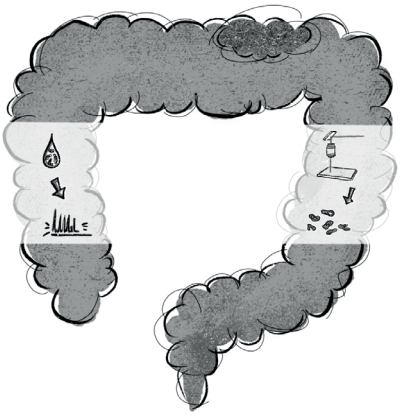
Improving diagnostic, prognostic and predictive biomarkers in colorectal cancer: the role of proteomics and stromatogenesis

Huijbers, A.



3

Case-controlled identification of colorectal cancer based on proteomic profiles and the potential for screening

Anouck Huijbers
Wilma E. Mesker
Bart J. Mertens
Marco R. Bladergroen
André M. Deelder
Yuri E.M. van der Burgt
Rob A.E.M. Tollenaar

Colorectal Disease 2014 Nov;16(11):907-13.

ABSTRACT

Aim: Colorectal cancer (CRC) screening programmes detect early cancers but unfortunately they have limited sensitivity and specificity. Mass spectrometry-based determination of serum peptide- and protein profiles provide a new approach for improved screening.

Method: Serum samples from 126 CRC pretreatment patients and 277 control individuals were obtained. An additional group of samples from 50 CRC patients and 82 controls was used for validation. Peptide and protein enrichments were carried out using reversed-phase C18 and weak-cation exchange magnetic beads (MBs) in an automated solid-phase extraction and spotting procedure. Profiles were acquired on a matrix-assisted laser desorption/ionization time-of-flight system. Discriminant rules using logistic regression were calibrated for the peptide and protein signatures separately, followed by combining the classifications to obtain double cross-validated predicted class probabilities. Results were validated on an identical patient set.

Results: A discriminative power was found for patients with CRC representative for all histopathological stages compared with controls with an area under the curve of 0.95 in the test set (0.93 for the validation set) and with a high specificity (94-95%).

Conclusion: The study has shown that a serum peptide and protein biomarker signature can be used to distinguish CRC patients from healthy controls with high discriminative power. This relatively simple and cheap test is promising for CRC screening.

INTRODUCTION

The lifetime risk of colorectal cancer (CRC) in the US population is 5–6% without screening, which is similar to the Netherlands (1-3). Early diagnosis reduces disease-related mortality (4). The number of patients diagnosed annually is still increasing, because of aging of the population and a small increase in the incidence at all ages. It is therefore expected that population screening programmes aiming at early detection of CRC will become more relevant. Currently the most promising screening tests used in population screening include the immunology-based faecal occult blood test (iFOBT), DNA markers in stool (sDNA), computed tomography colonography (CTC) and colonoscopy (4;5). The iFOBT uses antibodies to detect the globin portion of human hemoglobin. Multiple brands of tests are available and specificity and sensitivity values reported in literature vary widely from 70% to 94% (6;7). Current advice is annual screening with iFOBT. Screening based on sDNA involves the identification of genetic modifications in the initiation of a sequenced progression from adenoma to carcinoma. The sensitivity and specificity of the various sDNA tests range from 52% to 91% and from 93% to 97% (5). The guideline recommendation is to screen every 3 years. Virtual colonoscopy or CTC is reported to have overall sensitivities of 55-94%, depending on the size of the detected polyps, with high specificity (91-96%) (8). Guidelines advise a screening interval of 5 years. Serum carcinoembryonic antigen (CEA) estimation lacks sensitivity and specificity (9). Although not used for screening, colonoscopy has a specificity and sensitivity of at least 95% for large polyps, but the miss rate for polyps smaller than 5mm is 15–25% and 0–6% for polyps of 10 or more millimetres (10).

Although current screening methods are widely available, there is room for improvement and new developments of simple, cost-effective and noninvasive screening tests (11;12). The use of protein biomarkers for early detection of cancer is promising (13;14). Comparison of serum protein patterns or profiles has allowed the separation of patients with cancer from healthy individuals (15). Alternatively tissue can be used to identify protein biomarkers (16), but results obtained from body liquids and cancer tissue may not be the same. We have developed a one-step, fully automated and standardized solid-phase extraction (SPE) method using functionalized magnetic beads (MBs) to enrich for subsets of peptides and proteins in a high-throughput fashion, in combination with matrix-assisted laser desorption/ionization – time of flight (MALDI-TOF) read-out (16-18). In this study, we use a combination of two different types of paramagnetic beads (MBs) to increase the number of detected features, namely based on weak cation exchange (WCX) and reversed phase (RP) C18-functionalization. Previously, we have shown that the statistical combination of two thus obtained data sets

improves classification of samples in a case-control study of breast cancer patients (19). In the current study, we used MALDI-TOF MS to generate a protein and peptide signature of a serum sample in a case-control set-up aiming for the detection of CRC.

