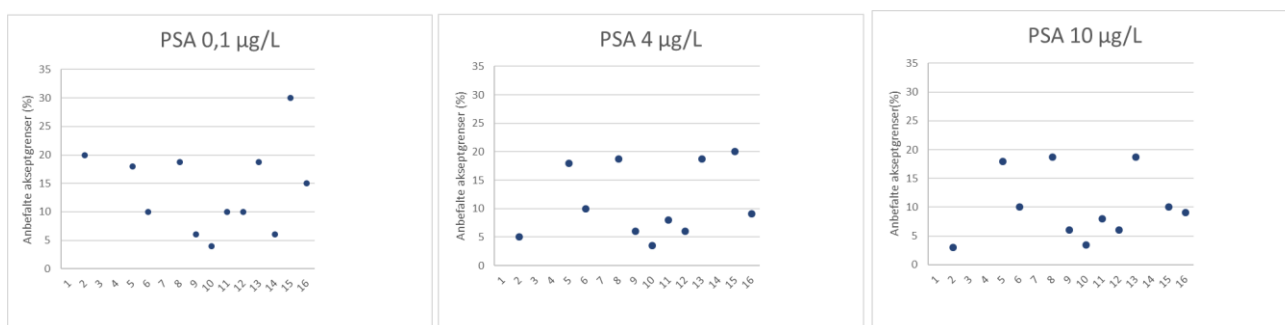
#### For HbA1c:

**Figur 1.** Av de 15 som har svart på spørsmålet har 11 anbefalt en akseptgrense for HbA1c for begge nivå. De fleste mener at en forskjell på mellom 3 og 5% er akseptabel for begge nivå.



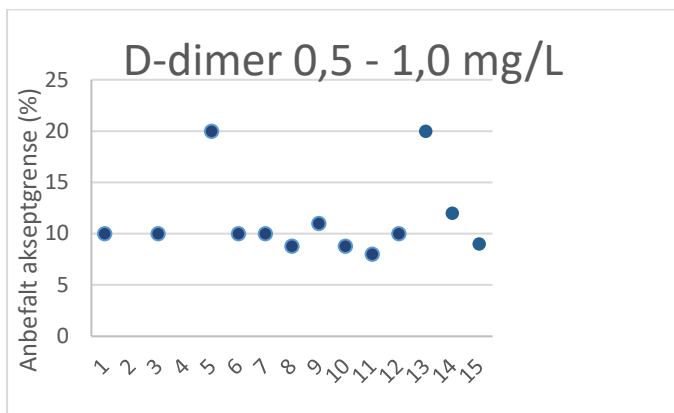
#### For PSA:

**Figur 2.** Av de 16 som har svart på spørsmålet har 11 anbefalt en akseptgrense for PSA for alle tre nivå. I det lave nivå svarer 6 av 12 at akseptgrensen bør være > 10%, rundt øvre referansegrense svarer 4 av 11 at akseptgrensen bør være > 10% og i det høye nivået svarer 3 av 11 at akseptgrensen bør være > 10%.



**For D-dimer:**

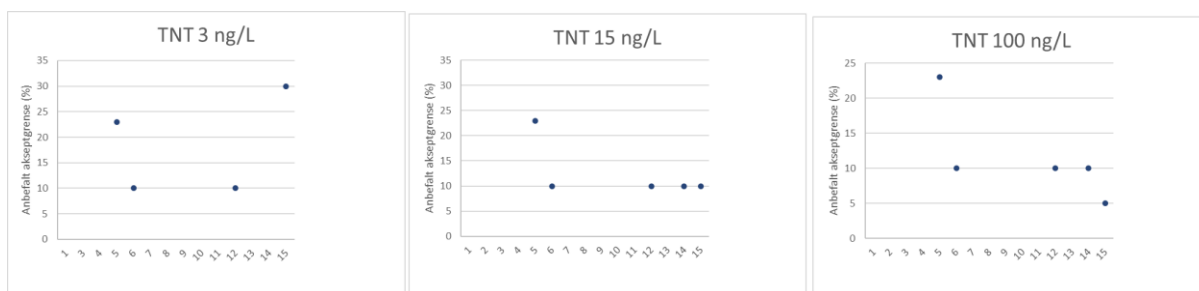
**Figur 3.** Av de 15 som har svart på spørsmålet har 13 anbefalt en akseptgrense, de fleste anbefaler en akseptgrense på 10%.



