

Kwaliteit met minder moeite; een evaluatie van kwaliteitscriteria op basis van biologische variatie

A.-K. BOER en G. van der SLUIJS VEER

Het analyseren van controlemonsters is een essentiële, maar kostbare en vooral arbeidsintensieve gelegenheid. In het verleden zijn diverse beoordelingscriteria ontwikkeld om een juiste balans te vinden tussen het ten onrechte af- en goedkeuren van bepalingenseries. In lijn met de Westgard-criteria, hebben we binnen ons laboratorium in 2001 nieuwe criteria ontwikkeld voor de bepalingen op de Roche Modular. Als basis werd gekozen voor de intra- en interindividuele biologische variatie. Als gevolg hiervan zijn de beoordelingscriteria van een zevental bepalingen verruimd naar een 3-, 4- of 5s regel. Deze studie laat retrospectief zien dat deze verruiming resulteert in een significante reductie van het aantal controlemonsters met 7%. Bovendien bleek het percentage controlemonsters dat moest worden herbepaald zelfs met 42% af te nemen. De gekozen methodiek leidt tot een grotere efficiëntie, zonder dat de kwaliteit van de bepalingen volgens externe kwaliteitscontroles van de SKZL afnam.

Trefwoorden: evaluatie; kwaliteitscontrole; biologische variatie

Het meten van controlemonsters vormt een belangrijke schakel in de kwaliteitsbewaking van een laboratorium. De kosten hiervan zijn aanzienlijk. Voor het goed- en afkeuren van kwaliteitscontroles zijn in de loop der tijd diverse beoordelingscriteria opgesteld (1). Alom bekend en nog steeds veel gebruikt, zijn de Levy-Jenning-plots, al of niet in combinatie met de zogenaamde 'standaard 2s-regel'. In 1981 publiceerde Westgard een veelgeciteerd artikel waarin een vaste combinatie van verschillende controleregels werd aanbevolen (2). Deze methodiek heeft nooit brede ingang gevonden omdat een geautomatiseerd systeem onontbeerlijk was om haar te kunnen toepassen. Bovendien was er geen optimale balans tussen ten onrechte goed- en afgekeurde bepalingenseries. Sinds 1994 heeft Westgard een systeem aanbevolen waarin op basis van de zogenaamde CLIA'88-regels per individuele bepaling een set controleregels werd geselecteerd die voornoemde balans wel optimaliseert (3). De methodiek is goed doordacht, alleen zijn de CLIA'88-regels opgesteld op basis van de opinie van

experts (4). Door de IFCC zijn meer objectieve kwaliteitscriteria voor bepalingen opgesteld op basis van de biologische variatie (5). Helaas wordt er daarbij geen aanbeveling gedaan over de in het dagelijks leven te hanteren kwaliteitscontroleregels. In Medisch Spectrum Twente is vanaf 1998 een geïntegreerde systematiek ontwikkeld: de Westgard-methodiek op basis van de biologische variatie. Begonnen werd om de 2s-criteria zoals die destijds werden gebruikt voor kwaliteitscontrole op de Hitachi 717 (Roche) om te zetten volgens de beschreven systematiek. In 1999 werd de Hitachi 717 vervangen door een Modular-P-configuratie (Roche). Aangezien de standaarddeviaties gemeten op de Hitachi 717 respectievelijk de Modular verschilden, zijn voor de Modular de beoordelingscriteria opnieuw vastgesteld. Op grond van de biologische intra- en interindividuele variatie en de gemeten standaarddeviaties bleek dat de marges van negen bepalingen ruimer gesteld konden worden dan de standaard 2s-criteria. Bovendien bleek op grond van betere reproduceerbaarheid dat de Modular-beoordelingscriteria van zeven bepalingen zelfs ruimer konden worden gedefinieerd dan de Hitachi-criteria. Dit betrof alanine-aminotransferase, amylase, bilirubine, cholesterol, creatinekinase, ijzer en lactaatdehydrogenase. In deze studie laten wij retrospectief zien dat het herbeziën van de beoordelingscriteria op grond van biologische variatie aan de hand van de richtlijnen volgens Westgard, kan resulteren in een 7%-afname van kwaliteitscontroles en een 42%-reductie van herhaalde kwaliteitscontroles, zonder verlies aan kwaliteit. Dit leidt tot een aanzienlijke tijds- en kostenbesparing.

