

Evaluatie van de externe kwaliteitscontrole voor het microscopisch beoordelen van bloedbeelden in de Regio Noord sinds 1992

H. de WIT¹, R. BEKKEMA², G.A. van den BERG¹, A.B. MULDER², M.R. HEINER-FOKKEMA^{1,2}, J.W. KOOPMANS¹ en J.W. SMIT³

Sinds 1992 wordt er in de regio Noord een extern kwaliteitscontrole programma verzorgd voor de handmatige differentiële bloedcelltellingen ('handdif'). Hierbij wordt een scoresysteem gebruikt waardoor het mogelijk is om een uitspraak te doen over de kwaliteit van deze handdif. Het blijkt dat de handdif van goede kwaliteit is. Wel is het vaak lastig om de uitslag van de 'dif' te vertalen in een uitspraak over achterliggende pathologie. In de loop der jaren valt een lichte kwaliteitsverbetering waar te nemen. Ziektebeelden met erythroïde pathologie scoren het laagst, gevolgd door lymfoïde en myeloïde pathologie. Het laten verrichten van de handdif door een beperkt aantal analisten lijkt een gunstig effect te hebben op de kwaliteit ervan.

De huidige celtelapparatuur is steeds beter in staat om een geautomatiseerde differentiële bloedcelltelling uit te voeren (1, 2). Het aantal handmatige microscopische differentiële tellingen (handdif) dat nog uitgevoerd wordt neemt dientengevolge dan ook sterk af. Hierdoor komt de kwaliteit van deze handdiffen onder druk te staan omdat analytisch personeel er steeds minder frequent mee in aanraking komt (3). Toch blijft in 10-20% van de analyses een microscopische beoordeling noodzakelijk. Bij veel laboratoria is de handdif nog steeds in de 24-uursdienst beschikbaar en dat vereist dat iedere dienstdoende analist deze handdif beheerst. In grote laboratoria kan dit betekenen dat 40 tot 60 analisten deze vaardigheid op peil dienen te houden.

In de Regio Noord wordt vanuit het Universitair Medisch Centrum Groningen (UMCG) een extern kwaliteitscontroleprogramma verzorgd ten behoeve van de handdif. In het najaar van 1992 is gestart met een scoresysteem. Hiermee is het mogelijk om de kwaliteit van de handdif in de regio te controleren, verbeteren en wellicht tot enige vorm van harmonisatie te komen.

De difrondzending

Zes keer per jaar worden vanuit het UMCG twee ongekleurde preparaten verstuurd naar de deelnemende laboratoria. Deze laboratoria rapporteren hun bevindingen weer aan het UMCG waar alle uitslagen worden

verwerkt. Een referentenpanel, bestaand uit klinisch chemici, een laboratoriumarts en gespecialiseerde analisten, beoordeelt vooraf alle bloedbeelden volgens de aanbevelingen van de CCKL-subcommissie Hematomorfologie en bepaalt criteria waaraan deelnemers geacht worden te voldoen (4). De samenstelling van het referentenpanel is alleen in 2005 gewijzigd door toetreding van een laboratoriumarts van het UMCG. De beoordeling van de preparaten wordt omgezet in een score. Score 1 heeft betrekking op afwijkingen in de dif die duiden op de achterliggende pathologie. Van de deelnemers wordt verwacht dat zij bij het beoordelen van het bloedbeeld een signaal af kunnen geven richting de aanvrager. Bij score 1 wordt gebruik gemaakt van een optelsysteem en de maximale score is 10 punten. Het is uitdrukkelijk niet de bedoeling een diagnose te stellen, maar wel dient men de aanvrager op het juiste spoor te zetten daar waar dat van toepassing is. Het referentenpanel streeft ernaar dat wanneer er een adequaat signaal wordt afgegeven naar de aanvrager er een voldoende wordt gescoord (≥ 6 punten). Als ook nog een juiste suggestie kan worden gegeven bij welke onderliggende pathologie de dif past dan wordt de maximale score van 10 punten toegekend. In score 2 komt de vaardigheid van het op juiste wijze uitvoeren van de handdif tot uiting. Er is gekozen voor een aftreksysteem waarbij men met 10 punten begint en elke afwijking in het bloedbeeld die wordt gemist of elke afwijking die ten onrechte wordt vermeld levert de deelnemer aftrek van punten op. In tabel 1 staat een voorbeeld weergegeven van de scores 1 en 2 van een preparaat van een patiënt met een chronische myeloïde leukemie.