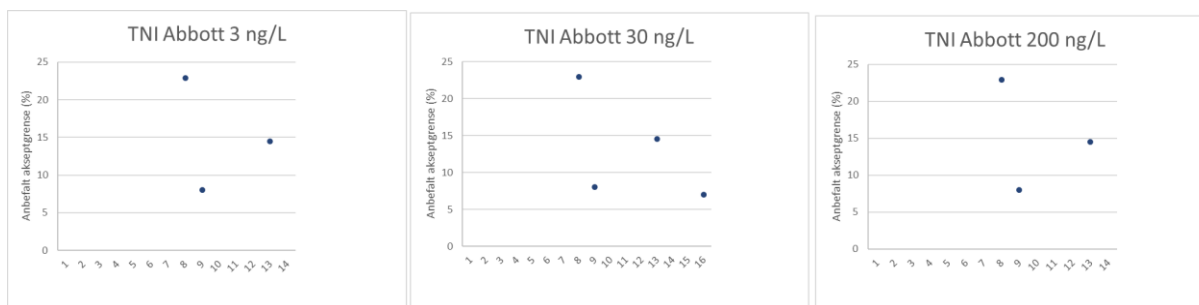
**For troponin analysene:**

**Figur 4.** For alle troponin metodene anbefaler de fleste at akseptgrensen for lot-forskjeller i lavt nivå er mellom 10 og 20%, i de to høyeste nivåene anbefaler flertallet en akseptgrense for lot-forskjeller på 10%.

**Troponin T (Roche Cobas)**

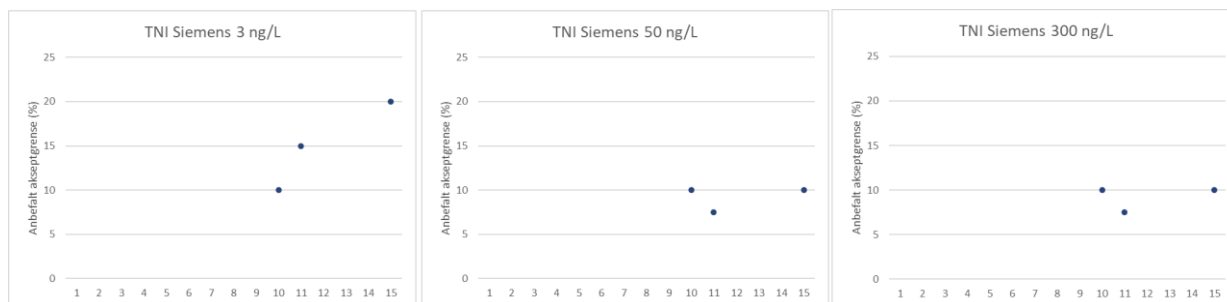


**Troponin I (Abbott)**





## Troponin I (Siemens)



Krav til maksimal differanse mellom to lot bør primært baseres på klinisk bruk av analysen, biologisk variasjon for analytten eller «state of the art» med utgangspunkt i de beste metodene tilgjengelig for rutinedrift i medisinske laboratorier (6).

De foreslåtte akseptgrensene tar utgangspunkt i prosjektgruppens kjennskap til analysemetoder i bruk i norske laboratorier. I tillegg er akseptgrensene vurdert og justert ut fra kvalitetskrav satt av fagmiljø, bruk av analysene i norske laboratorier, biologisk variasjon og laboratorienes svar på spørsmål om akseptgrenser beskrevet over. Et viktig spørsmål er hva som skal være innenfor akseptgrensene. Skal bare gjennomsnittsverdien til differansen eller skal også et konfidensintervall, være innenfor akseptgrensene? Prosjektgruppen har valgt at kravet skal være at også 90% konfidensintervall til gjennomsnittsdifferansen skal være innenfor akseptgrensene. Dersom kvalitetskrav settes til mindre enn  $3 \cdot SD_a$  (der  $SD_a$  er analysens repeterbarhet) vil 6 replikater gi en del falske alarmer (5). For å unngå for mange falske alarmer er derfor antall replikater økt til 12 for D-dimer analysen. Dersom D-dimer metoden laboratoriet benytter har et standardavvik (repeterbarhet) som er mindre enn  $1/3$  av akseptgrensen er det tilstrekkelig med 6 replikater.

For enkelte analyser kan det være hensiktsmessig å benytte konsentrasjonsavhengige kvalitetskrav. I tillegg er det ønskelig at kumulative endringer overvåkes og tas med i de vurderinger som gjøres.

