

LEIDRAAD INTERNE KWALITEITSCONTROLE

(herziening Leidraad Analytische Kwaliteitscontrole
met behulp van Six Sigma (2015))

INITIATIEF

Nederlandse Vereniging voor Klinische Chemie en Laboratoriumgeneeskunde (NVKC)

FINANCIERING

Dit project werd gefinancierd uit de Stichting Kwaliteitsgelden Medisch Specialisten (SKMS 2
projectnummer: Z69118127)



Colofon

LEIDRAAD INTERNE KWALITEITSCONTROLE © 2025

Nederlandse Vereniging voor Klinische Chemie en Laboratoriumgeneeskunde

Postbus 5056, 3502 JB Utrecht

Telefoon: 088 505 34 20

E-mail: kwaliteit@nvkc.nl

Website: <https://www.nvkc.nl/>

Inhoudsopgave

Hoofdstuk 1 Verantwoording	5
1.1 Samenstelling werkgroep	5
1.2 Belangenverklaringen	5
Hoofdstuk 2 Overzicht van de aanbevelingen	8
Hoofdstuk 3 Samenvatting	10
Hoofdstuk 4 Aanleiding en doel	11
4.1 Aanleiding	11
4.2 Doel	11
Hoofdstuk 5 Toepassingsgebied en beperkingen (scope)	12
Hoofdstuk 6 Samenhang met andere documenten	13
Hoofdstuk 7 Theoretische achtergrond	14
7.1 Algemeen	14
7.1.1 Risico-gebaseerde inrichting van de interne analytische kwaliteitscontrole	15
7.2 Analytische kwaliteitscontrole	16
7.2.1 Externe kwaliteitscontrole (EQA)	16
7.2.2 Interne kwaliteitscontrole (IQC)	16
7.3 Analytische kwaliteitseisen en selectie van tolerantiegrenzen	17
7.3.1 Analytische prestatiespecificatie (APS)	17
7.3.2 APS op basis van totale meetfout (TEa) versus meetonzekerheid (MU)	19
7.4 Vaststellen van de imprecisie	20
7.4.1 Individueel versus virtueel analyzer model	21
7.4.2 Grondslag SD-waarde	22
7.5 Westgard QC	22
7.6 Six Sigma	23
7.6.1 Inleiding	23
7.6.2 Tekortkomingen van TEa-gebaseerde sigmascores en alternatieven	24
7.6.3 Toepassingsgebieden Six Sigma, en beperkingen hierin	25
7.7 Risico-gebaseerde QC	26
7.8 <i>Patient-Based Real-Time Quality Control (PBRTQC)</i>	26
7.9 Borgen vergelijkbaarheid tussen analyzersystemen (binnen laboratoriumorganisatie)	27

Hoofdstuk 8 Praktische inrichting analytische kwaliteitscontrole	29
8.1 Stroomschema: leeswijzer	29
8.2 Keuze bewakingsmodel per test	29
8.3 Aard interne controlemateriaal	32
8.4 Berekening van de spreiding	32
8.5 Selectie van periodiciteit	33
8.6 Selectie van Sigma (<i>multi</i>)rules	33
8.7 <i>Bracketed Risk-Based Quality Control</i>	33
8.8 <i>Patient-Based Real-Time Quality Control</i> (lopende patiëntengemiddelden)	34
8.9 Conventionele methode (2SD-gebaseerd)	34
8.10 Risicoanalyse	34
8.10.1 Risico inschatting algemeen	34
8.10.2 Bepalen van het restrisico	35
8.11 Borgen vergelijkbaarheid tussen analyzersystemen (binnen laboratoriumorganisatie)	35
Hoofdstuk 9 Referenties	38
Hoofdstuk 10 Begrippen en afkortingen	43
Bijlage A Samenhang met NEN-ISO 15189:2023	47
Bijlage B Analytische spreiding en tolerantiegrenzen	48
B.1 Verschillende modellen voor tolerantiegrenzen	48
B.2 Tolerantiegrenzen en prestatie-eisen volgens de IVDR	50
Bijlage C APS en medicatiebewaking	51
Bijlage D Meetonzekerheid versus <i>total error</i> concept	53
Bijlage E Westgard QC	54
Bijlage F Het Six Sigma concept	55
F.1 Algemeen	55
F.2 Toepassing van Six Sigma in het klinisch-chemisch laboratorium	55
F.3 Bepalen van de sigmascore	56
F.4 Rekenvoorbeelden <i>Sigma Metrics</i> QC-procedures	56
F.4.1 Verificatie of validatie van een test	58
F.4.2 Toepassen van QC-regels op basis van berekende sigmascore	59
F.4.3 Niet-robuste tests met sigmascore < 3	59
Bijlage G Six Sigma gebaseerde kwaliteitscontrole met het <i>Moving Range</i> model	61
G.1 Berekeningen	62
G.1.1 Verschillen opeenvolgende waarden (<i>Moving Range</i>)	62

G.1.2 Individuele analyzer	63
G.1.3 Virtuele analyzer	64
G.1.4 Onderlinge samenhang van de verschillende spreidingsbronnen	66
G.1.5 <i>Moving Average</i> kaarten	66
Bijlage H <i>Patient-Based Real-Time Quality Control</i>	69
Bijlage I Stappenplan correctieve acties bij afwijkingen interne QC	72

Hoofdstuk 1 Verantwoording

Het initiatief van deze leidraad ligt bij de Nederlandse Vereniging voor Klinische Chemie en Laboratoriumgeneeskunde (NVKC). Voor het opstellen van deze leidraad is medio 2021 een werkgroep ingesteld. De leden van de werkgroep zijn door de NVKC gemandateerd voor deelname. De leidraad is opgesteld onder verantwoordelijkheid van het Cluster Kwaliteit van de NVKC. Mw. L. Sarkar en mw. N. van Gool-Verdaasdonk, beiden vakexpert QC bij Result Laboratorium, Dordrecht hebben met bruikbaar commentaar aan de totstandkoming van de leidraad bijgedragen. De leden van de NVKC, NVMM, NVZA en het CMI zijn uitgenodigd voor de knelpunteninventarisatie en de commentaarfase. Het bestuur van de NVKC heeft de leidraad geautoriseerd.

1.1 Samenstelling werkgroep

Dr. ir. N. de Jonge (voorzitter), klinisch chemicus, Haaglanden Medisch Centrum, Den Haag

Dr. W.P.H.G. Verboeket-van de Venne (secretaris), wetenschappelijk onderzoeker klinische chemie, Zuyderland Medisch Centrum, Heerlen/Sittard-Geleen

Dr. drs. W.P. Oosterhuis, arts klinische chemie, Reinier Haga Medisch Diagnostisch Centrum, Delft

Dr. H.H. van Rossum, klinisch chemicus, Antoni van Leeuwenhoek Ziekenhuis, Amsterdam

Dr. M. van Schrojenstein Lantman, data analyst, Stichting Kwaliteitsbewaking Medische Laboratoriumdiagnostiek (SKML), Nijmegen

Drs. M. Severens, Black Belt, SSSS, Maastricht

Dr. H.J. Vermeer, klinisch chemicus, Albert Schweitzer Ziekenhuis, Dordrecht

1.2 Belangenverklaringen

De leden van de werkgroep hebben schriftelijk verklaard of ze in de laatste drie jaar een (financieel ondersteunde) betrekking onderhielden met commerciële bedrijven, organisaties of instellingen die in verband staan met het onderwerp van de leidraad (conform Code ter voorkoming van oneigenlijke beïnvloeding door belangenverstremgeling, opgesteld door KNAW, KNMG, GR, CBO, NHG en OMS, versie januari 2012). Tevens is navraag gedaan naar persoonlijke financiële belangen, belangen door persoonlijke relaties, belangen d.m.v. reputatiemanagement, belangen vanwege extern gefinancierd onderzoek en belangen door kennisvalorisatie. Een overzicht van de belangen van de leden van de werkgroep en het oordeel over het omgaan met eventuele belangen vindt u in onderstaande tabel.

Achternaam werkgroeplid	Hoofdfunctie	Nevenwerkzaamheden	Persoonlijke financiële belangen	Persoonlijke relaties	Extern gefinancierd onderzoek	Intellectuele belangen en reputatie	Overige belangen	Actie
de Jonge	Klinisch chemicus, Haaglanden Medisch Centrum, Den Haag	Voorzitter LESA Rationeel aanvragen van laboratoriumdiagnostiek (NHG, NVKC, NVMM, SAN)(onbetaald, enkel onkostenvergoeding).	Geen.	Geen.	Geen.	Geen.	Geen.	Geen.
Verboeket-van de Venne	Wetenschappelijk onderzoeker klinische chemie (0,65 fte), Zuyderland Medisch Centrum, Heerlen/Sittard-Geleen; beleidsadviseur kwaliteit (0,15 fte), NVKC, Utrecht	Geen.	Geen.	Geen.	Geen.	Geen.	Geen.	Geen.
Oosterhuis	Arts klinische chemie, Reinier Haga Medisch Diagnostisch Centrum, Delft (personeel niet in loondienst)(25%); Morgagni BV (10%); Oosterhuis Medisch Advies (eenmanszaak)(afgezien van werkzaamheden RHMDC 25% projecten)	RHMDC: medisch inhoudelijke consultverlening (betaald). Morgagni BV: uitvoeren van gesubsidieerde projecten in de gezondheidszorg (betaald). Oosterhuis Medisch Advies: als zelfstandige werkzaam voor RHMDC (betaald), tevens werk voor commissies en werkgroepen binnen het vakgebied, peer reviews ed. (onbetaald).	Geen.	Geen.	N.v.t. Overige gesubsidieerde projecten staan niet in relatie tot dit project.	Geen.	Geen.	Geen.
Van Rossum	Laboratoriumspecialist klinische chemie, Antoni van Leeuwenhoek ziekenhuis, Amsterdam	Directeur en eigenaar Huvaros BV (betaald). Chief Scientific Officer (CSO) Self Safe Sure Blood collection BV (onbetaald). Lid van commissie voor het opstellen van het CLSI-document C65 Biochemical Tumormarker Testing.	Eigenaar Huvaros BV: vermarkt MA Generator waarmee instellingen voor lopende patiëntengemiddelden worden verkregen. Aandeelhouder SelfSafeSure Blood Collections BV.	Geen.	Er zijn van de volgende bedrijven in de IVD-sector financiële middelen verkregen voor onderzoek of	Geen.	Geen.	Geen.

		Lid herziening NIV Richtlijn Borstkanker.	Patenten: 1) PCT/NL2016/050315 Method for setting measuring equipment, computer program and measuring equipment 2) NL2019/079641 Blood collection device and method for the self-collection of blood by a user		indirect via Huvaros BV voor specifieke dienstverlening: Abbott, Beckman Coulter, Roche Diagnostics, Siemens en Sysmex.			
Van Schrojenstein Lantman	Softwareontwikkelaar SKML, Nijmegen; Radboudumc, Nijmegen; QA Apotheker, Forinchem Pharma, Gorinchem.	Adviseur sectie algemene chemie SKML, lid working group central database EQALM	Bedrijf Mart-works: gericht op grafische visualisatie, schrijven van scripts en ontwikkeling van scholingsmateriaal.	Geen.	Geen.	Geen.	Geen.	Geen.
Vermeer	Klinisch chemicus, Albert Schweitzer Ziekenhuis, Dordrecht; Beatrixziekenhuis, Gorinchem; Result Laboratorium CV, Dordrecht	Vakdeskundige Raad voor Accreditatie betreffende beoordelingen tegen ISO 15189.	Aandeelhouder Result Laboratorium CV; dit betreft een minderheidsbelang.	Geen.	Geen.	Geen.	Geen.	Geen.
Severens	Coach/trainer 6 Sigma en Statistiek van SSSS	Geen.	Geen.	Partner van klinisch chemicus Zuyderland Medisch Centrum.	Geen.	Geen.	Geen.	Geen.

Hoofdstuk 2 Overzicht van de aanbevelingen

Aanbeveling 1

De frequentie en inrichting van de interne kwaliteitscontrole voor individuele apparaten is risicogebaseerd; enerzijds op de statistische kans dat de APS (niet) wordt behaald, en anderzijds op de klinische ernst van een niet of laat gedetecteerde fout. Hierbij kan gebruik worden gemaakt van de sigmascore, de (op risicoanalyse gebaseerde) klinische consequenties voor patiënten en het aantal patiënten dat opnieuw moet worden opgeroepen bij een eventuele terugroepactie (*recall*). (zie 7.1.1)

Aanbeveling 2

Om de analytische prestatie van een test en de bewaking daarvan met interne kwaliteitscontrole zo goed mogelijk aan te laten sluiten bij de klinische behoefte, wordt voor iedere laboratoriumbepaling de minimale APS (*Analytical Performance Specification* (NL: analytische prestatiespecificatie)) vastgesteld waaraan deze dient te voldoen. (zie 7.3.1)

Aanbeveling 3

APS is de voorkeursterm boven kwaliteitsdoelen, analytische kwaliteitseisen, prestatie-karakteristieken, *performance standards*, *analytical goals* etc. (zie 7.3.1)

Aanbeveling 4

De APS wordt vastgesteld op basis van de in 2014 in Milaan vastgestelde criteria. (zie 7.3.1)

Aanbeveling 5

Indien pre-analytische en/of post-analytische factoren de onderzoeksresultaten significant beïnvloeden, kunnen ook hiervoor APS gedefinieerd worden. (zie 7.3.2)

Aanbeveling 6

Sigmascores (*Sigma Metrics*) worden periodiek geëvalueerd, bijvoorbeeld eens per 2-4 jaar en wanneer daar aanleiding toe is, bijvoorbeeld bij aanpassingen aan de IQC-inrichting. (zie 7.6.3)

Aanbeveling 7

Patient-Based Real-Time QC (PBRTQC) kan in sommige gevallen de analytische kwaliteitsbewaking versterken. Een dergelijk systeem kan worden ingezet in het geval van evidente tekortkomingen van of beperkingen aan het uitvoeren van IQC op basis van controlematerialen. (zie 7.8)

Aanbeveling 8

Het virtuele analyzermodel is toepasbaar indien de bias, onderlinge bias en imprecisie van alle analyzers tezamen binnen de TEa, respectievelijk MAU voor de totale virtuele analyzer vallen. (zie 8.11)

Aanbeveling 9

Indien een laboratoriumorganisatie niet kiest voor een virtuele analyzer, maar wel te maken heeft met meer analyzers en/of meer locaties, kan een alternatieve aanpak gekozen worden om individuele laboratoriumbepalingen dusdanig te bewaken dat de uitslagen uitwisselbaar blijven. Voor ieder afzonderlijk systeem dient dan per bepaling eigen doelwaarden en doel-standaarddeviaties te worden vastgesteld en adequaat bewaakt. (zie 8.11)

Hoofdstuk 3 Samenvatting

Deze leidraad biedt een praktische handleiding voor (medische) laboratoria voor het inrichten van de (interne) analytische kwaliteitscontrole.

De analytische kwaliteitscontrole is gericht op het vroegtijdig detecteren van meetfouten gedurende het meetproces, zodat deze tijdig kunnen worden gecorrigeerd. De werkwijze om deze fouten te herkennen bestaat uit het gebruik van statistische kwaliteitscontrole (*Statistical Quality Control, SQC*) op twee specifieke manieren: interne kwaliteitscontrole (*Internal Quality Control, IQC* en *Patient-Based Real-Time Quality Control, PBRTQC*) en externe kwaliteitscontrole (*External Quality Assessment, EQA*).

De norm NEN-EN-ISO 15189:2023 (Medische laboratoria – Bijzondere eisen voor kwaliteit en competentie) op basis waarvan medische laboratoria worden geaccrediteerd, stelt eisen aan het systeem van kwaliteitsbewaking.

Het risico op onjuiste testuitslagen wordt in de basis ondervangen door de inrichting en uitvoering van interne kwaliteitscontrole. Hierbij worden kwaliteitscontrolemonsters als surrogaat patiëntenmonsters gemeten waarbij de resultaten worden getoetst aan de hand van vooraf vastgestelde criteria.

Ook als de interne kwaliteitscontrole aan alle criteria voldoet, bestaat er een restrisico op onjuiste resultaten. De omvang van dit restrisico is afhankelijk van de gemaakte keuzes bij de inrichting van het interne kwaliteitscontrolesysteem, maar ook van de inrichting van de laboratoriumorganisatie (m.n. logistiek, aantal analyzers en locaties). Dit restrisico kan verder worden gereduceerd door externe kwaliteitscontrole.

Voor de praktische inrichting van analytische kwaliteitscontrole van kwantitatieve bepalingen kan gebruik worden gemaakt van het stappenplan dat in deze leidraad wordt beschreven. De gebruiker dient hierin op onderdelen zelf keuzes te maken.

Hoofdstuk 4 Aanleiding en doel

4.1 Aanleiding

In 2015 is de Nederlandse leidraad Analytische Kwaliteitscontrole met behulp van Six Sigma verschenen [NVKC, 2015]. Sindsdien zijn er verschillende wetenschappelijke publicaties en ISO-richtlijnherzieningen verschenen op het gebied van analytische kwaliteitscontrole [ISO/TC 212 – Medical laboratories and in vitro diagnostic systems; van Schrojenstein Lantman, 2024]. De hierin geïntroduceerde terminologie, ontwikkelingen en werkwijzen omtrent de monitoring, beoordeling en borging van analytische prestaties gaven aanleiding voor herziening van de leidraad. Uit een in 2023 uitgevoerde enquête onder klinisch chemische laboratoria in Nederland bleek dat er een grote variatie is in opzet en uitvoering van de analytische kwaliteitsbewaking [van Rossum et al., 2024]. Wij verwachten daarom dat deze herziene versie voorziet in de behoefte vanuit het werkveld, zowel voor duiding van de intussen verschenen publicaties en richtlijnen, als voor de meer praktische implementatie van werkwijzen m.b.t. interne analytische kwaliteitscontrole.

4.2 Doel

Het doel van deze leidraad is het bieden van een praktische handleiding voor (medische) laboratoria voor het inrichten van de interne analytische kwaliteitscontrole.

Deze leidraad geeft de relevante achtergrondinformatie omtrent analytische kwaliteitscontrole en de verschillende wijzen waarop deze kan worden bewaakt, inclusief het bepalen van grenzen en eisen, en de uitvoering hiervan. Elke vorm van analytische kwaliteitsborging die in deze leidraad wordt genoemd heeft unieke voor- en nadelen en vraagt aanpalend beleid, waarbij de uiteindelijk keuze afhankelijk is van de relevante testtoepassingen en de organisatie en logistiek van het laboratorium.

In de voorliggende leidraad wordt na een algemene inleiding over interne kwaliteitsbewaking in Hoofdstuk 7 een uitwerking gegeven van het theoretisch concept met betrekking tot de interne kwaliteitscontrole, waarbij naast Six Sigma ook *Bracketed Risk-Based Quality Control*, *Patient-Based Real-Time Quality Control* en conventionele methoden worden besproken. Hoofdstuk 8 geeft een praktische samenvatting met een stappenplan, waarmee de lezer een keuze kan maken uit de verschillende opties. Voor meer verdieping wordt verwezen naar de bijlagen.

Hoofdstuk 5 Toepassingsgebied en beperkingen (scope)

Deze leidraad richt zich primair op de inrichting van de analytische kwaliteitsbewaking van kwantitatieve meetmethoden met behulp van interne kwaliteitscontroles. Hierbij worden enkele handvatten gegeven voor de praktische inrichting o.b.v. het risico voor de patiënt(enzorg). De leidraad is zoveel mogelijk gebaseerd op resultaten van wetenschappelijk onderzoek; voor een deel is deze leidraad echter gebaseerd op modellen en aannames waarover niet altijd consensus bestaat. Omdat veel ontwikkelingen m.b.t. het onderwerp van de leidraad nog gaande zijn is het voor sommige onderdelen niet mogelijk om eenduidige adviezen te geven.

Voor het leveren van een betrouwbaar meetresultaat is een integrale kwaliteitsstrategie essentieel. Vakbekwaam personeel, juiste omgevingscondities, deugdelijke apparatuur en informatiesystemen, en een goede verbetercultuur zijn noodzakelijke voorwaarden om een goed resultaat te kunnen produceren. Deze leidraad beperkt zich tot de analytische kwaliteitscontrole in engere zin, waarbij wordt aangenomen dat aan de hiervoor benoemde randvoorwaarden wordt voldaan. De hierbij bedoelde aspecten zijn onderwerp van de NEN-EN-ISO 15189:2023 norm.

Hoofdstuk 6 Samenhang met andere documenten

In een aantal landen is de analytische kwaliteitsbewaking aan wet- en regelgeving onderhevig. In de Verenigde Staten bijvoorbeeld is hiervoor de CLIA (*Clinical Laboratory Improvement Amendments*) van toepassing [CLIA], in Duitsland de RiliBÄK (*Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen*) [2023]. Deze regelgeving is vooral gericht op het d.m.v. externe kwaliteitscontrole voorkomen van toepassing van ondermaatse bepalingen. In Nederland is in het kader van deregulering dergelijke regelgeving niet van toepassing. Voor geneesmiddelen wordt binnen de Europese Unie voor bioanalytische methodevalidatie ook gerefereerd aan de EMA/CHMP/ICH/172948/2019 M10 richtlijn. De eisen van deze richtlijnen kunnen worden gehanteerd als grenzen als handvat voor de kwaliteitscontrole.

De norm NEN-EN-ISO 15189:2023 (Medische laboratoria – Bijzondere eisen voor kwaliteit en competentie) [2023], op basis waarvan medische laboratoria worden geaccrediteerd, stelt eisen aan het systeem van kwaliteitsbewaking. Een overzicht van de belangrijkste eisen die in NEN-EN-ISO 15189:2023 worden gesteld aan de analytische kwaliteitsbewaking staat in [Bijlage A](#). Een belangrijke wijziging in de nieuwe versie van de ISO 15189 (NEN-EN-ISO 15189:2023) is dat expliciet wordt gesteld dat statistische technieken moeten worden toegepast om de resultaten te beoordelen (normelement 7.3.7.1). Daarnaast behoort de beoogde klinische toepassing van het onderzoek in overweging te worden genomen (normelement 7.3.7.2.a.1). De norm biedt uitdrukkelijk de mogelijkheid voor toepassing van trendanalyse, zoals voortschrijdend patiëntengemiddelde, indien er geen passend intern QC-materiaal beschikbaar is (normelement 7.3.7.2.c.1). Het is ook toegestaan om patiëntenmonsters te hertesten of gebruik te maken van monsters die getest zijn met een alternatieve methode die metrologisch traceerbaar is tot referenties van dezelfde of hogere orde (normelement 7.3.7.2.c.2). Deze werkwijze valt echter buiten de scope van deze leidraad. Het laboratorium kan zelf een passende methode kiezen, rekening houdend met eisen gesteld aan commuteerbaarheid, meetbereik en interferentie (normelement 7.2.7.6.1).

Als toevoeging op de ISO 15189 is recent ook de ISO 5649:2024 (Medical laboratories — Concepts and specifications for the design, development, implementation and use of laboratory-developed tests) gepubliceerd. Deze richtlijn beschrijft specifiek dat er interne als externe kwaliteitscontrole op eigen ontwikkelde tests (*lab developed tests*, LDT) moet plaatsvinden.

Andere voor medische laboratoria relevante kwaliteitsdocumenten, die werkwijzen voor de analytische kwaliteitsbewaking beschrijven, zijn de CLSI-documenten EP23 en C24 [CLSI]. In deze documenten wordt het potentiële risico voor de patiënt als basis voor de inrichting van de analytische kwaliteitsbewaking genomen. Er worden echter niet altijd concrete handvatten gegeven hoe dit uit te voeren. Daarnaast is er geen wetenschappelijke bewijslast beschikbaar dat dit daadwerkelijk klinische meerwaarde heeft.

Deze leidraad biedt een strategie om de gestelde eis in ISO 15189:2023 te vertalen naar concrete criteria.

Hoofdstuk 7 Theoretische achtergrond

7.1 Algemeen

Op het moment dat de geschiktheid van een analyzer of meetmethode voor de beoogde testtoepassingen is geverifieerd of gevalideerd kan een test in gebruik worden genomen. Om echter veilige en betrouwbare zorg te leveren is het noodzakelijk dat ook tijdens de gebruiksduur van het meetsysteem de kwaliteit van de gerapporteerde uitslagen gewaarborgd blijft.

De analytische kwaliteitscontrole is gericht op het vroegtijdig detecteren van meetfouten gedurende het meetproces, zodat deze tijdig kunnen worden gecorrigeerd. De werkwijze om deze fouten te herkennen bestaat uit het gebruik van statistische kwaliteitscontrole (*Statistical Quality Control, SQC*) op twee specifieke manieren: interne kwaliteitscontrole (*Internal Quality Control, IQC* en *Patient-Based Real-Time Quality Control, PBRTQC*) en externe kwaliteitscontrole (*External Quality Assessment, EQA*).

Het is daarbij ook van belang op welke wijze de uitkomsten van controles geïnterpreteerd worden: waar ligt de grens waarop een uitslag moet worden afgekeurd? Wanneer worden uitslagen van patiënten afgekeurd en moeten deze opnieuw worden bepaald (terugroepactie of *recall procedure*)? Wanneer is er sprake van een significante systematische fout (bijvoorbeeld door verloop of drift)?

Hierbij moeten de volgende aspecten in ogenschouw worden genomen:

- De toepassing van de test, rekening houdend met o.a. bepalingenvolume, logistiek en consequenties van foutieve klinische handeling;
- De vereiste analytische prestatiespecificaties (APS) die hieruit volgen;
- Risico op ontstaan van meetfouten;
- De mogelijkheid tot tijdige detectie hiervan [NEN-EN-ISO 15189:2023; CLSI EP23].

Bij interne kwaliteitscontrole (IQC) worden kwaliteitscontrolemonsters als surrogaat patiëntenmonsters gemeten waarbij de resultaten worden getoetst aan de hand van vooraf vastgestelde criteria (paragraaf 7.2). *Patient-Based Real-Time Quality Control* is, zoals de naam aangeeft, gebaseerd op resultaten van patiëntenmonsters (paragraaf 7.8).