Methodie

Oorspronkelijk werd binnen ons laboratorium de 2s-regel gehanteerd voor het goed- of afkeuren van kwaliteitscontroles. Als gevolg van de natuurlijke imprecisie in een meting betekent dit echter dat in 5% van de gevallen de 2s-regel wordt overschreden en de bepaling dus ten onrechte wordt afgekeurd. Het minder stringent maken van beoordelingscriteria (bijv. een 3s-regel i.p.v. 2s-regel) heeft weliswaar tot gevolg dat het aantal onterechte afkeuringen afneemt, maar heeft ook tot gevolg dat werkelijke afwijkingen minder snel worden ontdekt. Relatief kleine afwijkingen hebben echter nauwelijks of geen klinische consequenties. Met name wanneer de fout vele malen kleiner is dan de biologische variatie binnen en tussen individuen zal dit geen gevolgen hebben voor de diagnostiek.

Laboratorium Medisch Spectrum Twente, Postbus 50000, 7500 KA Enschede

E-mail: A.Boer@ziekenhuis-mst.nl

Toen op ons laboratorium de Hitachi 717 werd vervangen door de Modular, hebben we in eerste instantie dezelfde beoordelingscriteria voor kwaliteitscontroles gehanteerd. Toen echter bleek dat de precisies op beide apparaten verschilden, hebben we in 2001 de beoordelingscriteria voor de Modular opnieuw vastgesteld. In lijn met bovenbeschreven redeneratie hebben we de biologische variatie in de diverse bepalingen als uitgangspunt genomen. Aangezien de fout van een meting wordt bepaald door juistheid en precisie, hebben we op grond van de intra- en inter-individuele variatiecoëfficiënten (6) (VC_{intra} en VC_{inter}) een schatting gemaakt van beide componenten (7). De imprecisie kan worden gedefinieerd als 0,825 maal de intra-individuele variatiecoëfficiënt en de juistheid wordt bepaald door zowel de intra- als inter-individuele variatiecoëfficiënt volgens de formule $0,25 \sqrt{(VC_{intra})^2 + (VC_{inter})^2}$ (7). De resultante van precisie en juistheid bepaalt uiteindelijk de toegestane maximale fout in de diverse bepalingen. Voor enkele bepalingen (albumine, elektrolyten, glucose en HDL-cholesterol) bleek echter dat deze berekende maximale fout kleiner was dan technisch mogelijk. In deze gevallen hebben we de totale fout berekend aan de hand van de 'state of the art'-variatioecoëfficiënten, zoals deze door de SKZL destijds waren gedefinieerd (8). In tabel 1 staat een overzicht van de diverse bepalingen zoals we die destijds met de Modular verrichten, met daarbij de per bepaling berekende toegestane fout. Samen met de gemeten analytische precisie van de Modular hebben we deze maximaal toegestane fouten in het computerprogramma Westgard QC Validator ingevoerd (9), om zo rationele beoordelingscriteria van de diverse bepalingen vast te stellen (tabel 1). Dit computerprogramma is oorspronkelijk ontworpen om volgens de Westgard-systeematiek beoordelingscriteria vast te stellen om aan de CLIA'88-richtlijnen te kunnen voldoen, maar is ook –zoals wij hebben gedaan– uitermate geschikt om criteria vast te stellen op basis van biologische variatie. Als gevolg van bovenbeschreven aanpak werden de beoordelingscriteria voor alanine-aminotransferase, amylase, bilirubine, cholesterol, creatinekinase, ijzer en lactaatdehydrogenase versoepeld, werden de criteria voor triglyceriden en γ -glutamyltransferase gehandhaafd op 3s en bleven de overige criteria op 2s staan (tabel 1).