Bij terugrapportage van de uitslagen aan de deelnemers door het UMCG worden de scores van het eigen laboratorium en van alle andere deelnemers vermeld. Tevens worden daarbij de afwijkingen vermeld die in score 1 punten hebben opgeleverd en in score 2 tot aftrek van punten hebben geleid. Hemocytometriegegevens, cytochemische kleuringen, immuunfenotypering en cytogenetica indien van toepassing, worden eveneens gerapporteerd. Tenslotte worden in een commentaar de bevindingen van het referentenpanel weergegeven.

Twee keer per jaar wordt in de vorm van een minisymposium aandacht besteed aan de preparaten die in de voorgaande periode zijn rondgestuurd. Hierbij zijn naast afgevaardigden van alle deelnemende laboratoria ook leden van het referentenpanel aanwezig. Er wordt uitgebreid ingegaan op de resultaten van de rondzendingen en tevens komen verschillende onder-

Stichting KCL, afdeling Klinische Chemie, Leeuwarden¹, UMCG, Laboratoriumcentrum, Groningen², LabNoord, sector Martini Ziekenhuis, Groningen³

Correspondentie: dr. H. de Wit, Stichting KCL, afdeling Klinische Chemie, Postbus 850, 8901 BR Leeuwarden
E-mail: harry.de.wit@znb.nl

werpen over het beoordelen van bloedbeelden in het algemeen aan de orde. Een vast onderdeel is tevens een inleiding van een internist/hematoloog over een ziektebeeld passend bij één van de preparaten.

Vanaf 1992 tot en met december 2006 zijn in totaal 172 preparaten rondgezonden met zeer uiteenlopende afwijkingen. Er zijn 26 preparaten van gezonde personen verstuurd, 41 preparaten met een myeloïde pathologie, 59 preparaten met een lymfoïde pathologie, 18 preparaten met een erytroïde pathologie, 11 malaria-preparaten en 17 preparaten met diverse afwijkingen. Onder de categorie myeloïde pathologie vallen onder andere de myeloïde leukemieën, myelodysplasie en eosinofilie. Lymfoïde pathologie omvat bijvoorbeeld de lymfoïde leukemieën, M. Kahler, 'hairy cell' leukemie, mononucleosis infectiosa en andere. De erytroïde pathologie bestaat onder andere uit ijzerdeficiënties, sikkelcelanemie, megaloblastaire anemie en sferocytose.

Daar waar een trendanalyse is uitgevoerd is dat gedaan door middel van de Spearman-test en non-parametrische analyse is verricht met de Mann-Whitney-test, beide met behulp van het programma Analyse-it.

Evaluatie

Van de 20 laboratoria die in 1992 deelnamen aan de kwaliteitsrondzending waren er in 2006 nog 16 over. De evaluatie is gebaseerd op deze 16 laboratoria. Wanneer de gemiddelde scores van de deelnemers voor elk preparaat worden uitgezet in de tijd ontstaat een beeld van de kwaliteit van de handdijf in de regio (figuur 1). De gemiddelde score 1 over de gehele periode was