Daarnaast beschikken laboratoria over andere instrumenten om de analytische kwaliteit te bewaken die niet zozeer de analytische prestatie controleren, maar wel gebruikt kunnen worden om analytische fouten te voorkomen, dan wel voortijdig te detecteren. Hieronder worden verstaan: analyzeralarmeringen m.b.t. mogelijke afwijkingen van apparatuur, methode, monster of resultaat en de technische vrijgave (paragraaf 7.4) ofwel confirmatie van resultaten met behulp van o.a. beslissgrenzen, *deltachecks*, et cetera. Al deze instrumenten hebben specifieke eigenschappen en maken deel uit van de operationele analytische kwaliteitsbewaking van een analysemethode.

Ook als de interne kwaliteitscontrole aan alle criteria voldoet, bestaat er een restrisico op onjuiste resultaten. De omvang van dit restrisico is afhankelijk van de gemaakte keuze bij de inrichting van het interne kwaliteitscontrolesysteem, maar ook van de inrichting van de laboratoriumorganisatie (m.n. logistiek, aantal analyzers en locaties). Dit restrisico kan verder worden gereduceerd door externe kwaliteitscontrole. Daarmee is de optimale wijze van deelname aan en interpretatie van externe kwaliteitscontrole afhankelijk van de inrichting van de interne kwaliteitscontrole (zie Tabel 7.1). Omdat externe kwaliteitscontrole buiten de scope van deze richtlijn valt, gaan we hier niet verder inhoudelijk op deze relatie in. Het is de verantwoordelijkheid van laboratoria om de consequenties van hun keuzes in hun intern kwaliteitscontrolebeleid te vertalen naar inrichtingseisen voor hun extern kwaliteitscontrolebeleid, om zo te komen tot een integraal kwaliteitscontrolebeleid zoals bedoeld in ISO 15189:2023.

Tabel 7.1. Overzicht van eigenschappen van de verschillende instrumenten beschikbaar in medische laboratoria voor de analytische kwaliteitsbewaking [Naar: van Rossum, 2022]

	IQC	EQC	Patient-Based Real-Time Quality Control				Analyzer vlaggen
			Lopend gemiddelde	Limiet controles	Delta controles	Multivariabele controles	
Kwaliteitsborging / kwaliteitscontrole per fase							
Pre-analytische fase		Beperkt	Ja	Ja	Ja	Ja	Ja
Analytische fase	Ja	Ja	Ja	Ja	Ja	Ja	Ja
Post-analytische fase		Beperkt	Beperkt	Ja	Ja	Ja	Ja
Klinische validatie				Ja	Ja	Ja	
Kenmerken							
Continu	Nee	Nee	Ja	Ja	Ja	Ja	Ja
Risico op niet-commuteerbaarheid	Ja	Ja	Nee	Nee	Nee	Nee	N.v.t.
Mogelijkheid tot foutdetectie voordat de fout optreedt	Ja	Nee	Nee	Nee	Nee	Nee	Ja
Eigenschappen voor foutdetectie							
Detectie van een systemische fout (bias)	Ja	Ja	Ja	Ja	Ja	Ja	N.v.t.
Detectie van een imprecisie fout	Ja	Ja	Onbekend	Onbekend	Onbekend	Onbekend	N.v.t.
Detectie van een willekeurige fout	Moeilijk	Nee	Moeilijk	Ja	Ja	Ja	N.v.t.

7.1.1 Risico-gebaseerde inrichting van de interne analytische kwaliteitscontrole

Het risico van onjuiste uitslagen met mogelijk ongewenste patiëntuitkomsten hangt af van verschillende factoren. Bij de inrichting van de IQC die hiertoe als maatregel wordt ingericht behoort hiermee rekening te worden gehouden. Het risico bestaat uit zowel een statistisch (kans op afwijkend resultaat) als een medisch risico, waarbij ook rekening moet worden gehouden met de detectie en omvang en impact van een meetfout op de volgende correctie hiervan (bijvoorbeeld terugroepactie (*recall*)) (**Aanbeveling 1**). Neem eis ISO 15189 7.3.7.2.D als uitgangspunt: frequentie moet gebaseerd op kans en ernst. Definieer risico als ‘kans maal ernst’. De risico’s kunnen van laboratorium tot laboratorium verschillen, afhankelijk van o.a. toepassingsgebied, grootte, organisatie en werkwijze.

Bij de inrichting van de IQC spelen de volgende factoren een rol:

- Toepassing van de bepaling (medische urgentie, relevant meetgebied, risico foutieve uitslagen);
- Omvang van het laboratorium (aantal analyzers en locaties analyzers, logistieke processen, analysefrequentie, toekenning van patiëntenmonsters aan één of meer analyzers);
- Selectie QC-materiaal (concentratie, commuteerbaarheid);
- Selectie tolerantiegrenzen (afhankelijk van medisch risico, testtoepassing, zie hiervoor paragraaf 7.3 APS);
- Frequentie meting (afhankelijk van statistisch en medisch risico), en de impact van *recall* (tevens sturing aan momenten waarop beoordeling plaatsvindt);
- Risico van meetresultaat buiten tolerantiegrenzen door de analytische prestatie (paragraaf 7.6) te beoordelen ten opzichte van de tolerantiegrenzen (paragraaf 7.7, Six Sigma).

Hierin zijn geen twee laboratoria hetzelfde. De inrichting en uitvoering van de IQC zal moeten worden afgestemd op de situatie van het laboratorium. Echter, elk laboratorium neemt een test in gebruik nadat in de verificatie of validatie geen significante systematische afwijking is gevonden en de meetonzekerheid van het systeem binnen de APS is bevonden.

In de praktijk kan een meetfout het gevolg zijn van zowel incidentele fouten, imprecisie, of één of meer systematische oorzaken die in de loop van de tijd kunnen variëren (zoals wisselingen in reagens-lot, omgevingscondities en trends en drifts die nog binnen APS zijn). Op de lange termijn zijn deze gecombineerde factoren te kwantificeren als meetonzekerheid. Meetonzekerheid omschrijft de lange-termijn variatie (paragraaf 7.6).

7.2 Analytische kwaliteitscontrole

7.2.1 Externe kwaliteitscontrole (EQA)

De EQA omvat alle SQC-methoden, die periodiek worden uitgevoerd waarbij gebruik gemaakt wordt van de diensten/programma's van een externe organisatie (referentielaboratorium, organisator van externe rondzendingen, industrie). De EQA is gericht op vergelijking met en afstemming op definitieve methoden, referentiemethoden of consensuswaarden op de lange termijn en is ook toepasbaar voor harmonisatie van uitslagen tussen laboratoria [Perich et al., 2014]. Dit laatste is overigens alleen mogelijk voor commuteerbare EQA-materialen. In Nederland verzorgt de organisatie SKML externe rondzendingen voor alle medische laboratoria.

Een belangrijke voorwaarde om de juistheid te kunnen vastleggen is kalibratie van de analytische procedure. Van de gebruikte standaarden dient de toegekende waarde en de commuteerbaarheid bekend te zijn en indien mogelijk herleidbaar naar een primaire referentiemethode c.q. erkend referentiemateriaal [JCTLM]. EQA valt buiten het bestek van deze leidraad. Voor verdere beschouwing van de externe kwaliteitscontrole, zie Jansen et al. [2014] en Buchta et al. [2025].

7.2.2 Interne kwaliteitscontrole (IQC)

De interne kwaliteitscontrole is ingericht om de meetonzekerheid binnen de toegestane APS te behouden. In veel klinisch chemische laboratoria wordt dit uitgevoerd door middel van metingen in controlematerialen met als doel te beoordelen of de onderzoeksprocedures goed functioneren. Onder de onderzoeksprocedure wordt verstaan de combinatie van het analysesysteem (apparatuur), de gebruikte reagentia en standaarden.

Afhankelijk van de resultaten van deze controlemetingen worden de analyseresultaten vrijgegeven voor de diagnostiek.

In klinisch chemische laboratoria wordt voor IQC sinds de jaren '70-'80 van de vorige eeuw veelvuldig gebruik gemaakt van criteria en beoordelingsmethodieken die zijn beschreven door Levey en Jennings [1950] en Westgard et al. [1981]. Deze wijze van kwaliteitscontrole is gebaseerd op een statistische benadering, zoals oorspronkelijk door Shewhart beschreven [1931]. Cotlove et al. [1970] toonden aan dat er onvoldoende rekening werd gehouden met verschillen tussen biologische en analytische variatie voor de verschillende bepalingen.

Met behulp van de rond 1985 ontwikkelde Six Sigma methodiek kan de toepassing van kwaliteitscontroleregels worden verbeterd, om zowel tot optimaal resultaat (hoge foutdetectiekans en lage onterechte afkeuring) als tot optimaal gebruik van middelen te komen.

De laatste jaren wordt in toenemende mate gebruik gemaakt van andere (risico gebaseerde) technieken, zoals *Bracketed Risk-Based Quality Control* [Westgard et al., 2018] en *Patient-Based Real-Time Quality Control* [van Rossum, 2020; van Rossum et al., 2021]. Deze concepten worden in de leidraad benoemd en aanvullende informatie is te vinden in de [Bijlage H](#).

Merk op dat kwaliteitscontrole de kwaliteit van een bepalingmethode in het algemeen niet kan verbeteren. De controle is erop gericht afwijkingen van een normaal prestatieniveau van een bepaling te detecteren. Daarbij is de stabiliteit van een bepalingmethode vooral van belang (drift). Er zijn aanwijzingen dat frequente controle met frequente herkalibratie zelfs ten nadele kan zijn van de kwaliteit (bij onnodige kalibraties van een stabiele test) [Lim et al., 2023; Ramamohan et al., 2012].

[Bijlage I](#) geeft een handreiking voor correctieve acties bij afwijkende interne QC. Nadere uitwerking van dit stappenplan valt buiten het bestek van deze leidraad.

7.3 Analytische kwaliteitseisen en selectie van tolerantiegrenzen

7.3.1 Analytische prestatiespecificatie (APS)

Om de analytische prestatie van een test zo goed mogelijk aan te laten sluiten bij de klinische behoefte moet de analytische prestatiespecificatie (APS, *analytical performance specification*) worden vastgesteld ([Aanbeveling 2](#)). Onder APS wordt verstaan de toegestane (*allowable*) *bias*, *allowable* imprecisie, *allowable total error* (TEa) en *allowable measurement uncertainty* (MAu). Zie ook pagina 18. In de literatuur worden voor de analytische prestatiespecificatie diverse synoniemen gebruikt: kwaliteitsdoel, analytische kwaliteitseis, prestatiekenmerk, *performance standard*, *analytical goal*, enz. APS is in de literatuur de meest gangbare term. Deze term wordt ook in deze leidraad gehanteerd. Om spraakverwarring in het werkveld te voorkomen, adviseert de werkgroep om bij voorkeur de term APS te gebruiken ([Aanbeveling 3](#)). Het is van belang dit specifiek te benoemen omdat er aan laboratoriumbepalingen ook andere eisen worden gesteld zoals economische (lage kosten), beschikbaarheid (korte *turn-around-time*) en veiligheid (veilige uitvoer voor medewerkers). De APS voor een bepaald laboratoriumonderzoek hoeft niet in alle klinische situaties hetzelfde te zijn: in rurale of afgelegen gebieden kunnen onderzoeken met minder stringente APS toch significante gezondheidswinst betekenen, vergeleken met de situatie waarin deze test niet beschikbaar is. Voor een aantal veel aangevraagde bepalingen zijn APS voorgesteld [Panteghini, 2024].

In principe dienen APS voor een bepaalde laboratoriumtest ieder beoogd klinisch gebruik te omvatten en daarnaast voor iedere verschillende methode apart te worden gedefinieerd. Immers, als een laboratoriumorganisatie APS niet objectief definieert en controleert, kan dit tot grote variaties in de laboratoriumuitslagen leiden en daarmee potentiële schade bij de patiënt.

In de internationale literatuur wordt met APC (*analytical performance characteristics*, analytische prestatiekenmerken) de feitelijke prestatie bedoeld die in de praktijk is aangetoond; deze wordt afgemeten tegen de minimaal te behalen APS [Thelen en van Schroyen Lantman, 2024].

Analytische prestatiespecificaties vinden verschillende toepassingen en kunnen op basis daarvan ook verschillen [Jones, 2024]. Deze specificaties spelen een rol bij de keuze welke test op een laboratorium te introduceren, bij de validatie of verificatie ervan en bij het ontwerpen van de kwaliteitscontrole. APS worden ook gebruikt bij externe kwaliteitscontrole, waarbij verschillende programma's tot een andere keuze kunnen komen met betrekking tot de gehanteerde APS. Naast de insteek van de programma's (wel/geen verplichte naleving (*compliance*)) en de gehanteerde Milaanmodellen (Milaan model 1, 2 en/of 3; zie hieronder) kan ook de definitie van deze modellen verschillen (bijv. de definitie van *state of the art*). Het gevolg is een verscheidenheid aan gehanteerde APS tussen organisaties. Derhalve wordt geadviseerd om als laboratorium een eigen grondslag te hebben voor de gehanteerde specificaties, waarbij van de APS van externe kwaliteitscontrole organisaties gebruik kan worden gemaakt. Voor CE-gemarkeerde tests gelden volgens de IVDR specificaties waarbij de klinische toepasbaarheid moet zijn aangetoond. Het is niet uitgesloten dat de fabrikant zijn APS baseert op de meest gunstige (lees minimale) eisen ([Bijlage B](#)). Hiermee is het raadzaam om de APS van de IVD-fabrikant te verifiëren bij ingebruikname van het product, maar verdere instelling en monitoring van APS uit te voeren op basis van eigen inzicht.

In 1999 is onder auspiciën van de IUPAC, de IFCC en de WHO een hiërarchisch raamwerk ontwikkeld voor de benodigde (analytische) kwaliteitseisen binnen de laboratoriumgeneeskunde [Fraser et al., 1999; Fraser, 2015; Kenny et al., 1999]. Deze zijn in 2014 herzien tijdens een strategische conferentie in Milaan. De strikte hiërarchische indeling werd losgelaten; tests worden ingedeeld op basis van

testkarakteristieken en klinische toepassing [Ceriotti et al., 2017; Braga en Panteghini, 2021]. Hierbij wordt onderscheid gemaakt tussen drie verschillende modellen:

De APS op basis van de in Milaan in 2014 vastgestelde criteria [Ceriotti et al., 2017] worden:

- 1 Gebaseerd op het effect op de klinische uitkomst binnen een specifieke klinische vraagstelling
 - 1a Op basis van directe uitkomststudies
 - 1b Op basis van indirecte uitkomststudies
- 2 Gebaseerd op de biologische variatie van een analiet
- 3 Een prestatiedoel ingegeven door *state-of-the-art* technologie

Uitkomstseisen (model 1a en b) zijn relevant voor bepalingen waarbij een duidelijke relatie bestaat tussen het meten van de bepaling en het uitvoeren van een klinische actie, zoals bij beslisgrenzen benoemd in richtlijnen. Hierbij kan worden gekozen om een directe uitkomststudie uit te voeren (1a) of om het effect van de meetonzekerheid te simuleren (1b).

Hoewel kwaliteitseisen gebaseerd op klinische uitkomst het beste lijken te zijn, blijkt het in de praktijk voor de meeste tests onmogelijk om een directe relatie te leggen tussen testuitkomst en behandelresultaat. Het alternatief hiervoor, model 1b, maakt gebruik van data simulaties waarmee de impact van verschillende APC's op een behandelconsequentie voor een bepaalde populatie wordt bepaald [Horvath et al., 2024]. Model 1 geldt o.a. voor HbA1c en hemoglobine [Panteghini, 2024]. Voor de vaststelling van de APS door middel van model 1b is via internet een calculator beschikbaar, die op basis van de eigen populatie, gehanteerde afkapgrenzen en gewenste overeenstemming de APS kan berekenen [Çubukçu et al., 2023].

Als model 1 niet beschikbaar of relevant is, dan is model 2 toepasbaar voor bepalingen met een stabiele biologische variatie (BV). Hierbij wordt de APS berekend op basis van de binnen- en tussen-persoonsvariatie (respectievelijk CV_i en CV_g). Voor de meeste tests wordt dit model toegepast. De conventionele vergelijking combineert de systematische fout (bias: $0,25 \times (CV_i^2 + CV_g^2)^{1/2}$) en toevallige fout (imprecisie):

$$TEa = 1,65 \times (0,5 \times CV_i) + 0,25 \times (CV_i^2 + CV_g^2)^{1/2} \text{ (individuele meetwaarde, eenzijdig)}$$

Hierin is de korte-termijn imprecisie (CV_a) afhankelijk van de binnen-persoonsvariatie [Fraser et al, 1999]; $CV_a < 0,5 \times CV_i$ (zie ook paragraaf 7.3.2).

Een ander model, dat ook aansluit bij het model van meetonzekerheid (*measurement uncertainty*, MU), gaat alleen uit van de binnen-persoonsvariatie. Hierbij is de maximaal toegestane meetonzekerheid (MAu) gelijk aan de vergelijking van de CV_a, terwijl de 'expanded MAu' (de MAU), met een factor 2 de centrale 95% van de meetonzekerheid beschrijft [Sandberg et al., 2024]. Een aanname van het MU-concept is dat er geen significante bias is. De toelaatbare bias kan op basis van verschillende grondslagen worden gedefinieerd, zoals de toelaatbare bias op basis van biologische variatie, een bias die niet significant afwijkt van 0 of een andere analyzer, of correctie van bias op basis van ISO/TS 20914:2019 (Medical laboratories — Practical guidance for the estimation of measurement uncertainty) geoorloofd is, of een bias die niet leidt tot het moeten handhaven van andere Westgardregels op basis van de sigmascore [EFLM Biological Variation Database, ISO/TS 20914].

Betrouwbare gegevens voor biologische variatie zijn beschikbaar op de website van de EFLM [EFLM Biological Variation Database]. De hierboven benoemde vergelijking van TEa en CV_a maken gebruik van de APS op basis van *desirable* ofwel wenselijke kenmerken (een constante van 0,5 voor de imprecisie en 0,25 voor de bias). Afhankelijk van het risico en de gewenste mate van monitoring kunnen ook optimale en minimale APS worden bepaald met behulp van imprecisiefactoren van resp.

0,25 en 0,75, en biasfactoren van resp. 0,125 en 0,375 [Fraser et al., 1997]. Model 2 geldt o.a. voor kalium en kreatinine [Panteghini, 2024].

Model 3: *State of the art*. Als data van biologische variatie niet beschikbaar of niet haalbaar zijn omdat concentraties niet stabiel zijn (bijv. urineonderzoek), dan blijven de prestatie-eisen voor analytische prestatie slechts streefwaarden. Hierbij wordt model 3 gebruikt om de *state-of-the-art* analytische prestatie te gebruiken als houvast om op basis hiervan de analytische prestaties te bepalen. Model 3 geldt o.a. voor CRP en hCG [Panteghini, 2024].

Voor APS en medicatiebewaking wordt verwezen naar [Bijlage C](#).

Er zijn verschillende definities van *state of the art*:

- Een percentage van de laboratoria dat deze limiet haalt (bijv. in het kader van externe kwaliteitscontrole);
- De hoogste kwaliteit die op de markt verkrijgbaar is [Braga en Panteghini, 2020];
- De (ruime) definitie van de MDCG, een algemene *state of the art*, die niet de best beschikbare kwaliteit hoeft te zijn [Medical Device Coordination Group 2022-2].

Borrillo en Panteghini [2024] hebben beschreven hoe bij ontbreken van één van bovenstaande *state-of-the-art* varianten deze zelf kan worden berekend.

De vereiste APS worden vergeleken met de behaalde APC, om op basis hiervan te beoordelen wat de relatie is tussen de gewenste en de behaalde kwaliteit. Op basis hiervan kan worden beoordeeld hoe stringent de monitoring van de analytische prestatie op dagelijkse basis moet zijn. Hierbij moet bij de inrichting rekening worden gehouden met het klinische risico. Indien een test verschillende klinische toepassingen heeft, moet de meest strikte APS worden gebruikt.

Als de APS niet realistisch haalbaar is, dan wordt dit een doel (*analytical performance goal*, APG) voor toekomstige APS. In de tussentijd kan een minder stringente APS worden gekozen waaraan wel degelijk kan worden voldaan met de feitelijke APC (bijv. *state of the art*).

De APS worden vastgesteld op basis van de in Milaan in 2014 vastgestelde criteria [Ceriotti et al., 2017] ([Aanbeveling 4](#)).

Zie ook paragraaf 8.2 en [Bijlage B](#).

7.3.2 APS op basis van totale meetfout (TEa) versus meetonzekerheid (MU)

APS kan worden gebaseerd op het *total error* concept of op de toegestane meetonzekerheid. Deze concepten zijn principieel verschillend: de TEa geeft het bereik aan waarbinnen de meetuitkomst zou moeten vallen uitgaande van de juiste waarde; de meetonzekerheid MAU geeft het bereik aan waarbinnen de juiste waarde zich met een bepaalde waarschijnlijkheid bevindt uitgaande van de meetuitkomst. Dit bereik is afhankelijk van het gebruikte betrouwbaarheidsinterval, waarbij in de MAU (*allowable expanded measurement uncertainty*) een factor wordt gebruikt, bijvoorbeeld 2 of 3 voor een betrouwbaarheidsinterval van 95% respectievelijk 99%. In de metrologie (en de daaruit afgeleide toepassingsgebieden zoals ISO) is meetonzekerheid een gevestigd begrip. Binnen de klinische chemie begint dit begrip ingang te krijgen. Het concept TEa is daarentegen buiten de klinische chemie niet gangbaar.

Op dit moment maken medische laboratoria veelal gebruik van het TE-concept. In het TE-concept is het toegestane *error budget* (*total allowable error*, TEa) voor een laboratoriumbepaling gebaseerd op één van de EFLM Milaan APS modellen (*medical outcome*, biologische variatie of *state of the art*). Vervolgens wordt dit *error budget* opgedeeld in zowel een component voor de maximaal toelaatbare

imprecisie als een component voor de maximaal toelaatbare bias. Deze beide componenten worden bij elkaar opgeteld. In de literatuur is intussen veel discussie over de wijze waarop deze verschillende budgetcomponenten bij elkaar worden opgeteld in het TEa-concept [Oosterhuis en Coskun, 2018]. Verder gaat het TE-model ervan uit dat de bias in de tijd constant is terwijl deze in de praktijk fluctueert doordat de verschillende bronelementen die tezamen voor de bias zorgen elk fluctueren met een onderling verschillende frequentie en amplitude [Thelen en van Schrojenstein Lantman, 2024].

Zowel TEa als MAU kennen zowel imprecisie als bias. Het belangrijkste verschil ligt in het feit dat de bias binnen MAU wordt geacht kleiner te zijn dan de aanvaardbare bias en daarmee onderdeel uitmaakt van de lange termijn imprecisie [Thelen en Van Schrojenstein Lantman, 2024]. Wanneer de bias groter is dan toelaatbaar dient deze te worden gecorrigeerd waarbij de onzekerheid van die correctie onderdeel wordt van MU [zie ISO/TS 29014].

Bij het TEa concept worden bias en imprecisie als uitwisselbare componenten gezien van een gezamenlijk *error budget*; hoe kleiner de imprecisie des te meer ruimte er is voor bias. Voor de meeste toepassingen is een separaat budget voor systematische fouten (bias) nuttig. Daarom wordt in de recente literatuur vooral gewerkt met het bewaken van lange termijn-imprecisie (uRw) binnen een MAU APS. De bias moet kleiner zijn dan maximaal aanvaardbaar.

Naast het MAu concept is er nog steeds plaats voor het TEa concept. Dat speelt met name bij EQA: daar moet de bias worden beoordeeld in een dataset waarin ook imprecisie een rol speelt. Doordat de bias potentieel groter is dan de toelaatbare bias, kan het MAu model niet worden toegepast en kan er geen onderscheid gemaakt worden tussen bias en imprecisie.

Als het laboratorium ervoor kiest APS toch op het TEa-concept te baseren, betekent dit dat er gefundeerde en specifieke APS vastgesteld moeten worden waarbij geborgd wordt dat de analytische imprecisie (CVa) en de bias tezamen binnen het criterium vallen.

Ten slotte kunnen ook pre-analytische en/of post-analytische (zoals berekende testen, scoresystemen) onzekerheden een uitslag beïnvloeden. In elk geval waarbij er significante onzekerheid bestaat die klinische interpretatieproblemen kan geven in de uiteindelijk gerapporteerde uitslag, kunnen hiervoor APS ook gedefinieerd worden, zulks ter beoordeling van het laboratorium ([Aanbeveling 5](#)).

Zie ook [Bijlage D](#).

7.4 Vaststellen van de imprecisie

De spreidingsmaat voor imprecisie is de standaarddeviatie (SD); deze zal in veel gevallen als percentage (variatiëcoëfficiënt CV) worden uitgedrukt. De SD-waarde vormt de grondslag van het vaststellen van de kwaliteit van een meetmethode en voor de inrichting van de kwaliteitscontrole.

Er zijn verschillende manieren om de imprecisie vast te stellen, met oplopende variatie:

- herhaalbaarheid (*repeatability*, 'binnen run': onder gelijkblijvende condities);
- *intermediate reproducibility* ('tussen run': tussenvorm);
- reproduceerbaarheid (verschillende labs, verschillende analyzers, verschillende analisten).

SD en de prestatie van een methode wordt doorgaans vastgesteld onder condities van *intermediate reproducibility*. Dat betekent op één analyzer, over een bepaalde tijdsperiode (bijv. 3 maanden). Een vaste periode kan niet worden gegeven, maar men moet uitgaan van een representatieve periode met kalibraties en reagenswisselingen.

In essentie zijn er drie mogelijkheden voor bepaling van de SD:

- Op basis van alle metingen waarbij uitbijters met een bekende oorzaak (d.w.z. zonder de extremen waaronder geen patiëntenresultaten zijn gerapporteerd) zijn verwijderd. Dit is conform ISO TS20914; afwijkende QC-resultaten door technische fouten dienen uitgesloten te worden van de berekening [ISO/TS 20914, 6.2]. Er zouden echter wel patiëntenmonsters onder deze afwijkende condities bepaald kunnen zijn, terwijl niet besloten is tot een *recall* procedure. Daarmee representeren deze afwijkende QC-uitslagen wel de condities waaronder patiëntenmonsters zijn bepaald. Inclusie van deze QC-uitslagen geeft daarmee een correcte meetonzekerheid weer.
- Geschoond voor uitbijters en representatief voor de onder controle situatie. Een alternatieve werkwijze is uitbijters op basis van statistische criteria te verwijderen. Zonder deze uitbijters verkrijgt men een 'reële' SD. Dit correspondeert met de kwaliteit zoals deze zou moeten zijn bij een correcte, stabiele prestatie van de bepalingsmethode. Veel gebruikte QC-softwareprogramma's blijken verschillend om te gaan met uitbijters (Unity RealTime: geen uitbijterdetectie, alle afwijkende QC wordt automatisch verwijderd (wordt niet meegenomen); QCToday: eigen uitbijterdetectie, wordt niet meegenomen in de statistiek).
- *Moving Range*, geschoond voor uitbijters en invloeden van *shifts*.