Resultaten

Aangezien de Modular gekoppeld is aan ons laboratoriuminformatiesysteem GLIMS, worden kwaliteitscontroles (Biorad Liquichek 1 en 2) automatisch geregistreerd. Hierdoor was het mogelijk om het effect van bovenbeschreven methodiek op het gebruik van kwaliteitscontrole materiaal retrospectief te bestuderen. Om het aantal variabele factoren te minimaliseren hebben we de zes maanden voorafgaande aan de nieuwe criteria (ingesteld op 6 september 2001) vergeleken met de 5 maanden daaropvolgend. Deze periodes zijn zo gekozen dat gedurende de gehele periode dezelfde lotnummers van het kwaliteitscontrole materiaal (lotnummer 75111 en 75212) werden gebruikt. Op ons laboratorium is het gebruikelijk om

Tabel 1. Nieuwe beoordelingscriteria op basis van biologische variatie

Bepaling	Toegestane totale fout (%)	Beoordelingscriteria
Alanine-aminotransferase (ALAT) *	32	3s, 2x2s, R4s
Albumine	5,0 ^{SA}	2s
Alkalische fosfatase	12	2s
Amylase*	18	4s
Anorganisch fosfaat	10	2s
Aspartaat-aminotransferase (ASAT)	15	2s
Bilirubine (totaal) *	31	3,5s
Calcium	3,6 ^{SA}	2s, 4x1s
Chloride	3,2 ^{SA}	2s
Cholesterol*	9,0	4s
Creatinekinase (CK) *	30	5s
Glucose	8,4 ^{SA}	2s
γ -Glutamyltransferase (γ GT)	22	3s
HDL-cholesterol	12 ^{SA}	2s
IJzer*	31	3,5s
Kalium	5,8 ^{SA}	2s
Creatinine	6,9	2s
Lactaatdehydrogenase (LD) *	9,5	4s
Natrium	2,1 ^{SA}	2s
Triglyceriden	28	3s
Uraat	12	2s
Ureum	16	2s

Overzicht van bepalingen werden geanalyseerd op de Modular met de berekende toegestane totale fout op basis van biologische intra- en inter-individuele variatie (zie tekst voor details) en de daaruit berekende beoordelingscriteria.

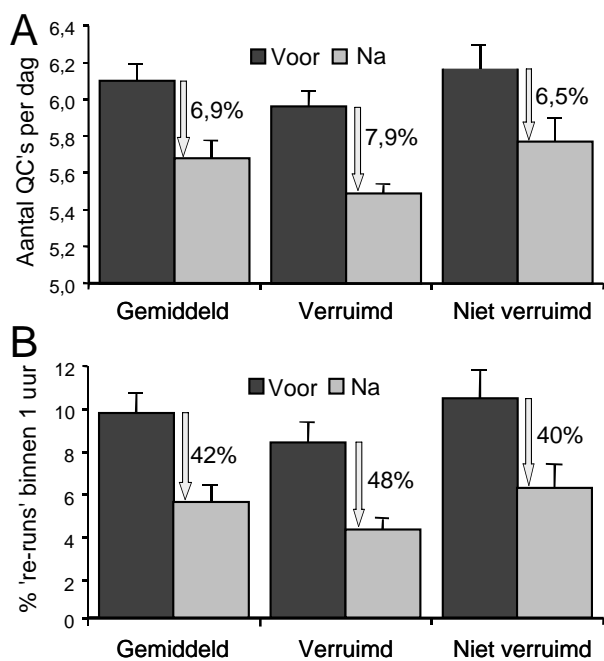
^{SA}Berekend op grond van 'state of the art'-variatioecoëfficiënt

* Bepaling met verruimde beoordelingscriteria

voor de Modular twee tot driemaal daags zowel Biorad Liquichek 1 als 2 te analyseren als kwaliteitscontrole. In de periode voorafgaande aan de nieuwe criteria werden er dan ook 6,1 kwaliteitscontroles per bepaling per werkdag uitgevoerd (figuur 1A). Na invoering van de nieuwe criteria daalde dit aantal tot 5,7 controles. Dit komt overeen met een significante afname van 6,9% ($p < 0,001$). Zoals verwacht valt dit deels toe te schrijven aan de minder stringente beoordelingscriteria voor alanine-aminotransferase, amylase, bilirubine, cholesterol, creatinekinase, ijzer en lactaatdehydrogenase. Figuur 1A laat zien dat het aantal kwaliteitscontroles voor deze specifieke bepalingen met 7,9% is afgenomen ($p < 0,001$). Opvallend is echter wel dat ook het aantal kwaliteitscontroles voor de overige bepalingen eveneens is afgenomen (6,5%, $p < 0,001$). Kennelijk resulteert de invoering van de nieuwe criteria niet alleen op statistische gronden tot een reductie van het aantal kwaliteitscontroles.