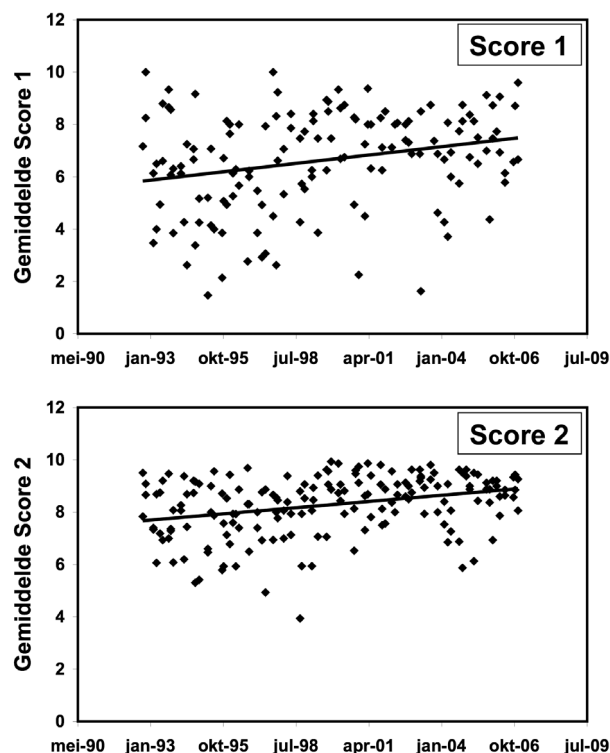
Tabel 1. Gegevens en de opbouw van scores 1 en 2 van een preparaat verstuurd op 19 mei 2004. De bijbehorende diagnose was chronische myeloïde leukemie. Bij cytogenetisch onderzoek werd een t(9;22) aangetoond en ook de PCR voor BCR-ABL was positief.

Vrouw van 38 jaar			
Hemocytometrie gegevens:	Leuko's	204,0	x10 ⁹ /l
	Ery	3,0	x10 ¹² /l
	Hb	4,7	mmol/l
	Ht	0,29	
	MCV	95	fl
	Trombo	813	x10 ⁹ /l
Score 1			
Het aanduiden van een linksverschuiving		4	punten
Het aanduiden van een hoog aantal basofiele granulocyten		2	punten
Het aanduiden van aanwezigheid van promyelocyten/blasten		2	punten
Het aanduiden van CML		2	punten
Score 2			
Een leukocytensubpopulatie-uitbijter		-1	punt
Het noemen van toxische korreling		-1	punt
Het noemen van kogelcellen		-1	punt
Het niet noemen van hypersegmentatie		-1	punt
Het niet noemen van een hoog aantal trombocyten		-1	punt
Het niet noemen van macrotrombocyten		-1	punt
Etc...			

6,6 en de gemiddelde score 2 was 8,3. Het percentage onvoldoende scores 1 en 2 (minder dan 6 punten) was respectievelijk 24% en 7%. De hoge gemiddelde score 2 en het lage percentage onvoldoenden wijzen er op dat de laboratoria in de Regio Noord een technisch goede handdijf afleveren. De signaalfunctie met betrekking tot de eventueel aanwezige pathologie (tot uitdrukking komend in de hoogte en het percentage onvoldoende scores 1) blijkt meer problemen op te leveren. De gemiddelde score 1 is laag en het percentage onvoldoenden is relatief hoog.

In beide grafieken van figuur 1 is een trendlijn weergegeven die een stijgend verloop laat zien. Dit duidt op een verbetering van de scores en een mogelijke kwaliteitsverbetering van de handdijf. Voor zowel score 1 als score 2 betreft het een significante stijging met r-waarden van respectievelijk 0,26 (p<0,002) en 0,32 (p<0,0001). Een andere aanwijzing voor een kwaliteitsverbetering valt op te maken uit figuur 2 waar het aantal onvoldoende scores van elke rondzending staat weergegeven in de tijd. De trendlijnen van zowel score 1 als score 2 laten een dalende trend zien. In beide gevallen betreft het een significante daling van het aantal onvoldoende scores; voor score 1 een r-waarde van -0,26 (p<0,002) en voor score 2 een r-waarde van -0,20 (p<0,009).