Naast verschillen tussen individuele analyzers of virtuele analyzersystematiek, zijn er ook verschillen op welke wijze de SD en/of doelwaarde voor controlematerialen wordt vastgesteld [van Rossum et al., 2024]. Zo zijn er laboratoria die deze voor ieder nieuw controlemateriaal opnieuw vaststellen. Andere laboratoria stellen de SD of CV eenmalig vast en nemen deze voor nieuwe controlematerialen over. Deze vaststelling betreft in het algemeen een iets langere termijn, rekening houdend met mogelijke drift, lot-tot-lot variatie en kalibratievariatie van de meetmethode. Ook kan een *gepoolde* SD van verschillende lotnummers controlematerialen worden gebruikt om een grotere dataset te verkrijgen. Hierbij dienen de sommen der kwadraten van de individuele SD's (of CV's, afhankelijk van de concentratieafhankelijkheid van de imprecisie).

Het afleiden van de SD uit een standaard Levey-Jennings grafiek is gevoelig voor uitbijters. Bovendien moeten data voor een correcte SD-berekening normaal verdeeld zijn. Dit is niet het geval als resultaten van meerdere analyzers worden gecombineerd en er sprake is van bias tussen analyzers. Verder is het voor het bepalen van een lange-termijn imprecisie noodzakelijk dat de materialen commuteerbaar zijn tussen de verschillende wisselende aspecten van het meetsysteem (kalibratoren, reagentia, *consumables*, etc).

Er zijn statistische methoden die – ook zonder uitbijtercorrectie – ongevoelig zijn voor afwijkende uitslagen om de SD te bepalen [Theodorsson, 2025]. Hierbij hoeven dus geen uitbijters uit de populatie te worden verwijderd om toch een goede en betrouwbare schatting te maken van de SD. Voorbeelden hiervan zijn Interkwartiel Range (binnenbox Box-plot [Box en Cox, 1964]), *Robust Standard Deviation*, *Average Moving Range*.

7.4.1 Individueel versus virtueel analyzer model

In veel laboratoria kunnen analyses op meer dan één analyzer worden uitgevoerd, soms ook verdeeld over diverse locaties. De analyzers kunnen gezamenlijk als een virtuele analyzer worden beschouwd.

Vanuit een laboratorium kan gekozen worden om de inrichting van SQC voor iedere analyzer individueel in te richten, met eigen doel en/of SD-waarden, of om te kiezen voor het virtuele analyzermodel [Badrick en Graham, 2018; Giannoli et al., 2024].

Het begrip ‘virtuele analyzer’ is nog niet volledig uitgekristalliseerd in de literatuur. In deze leidraad definiëren we het begrip als volgt:

Het virtuele analyzermodel is van toepassing voor laboratoria die voor dezelfde analyse meer dan één analyzer gebruiken, waar patiëntenmonsters aan een willekeurige analyzer worden aangeboden en de resultaten onder dezelfde rapportregel worden gerapporteerd.

De kwaliteitseisen betreffen dan de virtuele analyzer. Dit betekent dat aan de individuele analyzers stringenter eisen moeten worden gesteld.

Uit een recente enquête in Nederland is gebleken dat een vorm van het virtuele analyzermodel werd toegepast in 89% van de laboratoria en daarmee de meest gebruikte basis van SQC-inrichting was. Van de laboratoria paste 22% het virtueel analyzermodel locatie-overstijgend toe [van Rossum et al., 2024].

De condities waaronder in een virtueel analyzermodel wordt gemeten houdt in dat er sprake is van meer bronnen van variatie: verschillende analyzer, verschillende laboratoria, mogelijk zelfs verschillende typen analyzers (dan is er sprake van reproduceerbaarheid condities). Ongecorrigeerde (en oncorrigeerbare) niveauverschillen tussen analyzers zullen gaan bijdragen aan de imprecisie, zodra de imprecisie berekend wordt door alle uitkomsten te combineren.

De SD's die worden berekend met het virtuele analyzermodel zijn meestal hoger dan SD's van één analyzer. De prestatie-eisen zoals afgeleid voor één analyzer zijn daarmee niet zonder meer toepasbaar voor de virtuele analyzer. Analyzers worden individueel gecontroleerd op afwijkingen en de bijdragen van een individuele analyzer aan de totale SD dient bekend te zijn.

Correctie van bias bij een afwijkende analyzer behoort tot de mogelijkheid om de SD van de virtuele analyzer te verbeteren.

7.4.2 Grondslag SD-waarde

Tot slot zijn er laboratoria die de SD niet alleen baseren op de analytische methode-implicisie, maar ook rekening houden met klinische tolerantie van de betreffende meetmethode. Strikt genomen is dit geen berekende SD meer. Hierbij wordt in het algemeen de SD ruimer gesteld dan enkel op basis van de methode-implicisie. Veelal wordt deze methode in combinatie met standaard 2SD controleregels gebruikt. In feite heeft de SD geen directe relatie meer met de methode-implicisie op het betreffende meetniveau. Overwegingen voor deze systematiek is dat de Levey-Jennings grafieken (ogenschijnlijk) eenduidig te interpreteren zijn en dat er gebruik gemaakt kan worden van een standaard set aan 2SD controleregels. Dit is met name van belang indien gebruik wordt gemaakt van informatiesystemen die nog sterk zijn georiënteerd op de conventionele Westgard QC-regels. Beperkingen in QC-systemen en/of -software spelen soms ook een rol. Aangezien de methode feitelijk onjuist is (er is immers geen sprake van een standaarddeviatie), wordt deze door de werkgroep ontraden. Het is het dringende advies van de werkgroep om bij de inrichting van het SQC gebruik te maken van de daadwerkelijk behaalde analytische SD.

7.5 Westgard QC

De Westgardregels [Westgard et al., 1981] vormen een combinatie van bewakingsregels die zodanig zijn opgesteld, dat voldaan wordt aan twee voorwaarden [Westgard en Burnett, 1990]: 1) er is minder dan 5% foutpositieve resultaten van interne kwaliteitscontrole; (*probability for false rejection* $P_{fr} < 0,05$); 2) de sensitiviteit (of *power*) voor het aantonen van een significante fout (zowel van bias als van toevallige fout/implicisie) is bij voorkeur 90% (*probability for error detection* $P_{ed} > 0,90$).

De in 1981 beschreven *multirules* (beter bekend als Westgard regels [Westgard et al., 1981] zijn nog altijd algemeen gangbaar, hoewel Cotlove et al. reeds in 1970 beschreven dat de tolerantiegrenzen per bepaling afhankelijk zijn van biologische en analytische variatie.

In de oorspronkelijke publicatie van Westgard et al. [1981] werd de combinatie van bewakingsregels toegepast, als aan het ingangscriterium was voldaan, namelijk overtreding van de 1_{2s} regel. In latere versies van de bewakingsregels is deze voorwaardelijke regel vervallen, en worden powerfuncties berekend op basis van een combinatie van bewakingsregels zonder voorwaardelijk criterium. De meest gebruikte Westgard regels zijn:

- $1_{2s}, 1_{3s}$ de tolerantie limiet wordt overschreden bij ± 2 (resp. 3) SD;
- 2_{2s} overschrijding door 2 achtereenvolgende controlemetingen plus 2 SD of min 2 SD;
- R_{4s} overschrijding van 1 controlemeting plus 2 SD en een naastliggende min 2 SD;
- 4_{1s} 4 achtereenvolgende controlemetingen zijn meer dan plus 1 SD of min 1 SD.

Zie ook [Bijlage E](#).

7.6 Six Sigma

7.6.1 Inleiding

Een veelgebruikte benadering voor het bepalen van de analytische kwaliteit is het gebruik van Six Sigma. Six Sigma is zowel een meetinstrument, methode als managementstrategie. Als meetinstrument is de sigmascore (*Sigma Metric*) een maat voor de kwaliteit van een proces, product of dienst. Deze sigmascore, verkregen door de variabiliteit van het proces/product te vergelijken ten opzichte van de gedefinieerde tolerantie, beschrijft de kwaliteit van een proces aan de hand van de kans op het ontstaan van een fout buiten de tolerantiegrenzen. Een proces met een sigmascore van 6 geeft aan dat er zes standaarddeviaties aan elke kant van de vastgestelde tolerantie ‘passen’; hiermee is de waarschijnlijkheid voor het optreden van een klinisch relevante meetfout zeer klein, volgens het Six Sigma model namelijk minder dan 3,4 defecten per miljoen (*Defects Per Million Opportunities*, DPMO, Tabel 7.2). In dit model wordt onderscheid gemaakt tussen ideale omstandigheden zonder bias en routineomstandigheden met een aangenomen kwaliteitsverlies van 1,5 SD.

De sigmascore kan worden gebruikt als managementstrategie omdat processen met een hogere sigmascore een grotere tolerantie hebben voor een eventuele verstoring. Aangenomen wordt, dat er een relatie bestaat tussen de sigmascore en het optreden van klinisch significante afwijkingen, en dat daarom deze processen minder streng en vaak bewaakt hoeven te worden ten opzichte van processen met een kleinere sigmascore [Badrack en Theodorsson, 2024].

Tabel 7.2. Sigmascore en corresponderende foutscore (DPMO)

Sigmascore	DPMO zonder bias	DPMO bij een bias van 1,5 SD
1	317.400	697.700
2	45.400	308.770
3	2.700	66.810
4	63	6.210
5	0,57	233
6	0,002	3,4

Het concept en de toepassing van Six Sigma in de laboratoriumdiagnostiek is uitvoerig benoemd in de voorgaande versie van de leidraad [NVKC Leidraad Analytische Kwaliteitscontrole met behulp van Six Sigma, 2015], en in de literatuur [Gras en Philippe, 2007; Westgard en Westgard, 2017;

Schoenmakers et al., 2011]. De sigmascore wordt vastgesteld tijdens validatie/verificatie om de benodigde monitoring in te stellen. Conventioneel wordt hierbij gebruik gemaakt van de volgende vergelijking:

$$\text{Sigma} = (\text{TEa} - \text{bias}) / \text{CVa}$$

TEa: *total allowable error*

CVa: analytische variatie (langlopend)

In de praktijk wordt er gewerkt met deze actuele sigmascore zonder de relatie met DPMO (Tabel 7.3) daarbij te betrekken. Met behulp van Westgard (multi)regels kan op basis van de sigmascore de vereiste analytische QC-monitoring bepaald worden. Binnen de klinische chemie wordt aangehouden dat we van een robuuste test spreken als een (spontane) verstoring in de test van maximaal 1,5 SD nog steeds een analytisch en klinisch acceptabele uitslag genereert.

Tabel 7.3. Interpretatie van de sigmascore voor de dagelijkse praktijk

Sigmascore	Duiding
< 3	(zeer)matige analytische kwaliteit, niet met <i>multirules</i> QC te beheersen
3-4	geschikt voor beoogde doel, echter frequente controle noodzakelijk
4-6	geschikt voor beoogde doel en met eenvoudige controle beheersbaar
> 6	uitstekende kwaliteit

7.6.2 Tekortkomingen van TEa-gebaseerde sigmascores en alternatieven

Het is belangrijk te beseffen dat het Six Sigma concept volgens Westgard is gebaseerd op de *total allowable error* (TEa) waarbij het error-budget bestaat uit een maximaal budget voor de toegestane variatie dat opgeteld wordt bij het maximale budget voor de toegestane bias. Budgetten voor variatie en bias zijn in dit concept daarmee onafhankelijk van elkaar en worden cumulatief bij elkaar gevoegd om de TEa te berekenen. Reeds eerder is besproken dat dit concept feitelijk onjuist is doordat de maximale bias en variatie niet onafhankelijk zijn en daarmee een overschatting geven van de TEa (zie paragraaf 7.3.2). Immers, in het geval dat de TEa voor een laboratoriumbepaling overschat wordt, geeft dit tevens een overschatting van de daaruit berekende sigmascore en resulteert dit in minder stringente Westgard-QC-bewakingsregels en/of een lagere frequentie van IQC-bepalingen dan feitelijk van toepassing is voor die test.

Het gebruik van Six Sigma op basis van TEa is tevens afwijkend ten opzichte van andere toepassingen van Six Sigma, wat kan leiden tot interpretatiefouten [Coskun et al., 2019; Badrick en Theodorsson, 2024]. Hierin zijn de aannames dat de bias constant is, terwijl deze in de praktijk kan variëren [Thelen en van Schrojenstein Lantman, 2024]. Verder zijn assays niet altijd traceerbaar en EQA niet altijd commuteerbaar, waardoor in de praktijk laboratoria vaak de bias op nul zetten en met louter de CVa de sigmascore berekenen. Hierdoor worden prestaties mogelijk overschat.

Als alternatief kunnen laboratoria overwegen om de sigmascore op basis van de toelaatbare meetonzekerheid te bepalen. Dit kan aan de hand van de volgende vergelijking:

$$\text{Sigma} = (\text{MAU}) / \text{uRw}$$

MAU = *expanded allowable measurement uncertainty*

uRw = lange-termijnimprecisie op basis van reproduceerbaarheid

Merk hierbij op dat ten opzichte van de TEa-sigma score er gebruik wordt gemaakt van lange-termijn onzekerheid (*uncertainty under repeatability conditions*, u_{Rw}) zoals bepaald over meerdere reagenslots, kalibraties en episodes van groot en klein onderhoud. In het geval van een virtuele analyzer zit hier ook de tussen-apparaatvariatie bij; dit is de standaarddeviatie over elke bias van elk apparaat (u_{an}). De totale lange-termijnonzekerheid van de virtuele analyzer ($u_{Rw,v}$) wordt berekend op basis van u_{an} en de *gepoolde* binnen-apparaatvariatie (u_p) middels de volgende vergelijkingen:

$$u_p = \sqrt{\frac{u_{Rw,1}^2 + u_{Rw,2}^2 + \dots + u_{Rw,n}^2}{n}}$$

u_p = *gepoolde* binnen-apparaatvariatie

$u_{Rw1..n}$ = lange-termijnimpresie van analyzers 1-n

n = aantal analyzers

$$u_{Rw,v} = \sqrt{u_p^2 + u_{an}^2}$$

$u_{Rw,v}$ = lange-termijnonzekerheid van virtuele analyzer

u_p = *gepoolde* binnen-apparaatvariatie

u_{an} = tussen-analyzervariatie (als standaarddeviatie over de bias van elk apparaat binnen de virtuele analyzer)

Aangezien de toelaatbare meetonzekerheid louter op binnen-persoonsvariatie wordt bepaald kunnen de sigma scores ten opzichte van het TEa-model lager uitvallen. Een praktische compensatie hiervoor is het gebruiken van minimale prestatiespecificaties voor het bepalen van de meetonzekerheid ($MAU = 0,75 \times CVi$), en hiermee de toelaatbare *expanded measurement uncertainty* ($MAU = 2 \times MAu$) te bepalen.

Een nadere uitwerking is te vinden in [Bijlage F](#).

7.6.3 Toepassingsgebieden Six Sigma, en beperkingen hierin

De APS waarmee de TEa voor een bepaalde laboratoriumbepaling wordt vastgesteld, wordt op diverse bronnen c.q. data gebaseerd, zie ook paragraaf 7.3. Het is belangrijk te beseffen dat de keuze van APS van invloed is op de resulterende sigma score. De laboratoriumorganisatie dient haar overwegingen voor de gekozen APS, en daarmee het TEa/MUa en berekende sigma score, voor alle bepalingen inzichtelijk te kunnen maken.

De berekening van de sigma score kan voor meer situaties worden toegepast:

- verificatie of validatie van nieuwe analyzersystemen;
- selectie van Westgard (multi)-QC-regels voor bewaking van individuele laboratoriumtests op basis van een berekende sigma score;

- identificeren van niet-robuste tests en ontwerp van alternatieve bewakingsregels;
- bewaking van een virtuele analyzer.

Uitwerking van deze verschillende toepassingsgebieden is te vinden in [Bijlage B](#).

De sigmascore met bijbehorende Westgard (multi)-QC-regels wordt vastgesteld tijdens de introductie van een laboratoriumtest op basis van verificatie- of validatiedata. Echter, het is een kunstfout als er gedurende de levensduur van de analyzer of testsysteem nooit een herbeoordeling van de sigmascore zou plaatsvinden. De frequentie waarmee herberekening van de sigmascores op basis van de dan geldende prestatie-eisen moet gebeuren, is wetenschappelijk niet of nauwelijks vast te stellen. Geadviseerd wordt om de sigmascore periodiek te beoordelen, bijvoorbeeld eens per 2-4 jaar, of als anderszins aanleiding hiervoor is ([Aanbeveling 6](#)).

Bij bepalingen die in de loop van de tijd instabiel blijken, door welke oorzaak dan ook, zal een hogere frequentie aan de orde zijn. Het is goed mogelijk dat een herberekening leidt tot een lagere of hogere sigmascore en daarmee tot strengere of minder strenge IQC-regels en daardoor tot (minder) frequentere kalibraties, etc. Om te voorkomen dat dit proces zich telkens herhaalt (jojo-effect) kan het laboratorium besluiten om strengere Westgard (multi)-QC-regels te handhaven dan op grond van de sigmascore op dat moment noodzakelijk is. Tegelijkertijd leidt vaker meten of strengere Westgard regels tot frequentere alarmen die een actie behoeven (bijv. kalibratie). Vaker ingrijpen resulteert dan niet in een stabiel proces maar verhoogt vaak de spreiding (de ingreep is vaak onterecht).

De Six Sigma benadering richt zich voor het risico primair op de kwaliteit van het analytische proces ten opzichte van klinische tolerantiegrenzen en niet zozeer op de daadwerkelijke klinische omvang, ofwel het aantal patiënten of patiëntenmonsters dat wordt blootgesteld aan een bepaald risico. Daarnaast worden andere risicofactoren niet meegenomen, zoals bekende methode-instabiliteit, waarschijnlijkheid van fout, reële klinische impact op basis van foutieve klinische interventie, of tijd tussen het beschikbaar zijn van een resultaat en detectie daarvan. Het laboratorium kan deze restrisico's meenemen in de selectie van periodiciteit en evaluatiemomenten bij het gebruik van Six Sigma.

7.7 Risico-gebaseerde QC

Als vervolg op *Sigma Metrics* is er initieel vanuit Bio-Rad en later ook vanuit Westgard Inc. een methode ontwikkeld die, naast de sigmascore van het proces als risicofactor, ook rekening houdt met het aantal patiënten of patiëntenmonsters dat wordt blootgesteld aan het risico [van Rossum et al., 2021].

Een belangrijke parameter hierin is het maximaal aantal mogelijke fouten binnen een analytische run dat een klinische tolerantiegrens overschrijdt maar niet wordt gedetecteerd met een specifieke controle regel (in de literatuur aangeduid met de term $E(N_{ur})$ [Bayat, 2017]). Op basis van deze parameter wordt vervolgens voor ieder analytisch de maximale grootte van een analytische run vastgesteld; het maximaal aantal patiëntenmonsters tussen twee controlemomenten. Deze is medebepalend voor de keuze van statistische (Westgard) controleregels.

Deze werkwijze is mogelijk van meerwaarde maar is praktische implementatie is (nog) lastig voor ruime toepassing in veel laboratoria.

7.8 Patient-Based Real-Time Quality Control (PBRTQC)

Het lopend patiëntengemiddelde is in de jaren '60 van de vorige eeuw voor het eerst beschreven als *Average of Normals* (AoN). Sinds deze eerste beschrijving zijn er nieuwe algoritmen ontwikkeld en is

de toepassing beschreven voor het gebruik als lopend patiëntgemiddelde. Voorbeelden van algoritmen van lopend patiëntgemiddelde binnen medische laboratoria zijn: lopend gemiddelde, lopende mediaan, exponentieel gewogen lopend gemiddelde (EWMA), en de algoritmen ontwikkeld door Brian Bull (XbarM, XbarB). Deze algoritmen worden veelal ondersteund door software aanwezig op klinische laboratoria (bijv. analyzersoftware, bijv. hemocytometrie-apparatuur, laboratoriuminformatiesysteem).

Ondanks de beschikbaarheid van bovengenoemde instrumenten is de praktische implementatie een grote uitdaging gebleken. Door de recente ontwikkelingen in het simuleren van fout-detectie eigenschappen van lopende patiëntgemiddelden, en verkrijgen van laboratoriumspecifieke instellingen, zijn hierin recent aanzienlijk vorderingen gemaakt. Toch blijft het praktisch implementeren een tijdsintensief proces dat (net als de gebruikelijke SQC) ook onderhouden dient te worden. Wel kunnen er aanzienlijke kwaliteitswinsten worden behaald.

Het voordeel van dergelijke systemen is dat deze continue analytische kwaliteitsbewaking binnen een analytische run mogelijk maken door ieder nieuw verkregen resultaat te beoordelen. Hieronder vallen controleregels (limiet, delta en multivariabel) die worden gebruikt voor de selectie van een noodzakelijke handmatige technische confirmatie, als lopend patiëntgemiddelde op basis van verschillende soorten algoritmes [van Rossum, 2022]. Aangezien dergelijke systemen sterk afhankelijk zijn van de patiëntenpopulatie van een laboratorium, is het lastig een objectief inzicht te krijgen in de foutdetectie eigenschappen. Deze kunnen namelijk aanzienlijk verschillen tussen klinische laboratoria. Recentelijk zijn er *in vitro* simulatiemethoden ontwikkeld en beschikbaar gekomen voor klinische laboratoria, waarmee laboratoriumspecifieke PBRTQC-instellingen kunnen worden verkregen. Daarnaast kunnen op basis van deze systemen ook de foutdetectie-eigenschappen voor een specifiek laboratorium worden vastgesteld. Dit laatste is relevant voor het beoordelen of PBRTQC-procedures een relevante aanvulling vormen voor reeds bestaande IQC. Meestal worden PBRTQC-methoden gebruikt ter aanvulling van de IQC bij meetmethoden waarbij de SQC niet toereikend is voor een adequate analytische kwaliteitsbewaking. Dit is bijvoorbeeld het geval bij meetmethoden met een sigmascore ≤ 4 , continue meetmethoden waarbij er een reëel risico is op het snel ontstaan van meetfouten of meetmethoden waarbij er geen stabiel of commuteerbaar QC materiaal beschikbaar is [van Rossum, 2020]. Door middel van twee prospectieve studies is voor lopende patiëntgemiddelden aangetoond dat dergelijke systemen het mogelijk maken om klinisch relevante fouten voortijdig (voor geplande QC meting) te detecteren en te corrigeren [van Rossum en van den Broek, 2020; van Rossum en Kemperman, 2017; van Rossum et al., 2021]. Daarnaast is beschreven dat PBRTQC, mits goed gevalideerd en beschikkend over adequate foutdetectie-eigenschappen, ook kan worden gebruikt om de controlefrequentie van interne kwaliteitscontroles monsters te verminderen [Fleming en Katayev, 2015].

Patient-Based Real-Time QC (PBRTQC) kan in sommige gevallen de analytische kwaliteitsbewaking versterken. Een dergelijk systeem kan worden ingezet in het geval van evidente tekortkomingen van of beperkingen aan het uitvoeren van IQC op basis van controlematerialen ([Aanbeveling 7](#)).

7.9 Borgen vergelijkbaarheid tussen analyzersystemen (binnen laboratoriumorganisatie)

De vraag is wanneer en op welke gronden een *recall*-procedure zou moeten worden uitgevoerd (ofwel het hertesten van patiëntenmonsters bij QC-uitslagen buiten de acceptatiegrenzen). Modellen gebaseerd op biologische variatie zijn te strikt voor bepaalde analieten met geringe biologische variatie, zoals natrium. Minimale prestatiekenmerken worden dan afgeleid van de *state of the art*; een algemene methode hoe deze acceptatiegrenzen moeten worden afgeleid ontbreekt nu echter nog [Braga en Panteghini, 2020; Çubukçu et al., 2021].

In de toenemende ontwikkeling naar grotere laboratoriumorganisaties met meerdere locaties, is dit een relevant thema. Voor bepalingen die slechts op een locatie en/of instrument worden verricht, wordt de doelwaarde en doel-SD separaat bepaald, de bijbehorende Westgard regels afgeleid en wordt zowel de interne kwaliteitscontrole (IQC) als de externe kwaliteitscontrole (EQC) ingericht. Indien gekozen wordt voor het multi-analyzerconcept, is voor alle analyzers in een netwerk een identieke doelwaarde en doel-SD van kracht. Een andere optie is dat er voor iedere analyzer afzonderlijke doelwaarden en doel-SD's worden berekend, waarbij de performance van alle individuele analyzers dient te vallen in een vooraf gestelde bandbreedte of *error budget*. ISO 15189 heeft voor de vergelijkbaarheid van onderzoeksresultaten een specifiek normelement (7.3.7.4).

Hoofdstuk 8 Praktische inrichting analytische kwaliteitscontrole

8.1 Stroomschema: leeswijzer

Klinisch chemische laboratoria hebben behoefte aan praktische handvatten voor de inrichting van de analytische kwaliteitsbewaking met behulp van interne kwaliteitscontroles. Het onderstaande stroomschema (Figuur 8.1) wil daarin voorzien. Dit stroomschema is alleen bedoeld voor diagnostische verrichtingen waarbij het laboratorium een kwantitatieve uitslag genereert. In het stroomschema is een aantal belangrijke beslispunten opgenomen waarbij het laboratorium een keuze dient te maken. Dit is grafisch weergegeven met behulp van ruiten. Daarnaast zijn in logische volgorde alle stappen opgenomen waarbij het laboratorium per test criteria dient vast te stellen en/of berekeningen moet uitvoeren. Waar nodig is per stap steeds een verwijzing gedaan naar de desbetreffende paragraaf met de relevante theoretische achtergrond.