Om te achterhalen of de afname van het aantal kwaliteitscontroles wordt veroorzaakt doordat er minder kwaliteitscontrole materiaal (al dan niet) ten onrechte wordt herbepaald hebben we voor beide periodes het percentage 're-runs' berekend. Een 're-run' is hierbij gedefinieerd als een controlemeting die binnen een uur om welke reden dan ook opnieuw wordt uitge-

voerd. In figuur 1B is te zien dat in de periode voorafgaande aan de nieuwe criteria gemiddeld 9,8% van de kwaliteitscontrolematerialen werden herbepaald, terwijl er met de nieuwe criteria nog slechts 5,7% werd herbepaald. Dit komt overeen met een afname van 42% ($p < 0,001$). Bovendien zie je in deze figuur dat er zowel bij de bepalingen met stringenter criteria (48%) als bij de overige bepalingen (40%) een sterke afname te zien is van het percentage 'reruns'. Dergelijke spectaculaire reducties van het aantal kwaliteitscontroles en herbepalingen resulteren in een behoorlijke kostenbesparing, maar mogen daarentegen natuurlijk niet ten koste gaan van de kwaliteit. Om dit te achterhalen hebben we gekeken naar de externe kwaliteitsbewaking van de SKZL. Op grond van rondzendingen wordt als onderdeel van deze enquête voor alle bepalingen een intralaboratoriumvariatiecoëfficiënt, een score (het percentage resultaten binnen drie standaarddeviaties van de consensus) en een ranking (rangorde t.o.v. andere laboratoria) berekend. In tabel 2 is te zien dat de waarden van deze drie grootheden niet veranderen na invoering van de nieuwe beoordelingscriteria. Ook bepalingen met verruimde beoordelingscriteria blijken dezelfde waarden te scoren als vóór die tijd, wat aangeeft dat de versoepeling van de beoordelingscriteria geen kwaliteitsverlies tot gevolg heeft gehad.



Figuur 1. Invoering van ruimere beoordelingscriteria voor kwaliteitscontroles (QC's) resulteert in een afname van het aantal kwaliteitscontroles. A) Voor alle bepalingen in tabel 1 is het gemiddelde aantal QC's per werkdag berekend in de periode 'voor' en 'na' invoering van de ruimere beoordelingscriteria ('Gemiddeld'). Bovendien is het aantal QC's per werkdag apart berekend voor bepalingen met ruimere ('Verruimd') en niet ruimere ('Niet verruimd') beoordelingscriteria. B) Op vergelijkbare wijze is het percentage QC's berekend dat binnen een uur werd herhaald ('re-runs'). De pijlen duiden op significante afnames met een $p < 0,001$.

Conclusie en Discussie

Deze studie laat zien dat het beoordelen van kwaliteitscontroles op basis van intra-individuele en inter-individuele biologische variatie een verruiming van diverse beoordelingscriteria toelaat, zonder dat dit ten koste gaat van de kwaliteit. De grootste winst van deze nieuwe methodiek zit in de sterke reductie van het totaal aantal kwaliteitscontroles. Na introductie van de nieuwe criteria nam het totaal aantal kwaliteitscontroles van bepalingen met 7% af. Bovendien bleek dat het aantal kwaliteitscontroles dat werd herbepaald zelfs met 42% afnam. Hierbij moet wel worden opgemerkt dat deze reductie ook deels valt toe te schrijven aan de reductie van het aantal kwaliteitscontroles voor bepalingen met ongewijzigde beoordelingscriteria. Mogelijk leidt de invoering van nieuwe beoordelingscriteria tot een betere bewustwording van het gebruik van controle materiaal, zodat bijvoorbeeld niet alle bepalingen van een controle materiaal worden herhaald indien slechts enkele bepalingen door de beoordelingscriteria worden afgekeurd. De reductie van 7% aan kwaliteitscontroles levert niet alleen een vergelijkbare kostenreductie aan reagentia, maar draagt bovendien bij aan een grotere efficiency doordat er zo'n 42% minder kwaliteitscontroles hoeven worden herbepaald. Aangezien op ons laboratorium patiëntenmonsters pas geautoriseerd mogen worden als de kwaliteitscontroles zijn goedgekeurd, betekent dit bovendien dat in 42% van de gevallen dat er iets aan de hand is met de kwaliteitscontrole, de patiëntgegevens eerder bekend kunnen zijn.