Eén van de doelstellingen van de halfjaarlijkse dif-bespreking is dat er consensus bereikt wordt over de wijze van uitvoeren en rapporteren van de differentiatie. Dit zou tot uiting kunnen komen in een afname van de spreiding van de resultaten van de rondzending. In figuur 3 staat voor beide scores de standaarddeviatie weergegeven in de tijd van elke afzonderlijke rond-



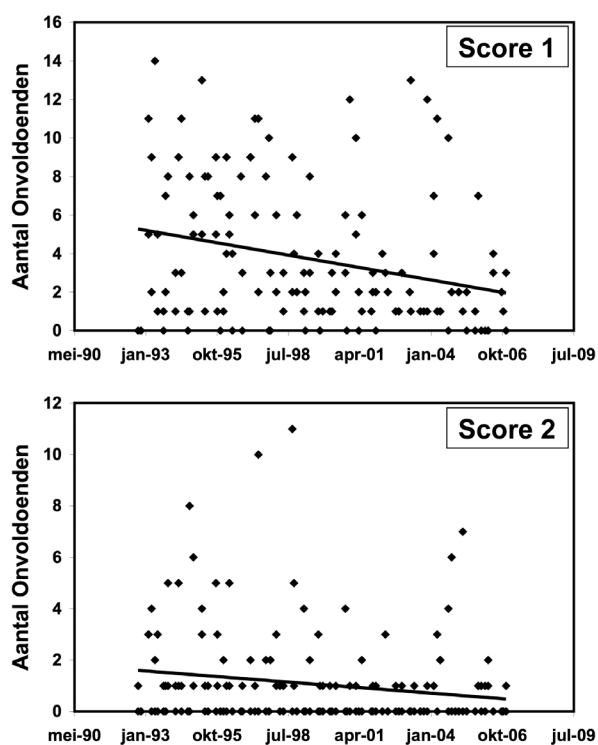
Figuur 1. Gemiddelde scores 1 en scores 2 van alle deelnemende laboratoria per rondgezonden preparaat sinds de start van de rondzending in oktober 1992. Elk punt in de grafiek geeft de gemiddelde score weer van 16 deelnemende laboratoria.

zending. Score 1 laat een licht dalende maar wel significante trend zien (r-waarde -0,21 ; $p < 0,02$). In score 2 is geen significante daling waar te nemen.

Uiteraard kan ook een overzicht verkregen worden van de prestaties van individuele laboratoria. Figuur 4 illustreert hiervan een drietal voorbeelden betreffende score 1. Laboratorium A is één van de betere deelnemers met een gemiddelde score van 8 en slechts 9% onvoldoende scores. Laboratorium B is een voorbeeld van een deelnemer waarbij enige verbetering op zijn plaats zou zijn. De gemiddelde score 1 van deze deelnemer over de gehele periode is 5,0 en het percentage onvoldoende scores is 49%. Beide laboratoria laten in de loop van de jaren geen kwaliteitsverbetering zien, afgeleid uit de afgebeelde trendlijn. Dit in tegenstelling tot laboratorium C dat een opmerkelijke stijging van score 1 laat zien (r-waarde: 0,43; $p < 0,0001$).

Eén van de deelnemers heeft in 2003 in het laboratorium een difgroep ingesteld bestaande uit 9 analisten die alle handdiffen verrichten. Vóór 2003 werd de handdif uitgevoerd door 40 dienstdoende analisten. Omdat dit laboratorium de difrondzending tevens als interne kwaliteitscontrole gebruikt is de kwaliteit vóór en na het instellen van de difgroep bestudeerd. In twee jaren voorafgaand aan 2003 werden 30 preparaten geëvalueerd waarbij 139 beoordelingen betreffende score 1 en 173 beoordelingen betreffende score 2 zijn betrokken. In 21% van de beoordelingen werd een onvoldoende score 1 behaald en de gemiddelde score 2 was 9,0. Na 2003 zijn 46 preparaten geëvalueerd waarbij 246 beoordelingen score 1 en 316 beoordelingen score 2 zijn betrokken. Het percentage onvoldoende scores 1 was nu 8% en de gemiddelde score 2 was 9,2. Dat betekende dat met name de signalering naar de aanvrager aanzienlijk is verbeterd (minder onvoldoende scores 1).