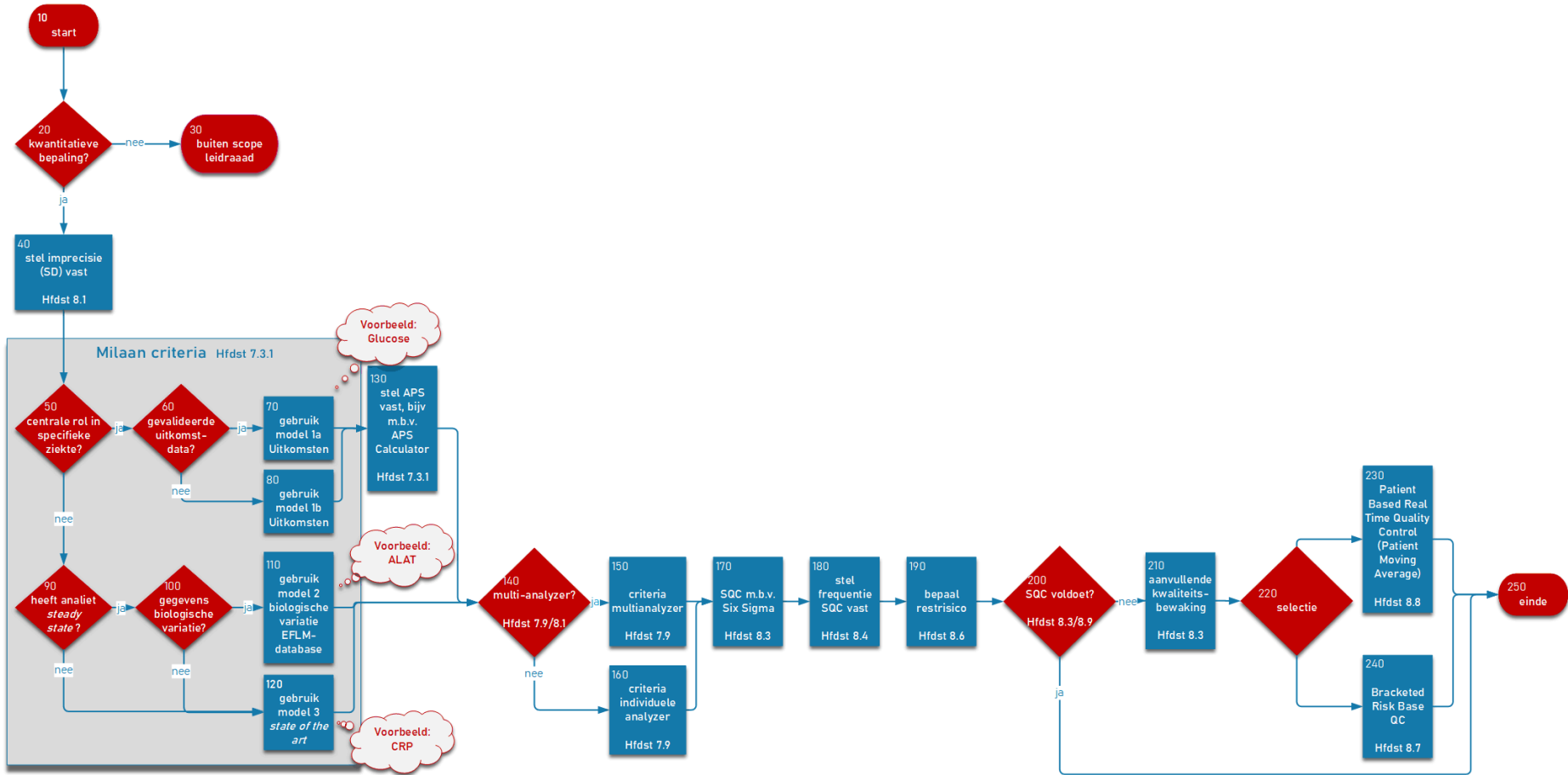
In dit hoofdstuk wordt een aantal onderdelen extra toegelicht, zoals de aard van het interne controlemateriaal, de wijze waarop de standaarddeviatie wordt vastgesteld en hoe voor een specifieke test na inrichting van de analytische kwaliteitscontrole het restrisico kan worden vastgesteld.

8.2 Keuze bewakingsmodel per test

In het begin van het stroomschema dient het laboratorium per test de APS vast te stellen op basis van objectieve criteria die ontleend zijn aan de modellen zoals voorgesteld op de *2014 Milan Strategic Conference*. Zoals bekend kan de APS voor de maximaal te accepteren meetonzekerheid (MU) van een diagnostische parameter gekozen worden op basis van 1) het effect van de analytische performance op de klinische uitkomst, 2) de biologische variatie, of 3) de *state-of-the-art* performance, te definiëren als de op dit moment technisch hoogst haalbare analytische prestatie tegen redelijke inspanningen en kosten.

De modellen gebruiken verschillende principes en er is geen sprake van een hiërarchie. Alles hangt af van de biologische en klinische karakteristieken van de meetgrootte (*measurand*) [Braga en Panteghini, 2021]. Praktisch betekent dit dat als er voor een test voor model 1 gekozen wordt, de APS wordt vastgesteld op basis van de impact van de spreiding rondom een klinische beslisgrens. Voor model 2 adviseren we gebruik te maken van de [EFLM Biological Variation Database](#). Ten slotte geldt voor model 3 dat de maximale technische analytische prestatie doorgaans vanuit de wetenschappelijke en/of vakliteratuur kan worden gehaald [Braga en Panteghini, 2021]. Soms moet echter worden volstaan met de *'state of our own art'* (de binnen het betreffende laboratorium maximaal haalbare analytische prestatie). Ter illustratie is voor een aantal bepalingen een uitwerking gegeven in Tabel 8.1.

Figuur 8.1. Stroomschema voor de inrichting van de analytische kwaliteitsbewaking met behulp van interne kwaliteitscontrole



Tabel 8.1. Voorbeeld voor het instellen van QC voor drie bepalingen op basis van Milaan 1 (uitkomst), Milaan 2 (biologische variatie) en Milaan 3 (*state of the art*). Afhankelijk van de afwezigheid van bias wordt gekozen voor MAU of TEa als basis voor APS. De vastgestelde lange termijnimplicie voor twee niveaus (N1, N2) QC-materiaal wordt hiermee vergeleken om tot sigmascores te komen. In geval van het voorbeeld van Milaan 2 heeft het laboratorium reeds PBRTQC in gebruik genomen en is *multirule* controle met IQC niet nodig.

APS op basis van	Voorbeeld	Aanwezigheid bias?	Vastgestelde imprecisie (lange termijn)	APS op basis van	APS	Sigma	Overweging	Monitoring kwaliteit
Milaan 1 – uitkomst	Glucose	n.s. (op basis van EQA)	N1 (4 mmol/L): 1,5 % N2 (10 mmol/L): 1,7 %	MAU	Nuchter glucose: 6,0 % [Panteghini, 2024]	N1: 4,0 N2: 3,5	N2 past niet bij beoogd gebruik 'nuchter', APS 6,9 % op basis van Milaan 2 leidt tot N2 sigma van 4,0	Op basis van sigma 4,0: Westgard <i>multirules</i> $1_{3s}/2_{2s}/R_{4s}/4_{1s}$ [Tabel 2 NVKC Leidraad Analytische Kwaliteitscontrole met behulp van Six Sigma, 2015]
Milaan 2 – biologische variatie	ALAT	1 % (niet groot genoeg om te corrigeren)	N1 (10 U/L): 6,0 % N2 (60 U/L): 3,0 %	TEa	18,5% [EFLM Biological Variation Database]	N1: 2,9 N2: 5,8	Lab kan geen sigma bewaken per QC-niveau: bewaak kleinste. Lab heeft PBRTQC breed in gebruik.	PBRTQC met Westgardregel $1_{2,5s}$
Milaan 3 – <i>state of the art</i>	CRP	n.s.	N1 (10 mg/L): 2,0 % N2 (100 mg/L): 1,8 %	MAU	12,6% [SKML Analyte Specifications, geraadpleegd op 25-02-2025]	N1: 6,3 N2: 7,0	> 6 sigma, mag 4,35 SD (of afgerond 4 SD)	Westgardregel $1_{4,35s}$

8.3 Aard interne controlemateriaal

Bij de keuze voor het controlemateriaal, dat wordt gebruikt voor de interne kwaliteitsbewaking, zijn vier overwegingen van belang:

- Alleen controlemateriaal, dat stabiel is en waarvan de matrix zo goed mogelijk overeenkomt met het type lichaamsmateriaal waarin de diagnostische bepalingen worden verricht, bijvoorbeeld in plasma, serum, volbloed, urine of liquor, is geschikt.
- Ten tweede dient het concentratieniveau van het te bewaken analyet in het controlemateriaal zoveel mogelijk overeen te komen met het niveau van de klinische beslisgrenzen waarvoor deze bepaling wordt gebruikt. Bij de keuze van geschikte controlematerialen dient ook rekening gehouden te worden met bijzondere patiëntpopulaties.
- Waar mogelijk heeft het de sterke voorkeur controlematerialen van derden te betrekken in plaats van controlematerialen van de leverancier te gebruiken. Hoewel de commuteerbaarheid van interne kwaliteitscontrole materialen vaak niet aangetoond is, is dit zelden een probleem voor de algemene klinische chemie [Kim et al., 2020]. Bij immunoassays en hematologische bepalingmethoden is de kans op non-commuteerbaarheid groter [Kim et al., 2020; van Schrojenstein Lantman et al., 2022; Vidali et al., 2021]. In gevallen waarin twijfel is over de commuteerbaarheid kunnen *gepoolde* patiëntmonsters als mogelijk alternatief gebruikt worden [van Schrojenstein Lantman et al., 2022, Braga en Panteghini, 2021]. Toepassing van PB(RT)QC kan hiernaast een alternatief bieden gezien deze methode van QC inherent commuteerbaar is [van Rossum, 2020].
- Indien er voor een bepaling geen commercieel controlemateriaal beschikbaar is, kan er zelf controlemateriaal bereid worden, bijv. op basis van samengevoegde (*gepoolde*) patiëntenmonsters, of door blanco materiaal te *spiken* (voor geneesmiddelen analyses is het namelijk niet altijd mogelijk om geschikte patiëntenmonsters in voldoende mate te verzamelen). Deze aanpak is ook bruikbaar indien een bepaling extra bewaking vereist in verband met suboptimale analytische prestatie of als er diagnostiek voor een specifieke patiëntenpopulatie plaatsvindt, bijv. dialysepatiënten. Het laboratorium dient de houdbaarheid en stabiliteit te valideren en materiaal te bereiden met klinisch relevante concentratieniveaus.

8.4 Berekening van de spreiding

Een basisonderdeel van de statische kwaliteitscontrole (SQC) is het vaststellen van de meetmethode-imprecisie in de vorm van de standaarddeviatie (SD) of procentueel met de variatiecoëfficiënt (CV-waarde) voor individuele kwaliteitscontrole materialen. Hiervoor worden in de praktijk verschillende methoden gebruikt.

De imprecisie of spreiding van het interne controlemateriaal dient per analyzer en concentratieniveau bepaald te worden. Naast de bepaling van de SD-waarde op basis van kwaliteitscontroles kunnen ook patiëntenmonsters worden gebruikt. Indien het interne controlemateriaal meerdere lotnummers omvat, zal de variatie voor ieder lot bekend moeten zijn om vervolgens een *gepoolde* variatie te kunnen berekenen. Als voorbeeld voor drie verschillende lots geldt de volgende kwadratensom [ISO/TS 20914]:

$$SD_{\text{tot}} = (SD_a^2 + SD_b^2 + SD_c^2 / 3)^{\frac{1}{2}}$$

NB: de vergelijking gaat uit van de situaties dat er per lotnummer ongeveer gelijkwaardige meetmomenten zijn. Als dit niet het geval is geniet gewogen *pooling* met inachtneming van de hoeveelheid monsters per *pool* de voorkeur.

De berekening van de standaarddeviatie (SD) kan per analyzer gebeuren of voor een virtuele analyzer met meerdere systemen. Voor een correcte SD-berekening zijn de data normaal verdeeld. Dat

betekent dat er in het geval van een virtuele analyzer geen significante bias [Stöckl et al., 1995; ISO/TS 20914:2019; Coskun, 2024] mag bestaan tussen verschillende analyzers aangezien dit een normale verdeling verstoort. Bovendien is de berekening van de SD zoals weergegeven in de bekende Levey-Jennings grafiek gevoelig voor uitbijters. Er zijn verschillende methoden beschikbaar voor uitbijteranalyse [Grubbs, 1950; 1969; Dixon, 1950; 1951; 1953]. Er bestaan ook methoden om de SD af te leiden, waarbij geen uitbijterselectie nodig is (zie [Bijlage G](#)).

8.5 Selectie van periodiciteit

Zowel het meetmoment op de dag, als de frequentie van bepalen van kwaliteitscontrolematerialen dient door het laboratorium te worden gedefinieerd; zie ook paragraaf 8.7. In algemene zin kan de periodiciteit op de volgende manieren worden vastgesteld:

- Op basis van de berekende sigmascore volgens de Six Sigma methodiek of *Bracketed Risk-Based* methodiek. De frequentie en te meten niveaus van de interne kwaliteitscontroles zijn voor deze methode getabelleerd [Schoenmakers et al., 2011].
- De *omvang* van de *run* tussen twee meetmomenten van de interne kwaliteitscontrole. Oftewel, het gaat hier om de impact van een afwijkende QC op de noodzaak van een eventueel overmeten van (veel) patiëntenmonsters of het starten van een *recall*-procedure.
- Praktische aspecten zoals het tegelijk kunnen meten van meerdere controles uit een hetzelfde controlemateriaal zoals bij een multi-parameter-analyzer. Voor vaststelling van de periodiciteit zullen in dit geval bepalingen met de laagste analytische kwaliteit de doorslag geven.
- De impact van klinische besluitvorming tussen verschillende aanvragers of patiëntcohorten. Voorbeeld: natrium. Natrium is een frequent gemeten analiet met een matige meetkwaliteit. Zo kan bijvoorbeeld de intensive care afdeling meer rigoureuze klinische beslissingen nemen op basis van het natrium ten opzichte van een nierpoli. Indien in dit voorbeeld er een dagelijks prikmoment van intensive care patiënten bestaat, kan worden overwogen om voor of na dit moment een kwaliteitscontrolemateriaal te meten zodat de tijd-tot-(eventuele) *recall* verkort wordt.

8.6 Selectie van Sigma (*multi*)rules

De specifieke controleregels, Westgard regels of *multirules* worden (samen met de periodiciteit) ingesteld op basis van de *Sigma Metrics*-systematiek [Schoenmakers et al., 2011]. Hierbij wordt gebruik gemaakt van op SD-gebaseerde controleregels (Westgard regels). De *Sigma Metric* methode heeft sinds het verschijnen van NVKC-leidraad uit 2015 ruim ingang gevonden in de klinisch chemische laboratoria en wordt hier niet verder toegelicht. Zie voor nadere uitwerking en voorbeelden in [Bijlage F](#).

8.7 Bracketed Risk-Based Quality Control

Zoals eerder besproken, wordt bij *Bracketed Risk-Based Quality Control* voor ieder analiet de maximale grootte van een analytische *run* vastgesteld op basis van 1) het patiëntenrisico, 2) het niet detecteren van een relevante analytische fout door een gekozen SQC opzet en 3) het aantal patiënten dat wordt blootgesteld aan dit risico. Westgard Inc. heeft recentelijk hiervoor een online calculator ontwikkeld waarmee laboratoria zelf, op basis van TE, eventuele bias en methode-imprecisie, de controleregels en maximale grootte van een analytische *run*, op basis van dit $E(N_{ur})$ model, kunnen bestuderen. Zie hiervoor: tools.westgard.com/frequency_calculator.shtml. Daarnaast biedt ook BioRad een dergelijke calculator aan.

De *Bracketed Risk-Based Quality Control* methode voorziet vooral in een ondersteunend instrument voor meetmethoden met een sigmascore tussen de 3,5 en de 6. Voor meetmethoden met sigmascores onder de 3,5 worden zulke kleine *run*-groottes verkregen dat dit veelal in de praktijk niet haalbaar en wenselijk is. Zo worden slechts tien patiëntenmonsters tussen twee controlemomenten getolereerd met meerdere controleniveaus en een Westgard *multirule* ($2_{2s} / 1_{3s} / R_{4s}$). Bij meetmethoden met een sigmascore groter dan 6 echter 1837 patiëntenmonsters met twee controleniveaus en een simpele Westgard regel (1_{3s}). Voor veel meetmethoden in gemiddelde klinische laboratoria zou dit dan betekenen dat er slechts sporadisch en over meerdere dagen, slechts een enkel controlemoment nodig zou zijn. De impact van een eventuele *recall*-procedure is daarmee echter wel groot.

Een andere uitdaging voor de praktische implementatie van *Bracketed Risk-Based Quality Control* is dat de maximale *run*-grootte voor ieder individueel analiet wordt vastgesteld. In de praktijk wordt echter veelvuldig gebruik gemaakt van controlematerialen waarmee tegelijkertijd de kwaliteitscontroles voor meerdere analieten tegelijk worden uitgevoerd op een multi-parameter-analyzer. Hierdoor is het om praktische redenen vaak juist wenselijk om op dezelfde momenten de controlemetingen voor deze analieten uit te voeren.

Concluderend is het inrichten van *Bracketed Risk-Based Quality Control* een inhoudelijk uitdagende en praktisch complexe uitdaging. Ook zijn er nog geen prospectieve studies geweest die de meerwaarde hiervan hebben aangetoond. Om deze reden is het voor de meeste laboratoria niet haalbaar en wenselijk om dit als basis voor de SQC te gebruiken. Aangezien deze leidraad de huidige ontwikkelingen in het veld wil benoemen, wordt er toch aandacht aan dit concept gegeven.

8.8 Patient-Based Real-Time Quality Control (lopende patiëntengemiddelden)

Voor het implementeren van lopende patiëntengemiddelden kan gebruik worden gemaakt van het stappenplan in [Bijlage H](#).

8.9 Conventionele methode (2SD-gebaseerd)

Deze methode noemen we conventioneel omdat dit historisch gezien de gebruikelijke manier was voor klinisch chemische laboratoria om de kwaliteit analytisch te bewaken. Concepten zoals de Milaan-modellen spelen hierin geen rol. Enkel een stringente bewaking op slechts analytische gronden is bij deze methode van toepassing.

Er zijn zeker situaties denkbaar waarin het laboratorium voor deze strenge bewakingsregel kiest. Dit kan het geval zijn voor bepalingen waarbij in de literatuur geen consensus bestaat over de impact van de spreiding op de klinische uitkomst of beslissing. Of waar fysiologisch geen gegevens beschikbaar zijn of van toepassing wat betreft de biologische variatie. En waar ten slotte in de literatuur geen duidelijke state-of-art gegevens voor de toelaatbare spreiding gepubliceerd zijn. De APS kan dan op alleen analytische gronden worden vastgesteld met een strenge analytische bewakingsregel.

8.10 Risicoanalyse

8.10.1 Risico inschatting algemeen

Voor het inschatten van het klinisch risico kunnen verschillende benaderingen worden toegepast. Enkele meer uitgewerkte methoden zijn de Six Sigma en de *Bracketed Risk-Based* methoden waarbij

de inrichting van de SQC afhankelijk is van de klinische tolerantie op basis van APS. Vanuit de Six Sigma methodiek wordt het risico ingeschat op basis van de sigmascore. Zoals genoemd in paragraaf 8.6 is de methode voor het bepalen van sigmascores uitgewerkt in [Bijlage F](#). Meer recentelijk worden risico's en eventueel noodzakelijke beheersmaatregelen beoordeeld o.b.v. meer generieke procesmatige methodieken voor de inschatting van het risico.

Bij een standaard 2SD-methodiek, waarbij reële analytische imprecisie (SD/CV) wordt gebruikt is de inrichting van de SQC in principe niet herleidbaar naar de klinische tolerantiegrenzen en klinisch risico. Echter, in feite wordt in deze situatie een zeer strenge analytische kwaliteitscontrole toegepast die praktisch haalbaar is; daarmee zal worden voldaan aan de haalbare klinische tolerantie dan wel *best practice* criteria. Dit zal wel resulteren in veel onnodige QC-afwijkingen.

8.10.2 Bepalen van het restrisico

Na het inrichten van de interne kwaliteitscontrole op basis van een standaard op risico-gebaseerde systematiek kan het zijn dat er een relevant restrisico wordt geconstateerd. Dit kan verschillende redenen hebben zoals: de beperking van SQC om, binnen de praktische uitvoerbaarheid, kleine fouten te detecteren die klinisch relevant zijn (bijv. meetmethoden met sigmascores <4), om bij testen met een bekend risico op acuut falen van de meetmethode een snelle, actuele detectie van meetfouten te kunnen realiseren, en de afwezigheid van stabiele en/of commuteerbare controlematerialen. Dergelijk restrisico kan aanleiding zijn om te kijken of PBRTQC-instrumenten dit risico kunnen verminderen of dat bijvoorbeeld aanvullende interne QC-metingen, zoals extra frequente metingen buiten de standaard methodiek of het toevoegen van extra controle materialen, nodig zijn.

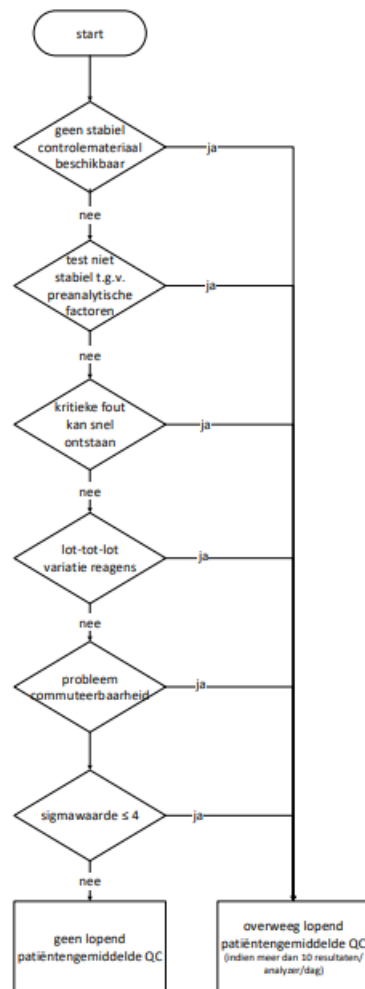
Een voorbeeld voor het vaststellen van restrisico en een beoordeling of lopende patiënten gemiddelden als instrument overwogen kunnen worden staat in [Figuur 8.2](#).

8.11 Borgen vergelijkbaarheid tussen analyzersystemen (binnen laboratoriumorganisatie)

Hieronder volgt een beschrijving onder welke voorwaarden een virtuele analyzer ingericht kan worden. Bij de initiële vaststelling van de doelwaarde van het IQC-level en de doel-SD is het van belang dat bij de verificatie van de individuele analyzers onderlinge bias zoveel mogelijk tot een minimum is gereduceerd. Indien één van de instrumenten toch een bias laat zien, dient de bias kleiner te zijn dan de toegestane (*allowable*) bias; de imprecisie dient kleiner te zijn dan toegestane imprecisie te zijn. Het virtuele analyzermodel is toepasbaar indien de bias, onderlinge bias en imprecisie van alle analyzers tezamen binnen de TEa, respectievelijk MAU voor de totale virtuele analyzer vallen. ([Aanbeveling 8](#)).

De initiële vaststelling van deze doelwaarden wordt verkregen door het ingelopen gewogen gemiddelde van alle analyzers die gebruikt worden voor een bepaalde bepaling. De overall-SD wordt berekend van alle analyzers een langlopende samengestelde variatiecoëfficiënt (CV) te nemen [ISO/TS 20914:2019]. Op basis hiervan kunnen de bijbehorende Westgardregels per bepaling worden geselecteerd. Bedenk dat de totale SD groter zal worden, ook al zijn alle analyzers binnen de norm. Er zal altijd een bijdrage zijn van tussen-analyzervariatie die niet met factoren kan worden gecorrigeerd. De APS is gebaseerd op de 1-analyzer situatie, voor een conglomeraat kan die norm lastig zijn. ISO 15189 heeft voor de vergelijkbaarheid van onderzoeksresultaten een specifiek normelement (7.3.7.4).

Figuur 8.2. Voorbeeld rest-risico en overweging bestudering lopende patiëntengemiddelden (MA=*Moving Average*) [Naar: van Rossum en van den Broek, 2019]



Belangrijk is dat de laboratoriumorganisatie zich dient te vergewissen van de continue geldigheid van de gehanteerde doelwaarde voor de IQC-levels en de doelwaarde-SD's. Hierbij geldt: hoe meer QC-data, des te groter de kans dat de doelwaarde 'juist' is en de doel-variantie realistisch is voor de uRw. Dit kan door periodieke herziening van de doelwaarden met eventueel actualisatie van de Westgardregels op basis van de opnieuw berekende sigmascore. Voor relevante bepalingen verdient het de voorkeur tevens een *Patient Moving Average* (PMA) in te richten zodat er parallel aan de resultaten van kwaliteitscontrole materiaal eveneens patiëntenmateriaal wordt meegenomen. Uiteraard geldt deze redenering eveneens voor individueel bewaakte analyzers, maar toepassing van PMA kan in het geval van een virtuele analyzer waaronder veel systemen en/of locaties 'hangen' een extra veiligheidsklep bieden (zie ook [Aanbeveling 7](#)).

In het geval van een virtuele analyzer wordt de frequentie waarmee IQC aan individuele apparaten wordt aangeboden gebaseerd op de sigmascore, overschrijding van controleregels en de consequenties inzake het aantal patiënten dat opnieuw opgeroepen moet worden bij een eventuele *recall*.

Redenerend vanuit (rest)risico's is er een aantal situaties te beschrijven dat aanleiding kan geven dat de door het laboratorium gestelde eisen van een virtuele analyzer niet worden gehaald.

- Reagensfout op een analyzer.
 - Eis: detectie door een afwijkende IQC.
- Reagensfout op alle analyzers
 - Eis: detectie door een trendbreuk van de IQC op alle analyzers
- Kalibratieprobleem op een analyzer
 - Eis: detectie door een afwijkende IQC op deze analyzer.
- Kalibratieprobleem op alle analyzers
 - Eis: detectie door een trendbreuk van de IQC op alle analyzers
- Hardware-probleem
 - Eis: detectie door een afwijkende IQC op deze analyzer.