De valgte akseptgrensene er vist i Tabell 1. I tilhørende Excel regneark er det også gitt mulighet for at deltakerne kan bruke egne akseptgrenser.

**Tabell 1.** Valgte analyser, anbefalt intervall for pasientprøvene, antall anbefalte replikater, akseptgrenser og kommentarer / referanser

Analyse	Velg pasientprøver mellom:		Minimum antall replikater:	Maksimalt systematisk avvik (%):	Kommentar / bakgrunn for valg av nivå:
PSA (prøve 1)	0,1 µg/L	0,5 µg/L	6	30	Residiv, evt. gjenværende kreftvev etter prostatektomi
PSA (prøve 2)	3,0 µg/L	5,0 µg/L	6	20	Nær øvre referansegrense, RCPA krav opp til 5,0 µg/L ± 0,4 µg/L
PSA (prøve 3)	10,0 µg/L	20,0 µg/L	6	20	Posttestsannsynlighet for kreft begynner å bli større enn pretestsannsynlighet RCPA <sup>a</sup> krav PSA > 5,0 µg/L: 8%
HbA1c (prøve 1)	40 mmol/mol	50 mmol/mol	6	5	Nær diagnostisk grense
HbA1c (prøve 2)	70 mmol/mol	80 mmol/mol	6	5	Økt risiko for komplikasjoner
Hs-Troponin T (prøve 1) (Roche)	5 ng/L	10 ng/L	6	20	Tidlig rule-out (LoD <sup>b</sup> -2*LoD)
Hs-Troponin T (prøve 2) (Roche)	10,5 ng/L	17,5 ng/L	6	10	99-percentil +/- ≈ 25%
Hs-Troponin T (prøve 3) (Roche)	150 ng/L	250 ng/L	6	10	Høye verdier, 200+/-50
Hs-Troponin I (prøve 1) (Abbott)	2 ng/L	5 ng/L	6	20	Tidlig rule-out (LoD <sup>b</sup> -2*LoD)
Hs-Troponin I (prøve 2) (Abbott)	20,0 ng/L	32,0 ng/L	6	10	99-percentil +/- ≈ 25%
Hs-Troponin I (prøve 3) (Abbott)	150 ng/L	250 mg/L	6	10	Høye verdier, 200+/-50
Hs-Troponin I (prøve 1) (Siemens)	2 ng/L	5 ng/L	6	20	Tidlig rule-out (LoD <sup>b</sup> -2*LoD)
Hs-Troponin I (prøve 2) (Siemens)	35,0 ng/L	58,0 ng/L	6	10	99-percentil +/- ≈ 25%
Hs-Troponin I (prøve 3) (Siemens)	150 ng/L	250 ng/L	6	10	Høye verdier, 200+/-50
D-dimer (pasientpool)	0,4 mg/L	1,0 mg/L	12	20	Nær beslutningsgrense på 0,5 mg/L

<sup>a</sup>RCPA = Royal College of Pathologists of Australasia, hentet fra <http://www.rcpaqap.com.au/docs/2014/chempath/ALP.pdf>

<sup>b</sup>Høyeste oppgitte LoD fra hver leverandør for reagens og instrument som brukes i Norge.

#### 4. Praktisk utførelse for laboratoriene

Testing av ny reagens-lot må gjøres mens man har nok antall tester igjen av eksisterende reagens-lot. Reagens-lot i bruk må være kalibrert i henhold til gjeldende anbefalinger og intern analytisk kvalitetskontroll må være innenfor egne akseptgrenser.

Prosjektgruppen anbefaler følgende prosedyre:

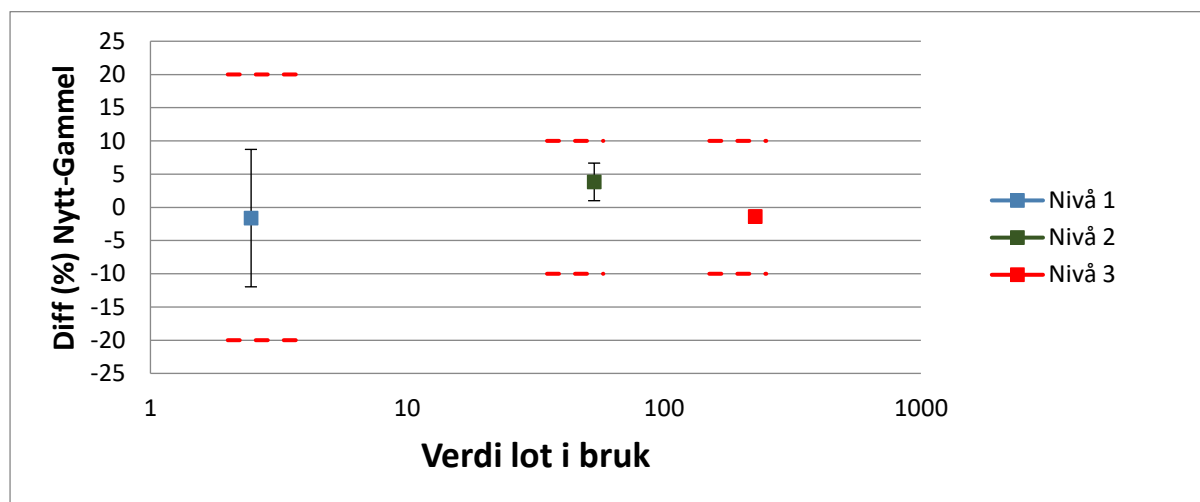
- Kalibrer ny reagens-lot.
- Velg pasientprøver i riktig nivå (se Tabell 1). Unngå pasientprøver som har synlig hemolyse, høy bilirubin eller lipemi (turbid prøve). Sjekk at det er nok prøvemateriale til å analysere anbefalt antall replikater. Dersom en pasientprøve ikke gir nok materiale til anbefalt antall analyseringer blandes flere pasientprøver godt før materialet fordeles. Kommenter i regnearket dersom to eller flere pasientprøver er blandet.
- Analyser pasientprøver med eksisterende reagens-lot og nytt reagens-lot. Analyser prøvene så tett som praktisk mulig i tid og unngå unødvendig oppbevaring av prøvene uten kork.
- Fyll inn nødvendige data og resultater i tilsendt Excel-mal. Dersom nye reagens-lot testes på flere instrumenter skal det bruket ett Excel-ark til hvert instrument.
- Vurder resultatene.
  - Dersom differansen mellom analysering med eksisterende og nytt reagens-lot inkludert 90%-konfidensintervall er innenfor krav til systematisk avvik kan ny reagens-lot godkjennes og tas i bruk. Analyser intern kvalitetskontroll og vurder disse.
- Send inn utfylte Excel-mal som vedlegg til epost til prosjektgruppen ved prosjektleder.

#### 5. Vurdering av resultater

Fyll inn data og resultater i Excel-mal laget av prosjektgruppen. Middelerdi, standardavvik, variasjonskoeffisient og standardfeil for middelerdien beregnes for hver prøve og hvert nivå. Differansen mellom gjennomsnittlig analyseresultat med reagens-lot i bruk og

gjennomsnitt med ny reagens-lot vises i diagram. I tillegg vises 90% konfidensintervall for differansen.

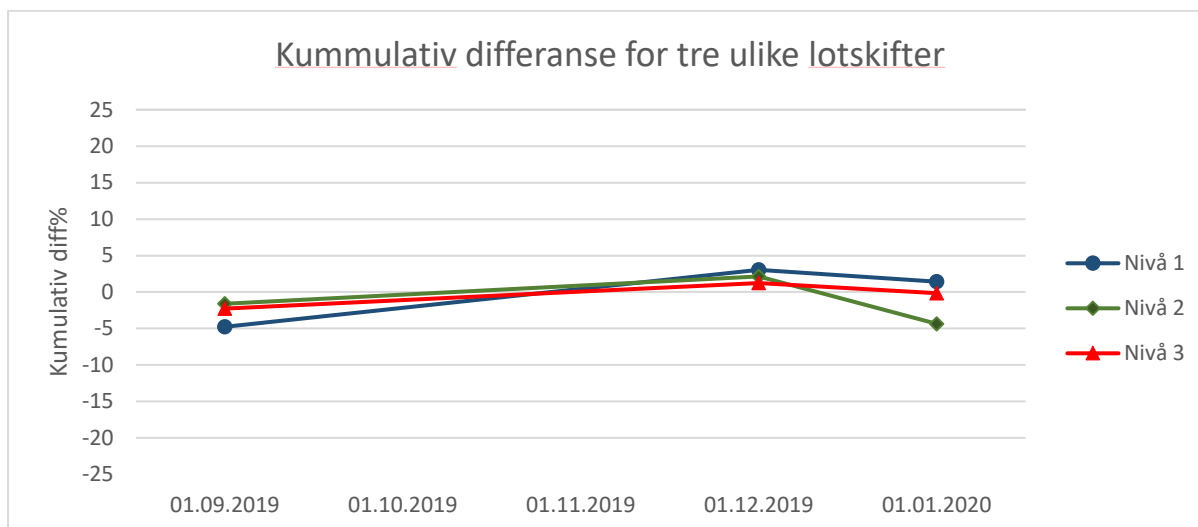
**Figur 5.** Eksempel på ett lot-skifte. Differansen mellom ny reagens-lot og reagens-lot i bruk inklusiv hele konfidensintervallet er innenfor akseptgrensene (stiplet rød linje) for alle tre prøvene, og den nye reagens-lotten kan dermed godkjennes.