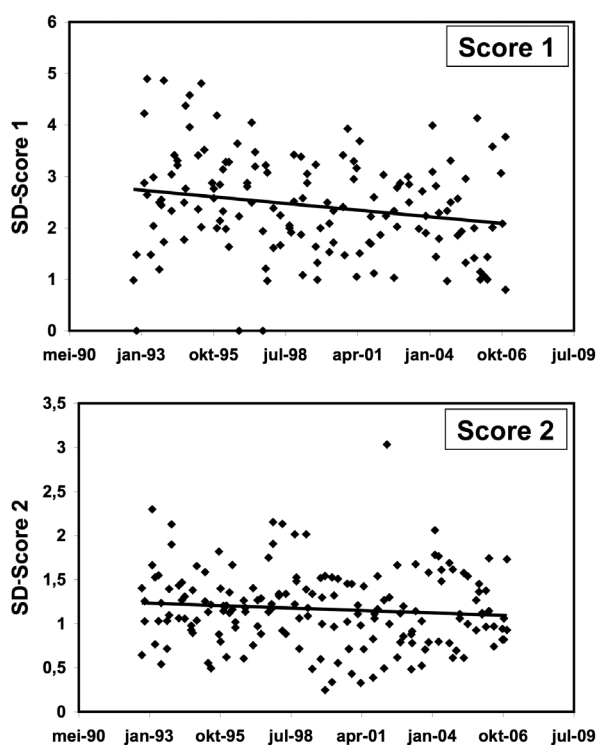
De preparaten van de rondzending zijn in te delen in



Figuur 2. Het aantal onvoldoende scores 1 en scores 2 van elk preparaat.

een aantal ziektebeeldcategorieën. De gemiddelde scores met de standaarddeviatie en het percentage onvoldoenden staan weergegeven in tabel 2. Opvallend is dat de laagste gemiddelde score en het hoogste percentage onvoldoenden vallen in de categorie erythroïde afwijkingen, gevolgd door de lymfoïde en de myeloïde pathologie. Tijdens een aantal halfjaarlijkse difbesprekingen is aandacht besteed aan de beoordeling van het rode bloedbeeld en dan blijkt inderdaad dat analisten dit als erg lastig ervaren. Bij de malariapreparaten was het percentage onvoldoende scores 13%. Dit betekent dat in 13% van de beoordeelde preparaten de aanwezigheid van malaria werd gemist. Hierbij dient echter wel te worden opgemerkt dat het in deze gevallen geen gerichte aanvraag naar malariadiagnostiek betrof. Bovendien doen sommige laboratoria de malariadiagnostiek niet zelf maar is dit ondergebracht bij het microbiologisch laboratorium.