Tabel 2. Invoering van ruimere beoordelingscriteria voor kwaliteitscontroles heeft geen effect op de kwaliteit

		Voor	Na	p-waarde
Intra-laboratorium-VC (%)	Gemiddeld	1,8	1,7	0,15
	Verruimd	1,8	1,6	0,22
	Niet verruimd	1,8	1,7	0,41
Score (%)	Gemiddeld	84	82	0,38
	Verruimd	79	77	0,70
	Niet verruimd	87	84	0,41
'Ranking' (%)	Gemiddeld	57	57	0,92
	Verruimd	52	51	0,89
	Niet verruimd	59	60	0,84

Overzicht van de gemiddelde intralaboratoriumvariatiecoëfficiënten, de scores (percentage resultaten binnen 3 standaarddeviaties van de consensus) en de 'ranking' (rangorde t.o.v. andere laboratoria) voor de diverse bepalingen zoals berekend in de externe kwaliteitsbewaking van de SKZL. Weergegeven zijn gemiddelden van de drie enquêtes voorafgaande aan en volgend op de invoering van de ruimere beoordelingscriteria. Bovendien zijn de waarden apart berekend voor bepalingen met ruimere ('Verruimd') en niet ruimere ('Niet verruimd') beoordelingscriteria. De p-waarden zijn berekend door de gemiddelden per test in beide perioden gepaard te vergelijken met een Student-t-test.

Literatuur

1. Stöckl D, Baadenhuijsen H, Fraser CG, Libeer JC, Petersen PH, Ricos C. Desirable routine analytical goals for quantities assayed in serum. *Eur J Clin Chem Clin Biochem* 1995; 33:157-169.
2. Westgard JO, Groth T. Design and evaluation of statistical control procedures: applications of a computer "quality control simulator" program. *Clin Chem* 1981; 27: 1536-1545.
3. Westgard JO. Selecting appropriate quality-control rules. *Clin Chem* 1994; 40: 499-501.
4. Medicare, Medicaid and CLIA programs; regulations implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA)-HCFA. Final rule with comment period. *Fed Regist* 1992; 57: 7002-7186.
5. Buttner J, Borth R, Broughton PMG, Bowyer RC. Approved Recommendation (1983) on Quality Control in Clinical Chemistry. Part 4. Internal Quality Control. *J Clin Chem Clin Biochem* 1983; 21: 877-884.
6. Ricos C, Alvarez V, Cava F, Garcia-Lario JV, Hernandez A, Jimenez CV, Minchinela J et al. Current databases on biological variation: pros, cons and progress. *Scan J Clin Invest* 1999; 59: 491-500.
7. Ricos C. External quality assesment: currently used criteria for evaluating performance in European countries, and criteria for future harmonization. *Eur J Clin Chem Clin Biochem* 1999; 34: 159-165.
8. Petersen PH, Ricos C, Stockl D, Libeer JC, Baadenhuijsen H, Fraser C, Thienpont L. Proposed guidelines for the internal quality control of analytical results in de medical laboratory. *Eur J Clin Chem Clin Biochem* 1996; 34: 983-999.
9. Westgard JO. A program for evaluating QC procedures. *Med Lab Observ* 1994; 26: 55-60.

Summary

Maintaining quality by doing less; an evaluation of quality controls based on biological variation. Boer A-K, Sluijs Veer G van der. Ned Tijdschr Klin Chem Labgeneesk 2004; 29: 38-41.

Analysis of control samples forms an essential, but expensive and labour-intensive link in the quality system of a laboratory. Several criteria have been developed to evaluate quality controls in order to find a balance in falsely rejected and approved controls. Based on the Westgard-system, our laboratory developed new evaluation criteria for quality controls on the Roche Modular in 2001. However, the intra-individual and interindividual biological variation were chosen as the basis for choosing the criteria. As a result, the rules of seven assays could be expanded. This study shows retrospectively that implementation of these criteria results in a 7% reduction of the use of quality controls. Furthermore, the percentage of quality controls that had to be re-analysed decreased with 42%. Therefore, our method contributes to increased efficiency, while external laboratory-quality controls did not show a reduction in analytical-quality.

Keywords: Quality control; biological variation