METHOD

Patients

Serum samples within the test set were obtained from 126 outpatients with CRC before treatment and from 277 healthy controls collected between October 2002 and December 2008. Validation specimens were obtained from 50 patients with CRC and 82 healthy controls. These were collected in the same way between January 2009 and May 2010. Histopathological examination of the surgical specimen reported the TNM stage (TNM Classification of Malignant Tumours fifth edition). Informed consent was obtained from all subjects and the study was approved by the Medical Ethical Committee of the Leiden Universal Medical Center.

Sample processing and MALDI-TOF measurement

The method of serum collection, sample and profile processing and data analysis has been described previously (16). The isolation of proteins and peptides from serum was performed using a kit based on magnetic bead purification with WCX- and RP C18 Mbs. using a standardized protocol. High-throughput SPE was followed by MALDI-TOF measurement on an Ultraflex II TOF/TOF spectrometer (Bruker Daltonics). In this way, so-called WCX- and RP C18-profiles were obtained, representing (small) protein and peptide signatures respectively.

Profile processing

For optimal data analysis, all WCX- and RPC18-profiles underwent baseline correction followed by alignment (19). A list of selected peaks (Table 1) was then compiled through the application of a peak selection procedure as previously described by our group (19). In this way, 57 peptides and proteins were selected in the WCX-profiles and 42 peptides in the RP C18-profiles. The summarized spectral measurements for each individual were then used within the discriminant analysis (19).

Statistical analysis

Discriminant rules were calibrated for the WCX and RP C18 data separately using logistic ridge regression (see Appendix S1) (19). A combined classification rule was calibrated using logistic regression on double cross-validation. Predictive performance of the calibrated combination was verified on the validation set. Error rates were

calculated as well as estimates of sensitivity and specificity, assigning each observation to the group for which the predicted class probability was highest, and the receiver operating characteristics (ROC) curve was plotted with the area under the ROC curve (AUC) indicating the ability to distinguish cancer from control samples (Fig. 1).

Table I Summary of all m/z-values used for statistical analysis of the peptide- and protein signatures from the reverse phase (RP C18) and weak cation exchange (WCX) profiles.

m/z value in WCX profile	Window (*m/z* units)*	*m/z*-value in RPC18 profile	No. of peaks†
1866	5	1020.6	3
1898	5	1077.7	3
1947	5	1206.7	3
2024	5	1211.8	3
2084	5	1260.6	3
2106	5	1263.7	3
2661	5	1348.9	3
2756	5	1350.8	3
2770	5	1368.0	3
2862	5	1418.6	3
2884	5	1434.7	3
2953	5	1440.7	3
3159	5	1449.9	3
3192	5	1465.8	3
3241	5	1481.8	3
3269	5	1518.9	3
3328	5	1536.9	3
3445	5	1561.9	3
3501	5	1563.0	3
3525	5	1606.0	3
3539	5	1615.8	3
3884	5	1627.0	3
3956	5	1691.1	3
3972	5	1740.1	3
3994	5	1778.1	3
4054	5	1865.2	3
4090	5	1896.0	3
4210	10	1934.2	3
4480	10	2021.3	3
4648	5	2271.1	3
4963	6	2378.4	3
5065	5	2553.1	3
5087	5	2602.5	3
5160	5	2616.5	3
5248	5	2753.7	3
5336	7.5	2768.4	3
5355	6	2931.5	3
5640	6	2937.8	3
5750	6	3156.8	4
5903	5	3190.6	4
5920	5	3261.7	4
6080	5	3954.2	4
6090	5		
6434	5		
6458	9		
6632	15		
6656	7		
7008	5		
7470	10		
7765	25		
7925	30		
8148	35		
8605	30		
8760	35		
8939	30		
9291	35		
10 270	30		

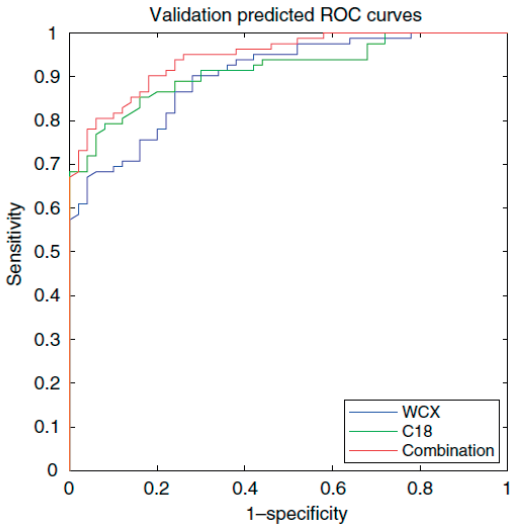


Figure 1 Receive operating characteristic (ROC) curves of the validation set based on weak cation exchange (WCX) and reverse phase (RP) C18 data sets separately and after combination. The area under the ROC curve is a measure of between-group separation (case–control).

RESULTS

Patients

There were 126 outpatients (76 men) with CRC before treatment of median age 65 (25-90) years. The control group included 277 normal individuals (110 men) of median age 56 (24-84) years. The validation set consisted of 50 pre-treatment CRC patients (28 men) of median age 68 (26-91) years and 82 controls (32 men) of median age 45 (21-75) years (Table 2).