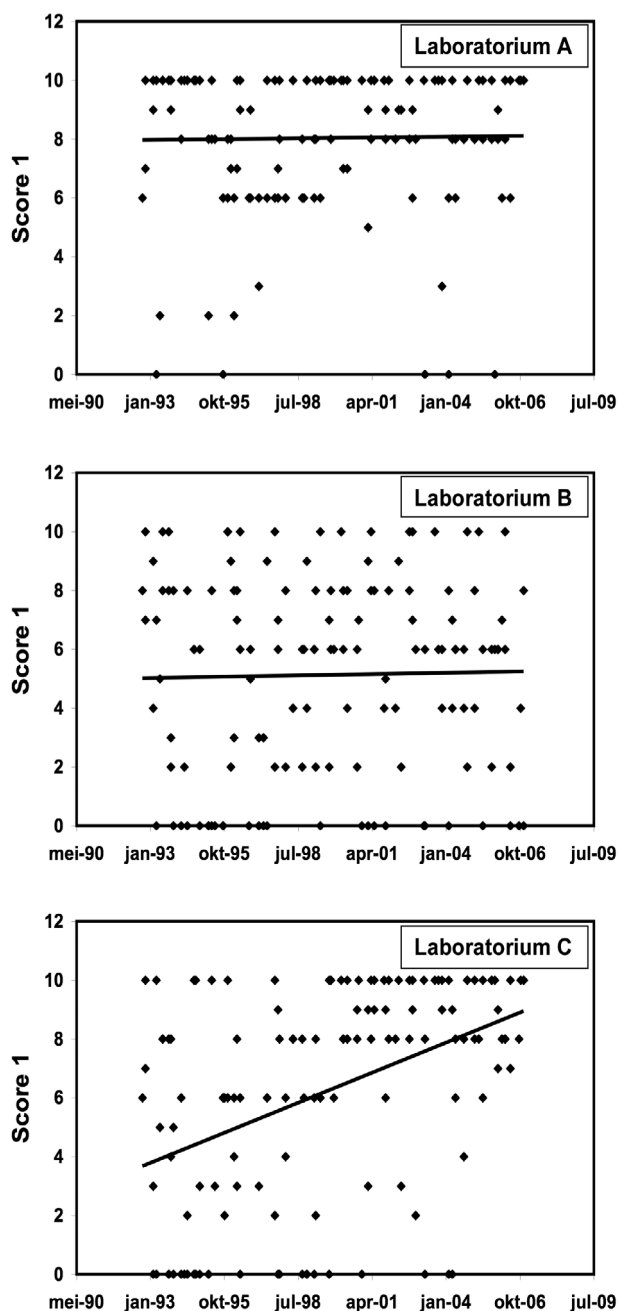
In een korte enquête gehouden onder de deelnemers blijkt dat men over het algemeen zeer tevreden is over de rondzendingen en de halfjaarlijkse difbijeenkomsten. Bij het beoordelen van de QC-preparaten en het rapporteren van de uitslagen wordt hieraan vaak extra aandacht besteed. Slechts vier deelnemers rapporteren de uitslag van een willekeurige analist. Vijf van de zestien laboratoria gebruiken het preparaat tevens als interne kwaliteitscontrole en nagenoeg alle laboratoria (15) gebruiken de rondzending voor (bij)scholing van analisten. Het aantal analisten dat per deelnemend laboratorium de handdif uitvoert was in 1993 gemiddeld 32 (range 13-70). In 2000 was dit gemiddeld 28 analisten per laboratorium (range 13-70) en in 2006 daalde dit gemiddelde tot 23 analisten (range 9-52). Bij nagenoeg alle laboratoria is de dif in de 24-uurs dienst beschikbaar en bij slechts twee deelnemers is de beschikbaarheid beperkt tot kantooruren.



Figuur 3. Standaarddeviatie van de gemiddelden van score 1 en score 2 van elk preparaat.

Tabel 2. Resultaten van scores 1 en 2 waarbij de rondzendingen gegroepeerd zijn in verschillende categorieën ziektebeelden

Pathologie	Score 1			Score 2		
	Gem. score	St. dev.	% Onvold.	Gem. score	St. dev.	% Onvold.
Gezond	n.v.t.	n.v.t.	n.v.t.	9,3	0,8	2
Myeloïd	7,1	2,0	21	7,9	1,2	10
Lymfoïd	6,5	2,5	25	8,4	1,2	6
Erytroïd	5,9	3,1	37	7,4	1,5	15
Malaria	6,4	2,6	13	8,1	1,3	5
Diversen	6,6	2,3	28	8,3	1,2	2