Om te zorgen dat de IQC een goede schatter blijft voor de betrouwbaarheid van een laboratoriumbepaling, wordt gebruik gemaakt van EQC. In het geval van een virtuele analyzer bestaat er in principe geen hiërarchie tussen de analyzers, er bestaan dus geen zogenaamde moeder-dochter apparaten. Dat betekent dat iedere analyzer *at random* blootgesteld dient te worden aan een EQC. Dit betekent dat er voor een bepaalde laboratoriumbepaling in theorie slechts een set EQC gebruikt hoeft te worden met de premisse dat de laboratoriumorganisatie ervoor zorgt dat elke ronde van een ringonderzoek at random op een individuele analyzer wordt bepaald zodat na verloop van tijd iedere analyzer meegenomen wordt in een EQC-ronde. Om te voldoen aan het principe *one fit for purpose, all fit for purpose*, heeft dit beleid enkele praktische consequenties:

- Voor iedere bepaling dienen beslisranges te worden vastgesteld waarbinnen een EQC-uitslag dient te vallen.
- Per bepaling zullen de analyzers zoveel mogelijk gebruik moeten maken van identieke reagens- en kalibratielots.
- Er moet een stappenplan voorhanden zijn zodra een EQC-resultaat buiten de beslisrange valt vanwege het omgekeerde principe *one wrong, all wrong*. In dit stappenplan is onderzoek naar de oorzaak van imprecisie/bias van het betreffende systeem opgenomen en vervolgens een correlatieonderzoek dat zich uitstrekt over alle analyzers. Eventuele bias moet zoveel mogelijk worden gemitigeerd of gecorrigeerd.
- Indien een laboratoriumorganisatie niet kiest voor een virtuele analyzer maar wel te maken heeft met meerdere analyzers en/of multilocaties, kan een alternatieve aanpak gekozen worden om individuele laboratoriumbepalingen dusdanig te bewaken dat de uitslagen uitwisselbaar blijven. Voor ieder afzonder systeem dient dan per bepaling eigen doelwaarden en doel-SD's vastgesteld te worden ([Aanbeveling 9](#)). Oftewel, alle resultaten per bepaling worden gecontroleerd tegen deze vooraf vastgestelde doelwaarde en doel-SD's die als het ware fungeren als gouden standaards. De doelwaardes en doel-SD's voor alle verschillende bepalingen kunnen ontleend kunnen worden aan: 1) op biologische variatie gebaseerde kwaliteitsspecificaties, 2) EQC-data voor de desbetreffende methode en analyzertype, 3) verificatiedata van het initiële analyzersysteem, waarbij feitelijk een moeder-dochter-principe wordt geïntroduceerd. NB: zodra de doelwaarden en doel-SD's voor de verschillende bepalingen worden ontleend worden aan de gemiddelde (IQC)-resultaten van meerdere systemen, is er sprake van toepassing van een virtuele analyzer.

In de dagelijkse praktijk kan de prestatie van alle individuele laboratoriumbepalingen in de tijd grafisch gevolgd worden door gebruik te maken van zogenaamde *overlay* Levey-Jennings plots waarbij voor elke analyzer in een oogopslag steeds duidelijk is welk systeem een overschrijding te zien geeft buiten de gestelde bewakingsgrenzen. Immers, door gebruik te maken van *overlay*-plots wordt voor een specifieke laboratoriumtest toch overall kwaliteitsgrenzen gevisualiseerd waarbinnen iedere analyzer deze bepaling moet meten.

Hoofdstuk 9 Referenties

- Abernethy RB. The New Weibull Handbook, Fifth Edition, 2004 (H.2-7).
- Badrick, T, Graham, P. Can a combination of average of normals and “real time” external quality assurance replace internal quality control? Clin Chem Lab Med 2018; 56: 549-553.
- Badrick T, Theodorsson E. Six Sigma - is it time to re-evaluate its value in laboratory medicine? Clin Chem Lab Med. 2024; 62: 2398-2400.
- Barger TE, Zhou L, Hale M, Moxness M, Swanson SJ, Chirmule N. Comparing exponentially weighted moving average and run rules in process control of semiquantitative immunogenicity immunoassays. AAPS J. 2010; 12: 79-86.
- Bayat H. Selecting multi-rule quality control procedures based on patient risk. Clin Chem Lab Med. 2017; 55: 1702-1708.
- Bayat H. Expected long-term defect rate of analytical performance in the medical laboratory: Assured Sigma versus observed Sigma. Biochem Med (Zagreb). 2018; 28: 020101.
- Borrillo F, Panteghini M. State-of-the-art model for derivation of analytical performance specifications: how to define the highest level of analytical performance technically achievable. Clin Chem Lab Med. 2024; 62: 1490-1496.
- Box GEP, Cox, DR. An analysis of transformations. Journal of the Royal Statistical Society, Series B. 1964; 26: 211-252.
- Braga F, Panteghini M. Derivation of performance specifications for uncertainty of serum C-reactive protein measurement according to the Milan model 3 (state of the art). Clin Chem Lab Med. 2020; 58: e263-e265.
- Braga F, Panteghini M. Performance specifications for measurement uncertainty of common biochemical measurands according to Milan models. Clin Chem Lab Med. 2021; 59: 1362-1368.
- Braga F, Pasqualetti S, Aloisio E, Panteghini M. The internal quality control in the traceability era. Clin Chem Lab Med. 2020; 59: 291-300.
- Buchta C, Marrington R, De la Salle B, Albarède S, Badrick T, Berghäll H, Bullock D, Coucke W, Delatour V, Geilenkeuser WJ, Griesmacher A, Henriksen GM, Huggett JF, Luppá PB, Pelanti J, Pezzati P, Sandberg S, Spannagl M, Thelen M, Restelli V, Perrone LA. Behind the scenes of EQA - characteristics, capabilities, benefits and assets of external quality assessment (EQA). Clin Chem Lab Med. 2025 Jan 6. doi: 10.1515/cclm-2024-1290. Epub ahead of print. PMID: 39753241.
- Burnett O, Ayers GJ. Quality control of drug assay. In: Richens A, Marks V, eds. Therapeutic Drug Monitoring. Edinburgh: Churchill Livingstone, 1981; 183-200.
- Carlson RO, Amirahmadi F, Hernandez JS. A primer on the cost of quality for improvement of laboratory and pathology specimen processes. Am J Clin Pathol. 2012; 138: 347-354.
- Cattaneo D, Panteghini M. Analytical performance specifications for measurement uncertainty in therapeutic monitoring of immunosuppressive drugs. Clin Chem Lab Med. 2023; 62: e81-e83.
- Cerioti F, Fernandez-Calle P, Klee GG, Nordin G, Sandberg S, Streichert T, Vives-Corróns JL, Panteghini M. Criteria for assigning laboratory measurands to models for analytical performance specifications defined in the 1st EFLM Strategic Conference. Clin Chem Lab Med. 2017; 55: 189-194.
- CLIA. Clinical Laboratory Improvement Amendments. Available at: <https://www.cms.gov/medicare/quality/clinical-laboratory-improvement-amendments> [Accessed 8 November 2024].

- CLSI. Clinical and Laboratory Standards Institute. C24 Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions, 4th Edition.
- CLSI. Clinical and Laboratory Standards Institute. [EP23 Laboratory Quality Control Based on Risk Management](#) [Accessed 8 November 2024].
- Coskun A. Bias in laboratory medicine: The dark side of the moon. *Ann Lab Med.* 2024; 44: 6-20.
- Coskun A, İnal BB, Serdar M. Measurement uncertainty in laboratory medicine: the bridge between medical and industrial metrology. *Turk J Biochem* 2019; 44: 121–125.
- Cotlove E, Harris EK, Williams GZ. Biological and analytical components of variation in longterm studies of serum constituents in normal subjects. III. Physiological and medical implications. *Clin Chem* 1970; 16: 1028-1032.
- Çubukçu HC, Vanstapel F, Thelen M, Bernabeu-Andreu FA, van Schrojenstein Lantman M, Brugnoli D, Mesko Brguljan P, Milinkovic N, Linko S, Vaubourdolle M, O'Kelly R, Kroupis C, Lohmander M, Šprongl L, Panteghini M, Boursier G; European Federation of Clinical Chemistry, Laboratory Medicine EFLM Working Group Accreditation, ISO/CEN standards WG-A/ISO. Improving the laboratory result release process in the light of ISO 15189:2012 standard. *Clin Chim Acta* 2021; 522: 167-173.
- Çubukçu HC, Vanstapel F, Thelen M, van Schrojenstein Lantman M, Bernabeu-Andreu FA, Meško Brguljan P, Milinkovic N, Linko S, Panteghini M, Boursier G. APS calculator: a data-driven tool for setting outcome-based analytical performance specifications for measurement uncertainty using specific clinical requirements and population data. *Clin Chem Lab Med.* 2023; 17: 597-607.
- Dixon WJ. Analysis of extreme values. *Ann Math Statist* 1950; 21: 488-506.
- Dixon WJ. Processing data for Outliers. *Biometrics* 1953; 9:74-89.
- Dixon WJ. Ratios involving extreme values. *Ann Math Statist* 1951; 22: 68-78.
- EFLM Biological Variation Database. Available at: <https://biologicalvariation.eu> [Accessed 8 November 2024].
- Fleming JK, Katayev A. Changing the paradigm of laboratory quality control through implementation of real-time test results monitoring: For patients by patients. *Clin Biochem.* 2015; 48: 508-513.
- Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. *Ann Clin Biochem.* 1997; 34: 8-12.
- Fraser CG, Kallner A, Kenny D, Petersen PH. Introduction: strategies to set global quality specifications in laboratory medicine. *Scand J Clin Lab Invest.* 1999; 59: 477-478.
- Fraser CG. Desirable standards of performance for therapeutic drug monitoring. *Clin Chem* 1987; 33: 387-389.
- Fraser CG. The 1999 Stockholm Consensus Conference on quality specifications in laboratory medicine. *Clin Chem Lab Med.* 2015; 53: 837-840.
- Giannoli JM, Bernard M, L'Hirondel J, Heim A, Badrick T. A model for managing quality control for a network of clinical chemistry instruments measuring the same analyte. *Clin Chem Lab Med* 2024; 63: 853-860.
- Gras JM, Philippe M. Application of the Six Sigma concept in clinical laboratories: a review. *Clin Chem Lab Med.* 2007; 45: 789-796.
- Grote-Koska D, Brand K. Using ISO/TS 20914:2019 to calculate the measurement uncertainty of immunosuppressive drugs in a clinical laboratory. *Scand J Clin Lab Invest.* 2020; 80: 309-312.
- Grubbs FE. Procedure for detecting outlying observations in samples. *Technometrics* 1969; 11: 1-21.
- Grubbs FE. Sample criteria for testing outlying observations. *Ann Math Statist* 1950; 21: 27-58.
- Horvath AR, Bell KJL, Ceriotti F, Jones GRD, Loh TP, Lord S, Sandberg S; Task Group Analytical Performance Specifications based on Outcomes of the European Federation of Clinical Chemistry and Laboratory

- Medicine. Outcome-based analytical performance specifications: current status and future challenges. *Clin Chem Lab Med*. 2024; 62: 1474-1482.
- Horvath AR, Bossuyt PM, Sandberg S, John AS, Monaghan PJ, Verhagen-Kamerbeek WD, Lennartz L, Cobbaert CM, Ebert C, Lord SJ; Test Evaluation Working Group of the European Federation of Clinical Chemistry and Laboratory Medicine. Setting analytical performance specifications based on outcome studies – is it possible? *Clin Chem Lab Med* 2015; 53: 841-848.
- [ICH guideline M10 on bioanalytical method validation \(EMA/CHMP/ICH/172948/2019\)](#) [Accessed 3 February 2025].
- [ISO/TC 212 – Medical laboratories and in vitro diagnostic systems](#) [Accessed 6 December 2024].
- [ISO/TS 20914:2019 – Medical laboratories – Practical guidance for the estimation of measurement uncertainty](#) [Accessed 10 December 2024].
- [ISO 5649:2024 - Medical laboratories – Concepts and specifications for the design, development, implementation and use of laboratory-developed tests](#) [Accessed 31 January 2025].
- Jansen R, Jassam N, Thomas A, Perich C, Fernandez-Calle P, Faria AP, Correia H, Barth JH, Weykamp C, Cobbaert C, Thelen M, Ricós C. A category 1 EQA scheme for comparison of laboratory performance and method performance: An international pilot study in the framework of the Calibration 2000 project. *Clin Chim Acta*. 2014; 432: 90-98.
- Jansen RTP, Laeven M, Kardol W. Internal quality control system for non-stationary, non-ergodic analytical processes based upon exponentially weighted estimation of process means and process standard deviation. *Clin Chem Lab Med* 2002; 40: 616–624.
- JCTLM. The Joint Committee for Traceability in Laboratory Medicine. Available at: <https://jctlm.org> [Accessed 8 November 2024].
- Jones GRD. Using analytical performance specifications in a medical laboratory. *Clin Chem Lab Med*. 2024; 62: 1512-1519.
- Kazmierczak SC. Laboratory quality control: using patient data to assess analytical performance. *Clin Chem Lab Med*. 2003; 41: 617-627
- Kenny D, Fraser CG, Hyltoft Petersen P, Kallner A. Consensus agreement. *Scand J Clin Lab Inv*. 1999; 59: 585.
- Kim S, Chang J, Kim S-K, Park S, Huh J, Jeong T-D. Sample size and rejection limits for detecting reagent lot variability: analysis of the applicability of the Clinical and Laboratory Standards Institute (CLSI) EP26-A protocol to real-world clinical chemistry data. *Clin Chem Lab Med* 2020; 59: 127-138.
- Kinns H, Pitkin S, Housley D, Freedman DB. Internal quality control: best practice. *J Clin Pathol*. 2013; 66: 1027-1032.
- Larbi JR, 2022. [Ensuring Compliance for your IVD's Performance Evaluation – Part 1 \(celegence.com\)](#)
- Larsen ML, Fraser CG, Petersen PH. A comparison of analytical goals for haemoglobin A1c assays derived using different strategies. *Ann Clin Biochem* 1991; 28: 272-278.
- Levey S, Jennings ER. The use of control charts in the clinical laboratory. *Am J Clin Pathol*. 1950; 20: 1059-1066.
- Lim CY, Ow Yang S, Markus C, Loh TP. Calibration frequency and analytical variability of laboratory measurements. *Clin Chim Acta*. 2023; 539: 87-89.
- Linnet K. The exponentially weighted moving average (EWMA) rule compared with traditionally used quality control rules. *Clin Chem Lab Med*. 2006; 44: 396-399.
- Loh TP, Smith AF, Bell KJL, Lord SJ, Ceriotti F, Jones G, Bossuyt P, Sandberg S, Horvath AR. Setting analytical performance specifications using HbA1c as a model measurand. *Clin Chim Acta*. 2021; 523: 407-414.

MDCG (Medical Device Coordination Group 2022-2): Developed stage of current technical capability and/or accepted clinical practice in regard to products, processes and patient management, based on the relevant consolidated findings of science, technology and experience. [Guidance - MDCG endorsed documents and other guidance - European Commission](#).

Montgomery DC, Introduction to Statistical Quality Control, 5th edition, John Wiley & Sons, Inc., 2005

NEN-EN-ISO 15189:2023. Available at: <https://www.nen.nl/nen-en-iso-15189-2023-en-305066> [Accessed 8 November 2024].

Nevalainen D, Berte L, Kraft C, Leigh E, Picaso L, Morgan T. Evaluating laboratory performance on quality indicators with the six sigma scale. Arch Pathol Lab Med. 2000; 124: 516-519.

[NVKC Leidraad Analytische Kwaliteitscontrole met behulp van Six Sigma](#). April 2015.

Oosterhuis WP, Bayat H, Ambruster D, Coskun A, Freeman KP, Kallner A, Koch D, Mackenzie F, Migliarino G, Orth M, Sandberg S, Sylte MS, Westgard S, Theodorsson E. The use of error and uncertainty methods in the medical laboratory. Clin Chem Lab Med 2018; 56: 209-219.

Oosterhuis WP, Coskun A. Sigma metrics in laboratory medicine revisited: We are on the right road with the wrong map. Biochem Med (Zagreb). 2018; 28: 020503.

Oosterhuis WP, Sandberg S. Proposal for the modification of the conventional model for establishing performance specifications. Clin Chem Lab Med 2015; 53: 925-937.

Oosterhuis WP. Gross overestimation of total allowable error based on biological variation. Clin Chem 2011; 57: 1334-1336.

Panteghini M, Sandberg S. Total error vs. measurement uncertainty: the match continues. Clin Chem Lab Med 2016; 54: 195-196.

Panteghini M. Analytical performance specifications for combined uncertainty budget in the implementation of metrological traceability. Clin Chem Lab Med 2024; 62: 1497-1504.

Perich C, Ricós C, Alvarez V, Biosca C, Boned B, Cava F, Doménech MV, Fernández-Calle P, Fernández-Fernández P, García-Lario JV, Minchinela J, Simón M, Jansen R. External quality assurance programs as a tool for verifying standardization of measurement procedures: Pilot collaboration in Europe. Clin Chim Acta. 2014; 432: 82-89.

Ramamohan V, Abbott J, Klee GG, Yih Y. Application of mathematical models of system uncertainty to evaluate the utility of assay calibration protocols. Clin Chem Lab Med. 2012; 50: 1945-1951.

RiliBÄK Aktualisierung der Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen. Dtsch Arztebl 2023; 120: A-994 / B-858.

Sandberg S, Coskun A, Carobene A, Fernandez-Calle P, Diaz-Garzon J, Bartlett WA, Jonker N, Galior K, Gonzales-Lao E, Moreno-Parro I, Sufrate-Vergara B, Webster C, Aarsand AK, on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine Task Group for the Biological Variation Database. Analytical performance specifications based on biological variation data - considerations, strengths and limitations. Clin Chem Lab Med. 2024; 62: 1483-1489.

Schoenmakers CH, Naus AJ, Vermeer HJ, van Loon D, Steen G. Practical application of Sigma Metrics QC procedures in clinical chemistry. Clin Chem Lab Med. 2011; 49: 1837-1843.

Shewhart WA. Economic control of quality of manufactured product. New York: Van Nostrand Company; 1931.

Stankovic AK, Romeo P. The role of in vitro diagnostic companies in reducing laboratory error. Clin Chem Lab Med. 2007; 45: 781-788.

Stewart MJ, Fraser CG. Desirable performance standards for assays of drugs. Ann Clin Biochem. 1989; 26: 220-226.

- Stöckl D, Baadenhuijsen H, Fraser CG, Libeer J-C, Petersen PH, Ricós C. Desirable routine analytical goals for quantities assayed in serum. Discussion paper from the members of the external quality assessment (EQA) working group A on analytical goals in laboratory medicine. *Eur J Clin Chem Clin Biochem* 1995; 33: 157-169.
- Thelen MHM, van Schrojenstein Lantman M. When bias becomes part of imprecision: how to use analytical performance specifications to determine acceptability of lot-lot variation and other sources of possibly unacceptable bias. *Clin Chem Lab Med*. 2024; 62: 1505-1511.
- Theodorsson E. Understanding and expressing variation in measurement results. *J Lab Precis Med* 2025;10:3 | <https://dx.doi.org/10.21037>.
- van Rossum HH. Moving average quality control: principles, practical application and future perspectives. *Clin Chem Lab Med* 2019; 57: 773-782.
- van Rossum HH, Bietenbeck A, Cervinski MA, Katayev A, Loh TP, Badrick TC. Benefits, limitations, and controversies on patient-based real-time quality control (PBRTQC) and the evidence behind the practice. *Clin Chem Lab Med* 2021; 59: 1213-1220.
- van Rossum HH, Kemperman H. Implementation and application of moving average as continuous analytical quality control instrument demonstrated for 24 routine chemistry assays. *Clin Chem Lab Med* 2017; 55: 1142-1151.
- van Rossum HH, van den Broek D. Design and implementation of quality control plans that integrate moving average and internal quality control: incorporating the best of both worlds. *Clin Chem Lab Med* 2019; 57: 1329-1338.
- van Rossum HH, van den Broek D. Ten-month evaluation of the routine application of patient moving average for real-time quality control in a hospital setting. *J Appl Lab Med* 2020; 5: 1184-1193.
- van Rossum HH, van Schrojenstein Lantman M, Severens M, Vermeer HJ, Verboeket-van de Venne WPHG, Oosterhuis W, de Jonge N. Quality control in the Netherlands; today's practices and starting points for guidance and future research. *Clin Chem Lab Med* 2024; 62: 2177-2184.
- van Rossum HH. Technical quality assurance and quality control for medical laboratories: a review and proposal of a new concept to obtain integrated and validated QA/QC plans. *Crit Rev Clin Lab Sci* 2022; 59: 586-600.
- van Rossum HH. When internal quality control is insufficient or inefficient: Consider patient-based real-time quality control! *Ann Clin Biochem*. 2020; 57: 198-201.
- van Schrojenstein Lantman M. [How measurement uncertainty impedes clinical decision-making](#). Thesis, 2024.
- van Schrojenstein Lantman M, Can Çbukçu H, Boursier G, Panteghini M, Bernabeu-Andreu FA, Milinkovic N et al. An approach for determining allowable between reagent lot variation. *Clin Chem Lab Med* 2022; 60: 681-688.
- Vidali M, Carobene A, Apassiti Esposito S, Napolitano G, Caracciolo A, Seghezzi M, Previtali G, Lippi G, Buoro S. Standardization and harmonization in hematology: Instrument alignment, quality control materials, and commutability issue. *Int J Lab Hematol*. 2021; 43: 364-371.
- Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem*. 1981; 27: 493-501.
- Westgard JO, Bayat H, Westgard SA. Planning Risk-Based SQC schedules for bracketed operation of continuous production analyzers. *Clin Chem*. 2018; 64: 289-296.
- Westgard JO, Burnett RW. Precision requirements for cost-effective operation of analytical processes. *Clin Chem* 1990; 36: 1629-1632.
- Westgard JO, Westgard SA. Six Sigma Quality Management System and Design of Risk-based Statistical Quality Control. *Clin Lab Med*. 2017; 37: 85-96.
- Westgard JO. Internal quality control: planning and implementation strategies. *Ann Clin Biochem*. 2003; 40: 593-611.

Hoofdstuk 10 Begrippen en afkortingen

In de leidraad worden de begrippen ‘bepaling’ (synoniem: test, analyse) en ‘apparaat’ (synoniem: toestel, instrument, analyzer) gehanteerd. In ISO 15189 worden deze begrippen samen aangeduid als ‘onderzoeksprocedure’. Vanwege de aansluiting bij wetenschappelijke literatuur is ervoor gekozen om in sommige gevallen gebruik te maken van Engelstalige begrippen.

Term	Definitie
APC, <i>Analytical Performance Characteristics</i>	Analytische prestatiekarakteristieken
APG, <i>Analytical Performance Goal</i>	Analytisch prestatiedoel
APS, <i>Analytical Performance Specification</i>	Analytische prestatiespecificatie. Maximaal toelaatbare bias, imprecisie en meetonzekerheid c.q. <i>total allowable error</i> .
Bias, systematische meetfout	Gemiddelde van een oneindig aantal herhaalde metingen van een te meten stof, verminderd met de werkelijke waarde van de te meten grootte.
BV, <i>Biological Variation</i>	Biologische variatie
Commuteerbaarheid	Mate van overeenstemming tussen meetresultaten voor een vastgestelde concentratie referentiemateriaal, verkregen volgens twee meetmethoden. Het betreffende materiaal is meestal een kalibrator.
Controleregel	Een beslissend criterium voor het interpreteren van controlegegevens. Vaak wordt hierbij gebruik gemaakt van zg. <i>control charts</i> , een grafische methode waarbij beoordeeld wordt of een proces wel of niet voldoet aan de statistische kwaliteitscontrole. Een voorbeeld van een controleregel is een overschrijding van een vooraf omschreven afwijking (bijvoorbeeld 3 SD).
CLSI	Clinical and Laboratory Standards Institute
CV, <i>Coefficient of Variation</i>	Variatiecoëfficiënt: de relatieve standaarddeviatie, dat wil zeggen de standaarddeviatie uitgedrukt als percentage van het gemiddelde (\bar{x}). $CV (\%) = (SD / \bar{x}) \times 100$.
CVa	Analytische variatie (imprecisie)
CVb	Biologische variatie. Hierbij wordt geen onderscheid gemaakt tussen intra-individuele en interindividuele variatie
CVi	Intra-individuele of binnen-persoons biologische variatie (ook wel aangeduid als <i>within-subject</i> variatie of CVw)

CVg	Interindividuele of tussen-persoon biologische variatie (ook wel aangeduid als <i>between-subject</i> variatie)
EFLM	European Federation of Clinical Chemistry and Laboratory Medicine
E(N _{ur})	Het maximaal aantal mogelijke fouten dat niet wordt gedetecteerd met specifieke controle regels die klinische tolerantiegrens overschrijdt
EQA, <i>External Quality Assessment</i>	Externe kwaliteitscontrole
Gemiddelde (\bar{x})	Het rekenkundig gemiddelde van een reeks waarden
IFCC	International Federation of Clinical Chemistry and Laboratory Medicine
Imprecisie	Reproduceerbaarheid, herhaalbaarheid. Er wordt onderscheid gemaakt tussen korte en lange termijnsimplicite. <i>Repeatability</i> = kort, <i>intermediate</i> = inclusief kalibratie en reagenswisselingen, <i>reproducibility</i> is over locaties en meetprocedures heen.
IQC, <i>Internal Quality Control</i>	Interne kwaliteitscontrole
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
IVDR	Guidance on Classification Rules for in vitro Diagnostic Medical Devices under Regulation (EU) 2017/746
Levey-Jennings grafiek	Grafiek waarin opeenvolgende resultaten van kwaliteitscontrolemonsters zijn getekend. In de grafiek worden gemiddelde waarden van de resultaten en standaarddeviaties aangegeven.
MAu, <i>allowable standard measurement uncertainty</i>	Toegestane meetonzekerheid
MAU, <i>allowable expanded measurement uncertainty</i>	Toegestane 'opgerekte' meetonzekerheid
Meetfout	Het verschil tussen een meetresultaat en de werkelijke waarde van de te meten grootte. De relatieve meetfout wordt berekend door de meetfout te delen door de streefwaarde.
MU, <i>measurement uncertainty</i>	Meetonzekerheid
Nauwkeurigheid, <i>accuracy</i>	De mate van overeenstemming tussen het resultaat van een meting en de werkelijke waarde van de te meten grootte