*Kvalitetskrav: differanse mellom gjennomsnitt av målinger med reagens-lot i bruk og gjennomsnitt av målinger av samme prøvemateriale med ny reagens-lot skal være innenfor akseptgrense, hele 90% konfidensintervall skal være innenfor akseptgrensene.*

Vurder også alltid diagram for kumulativ differanse.

**Figur 6.** Eksempel på kumulativ differanse for tre ulike lotskifter. I tillegg til å vurdere hvert enkelt lotskifte gir plottet under mulighet for å vurdere flere reagensskifter samlet.



### 5.1. Hva kan man gjøre når akseptgrensene overskrides?

Det kan være flere årsaker til at akseptgrensene overskrides. Dersom bare en av prøvene har en differanse mellom ny reagens-lot og reagens-lot i bruk der gjennomsnittet eller deler av konfidensintervallet er utenfor akseptområdet kan forsøket utvides med en eller flere nye pasientprøver. Årsaken kan også være kalibreringsproblemer med reagens i bruk og / eller nytt reagens, eller reagens-lot i bruk kan ha driftet siden forrige kalibrering. En mulig løsning kan være å recalibrere reagens-lot i bruk og nytt reagens, sjekk da eventuelle pipetter som brukes dersom kalibrator er tørrstoff som løses i en væske.

En annen mulighet er at vi har valgt akseptgrenser som er for strenge i forhold til det produsentene leverer, vi har ikke etterspurt produsentenes oppgitte maksimale lot-til-lot reagensvariasjon.

Dersom to eller flere av de valgte pasientprøvene viser en differanse utenfor akseptgrensene som vurderes som en reell og uakseptabel lot-til-lot variasjon må produsent av reagens og analyseinstrument kontaktes. Følg laboratoriets interne rutiner. Analysen kan ikke faktoreres kun på bakgrunn av testing for lot-til-lot variasjon med vår modell, dersom man ønsker å faktorisere bør det gjøres en regresjonsanalyse med flere prøver.

## 6. Innsending av resultater til Noklus i prosjektperioden

Alle som deltar i prosjektet sender i prosjektperioden på to år alle gjennomførte forsøk til prosjektleder. I prosjektperioden kan deltakerne kontakte Noklus for å få informasjon om resultater delt fra andre deltakere.

Når prosjektperioden er over skal prosjektgruppen innhente erfaringer fra deltakerne. Sammen med prosjektgruppens egne erfaringer besluttes det om det skal etableres et tilbud til norske laboratorier om å delta i et nasjonalt reagens-lot overvåkingsprogram som ledd i kvalitetssikring av norske laboratorieanalyser.

## 7. Referanser

1. Institute CaLS. EP26-A User Evaluation of Between-Reagent Lot Variation; Approved Guideline. CLSI; 2013.
2. Chen X, Wang J, Zhang W, Xie E, Zhang B, Xu HG. Failure of internal quality control in detecting significant reagent lot shift in serum creatinine measurement. *J Clin Lab Anal.* 2019:e22991.
3. Thompson S, Chesher D. Lot-to-Lot Variation. *Clin Biochem Rev.* 2018;39(2):51-60.
4. Miller WG, Erek A, Cunningham TD, Oladipo O, Scott MG, Johnson RE. Commutability limitations influence quality control results with different reagent lots. *Clin Chem.* 2011;57(1):76-83.
5. Asberg A, Solem KB, Mikkelsen G. Determining sample size when assessing mean equivalence. *Scand J Clin Lab Invest.* 2014;74(8):713-5.
6. Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: Consensus Statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med.* 2015;53(6):833-5.