Figuur 4. Scores 1 van drie individuele laboratoria. Laboratorium A heeft een gemiddelde score 1 van 8 en 9% onvoldoenden en laboratorium B een gemiddelde score 1 van 5,0 en 49% onvoldoenden. Laboratorium C maakt een opmerkelijke progressie door in de loop der jaren ($r=0,43$; $p<0,0001$)

Discussie

Gesteld kan worden dat de kwaliteit van de handdif goed is als daarmee bedoeld wordt de beoordeling van celtypen en de afwijkingen daarin. Echter het vertalen van de afwijkingen in een signaal betreffende de achterliggende pathologie blijkt veel lastiger te zijn. Juist dit laatste punt geeft de handdif meerwaarde ten opzichte van de geautomatiseerde differentiële telling. In feite wordt met een technisch juiste beoordeling (tot uitdrukking komend in score 2) de uitslag van de celltellers gecontroleerd terwijl het in principe mogelijk moet zijn om met een handdif de aanvrager een aanwijzing voor mogelijke pathologie te geven (score 1). In de loop van de tijd is een zekere kwaliteitsverbetering waar te nemen. Dat blijkt uit een verbetering van scores 1 en 2 en uit een afname van het aantal onvoldoende scores. Over de oorzaken van deze kwaliteitsverbetering valt slechts te speculeren. Wellicht is door alle aandacht voor de handdif en de halfjaarlijkse bespreking van de resultaten de kwaliteit verbeterd. Een aanwijzing hiervoor is de afname van de standaarddeviatie van score 1 in de loop van de tijd. Er is meer overeenstemming ontstaan over het vertalen van de uitslag van de handdif in een signaal naar de aanvrager. Een andere oorzaak kan liggen in het feit dat er een tendens bestaat om het aantal analisten dat de handdif uitvoert te beperken en dat daarmee de expertise beter op peil wordt gehouden.

Omdat er in de afgelopen jaren ruim 170 preparaten zijn rondgezonden is het voor individuele laboratoria mogelijk om een goed beeld te krijgen van de kwaliteit van hun eigen handdif. Er kan bijvoorbeeld gekeken worden naar de scores bij verschillende pathologieën en indien noodzakelijk gericht scholing worden gegeven.

Wanneer preparaten worden ingedeeld in verschillende pathologiecategorieën dan blijkt erytoïde pathologie de meeste problemen op te leveren gevolgd door lymfoïde en myeloïde pathologie. Analisten geven tijdens de difbijeenkomsten ook aan dat men de beoordeling van het rode bloedbeeld moeilijk vindt. Herhaalde aandacht hiervoor heeft enige verbetering in de scores tot gevolg gehad maar het aantal preparaten is te klein om hierin een significante trend te ontdekken. Bij de lymfoïde pathologie is er wel een significante verbetering waar te nemen ($p<0,04$). Voor 2000 was de gemiddelde score 1 van de lymfoïde pathologie $5,8 (\pm 2,4)$ en na 2000 was dit $7,2 (\pm 1,5)$. Na 2000 is herhaaldelijk aandacht besteed aan de lymfocytenmorfologie tijdens de difbijeenkomsten.

In 2000 heeft een werkgroep van de Vereniging Hematologische Laboratoriumdiagnostiek (VHL) aanbevelingen geformuleerd omtrent indicaties voor het verrichten van een microscopische differentiatie. Deze werkgroep heeft aangegeven dat er slechts een beperkt aantal indicaties zijn tot het uitvoeren van een cito microscopische dif. Tevens zou het aantal analisten dat de handdif zou moeten kunnen uitvoeren beperkt moeten worden om een voldoende kennisniveau te handhaven. Een beperkt aantal laboratoria in de Regio Noord hebben in navolging van dit advies een difgroep ingesteld. Eén van de laboratoria heeft daarbij de kwaliteit geëvalueerd. Vastgesteld kon worden dat er een aanwijsbare kwaliteitsverbetering heeft plaatsgevonden. Vooral signalering van pathologie liet een duidelijke verbetering zien. Het percentage onvoldoendes in score 1 daalde van 21% naar 8%. Beperking van het aantal analisten dat de handdif uitvoert doet de expertise toenemen.