Onderzoeksprocedure	Meting van een analiet om klinisch bruikbare informatie op te leveren
PBRTQC, <i>Patient-Based Real-Time Quality Control</i>	Actuele kwaliteitscontrole op basis van patiëntresultaten
P_{ed} , <i>probability for error detection</i>	Een prestatiekenmerk van een QC-procedure die beschrijft hoe vaak een analysereeks wordt verworpen op basis van fouten die niet gerelateerd zijn aan de onnauwkeurigheid van de meting. Idealiter is de P_{ed} 1,00 voor klinisch relevante fouten. In de praktijk streven we naar een $P_{ed} > 0,90$ bij het selecteren en ontwerpen van QC-procedures.
P_{fr} , <i>probability for false rejection</i>	Een prestatiekenmerk van een QC-procedure die beschrijft hoe vaak een analysereeks wordt verworpen indien er geen fouten optreden, met uitzondering van de onnauwkeurigheid van de meting. Idealiter is de P_{fr} 0,00. In de praktijk streven we naar een $P_{fr} < 0,05$.
Precisie, <i>precision</i>	De mate van overeenstemming tussen onafhankelijke testresultaten verkregen onder vastgestelde voorwaarden
QC, <i>Quality Control</i>	Kwaliteitscontrole. Een algemene term die verwijst naar bewaking en beoordeling van laboratoriumprocessen om zodoende problemen te identificeren en kwaliteit te handhaven.
Repeteerbaarheid, <i>repeatability</i>	Precisie onder herhaalbare omstandigheden; waarbij onafhankelijke testresultaten worden verkregen met dezelfde methode bij identiek analysemateriaal in hetzelfde laboratorium, uitgevoerd door dezelfde persoon, met dezelfde apparatuur in een kort tijdsbestek
Reproduceerbaarheid, <i>reproducibility</i>	Precisie onder reproduceerbare omstandigheden; waarbij de testresultaten verkregen worden met dezelfde methode bij identiek analysemateriaal in verschillende laboratoria, uitgevoerd door verschillende medewerkers, met verschillende apparatuur
Sigmascore, <i>Sigma Metric</i>	Een numerieke waarde die de prestatie van een bepaling karakteriseert. Met andere woorden: hoe vaak past de standaarddeviatie of de sigmascore in de tolerantiegrens of kwaliteitseis van een bepaling?
SKML	Stichting Kwaliteitsbewaking Medische Laboratoriumdiagnostiek
SQC, <i>Statistical Quality Control</i>	Statistische kwaliteitscontrole: die aspecten van de kwaliteitscontrole waarbij statistiek wordt

	toegepast, in tegenstelling tot het brede scala van kwaliteitscontrole die uit veel meer procedures bestaat (preventief onderhoud, apparaat controle, validatie van bepalingen). Statistische kwaliteitscontrole wordt vaak gebruikt om routinematig de kwaliteit van een methode te bewaken en een verandering in kwaliteit te signaleren.
Standaarddeviatie (SD)	Een maat voor de spreiding van de meetwaarden rond hun gemiddelde, uitgaande van een normale verdeling
TE, <i>total error</i>	Totale meetfout: het netto- of gecombineerde effect van toevallige en systematische meetfouten
TEa, <i>total allowable error</i>	Tolerantiegrens: een analytische kwaliteitseis die aangeeft welke toevallige fout (<i>imprecision</i>) en welke systematische fout (onnauwkeurigheid, <i>inaccuracy</i>) nog aanvaardbaar is in een enkele meting of een resultaat van een meting.
Toevallige meetfout, <i>random error</i>	Het verschil tussen een meetresultaat en het gemiddelde van een oneindig aantal herhaalde metingen van een te meten stof.
uRw	Lange-termijn imprecisie

Bijlage A Samenhang met NEN-ISO 15189:2023

NEN-EN-ISO 15189:2022 stelt eisen aan de(interne) kwaliteitscontrole. De belangrijkste wijzigingen t.o.v. NEN-EN-ISO 15189:2012 betreffen de onderstreepte normteksten uit NEN-EN-ISO 15189:2023, die zijn overgenomen met toestemming van NEN te Delft –www.nen.nl):

7.3.7 De validiteit van onderzoeksresultaten waarborgen

7.3.7.1 Algemeen

Het laboratorium moet een werkwijze hebben om de validiteit van resultaten te monitoren. De hieruit voortvloeiende data moeten dusdanig worden geregistreerd dat trends en verschuivingen kunnen worden waargenomen en, waar praktisch uitvoerbaar, moeten statistische technieken worden toegepast om de resultaten te beoordelen. Dit monitoringsproces moet worden gepland en beoordeeld.

7.3.7.2 Interne kwaliteitscontrole (IQC)

- a. Het laboratorium moet beschikken over een IQC-werkwijze voor het monitoren van de voortdurende validiteit van onderzoeksresultaten, overeenkomstig gespecificeerde criteria, die verifieert of de beoogde kwaliteit wordt bereikt en die geldigheid van de resultaten voor klinische besluitvorming waarborgt.
 1. De beoogde klinische toepassing van het onderzoek behoort in overweging te worden genomen, aangezien de prestatiespecificaties voor dezelfde te meten grootte in verschillende klinische omstandigheden kunnen verschillen.
 2. De werkwijze behoort het ook mogelijk te maken om voor de onderzoeksmethode verschillen in kalibratoren of reagentia van verschillende lotnummers, of beide, te detecteren. Om dit mogelijk te maken behoort de werkwijze te voorkomen dat IQC-materiaal op dezelfde dag /run van lot wisselt als het reagens en/of de kalibrator.
 3. ...
- b. ...
- c. Indien er geen passend IQC-materiaal beschikbaar is, moet het laboratorium het gebruik van andere methode voor IQC overwegen. Voorbeelden van dergelijke andere methoden zijn:
 1. trendanalyse van patiëntenresultaten, bijv. met voortschrijdend gemiddelde van patiëntenresultaten of het percentage monsters met resultaten lager of hoger dan bepaalde waarden of geassocieerd met een diagnose;
 2. vergelijking van de resultaten van patiëntenmonsters volgens een gespecificeerd schema met de resultaten voor patiëntenmonsters die zijn onderzocht met een alternatieve werkwijze waarvan uit validatie blijkt dat de kalibratie metrologisch traceerbaar is tot referenties van dezelfde of een hogere orde als gespecificeerd in ISO 17511;
 3. opnieuw testen van bewaarde patiëntenmonsters.
- d. ...
- e. ...
- f. ...
- g. ...

Bijlage B Analytische spreiding en tolerantiegrenzen

B.1 Verschillende modellen voor tolerantiegrenzen

Na de conferentie in Milaan in 2014 zijn criteria uitgewerkt om tests in te delen in de drie modellen: gebaseerd op uitkomst, op biologische variatie of op de *state of the art*.

Voor het op uitkomst gebaseerde model moet aan enkele eisen worden voldaan: In principe moet de test een centrale en goed gedefinieerde rol hebben in de besluitvorming van een specifieke ziekte of klinische situatie en er moet sprake zijn van eenduidige beslisgrenzen. Ook moeten de resultaten direct van invloed zijn op de behandeling van de patiënt en daarmee ook op het resultaat van de behandeling.

Directe uitkomststudies op basis van dubbelblinde gerandomiseerde gecontroleerde onderzoeken zijn echter schaars doordat deze een grote uitdaging vormen. Als alternatief kunnen indirecte uitkomststudies worden gebruikt, op basis van de impact van de analytische uitvoering van de test op klinische classificaties of medische beslissingen en daarmee op de waarschijnlijkheid van uitkomsten (bijvoorbeeld door simulatie of beslissingsanalyse). Deze indirecte uitkomststudies zijn beter haalbaar en zullen naar verwachting meer beschikbaar komen [Horvath et al., 2015].

Voorbeelden van bepalingen met tolerantiegrenzen gebaseerd op uitkomst-studies zijn: totaal-, HDL- en LDL-cholesterol, glucose, albumine (plasma), hemoglobine, trombocyten, TSH en HbA1c [Ceriotti et al., 2017; Loh et al., 2021, Panteghini, 2024].

Biologische variatie

In dit model wordt de analytische variatie ten opzichte van de biologische variatie geminimaliseerd. Merk op dat er in dit geval geen directe verband is met de klinische toepassing van de test. Een lage analytische 'ruis' vergeleken met de biologische variatie zal echter de klinische interpretatie van de resultaten vergemakkelijken.

In principe zijn bij dit model er twee verschillende situaties:

- een meetgrootte wordt op een bepaald concentratieniveau in het serum/plasma gehouden, anders treden er symptomen op (d.w.z. de meetgrootte staat onder strikte homeostatische controle, bijv. natrium);
- de situatie waarin een meetgrootte de facto een stabiele concentratie heeft, maar afwijkingen van deze concentratie op zichzelf geen symptomen zullen veroorzaken (bijv. kreatinine).

De conventionele vergelijking combineert termen voor systematische fout (bias) en toevallige fout (imprecisie) voor de maximum (*allowable*) fout:

$$TEa = 1,65 \times (0,5 \times CVi) + 0,25 \times (CVi^2 + CVg^2)^{1/2}$$

Dit conventionele lineaire model is om een aantal redenen bekritiseerd [Oosterhuis, 2011; Oosterhuis et al., 2018]:

- Zowel maxima van toegestane bias als imprecisie worden opgeteld om TEa te berekenen. Dit is een pragmatische oplossing die voor het eerst werd voorgesteld voor toepassing bij externe kwaliteitscontrole [Oosterhuis en Sandberg, 2015]. De theoretische basis hiervoor ontbreekt. Twee maximaal toegestane fouten worden opgeteld, afgeleid onder condities die elkaar wederzijds uitsluiten, namelijk van respectievelijk geen bias (voor imprecisie) en geen imprecisie (voor de bias term). De som zal een leiden tot een overschatting van TEa en te ruime limieten.

- De specificatie voor imprecisie is volgens verschillende modellen $CVa < 0,5 \times CVb$ (met CVb de biologische variatie) [Stöckl et al., 1995]. In het geval van diagnose kan dit worden geschreven als $CVa < 0,5(CVi^2 + CVg^2)^{1/2}$. In het geval van monitoring wordt alleen de binnen-persoonsvariatie CVi meegenomen. De maximaal toegestane bias werd afgeleid als: $bias = 0,25 \times CVb$ of $0,25 \times (CVi^2 + CVg^2)^{1/2}$. In het conventionele model wordt deze bias-term toegepast in het geval van monitoring, hoewel deze uitdrukking is afgeleid van een referentiebereikmodel en alleen van toepassing is op diagnose.
- Het conventionele lineaire TEa-model houdt alleen rekening met de analytische variatie en de bias. Wanneer een tolerantiegrens wordt gedefinieerd – uitgaande van een Gaussische verdeling – met een vast aantal testresultaten buiten de limiet (bijv. 5% bij $z = 1,64$), tonen de combinaties van bias en imprecisie (CVa) die aan deze voorwaarde voldoen een rechte lijn met een helling van $-z$. Aan het ene uiterste van de lijn hebben we $bias = 0$, met $CVa = TEa/z$; aan het andere uiterste hebben we $CVa = 0$ (een hypothetische waarde) en $bias = TEa$. De rechte lijn heeft het kenmerk dat de tolerantiegrens TEa , een combinatie van bias en imprecisie, een constante is die gelijk is aan de bias bij $CVa = 0$. Deze lineariteit is van fundamenteel belang voor de Westgardregels. Dit model is echter alleen geldig als imprecisie en bias de enige betrokken variabelen zijn, of met andere woorden, wanneer de distributie van testresultaten Gaussisch is en volledig wordt gedefinieerd door de analytische bias en imprecisie alleen. Dus houdt het TE-model geen rekening met biologische variatie en andere aanvullende oorzaken van variatie die gelden voor patiëntenmonsters. Biologische variatie, hoewel niet relevant bij het monitoren van variatie in controlemonsters, is zeker relevant bij het omgaan met patiëntenmonsters [Panteghini en Sandberg, 2016]. Er kan wel rekening gehouden worden met biologische variatie door deze in het model op te nemen. Hierdoor verandert de relatie tussen imprecisie en bias echter van een rechte lijn naar een curve [Larsen et al., 1991]. Het gevolg is dat de waarde van TEa geen constante meer is maar verandert afhankelijk van de combinatie van bias en imprecisie. Anders gesteld, verandering van imprecisie en bias die hetzelfde effect hebben op de totale fout TE , hebben een verschillend effect op de uitkomsten van patiëntenmonsters.
- $CVa < 0,5 \times CVi$ is een conditie die eigenlijk niet geborgd kan worden door interne kwaliteitscontrole (bijvoorbeeld leidend tot een sigmascore lager dan 3:

Sigmascore: $TEa = bias + 1,65(0,5CVb)$ ->

$$\begin{aligned}
 \text{Sigma} &= (TEa - bias)/CVa \\
 &= (bias + 1,65(0,5 \times CVb))/CVa && \text{neem aan bias} = 0 \\
 &= (0,25 \times CVb + 0,825 \times CVb)/CVa \\
 &= 1,075 \times CVb/CVa && \text{neem aan CVa} = 0,5 \times CVb \\
 &= 2,15 \text{ CVa}/CVa
 \end{aligned}$$

Sigmascore = 2,15

Om die reden kunnen de coëfficiënten van 0,5 en 0,25 voor imprecisie en bias in deze vergelijking in twijfel worden getrokken.

State of the art

Er bestaat geen consensus over, hoe c.q. welke *state of the art* moet worden toegepast. Volgens Borrillo en Panteghini [2024] wordt als *state of the art* de test beschouwd, welke de beste kwaliteit heeft van de methoden die op de markt beschikbaar zijn. Deze kwaliteit wordt beschouwd als *desirable*. De kwaliteit – uitgedrukt als meetonzekerheid (imprecisie) - die vervolgens als minimum

nog acceptabel is, ligt 50% hoger. Volgens dit model is de minimum APS voor CRP 5,6%, voor ferritine 5,7%

B.2 Tolerantiegrenzen en prestatie-eisen volgens de IVDR

Voor CE-gemarkeerde tests gelden volgens de IVDR-specificaties waarvan de klinische toepasbaarheid aangetoond moet zijn. Naar verwachting worden deze prestatie-eisen nogal ruim, want meeste tests zullen in aanmerking komen voor een CE-markering.

Er zijn drie pijlers van een prestatie-evaluatie volgende de IVDR: de wetenschappelijke validiteit van de analiet, de analytische prestatie en de klinische prestatie. Daarmee bestaat een prestatie-evaluatie uit de beoordeling en analyse van gegevens om de wetenschappelijke validiteit, de analytische en toepasbare manier van de klinische prestatie van een meetmethode vast te stellen of te verifiëren [Larbi, 2022].

Wetenschappelijke validiteit van een analiet is in principe de associatie van analiet met een klinische aandoening. Bij patiënten met diabetes zal wetenschappelijk bewezen moeten worden dat het HbA1c een maat is van hun bloedsuikerspiegel gedurende een bepaalde periode.

Analytische prestaties zijn de prestaties van het apparaat zelf om daadwerkelijk de analiet te detecteren die u wilt meten (juistheid, precisie etc.). De klinische prestaties is de correlatie van deze testresultaten met een bepaalde klinische conditie (bijv. sensitiviteit en specificiteit). Deze evaluatie wordt vergeleken met de *state of the art* van het vakgebied. De (ruime) definitie van de MDCG hiervan is [Medical Device Coordination Group 2022-2]: *Developed stage of current technical capability and/or accepted clinical practice in regard to products, processes and patient management, based on the relevant consolidated findings of science, technology and experience. Note: The state-of-the-art embodies what is currently and generally accepted as good practice in technology and medicine. The state-of-the-art does not necessarily imply the most technologically advanced solution. The state-of-the-art described here is sometimes referred to as the “generally acknowledged state-of-the-art.”*

Het wordt de vraag welke rol deze IVDR-APS zal gaan innemen, en of deze APS leidend zal worden of dat deze door het vakgebied als te ruim zal worden beschouwd. Ten tijde van het opstellen van deze leidraad is het nog te vroeg om hierover een uitspraak te doen.

Als toevoeging op de ISO 15189 is recent ook de ISO 5649 gepubliceerd [2024]. Deze richtlijn beschrijft specifiek dat er interne als externe kwaliteitscontrole op eigen ontwikkelde tests (*lab developed tests*, LDT) moet plaatsvinden.

Bijlage C APS en medicatiebewaking

Het monitoren van geneesmiddelen (*therapeutic drug monitoring*, TDM) wordt toegepast om individuele doseringsschema's te optimaliseren. Dit heeft meestal betrekking op geneesmiddelen met een smalle therapeutische breedte zoals digoxine, anti-epileptica, anti-infectieuze middelen en immunosuppressieve medicatie. Het vaststellen van APS voor geneesmiddelen heeft minder aandacht gekregen dan voor andere bepalingen [Cattaneo en Panteghini, 2023]. De Milaan consensus heeft betrekking op bepalingen van lichaamseigen stoffen. Onder biologische variatie wordt verstaan de fysiologische variatie rond een homeostatisch instelpunt. Het model dat uitgaat van de biologische variatie is daarmee niet van toepassing op medicatie en andere toegediende stoffen.

Een eerder model voor APS van medicatie gaat uit van een precisie die – op het niveau van 50% van de concentratie van de ondergrens van het therapeutisch interval (C_L) – een onderscheid kan maken tussen afwezigheid van het middel en de ondergrens van het therapeutisch interval [Burnett en Ayers, 1981]. Op basis hiervan wordt afgeleid:

$$CVa \leq 0,17 \times C_L$$

met C_L de concentratie t.h.v. de ondergrens van het therapeutisch interval

Nadeel van deze methode is dat de grenzen van het therapeutische interval empirisch zijn vastgesteld, en zelf weer afhankelijk zijn van de bepalingsmethode.

Een ander model is gebaseerd op de farmacokinetiek van een geneesmiddel [Fraser, 1987; Stewart en Fraser, 1989]:

$$CVa \leq 0,25 [(2^{T/t} - 1)/(2^{T/t} + 1)] \times 100\%$$

met T= tijdsinterval tussen doseringen (uur); t is de gemiddelde halfwaardetijd (uur)

Op basis van dit model werden de volgende APS afgeleid [Stewart en Fraser, 1989]:

Medicatie	APS (%)
Theofylline	11,1
Digoxine	5,3
Carbamazepine	6,4
Fenobarbital	2,2
Fenytoïne	3,6
Valproaat	6,4

Hetzelfde model werd gebruikt om APS af te leiden voor immunosuppressiva. Hierbij werd aanvullend nog een minimum APS gehanteerd van 150% van deze waarde [Cattaneo en Panteghini, 2023]:

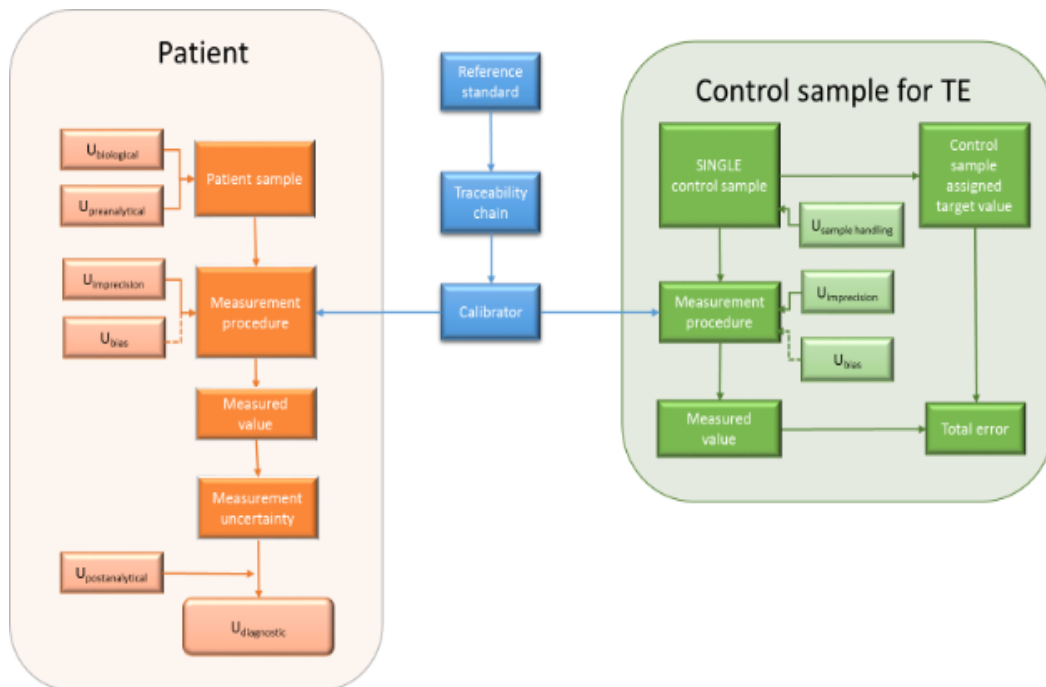
Medicatie	Frequentie toediening (uur)	Halveringstijd (uur)	APS (%)	minimum APS (%)
Cyclosporine	12	8,4	11,40	17,02
Everolimus	12	25,6	4,03	6,04
Sirolimus	24	60,0	3,44	5,17
Tacrolimus (snelle afgifte)	12	12,1	8,27	12,40
Tacrolimus (gereg. afgifte) (Envarsus XR)	24	30,0	6,76	8,51

In een recente publicatie met betrekking tot immunosuppressiva werd een algemene *state of the art* CVa van $\leq 10\%$ (bij voorkeur $\leq 6\%$) aanbevolen. Met LCMS/MS-methoden van de onderzochte immunosuppressiva blijkt deze kwaliteit voor de gemeten medicatie over het algemeen haalbaar [Grote-Koska en Brand, 2020].

Het Europees Geneesmiddelenbureau (European Medicines Agency, EMA) en de US Food and Drug Administration (FDA) eisen voor analytische variatie van een chromatografische methode zijn 15% en 20% op het niveau van LLOQ (*lower limit of quantification*, onderste kwantificatielimiet). EMA/FDA hanteren op deze wijze een minimumnorm; waar nodig kunnen strengere normen worden gehanteerd.

Bijlage D Meetonzekerheid versus *total error* concept

Figuur D.1. Schematische relatie meetonzekerheid (*measurement uncertainty*, MU) en *total error* (TE) [Oosterhuis et al., 2018].



Bij meting van het patiëntenmonster is er geen referentiepunt, behalve de kalibratie en via deze en de traceerbaarheidsketen de referentiestandaard. Bij IQC geldt het TE-model, als men aanneemt dat het (surrogaat) referentiepunt de interne QC-standaard is waaraan de afwijking van het meetresultaat wordt afgemeten.

De tegenstelling tussen beide modellen zou men deels kunnen nuanceren [Oosterhuis et al., 2018]. Hoewel de onzekerheid van het meetresultaat van patiëntenmonsters is onderworpen aan het MU-model, kan men bij het meetresultaat van kwaliteitscontrolemonsters het TE-model toepassen. Dit zou kunnen door als (surrogaat) werkelijke waarde de richtwaarde van de kwaliteitscontrole te nemen.

Bijlage E Westgard QC

Westgard QC-regels

Regel	Criteria	Probleem	Voorbeeld
1 _{2s}	Meting overschrijdt 2 standaarddeviaties boven of onder het gemiddelde	Juistheid en/of imprecisie	
1 _{3s}	Meting overschrijdt 3 standaarddeviaties boven of onder het gemiddelde	Juistheid en/of imprecisie	
2 _{2s}	2 achtereenvolgende metingen overschrijden 2 standaarddeviaties aan dezelfde zijde van het gemiddelde	Juistheid en/of imprecisie	
R _{4s}	2 metingen in dezelfde reeks verschillen 4 standaarddeviaties of meer (bijvoorbeeld meer dan 2 SD boven, resp. onder het gemiddelde)	Imprecisie	
4 _{1s}	4 achtereenvolgende metingen wijken aan dezelfde kant (boven of onder het gemiddelde) meer dan 1 SD af van het gemiddelde	Juistheid	
10 _x	10 achtereenvolgende metingen geven resultaten aan dezelfde zijde t.o.v. gemiddelde (onder of boven)	Juistheid	

Bron: wikipedia; Afbeeldingen: Mikael Häggström - Own work, CC0, <https://commons.wikimedia.org/w/index.php?curid=94876598>

Bijlage F Het Six Sigma concept

F.1 Algemeen

In 2008 is dit concept geïntegreerd in de *lean* productie methodologie. Het doel van deze methodologie, inmiddels bekend als *Lean Six Sigma*, is optimale waardecreatie voor de klant en minimale verspilling.

Nader beschouwd is Six Sigma toepasbaar als methode, managementstrategie of meetinstrument. Kern van de methode is de DMAIC-verbetercyclus (*Define, Measure, Analyze, Improve, Control*). Als managementstrategie is Six Sigma een gestructureerde methode om procesverbetering binnen organisaties te organiseren en vast te houden.

De 1,5 sigma correctie is bij Six Sigma niet gekoppeld aan een systematische fout maar aan de begrippen korte en lange termijn. Experimenten, vrijgaven, validaties worden onder gecontroleerde omstandigheden uitgevoerd. Random variabelen (die op korte termijn stabiel zijn) gaan op lange termijn variëren en voegen daardoor extra spreiding aan de analyse toe. De korte termijn sigmascore wordt hiervoor met 1,5 sigma verlaagd en in de praktijk de Z-bench genoemd. Voorbeeld: een validatie experiment geeft een sigmascore van bijv. 6 (korte termijn) en de hieruit lange termijn sigmascore is $6 - 1,5 = 4,5$. Vanaf deze waarde mag met bias gewerkt worden.

F.2 Toepassing van Six Sigma in het klinisch-chemisch laboratorium

Six Sigma heeft sinds 2000 ook ingang gevonden binnen de klinisch-chemische laboratoria [Nevalainen et al., 2000; Kazmierczak, 2003; Gras en Philippe, 2007; Stankovic en Romeo, 2007; Schoenmakers et al., 2011; Carlson et al., 2012]. Schoenmakers et al. [2011] beschreven een praktische werkwijze voor de selectie van Westgard regels, gebaseerd op de analytische kwaliteit van een bepaling uitgedrukt in de sigmascore. Voor recente overzichtsartikelen wordt verwezen naar Westgard en Westgard [2017], Bayat [2018], Kinns et al. [2013] en Oosterhuis en Coskun [2018]. Deze benadering wordt in deze leidraad ook gevolgd.