Een recente ontwikkeling is de intrede van digitale beeldverwerking bij de microscopische differentiatie. Uit diverse recente publicaties blijkt dat dit mogelijk een zinvolle aanvulling is op de huidige diagnostische mogelijkheden (5-7). Het effect van digitale beeldverwerking op de kwaliteit van de handdif zal in de toekomst moeten blijken. Aan de ene kant zouden dergelijke systemen de analist ondersteuning kunnen bieden bij automatisering, rubricering van cellen en archivering van uitslagen. Aan de andere kant zou de expertise van de analist door beperking van het uitvoeren van de handdif opnieuw onder druk kunnen komen te staan. De toekomst zal uitwijzen welke rol er is voor deze beeldverwerkende technieken.

Samenvattend kan geconcludeerd worden dat de handdif in de Regio Noord in technisch opzicht van goede kwaliteit is. Het vertalen van de dif in een signaal naar de aanvrager kan nog wel eens een probleem zijn. Kwaliteitsverbetering valt te verwachten van het onderbrengen van de handdif bij een beperkt aantal analisten.

Literatuur

1. Ward PCJ. The CBC at the turn of the millennium: an overview. *Clin Chem* 2000; 46 (8): 1215-1220.
2. Pierre RV. Peripheral blood film review: the demise of the eyecount leukocyte differential. *Clin Lab Med* 2002; 22 (1): 279-297.
3. Meer W van der, Willekens FLA, Gelder W van, Willems JL. De virtuele differentiatie van bloedcellen als hulpmiddel om de controverse van de beoordeling van de staafkernige granulocyt te toetsen. *Ned Tijdschr Klin Chem Labgeneesk* 2004; 29: 313-316.
4. Anonymus. Aanbevelingen CCKL subcommissie hematologie, januari 1990. Bijlage *Tijdschr NVKC* 1991: 57-67.
5. Triepels RH, Huiskes-Bloemen I, Doelman CJA. Digitale beeldverwerking bij de differentiatie van bloedcellen: een aanwinst voor de klinische chemie? Een evaluatie van de diffMaster Octavia. *Ned Tijdschr Klin Chem Labgeneesk* 2005; 30: 268-271.
6. Ceelie H, Dinkelaar RB, Gelder W van. Examination of peripheral blood films using automated microscopy; evaluation of DiffMaster Octavia and Cellavision DM96. *J Clin Pathol* 2007; 60 (1): 72-79.
7. Gelder W van, Ceelie H, Berendes PB, Prins J, Dinkelaar RB. Verbeterde efficiency van hemocytometrische analyse door een combinatie van XE2100, SIS en DM96 (Sysmex). *Ned Tijdschr Klin Chem Labgeneesk* 2006; 31: 207-208.

Summary

Wit H de, Bekkema R, Berg GA van den, Mulder AB, Heiner-Fokkema MR, Koopmans JW, Smit JW. Evaluation of an external quality control program for the microscopic review of peripheral blood smear. Ned Tijdschr Klin Chem Labgeneesk 2008; 110: 114.

For several years an external quality-control program is organized in the northern part of the Netherlands for the manual peripheral blood smear review. A score is used in which we assess the quality of this manual review. The quality of the manual review is good although it is difficult to translate the results in a pathological condition. Through the years a slight improvement in quality can be demonstrated. The score of erythroid pathology is relatively poor followed by lymphoid and myeloid pathology. Performance of the manual blood smear review by a limited number of technicians seems to improve quality.