Het Six Sigma concept zoals toegepast in de klinische chemie is gebaseerd op acceptatiegrenzen die klinisch gedefinieerd zijn. In de praktijk worden grenzen vaak afgeleid van de biologische variatie. Strikt genomen moeten de grenzen conform de Milaan criteria worden vastgesteld. Deze benadering heeft bepaalde consequenties. Bij ruime klinische grenzen t.o.v. de analytische prestaties van een test (bijv. enzymen) kan het voorkomen dat een test niet meer *in control* is, maar dat deze afwijking klinisch niet relevant is en – binnen het Six Sigma concept - genegeerd kan worden. Ook komt het voor dat een test niet voldoet aan de prestatie-eisen op basis van de biologische variatie (bijv. natrium). De test heeft daarmee een sigmascore die te laag is om IQC op af te stemmen.

Men kan ook kiezen voor de volgende benadering: de IQC wordt uitsluitend ingericht op basis van de analytische prestaties. Dit betekent in dit verband, dat de kwaliteit wordt bewaakt op basis van de prestaties bij validatie/verificatie. Daarmee wordt geborgd dat de analyse *in control* is. Pas wanneer er sprake is van een significante technisch-analytische afwijking wordt, op basis van een klinisch afgeleide prestatielimiet, beoordeeld of de afwijking ook klinisch significant is. Is dit het geval, dan wordt daarop actie ondernomen, bijvoorbeeld met een *recall* procedure.

F.3 Bepalen van de sigmascore

De berekening van de sigmascore is afhankelijk van de proces-inrichting van het laboratorium. Indien het uitgangspunt is dat alle analyzers bewaakt worden met inachtneming van dezelfde grenzen (het z.g. virtuele analyzerprincipe), dan moeten de APS ook voor de virtual analyzer vastgesteld worden.

De APC wordt in dit geval bepaald in overeenstemming met de ISO/TS 20914:2019 (Medical laboratories — Practical guidance for the estimation of measurement uncertainty). De voorwaarde voor het gebruik van meetonzekerheid is de afwezigheid van kritische bias, wat bepaald kan worden op basis van Milaan-consensusmodellen (1) uitkomst, (2) biologische variatie [EFLM Biological Variation Database], of (3) *state of the art*. De bias kan met behulp van EQA of commuteerbaar materiaal zoals patiëntenpool materiaal bepaald worden. Hiervoor wordt minimaal een half jaar aan data afkomstig van routinematige metingen van intern kwaliteitscontrolemateriaal, idealiter controlematerialen van derde partijen [Braga et al., 2020] gebruikt. Als eerste stap wordt per analyzer en level en lot van controlemateriaal bepaald wat de gemiddelde ligging en de spreiding van dit controlemateriaal is. De spreiding van de gemiddelde waarden tussen de analyzers wordt bepaald als de tussen-analyzervariatie. De gevonden variatie binnen analyzers wordt *gepoold* om de *gepoolde* binnen-analyzervariatie te berekenen. Vervolgens wordt de som der kwadraten van de tussen-analyzervariatie en de binnen-analyzervariatie berekend, waarvan de wortel de meetonzekerheid op virtual-analyzer niveau beschrijft.

De sigmascore van de virtual analyzer wordt vervolgens berekend door de toelaatbare meetonzekerheid te delen door de meetonzekerheid van de virtual analyzer.

In het geval dat de bias tussen apparaten niet kan worden verholpen, of als de voorkeur uitgaat naar het instellen van SQC per analyzer, dan kan de sigmascore ook bepaald worden per analyzer. De tussen-apparaatvariatie kan periodiek worden geëvalueerd, met commuteerbare interne kwaliteitscontrolematerialen (IQC-component 2) of met behulp van externe kwaliteitscontroledata. Per analyzer wordt de gemiddelde ligging van het controlemateriaal bepaald, alsmede de variatie binnen de analyzer.

Bij het bepalen van de sigmascore per apparaat wordt gebruik gemaakt van het MAU of TEa concept, waarbij de bias van de individuele apparaten m.b.v. EQA of andere commuteerbare materialen (zoals patiëntenpool materiaal) wordt vastgesteld. Daarnaast wordt de binnen-apparaatvariatie bepaald m.b.v. IQC-component 2, waarna de sigmascore berekend kan worden door de gevonden bias van de toelaatbare analytische prestatie eisen (APS) af te trekken, en het resulterende getal te delen door de binnen-apparaatvariatie. Vervolgens wordt per apparaat en per QC-niveau een sigmascore berekend, waarna TEa of andere APS kan worden vastgesteld.

F.4 Rekenvoorbeelden *Sigma Metrics* QC-procedures

Zoals beschreven in paragraaf 7.4 wordt TEa berekend op basis van de biologische variatie (BV). Bij een bepaalde sigmascore is er een lineair verband tussen bias en de imprecisie (CVa):

$$-\text{sigma} \times \text{CVa} + \text{TEa} = \text{bias}$$

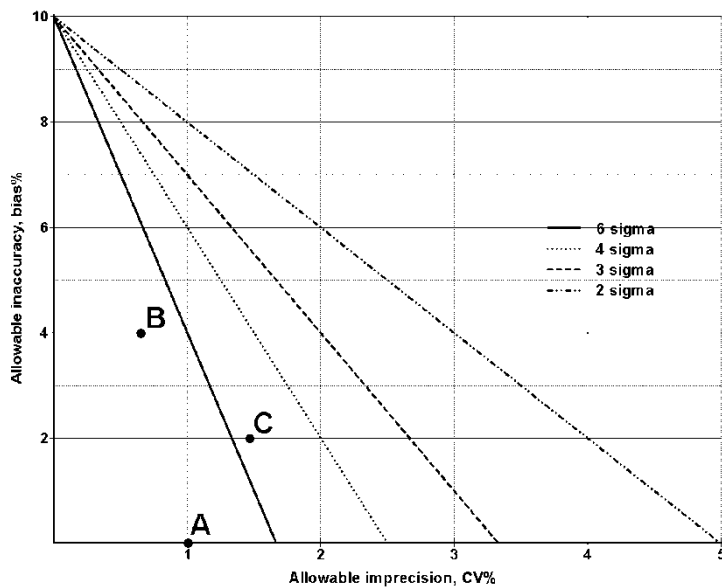
Dit beschrijft een rechte lijn (zie Figuur F.1) met

$$\text{Richtingscoëfficiënt} = -\text{sigma}$$

$$\text{As-afsnede} = \text{TEa}.$$

Bij bias=0 geldt: $\text{CV} = \text{TEa} / \text{sigma}$

Figuur F.1. Verband tussen bias en imprecisie voor een meetproces met een TEa van 10% [naar: Schoenmakers et al., 2011].



In dit voorbeeld is TEa gesteld op 10%. Merk op dat bij CVa=0 geldt bias = TEa.

Punt A geeft een bepaling weer zonder bias en met een analytische variatie van 1%. Berekening van de sigmascore geeft een waarde van $(10 - 0)/1 = 10$.

Punt B geeft een bepaling weer met een bias van 4% en een analytische variatie van 0,6%. De sigmascore is $(10 - 4)/0,6 = 10$.

Point C geeft een bepaling weer met een bias van 2% en een analytische variatie van 1,5%. De sigmascore is $(10 - 2)/1,5 = 5,3$.

Merk op dat de sigmascore 1,65 is, als precies wordt voldaan aan voorwaarde TE=TEa.

Noot

Aanname: geen bias (bias = 0).

$$TEa = 0 + 1,65 \times CVa \text{ (allowable)}$$

$$\text{Sigma} = TEa - \text{bias} / CVa$$

Met TEa=0=1,65 x CV en CVa = CVa (allowable):

$$\text{Sigma} = 1,65 \times CV(\text{allowable}) / CV(\text{allowable}) = \underline{1,65}$$

Tabel F.1. Berekening van TEa en sigmascores voor verschillende bepalingen.

CVi: intra-individuele variatie; CVg: inter-individuele variatie; CVa: analytische variatie (volgens Schoenmakers et al., 2011); TEa: *total allowable error* (gebaseerd op *desirable* APS, volgens de EFLM Biological Variation Database, geraadpleegd op 27-02-2025).

Bepaling	CVi	CVg	TEa	CVa	Sigma score
Chloride	1,0	1,3	1,2	4,7	0,3
Natrium	0,5	0,7	0,6	1,1	0,5
Calcium	1,8	2,7	2,3	2,0	1,2
Magnesium	2,6	5,9	3,8	3,1	1,2
Ceruloplasmine	5,0	15,1	8,1	4,0	2,0
Glucose	4,6	8,1	6,1	2,9	2,1
Albumine	2,5	4,1	3,3	1,5	2,2
Kreatinine	4,4	16,2	7,8	2,8	2,8
Cholesterol	5,3	15,3	8,4	2,8	3,0
LDH	4,4	11,8	6,8	2,2	3,1
Totaal eiwit	2,6	3,5	3,2	0,9	3,6
ASAT	8,4	19,4	12,2	2,9	4,2
Kalium	3,9	5,3	4,9	1,1	4,5
Transferrine	3,9	13,9	6,8	1,3	5,2
HDL-cholesterol	5,7	20,0	9,9	1,9	5,2
LDL-cholesterol	7,7	20,2	11,8	2,1	5,6
Lipase	7,6	25,2	12,9	2,1	6,1
Triglyceriden	19,7	32,8	25,8	4,0	6,5
Ureum	13,3	20,6	17,1	2,6	6,6
Alkalische fosfatase (lever)	6,0	21,0	10,4	1,5	6,9
IgM	5,9	48,5	17,1	2,4	7,1
Urinezuur	8,1	22,4	12,6	1,7	7,4
Fosfaat	7,7	10,7	9,6	1,2	8,0
Haptoglobine	8,6	39,0	17,1	2,0	8,6
IgG	3,5	17,1	7,3	0,8	9,1
gGT	8,2	45,2	18,2	2,0	9,1
Ferritine	12,9	132,0	43,8	4,6	9,5
IgA	7,5	32,5	14,5	1,3	11,2
Amylase	6,6	24,2	11,7	1,0	11,7
CRP	33,7	86,4	51,0	3,7	13,8
Bilirubine totaal	20,2	24,6	24,6	1,7	14,5
ALAT	11,2	35,2	18,5	1,1	16,8
CK	14,1	32,9	20,6	1,2	17,2
IJzer	27,6	26,7	32,4	1,2	27,0

F.4.1 Verificatie of validatie van een test

Deze situatie doet zich bijvoorbeeld voor bij introductie en bij validatie van een test. Dit gebeurt doorgaans met kwaliteitscontrolemonsters en gedurende een langere periode, zodat variaties door reagens wisselingen en onderhoud van apparatuur ook in rekening worden gebracht. Bias wordt in dit geval doorgaans niet opgenomen in de berekening aangezien het concept van meetonzekerheid hier feitelijk van toepassing is, namelijk dat er eigenlijk geen bias mag zijn aangezien hiervoor dient te

worden gecorrigeerd. Het laboratorium kan ook eigen referentiewaarden vaststellen, waarmee de (methodeafhankelijke) bias op een andere manier wordt ingecalculleerd. Bias zal doorgaans geen invloed op de afweging hebben of een test goed genoeg is om in de praktijk toegepast te worden; de imprecisie is daarvoor juist wel belangrijk. De sigmascore is: $(1,65 \times 0,5 \times CVi) / CVa$. Zodra de sigmascore > 3 , zou de test in theorie kunnen worden gebruikt, mits er geen andere redenen bestaan die maken dat de test niet geschikt is voor de klinische toepassing. Bij een sigmascore < 3 is de kwaliteit niet zonder meer voldoende. Uiteraard kan voor dergelijke niet-robuste tests gelden dat dit nu eenmaal de *state-of-art* is en toch een relevante test voor de klinische zorg blijven. De natriumbepaling is hiervan een uitstekend voorbeeld.

F.4.2 Toepassen van QC-regels op basis van berekende sigmascore

In het algemeen is het aan te bevelen om de interne kwaliteitscontrole te ontwerpen per analytische procedure en indien mogelijk per controle materiaal en level. Hierbij wordt het gemiddelde en de bijbehorende standaarddeviatie berekend uit minimaal 20 resultaten (de initiële SD), bij voorkeur over meer resultaten en een lange periode. In het laatste geval zal de SD groter zijn. Bij het ontwerpen van de interne QC met behulp van de Westgard (multi)QC-regels dient die bewakingsregel of combinatie van bewakingsregels te worden gebruikt waarbij de kans op onterecht afkeuren van controlemetingen zo laag mogelijk is (in ieder geval lager dan 5%; $P_{fr} < 0,05$) en de detectie van echte fouten zo hoog mogelijk is (in ieder geval hoger dan 90%; $P_{ed} > 0,90$).

Bij het berekenen van de sigmascore voor een analytisch proces dient allereerst de tolerantie in het resultaat te worden vastgesteld. Veel gebruikte bronnen voor de tolerantie (TE) zijn de gegevens omtrent biologische variaties (EFLM Biological Variation Database; TEa) en de TE uit de rondzendingen van de SKML op basis van *state-of-art* (TEsa). Deze score is een beoordeling van het analytische proces. Sigmascores > 6 zijn ideale processen en kunnen bewaakt worden met een enkelvoudige bewakingsregel welke voldoet aan de regel $1_{s4,35}$. In de praktijk is de 1_{s4} regel geschikt. Bij sigmascores 3-6 dient de IQC ontworpen te worden zoals beschreven waarbij vastgelegd wordt de bron van de TE, de P_{fr} , P_{ed} , de frequentie en het aantal levels. Bij processen met sigmascore < 3 dienen maatregelen te worden genomen om de bewaking mogelijk te maken op basis van de Westgard (multi)QC-regels.

Er kan verwezen worden naar het artikel van Schoenmakers et al. [2011] waarin een tabel is opgenomen met sigmascores (scores 3-6) waarbij voor iedere score de bijbehorende Westgard (multi)QC-regels, aantal levels en IQC-metingen, inclusief de kansen op zowel terechte als onterechte foutdetectie getabelleerd zijn. Dit is een handig en praktisch overzicht dat behulpzaam is voor toepassing van *Sigma Metrics* in het eigen laboratorium. Daarnaast zijn er intussen diverse software tools die op basis van ingebrachte TEa en CVa dergelijke calculaties verrichten en adviseren welke Westgard (multi)QC-regels en controlelevels toegepast moeten worden. Ook in ondersteunende softwarepakketten zoals Unity (BioRad) of softwarepakketten van Randox zijn deze functionaliteiten bijvoorbeeld ingebouwd.

F.4.3 Niet-robuste tests met sigmascore < 3

Het ontwerpen van de IQC is alleen gericht op het efficiënt organiseren van de bewaking van een analytisch proces waarbij de uitgangspunten zijn een $P_{fr} < 0,05$ en een $P_{ed} > 0,90$. Bij sigmascores < 3 is dit niet meer mogelijk en kan de bewaking alleen gericht zijn op maximale QC op basis van de analytische prestaties alleen. Het TEa-concept gaat dus niet meer op tenzij het analytisch proces wordt verbeterd (kleinere CVa).

Bij gebruik van de BV als bron van de TEa kan deze worden verruimd, bijvoorbeeld door te kiezen voor een minimale performance in plaats van een wenselijke of optimale performance [Fraser et al., 1997]. Mocht dit niet toereikend zijn om een sigmascore > 3 te verkrijgen, dan kan als bron voor de TE op basis van biologische variaties over worden gegaan op de *state of the art* (TEsa) bijvoorbeeld vanuit de EQC-resultaten.

Mocht dit nog steeds geen soelaas bieden, dient een beredeneerd alternatief en een op maat gemaakte aanpak gerealiseerd te worden. In de praktijk betekent dit vaak het bewaken op basis van de analytische prestatie (*state of our own art*).

Bijlage G Six Sigma gebaseerde kwaliteitscontrole met het *Moving Range* model

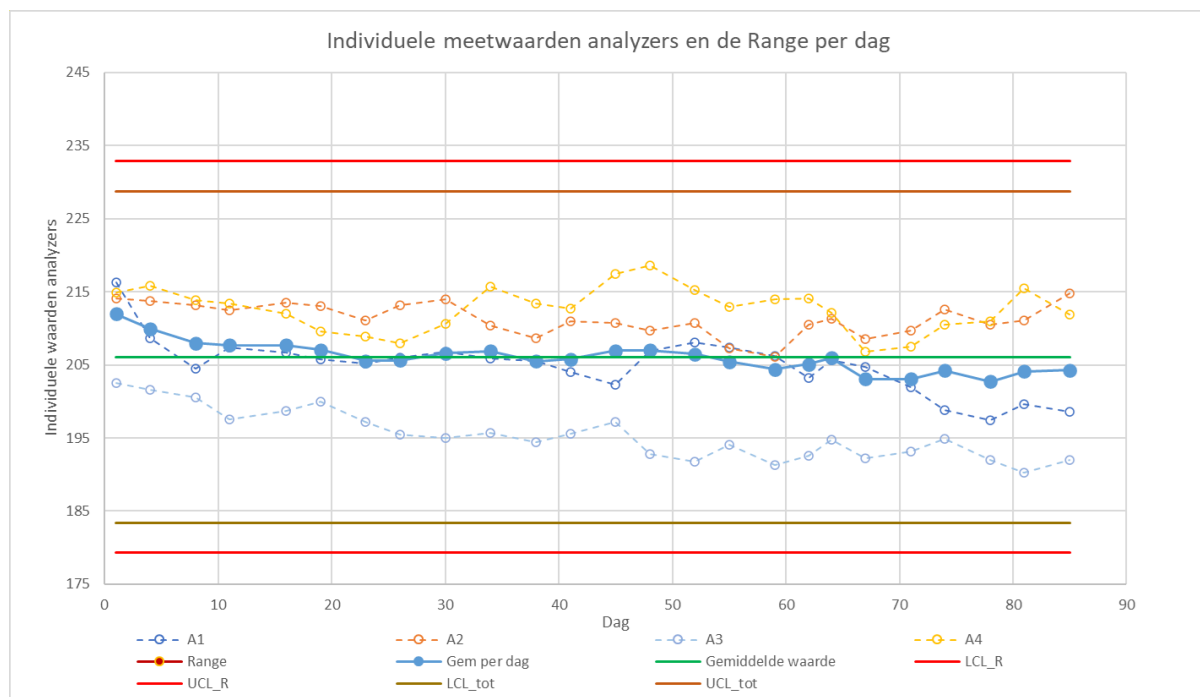
Binnen de klinische chemie wordt het Six Sigma model vooral toegepast zoals het door Westgard [2003] is beschreven. Buiten de klinische chemie wordt het Six Sigma model ook veel toegepast. Daarbij wordt echter andere statistiek toegepast, zoals de EWMA (*Exponentially Weighted Moving Average*). Van deze laatste methode is de toepasbaarheid ook in de klinische chemie aangetoond [Jansen et al., 2002; Linnet, 2006]. Hoewel de toepassing wordt belemmerd door een beperkte beschikbaarheid in de huidige kwaliteitscontrole software, wordt deze methodiek toch hier toegelicht.

De navolgende uitwerking van deze de Six Sigma methodiek gaat uit van het bewaken van de analytische prestatie van een klinisch chemische analyzer, dus zonder in eerste instantie rekening te houden met de biologische variatie en daarvan afgeleide grenzen. Deze controlelimieten – gebaseerd op klinisch relevante limieten - kunnen natuurlijk ook in dit model toegevoegd worden.

Voorbeeld virtuele analyzer met gebruik van het *Moving Range* model

Het uitgangspunt in de navolgende benadering is, dat gebruik wordt gemaakt van een virtuele analyzer. De controle-metwaarden van de verschillende analyzers worden samengevoegd (onder de aanname dat deze niet gecorreleerd zijn, ofwel statistisch onafhankelijk van elkaar zijn).

Met dit model worden één gemiddelde en gelijke grenswaarden voor alle analyzers vastgesteld en deze doel- en grenswaarden worden voor de hele organisatie gebruikt. In onderstaande figuur zijn de individuele punten van 4 analyzers voor ferritine en de bijbehorende controlegrenzen weergegeven.



In de berekening worden 3 verschillende vormen van spreiding gebruikt:

- $SD_{\text{korte termijn}}$: een andere term voor de korte termijn of analytische spreiding is de dag-tot-dag spreiding. Deze spreiding is gebaseerd op de *random* spreiding binnen één analyzer waarvoor niet gecorrigeerd kan worden. Deze wordt berekend met behulp van de *Moving Range*. Deze analytische spreiding is de meest basale prestatie zonder additionele variatie of drift, is voor elke analyzer uniek en blijkt door kwaliteitsverschillen toch te kunnen variëren tussen de verschillende analyzers.
- $SD_{\text{lange termijn}}$: dit is een lange termijn spreiding per analyzer die bestaat uit *shifts* en trends binnen een analyzer in combinatie met *random* spreiding. Een term die niet in de lange termijn standaardafwijking wordt meegenomen (of kan worden meegenomen) is een algehele drift van een laboratorium ten opzichte van andere laboratoria. $SD^2_{\text{lange termijn}} = (SD^2_{\text{totaal}} - SD^2_{\text{korte termijn}})$.
- SD_{totaal} : dit is de totale spreiding die opgebouwd is uit de *pooled* lange termijn spreiding en de bias verschillen tussen de analyzers ($SD^2_{\text{totaal}} = SD^2_{\text{lange termijn}} + SD^2_{\text{tussen analyzers}}$).

De buitenste rode lijnen zijn gebaseerd op SD_{totaal} en een *coverage* factor van 3. De binnenste donkere rode lijnen is de $SD_{\text{lange termijn}}$, ook met een *coverage* factor van 3. Als de sigmascore het toelaat (grenzen liggen buiten SD_{totaal} van de analyzers) kan voor de vaststelling van de sigmascore bijvoorbeeld de minst goede analyzer genomen worden (met de hoogste $SD_{\text{korte termijn}}$). De sigmascore is bepalend hoe met een *out-of-control* signaal wordt omgegaan.

G.1 Berekeningen

G.1.1 Verschillen opeenvolgende waarden (*Moving Range*)

Bij een Levey-Jennings grafiek [Levey en Jennings, 1950] is de randvoorwaarde dat de controlepunten onafhankelijk van elkaar zijn. Een vereiste is echter, dat de datapunten systematisch in de tijd worden bepaald en een tijdreeksanalyse mogelijk is. Verondersteld wordt dat binnen deze tijdsperiode van opeenvolgende controles het proces stabiel is en dat ook de gemeten spreiding binnen deze periode stabiel en *random* verdeeld is. Vaak wordt een frequentie genomen om een- of tweemaal per dag een controle uit te voeren. Dit betekent dat de data op basis van 12 of 24 uur gesorteerd kunnen worden en voor die betreffende periode de spreiding kan worden berekend.

Een centrale parameter om de spreiding te kunnen berekenen is de gemiddelde *Moving Range* (MR_{gem}). Dit zijn de absolute verschillen van steeds twee opeenvolgende controlewaarden bij een vast tijdsinterval tussen de controles (binnen één apparaat/rotor). Aangezien de opeenvolgende controlewaarden onafhankelijk van elkaar zijn is de gemiddelde waarde van deze absolute verschillen een maat voor de spreiding. De berekende spreidingswaarde is de laagste spreiding die voor de analyzer kan worden berekend ofwel binnen het interval van de controles is de berekende (korte termijn) spreidingswaarde de best haalbare waarde. De berekende waarde is afhankelijk van de frequentie van de gemeten controles. Wordt de periode tussen de controles kleiner dan kan de (korte termijn) spreiding eveneens afnemen. De controlefrequentie dient afgestemd te worden om afwijkingen binnen een analyzer te detecteren en de kosten van procescontrole.

Een tweede belangrijke functie van de MR_{gem} is om te controleren of de aanname dat de controlewaarden onafhankelijk van elkaar zijn klopt. Als de $SD_{MR_{\text{gem}}}$ en SD_{ANOVA} niet aan elkaar gelijk zijn dan zijn de waarden niet onafhankelijk van elkaar en betekent dit dat ze door andere factoren zoals lotverschillen, veroudering van reagens, etc. worden beïnvloed.

De korte termijn spreiding wordt berekend m.b.v. de gemiddelde *Moving Range* (MR_{gem}). De ratio van de MR_{gem} met de constante d_2 (1,128 voor twee opvolgende waarden $n=2$) geeft een schatting van de dag-tot-dag of korte-termijn SD:

$$SD_{\text{korte termijn}} = MR_{\text{gem}} / 1,128$$

Bij een stabiele analyzer (zonder trends of *shifts*) is de berekende MR-spreiding gelijk aan de spreiding van de lange termijn / populatie (ANoVA of berekend met Levey-Jennings). Zijn er trends of shift binnen de Levey-Jennings controlekaart zichtbaar, dan is de korte-termijn spreiding berekend met de *Moving Range* lager dan de lange termijn SD bepaald met Levey-Jennings of ANoVA. De korte-termijn SD kan ook hoger zijn dan de ANoVA-SD. Dit betekent dat de analyzer extra spreiding toevoegt wat kan wijzen op een technisch probleem binnen de analyzer.

Een voordeel van de spreiding berekend met de *Moving Range* is dat bij een *out-of-control* situatie onderscheid kan worden gemaakt tussen een toename van de analytische (korte-termijn) spreiding en een *shift* of trend.

G.1.2 Individuele analyzer

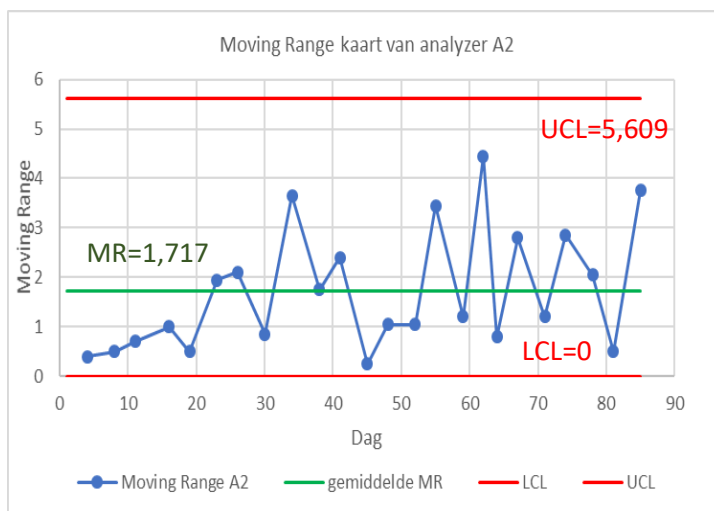
Met de *Moving Range* kunnen de controlegrenzen voor de individuele controlekaart worden afgeleid.

De controlegrenzen voor de MR-kaart worden als onderstaand afgeleid. Hierbij worden dus de absolute verschillen tussen opeenvolgende waarden in de controlekaart uitgezet. Dit is een maat voor de korte-termijn spreiding. Voor $n=2$, verschil opeenvolgende waarden:

$$UCL = MR_{\text{gem}} \times B_4 \quad (\text{schatting voor } m + 3SD) \quad (B_4=3,267 \text{ voor } n=2)$$

$$LCL = MR_{\text{gem}} \times B_3 \quad (\text{schatting voor } m - 3SD) \quad (B_3=0 \text{ voor } n=2)$$

In onderstaande figuur is op de horizontale as de dag weergegeven waarop de controle is uitgevoerd en de verticale as is de *Moving Range* waarde weergegeven. De bovenste controlegrens (UCL: *Upper Control Limit*), de gemiddelde *Moving Range* en de onderste controlegrens (LCL: *Lower Control Limit*) zijn weergegeven.



In dit voorbeeld: $MR_{\text{gem}} = 1,717$; $LCL = 0 \times 1,717 = 0$; $UCL = 3,267 \times 1,717 = 5,609$

Deze controlekaart wordt alleen gebruikt bij het vaststellen van de korte termijn- of analytische spreiding van een analyzer van de virtuele analyzer (*coverage factor* = 3).

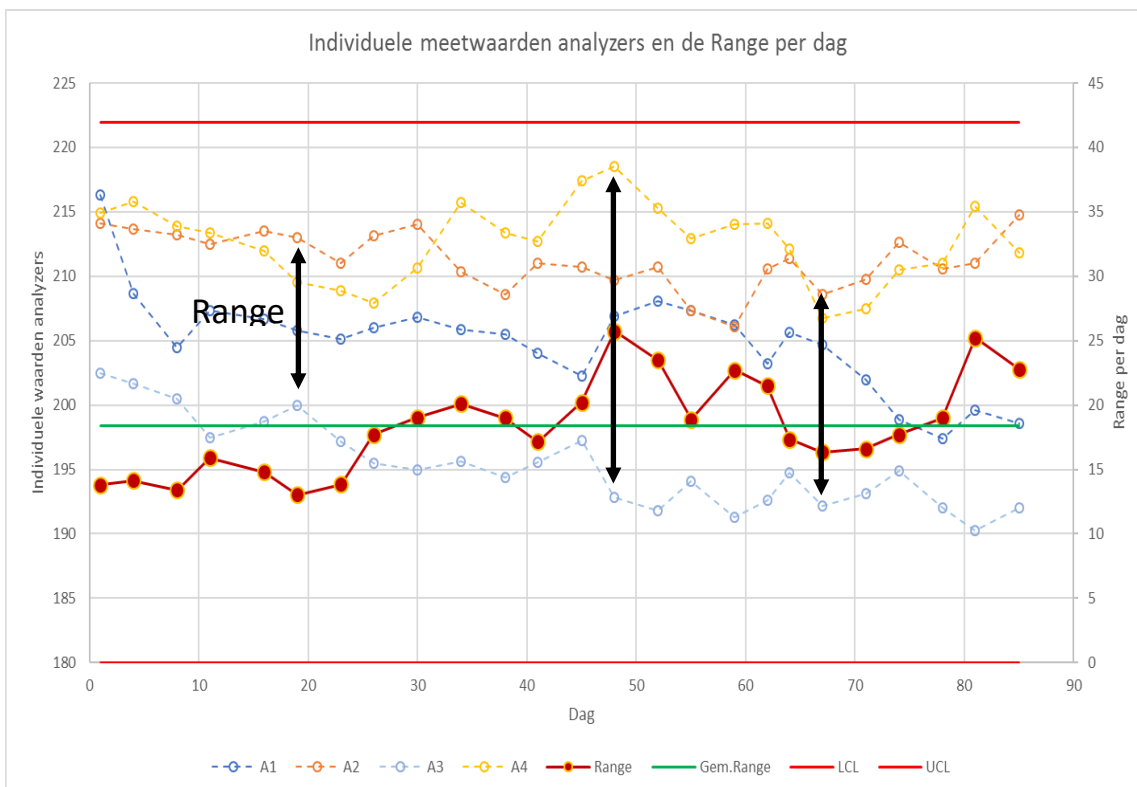
Bij het gebruik van meerdere analyzers kunnen de individuele $SD_{\text{korte termijn}}$ waarden worden *gepooled* in een $SD_{\text{korte termijn}}$ (*pooled*). De $SD_{\text{korte termijn}}$ wordt gebruikt voor het vaststellen van de controlegrenzen van de standaard controlekaart ($3 \times SD_{\text{korte termijn}}$). Deze kaart is hierdoor gevoelig voor het detecteren van uitbijters, *shifts* en trends.

De $SD_{\text{korte termijn}}$ komt overeen met de $SD_{\text{korte termijn}}$ die bij de validatie en/of vrijgave van apparatuur is bepaald.

G.1.3 Virtuele analyzer

Bij een virtuele analyzer wordt het maximale verschil (*range*) bepaald tussen de analyzers die alle op dezelfde dag zijn bepaald.

Uit de gemiddelde *range* (per dag) kan een schatting gemaakt worden over de spreiding tussen de verschillende analyzers. De spreiding van deze parameter is een maat voor de combinatie van korte- en lange termijn spreiding. De spreiding is een goede schatting van de totale spreiding van de virtuele analyzer.



Voorbeeld berekening bovenste (UCL) en onderste (LCL) controlegrens:

Gemiddelde *range* = 18,38

$$LCL = 18,38 \times D3(n=4) = 0$$

$$UCL = 18,38 \times D4(n=4) = 18,38 \times 2,282 = 41,94$$

SD_{totaal} wordt berekend door de gemiddelde *range* te delen door $d2(n=4)$:

$$SD_{\text{totaal}} = R_{\text{gem}} / d2(n=4) = 18,38 / 2,059 = \underline{8,927}$$

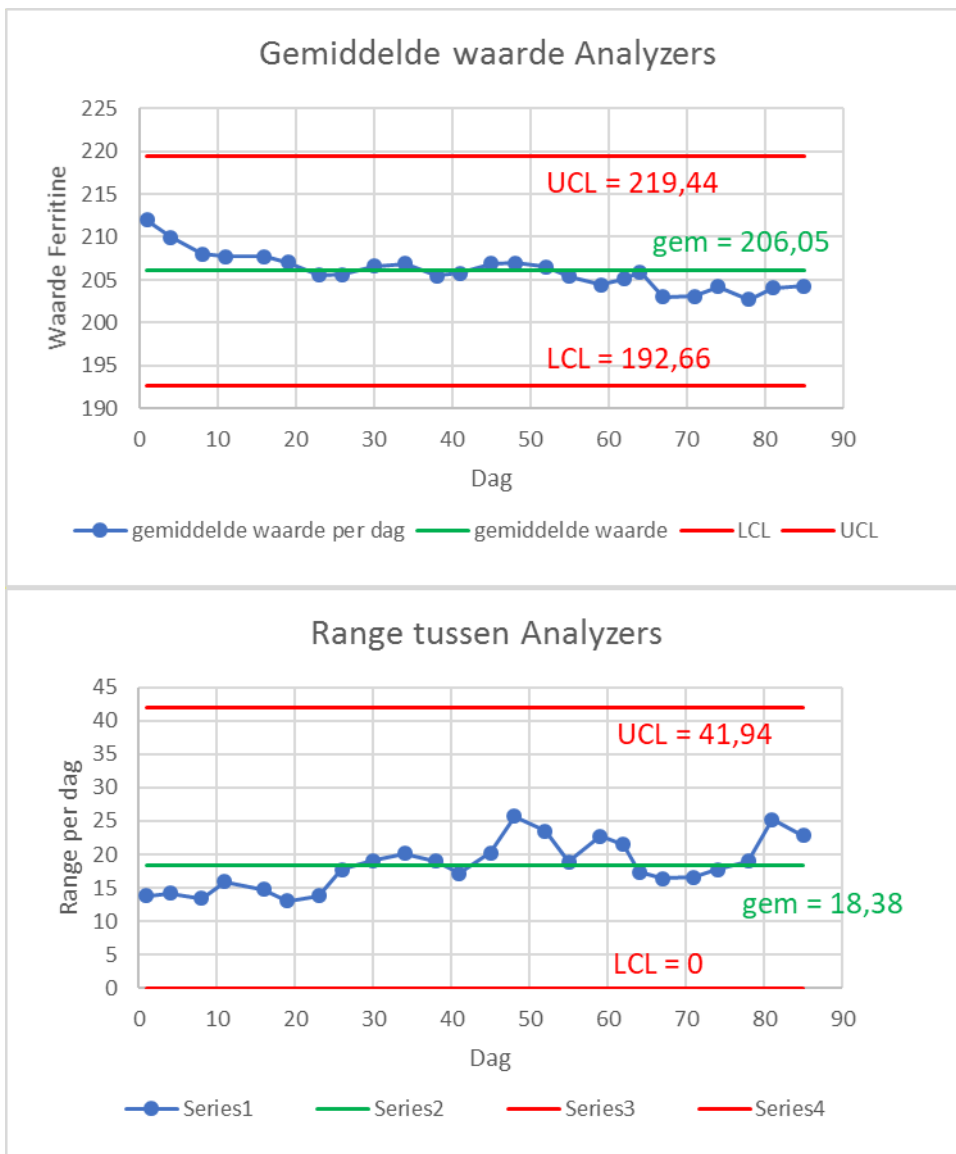
Een andere optie om SD_{totaal} te bepalen is m.b.v. de spreiding berekend door alle data van de analyzers te *poolen*. De randvoorwaarde is wel dat de data normaal zijn verdeeld (te controleren met bijv. de Benard approximation of the Median Rank (MR) test [Abernethy, 2004]).

Deze controlekaart wordt alleen gebruikt bij het vaststellen van de controlegrenzen van de virtuele analyzer (*coverage factor* = 3).

De gemiddelde waarde over alle analyzers van de virtuele analyzer kan ook worden bepaald. Als wordt verondersteld dat alle analyzers van de virtuele analyzer onafhankelijk van elkaar zijn, dan zou de gemiddelde waarde constant moeten zijn. De controlegrenzen die voor deze gemiddelde *Moving Average* kaart kunnen als volgt worden berekend:

$$\text{Bovenste- en onderste controlegrens (UCL, LCL)} = MA \pm (A_2 \times R_{\text{gem}})$$

Waarbij: $(A_2 \times R_{\text{gem}}) = (\text{schatting voor gemiddelde waarde} \pm 3SD)$



Voorbeeld:

Gemiddelde *range* (R_{gem}) = 18,38; met $A_2(n=4) = 0,729$:

$$R_{\text{gem}} \times A_2(n=4) = 0,729 \times 18,38 = 13,4;$$

Gemiddelde waarde analyzers = 206,06 dus:

$$\text{UCL} = 206,05 + 13,4 = 219,45 \text{ en}$$

$$\text{LCL} = 206,05 - 13,4 = 192,65.$$

G.1.4 Onderlinge samenhang van de verschillende spreidingsbronnen

De totale spreiding van een analiet wordt berekend op basis van zowel korte- als lange termijn spreiding per analyzer en de verschillen tussen analyzers. De buitenste controlegrenzen zijn afkomstig van $3SD_{\text{totaal}}$; de binnenste controlegrenzen zijn afkomstig van de *pooled* $3SD_{\text{langetermijn}}$ van de 4 analyzers.

Voor individuele analyzers geldt:

$$SD_{\text{lange termijn}} = \sqrt{SD_{\text{korte termijn}}^2 + SD_{\text{trends/shifts}}^2}$$

$SD_{\text{korte termijn}}$ = identiek aan spreiding berekend onder 1 en vergelijkbaar met $SD_{\text{analytisch}}$.

Voor een combinatie van analyzers geldt:

$$SD_{\text{tussen analyzers}} = \text{afgeleid van } \sqrt{SD_{\text{totaal}}^2 - SD_{\text{lange termijn}}^2}$$

SD_{totaal} kan geschat worden door de spreiding over alle analyzers over een langere periode te bepalen mits de data normaal zijn verdeeld. Eveneens kan de SD_{totaal} geschat worden uit de gemiddelde *range* per tijdseenheid tussen alle analyzers.

Merk op als SD_{totaal} wordt gecorrigeerd voor $SD_{\text{lange termijn}}$, dan is dat de $SD_{\text{tussen analyzers}}$ en geeft dit een schatting voor de bias tussen de analyzers. Voor het berekenen van de bias tussen analyzers wordt verwezen naar de *Moving Average* kaarten.

G.1.5 Moving Average kaarten

De Cumulatieve Som (Cusum) en *Exponential Weighted Moving Average* (EWMA) kaarten zijn *Moving Average* kaarten [Jansen et al., 2002; Linnet, 2006]. In tegenstelling tot de *multirules* methode van Westgard, welke alleen maar de waarschijnlijkheid aangeeft van het voorkomen van een bepaalde afwijking (*shift*), wordt de afwijking in de EWMA-kaart in de betreffende meeteenheden weergegeven. De EWMA-kaart heeft verschillende voordelen, o.a. de lengte van de *Moving Average* kan eenvoudig worden gevarieerd (de kaart kan wisselen tussen een “normale” Shewhart controlekaart (controlegrenzen gebaseerd op $SD_{\text{kortetermijn}}$) en *Moving Average* kaart). Verder kan de *Average Run Length* (gemiddeld aantal controlepunten voordat een afwijking zichtbaar wordt) worden ingesteld. Tevens kan de grootte van de *shift* die relevant is voor proces worden ingesteld (kleinere *shifts* worden niet meegenomen en worden behandeld als ruis). Dit betekent dat met het instellen van deze parameters dezelfde methode voor alle bepalingen kan worden toegepast.

In de onderstaande tabel is voor de verschillende controlekaarten de *Average Run Length* (ARL) berekend. In deze tabel is duidelijk te zien dat de EWMA/CUSUM-kaart voor kleinere *shifts* (bijv. 1

sigma *shift*) duidelijk beter presteert dan de Shewhart kaart. Voor geen en grote *shifts* is de prestatie van de verschillende controlekaarten vergelijkbaar. De prestaties van de *multirules* van Westgard geeft vergelijkbare waarden als de EWMA-kaarten [Barger et al., 2010].

In de tabel wordt de *shift* of afwijking uitgedrukt in de standaardafwijking gecorrigeerd voor de lengte van de *Moving Average* (MA) weergegeven. De controlegrenzen van de kaarten zijn ingesteld voor een ARL(0) van 500 (bij een stabiel proces zonder *shift* is er gemiddeld een alarm per 500 datapunten).

Shift {in multiples of σ / \sqrt{n} }	Standard control chart	CUSUM chart $h=5, k=0,5$	EWMA chart $l=0,142, k=2,79$	Multyrules Westgard $1_{3s}+2$ of $3_{2s}+4_{1s}+8_x$
0,00	370,4	465,4	370,4	111
0,25	282,1	139,5	102,9	
0,50	155,2	38	31,1	
0,75	81,2	17	15,2	
1,00	43,9	10,4	9,6	11
1,25	25	7,4	6,9	
1,50	15	5,7	5,5	
1,75	9,5	4,7	4,5	
2,00	6,3	4	3,9	4
2,25	4,4	3,5	3,4	
2,50	3,2	3,1	3	
2,75	2,5	2,8	2,7	
3,00	2	2,6	2,5	2

In de praktijk is het verwijderen van uitbijters van belang bij het gebruik van *Moving Average* kaarten. Bias wordt gedurende langere tijd door uitbijters verstoord. Hierdoor kan de exacte bias onvoldoende worden vastgesteld.

Moving Average kaarten zijn een essentieel onderdeel van het werken met het virtuele analyzer model. Aangezien de $SD_{\text{korte termijn}}$ vaak de grootste bron van spreiding is, is het vaststellen van een bias en een verandering in de bias tussen de verschillende analyzers haast onmogelijk. *Moving Average* kaarten helpen om bias tussen analyzers met een kleiner betrouwbaarheidsinterval ($\frac{1}{\sqrt{n} (\text{lengte van de MA})}$) te bepalen. De hogere betrouwbaarheid is ook waardevol bij het interpreteren van EQC-uitslagen een analyzer binnen het cluster van alle analyzers binnen het laboratorium (de $SD_{\text{korte termijn}}$ is vaak de grootste spreidingsbron).

Bij de berekening van de verschillende controlegrenzen en spreiding kan gebruikt worden van de volgende tabel met constanten [Montgomery, 2005]:

n	d2		B3	B4	B5	B6		A	A2	A3
1										
2	1,128		0	3,267	0	2,606		2,121	1,88	2,659
3	1,693		0	2,568	0	2,276		1,732	1,023	1,954
4	2,059		0	2,266	0	2,038		1,5	0,729	1,628
5	2,326		0	2,039	0	1,964		1,342	0,577	1,427
6	2,534		0,03	1,97	0,029	1,874		1,225	0,483	1,287
7	2,704		0,118	1,832	0,113	1,806		1,134	0,419	1,182
8	2,847		0,185	1,815	0,139	1,751		1,061	0,373	1,099
9	2,97		0,239	1,761	0,232	1,707		1	0,337	1,032
10	3,078		0,284	1,716	0,276	1,649		0,949	0,308	0,975

Bijlage H *Patient-Based Real-Time Quality Control*

Stappenplan voor het inrichten van *Patient-Based Real-Time Quality Control* of lopende patiëntengemiddelden:

1) Stel vast voor welke analieten een lopend patiëntengemiddelde van toegevoegde waarde zou kunnen zijn

Algemene eigenschappen voor analieten waarbij lopend patiëntengemiddelde van meerwaarde zouden kunnen zijn, zijn analieten met een hoog bepalingsvolume en een lage variatie in uitslagen. Er zijn twee strategieën beschreven:

- A) Het toevoegen van lopend patiëntengemiddelde als IQC ten behoeve van bepalingen waarvoor de traditionele interne kwaliteitscontrole tekortschiet. Dit betreft bijvoorbeeld
 - i) analieten met een lage sigmascore;
 - ii) een risico op het acuut falen van een meetsysteem (bijv. het falen van een ISE electrode);
 - iii) testen zonder, of met instabiel of niet commuteerbaar controlemateriaal (bijv. hemocytometrie QC).
- B) Het gebruiken van lopende patiëntengemiddelden om de inzet van controlemetingen efficiënter in te richten.

Het is het meest praktisch om te starten met het implementeren van lopende patiëntengemiddelden voor de meest kritische tests. Het stroomschema zoals weergegeven in Figuur 8.2 kan hiervoor als uitgangspunt worden gebruikt.

2) Optimalisatie en validatie instelling voor lopende patiëntengemiddelden

Dit is over het algemeen het meest uitdagende deel. Twee essentiële criteria voor een lopend patiëntengemiddelde procedure zijn:

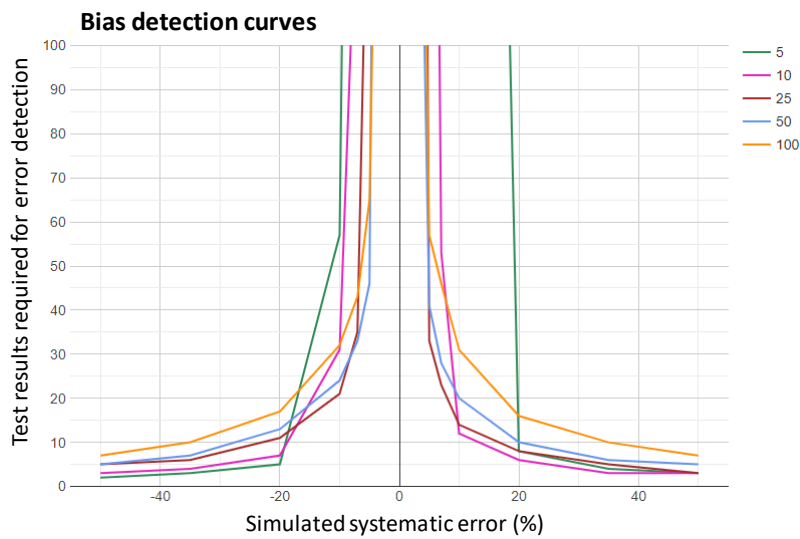
- i) het aantal alarmen dat wordt verkregen is beheersbaar;
- ii) het is bekend welke fouten wel en welke fouten niet kunnen worden gedetecteerd.

Het eerste is een praktische vereiste om alarmmoeheid te vermijden en het laatste is om alleen die lopend patiëntengemiddelde te gebruiken die bijdragen aan de algehele kwaliteitsborging van een test en die misschien kunnen worden gebruikt om het aantal IQC-metingen te verminderen.

In het verleden werden *trial-and-error*-methoden gebruikt die alleen de variatie van gemiddelde waarden beschrijven, en zich alleen richten op het hebben van een beheersbaar aantal alarmen. Dit is in de praktijk waarschijnlijk nog steeds de meest gebruikte methode. Bij deze methoden is er geen geobjectiveerd inzicht in de foutdetectieprestaties van de lopend patiëntengemiddelde en of ze überhaupt van waarde zijn.

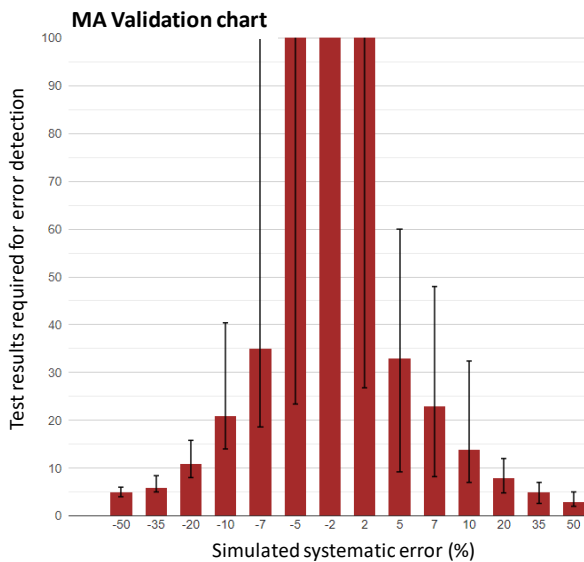
Recent ontwikkelde simulatiemethoden voor foutdetectie stellen laboratoria in staat om laboratoriumspecifieke procedures te optimaliseren en te valideren door simulaties uit te voeren op hun laboratoriumspecifieke datasets. Deze inzichten maken echte optimalisatie voor het detecteren van de meest relevante fouten mogelijk, evenals het valideren van hun prestaties voor het detecteren van alle fouten van belang. Deze resultaten worden gerapporteerd in zogenaamde bias detectiecurves (voor optimalisatie) en *validation charts* (voor validatie) [van Rossum, 2019].

Figuur H.1. Bias detectiecurves voor optimalisatie van een lopend patiëntengemiddelde voor calcium.



Bias detectiecurves zijn weergegeven voor 5 procedures voor berekening van lopende patiëntengemiddelden na het uitvoeren van errordetectie simulaties. De weergegeven lopende patiëntengemiddelden in de figuur berekenen het gemiddelde van de laatste 5, 10, 25, 50 of 100 calciumresultaten. De gekleurde lijnen geven het mediaan aantal patiëntenresultaten weer dat nodig is voor errordetectie. Deze bias detectiecurves kunnen vervolgens worden gebruikt om het optimale algoritme te selecteren voor lopend patientgemiddelde.

Figuur H.2. Moving Average validation chart van een lopend patiëntengemiddelde voor calcium.



De validation chart is verkregen voor een lopende patiëntengemiddelde voor calcium dat het gemiddelde berekent van de laatste 25 patiëntenresultaten. De rode kolommen geven het mediane aantal patiëntenresultaten weer dat nodig is voor foutdetectie en de error-bars het 95% interval van het aantal testen dat nodig is geweest voor foutdetectie. De Moving Average validation chart wordt gebruikt voor het valideren van de errordetectie eigenschappen van een lopend patiëntengemiddelde. Als voorbeeld; figuur H.2. laat zien dat een fout van 10% met een kans van

50% wordt gedetecteerd in 14 patiënten resultaten, en met een kans van 97,5% in 33 patiëntenresultaten. Voor het detecteren van grotere fouten zijn minder patiënten resultaten nodig.

Het is aan te raden om de geschiktheid van de verkregen instellingen te bevestigen in een echte operationele omgeving vóór de ingebruikname. Belangrijk doel is of de lopende patiënten gemiddelden-instellingen daadwerkelijk een beheersbaar aantal alarmen geven. Pas op basis hiervan de instellingen (controlegrenzen) zo nodig aan.

3) ICT-inrichting voor uitvoeren van lopende patiënten gemiddelden, grafieken daarvan en alarmbeheer

Dit moet ten minste het volgende omvatten

- a) Programmeren van lopende patiëntengemiddelden-instellingen (truncatiegrenzen, algoritme en controlegrenzen) in de gebruikte software (kan apparaat software zijn (bijv. hemocytometrie), middleware of LIS-systeem en gebruik op apparaat niveau.
- b) 24/7 *Real-time* alarmering van analisten. Zeker als het doel van de lopende patiëntengemiddelden is om een continue analytische kwaliteitsbewaking te realiseren is een adequate continue alarmering essentieel.
- c) Het weergeven van de lopende patiëntengemiddelden in relevante grafieken voor beoordeling. Het is aan te bevelen om de lopende patiëntengemiddelden in een andere grafiek dan de QC te tonen. Door de grote hoeveelheid lopend patiëntengemiddelde resultaten, kan het lastig zijn de resultaten van de normale QC te zien.

4) Ontwerp en implementeer een alarmprotocol

Ontwerp een protocol voor analisten voor het beheren van lopende patiëntengemiddelde-alarmen. Enkele zaken die kunnen worden overwogen:

- a) QC z.s.m. uitvoeren: zou een enorm inzicht moeten geven of er een grotere kritieke fout is (of niet).
- b) Dupliceer uitslagen (bijv. 4) op een dubbel analyzer systeem; maakt detectie mogelijk van tijdelijke assay fouten of fouten die moeilijk te detecteren zijn met IQC (bijv. tests met een lage sigmascore).
- c) Recente uitslagen controleren (bijv. afzonderlijke patiënten met extreme uitslagen). Maak eventueel een beslisregel om deze patiënten uit te sluiten van toekomstige PBRTQC-berekeningen.

Bijlage I Stappenplan correctieve acties bij afwijkingen interne QC

Onderstaand schema is een handreiking voor correctieve acties bij afwijkende interne QC. Er zit een aantal afwegingen in, zoals niet oneindig kalibreren.