Table 2 Patient characteristics.

	Test set		Validation set	
	Patients	Controls	Patients	Controls
n	126	277	50	82
Median age (years)	65	56	68	45
Age range (years)	25–90	24–84	26–91	21–75
Male (%)	76 (60)	110 (40)	28 (56)	32 (39)
Female (%)	50 (40)	167 (60)	22 (44)	50 (61)
Stage (Dukes)				
A (%)	22 (18)	NA	7 (14)	NA
B (%)	52 (41)	NA	15 (30)	NA
C (%)	28 (22)	NA	18 (36)	NA
D (%)	24 (19)	NA	10 (20)	NA

NA, not applicable.

Peptide- and protein signatures

In total 535 serum samples (test- and validation set) were processed with two types of magnetic beads. MALDI-TOF profiles were obtained in quadruplicate, yielding 2140 WCX- and 2140 RPC18-profiles. Profiles were baseline-corrected, aligned and of the mean of the four quadruplicates was calculated. Low-quality profiles (approximately 1%) resulting from a failure in sample processing or failed MALDI-spotting were excluded from analysis. The strategy for data analysis and statistical evaluation is shown in Fig. 2. First, all peptide- and protein profiles, obtained from RP C18 and WCX workup procedures were aligned to the *m/z*-axis. Then 42 and 57 peaks (summarized in Table 1) were selected from the RP C18- as well as WCX profiles indicating patient samples (in green) and controls (in blue). In this way, two data sets were obtained that were used for statistical analysis. In the combination plot of the patient samples the correctly classified cases are in green, whereas the remaining wrongly classified cases are in red. From this plot it can be seen that 18 of the 50 cases were incorrectly classified. The combined classification results of the control samples show that all were correctly classified (in blue) (Fig. 2, bottom right). From this plot it can be seen that only 4 of the 82 control samples were incorrecty classified as "cases" (in red). The clinical background of incorrectly classified patient and control samples was further evaluated and the results are summarized in Table 3.

Figure 2 Overview of study design and classification results. On the left-hand side typical examples of peptide (reverse phase, RP C18) and protein (weak cation exchange, WCX) profiles are shown, cases are in green and controls in blue. From all RP C18 and WCX profiles 42 and 57 peaks, respectively, were selected for statistical analysis. The results for the validation set are plotted on the right-hand side and further explained in the Results.

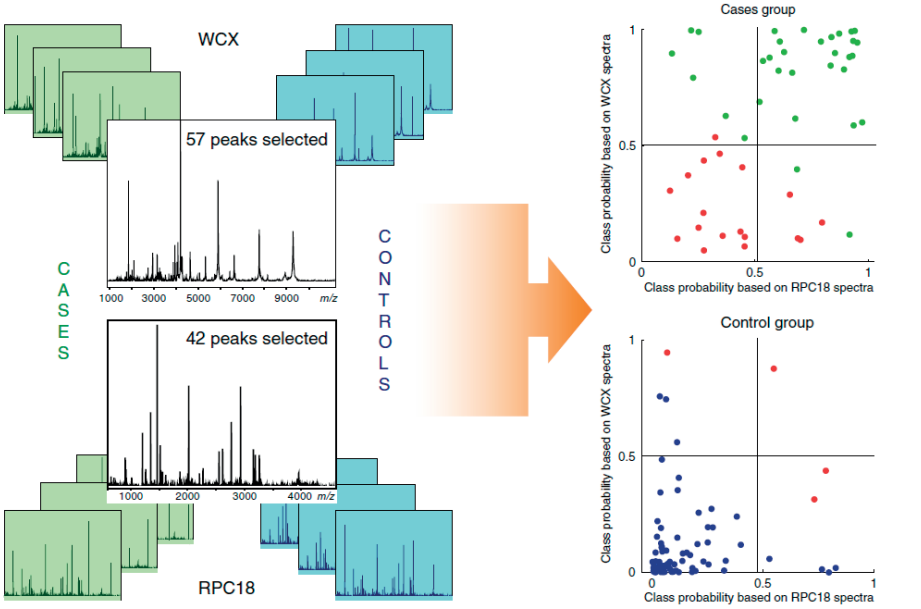


Table 3 Characteristics of misclassified cases for different cut-off values. 'Misclassified cases' in the case group are patients with colorectal cancer (CRC) with the specific cut-off value who were misclassified as controls, whereas 'Misclassified cases' in the control group represent controls with the same cut-off value were misclassified as CRC patients.

	Cut-off value							
	0.5		0.3		0.2		0.1	
	Case	Control	Case	Control	Case	Control	Case	Control
(a) Characteristics of misclassified cases and controls in the test set (total number of cases and controls 126 and 277)								
Total misclassified (%)	23 (18)	16 (5.8)	13 (10)	36 (13)	10 (7.9)	51 (18)	4 (3.2)	71 (26)
Sensitivity	0.817		0.896		0.921		0.968	
Specificity	0.942		0.870		0.816		0.754	
Gender								
Male (%)	19 (83)	9 (56)	12 (92)	17 (47)	9 (90)	21 (41)	4 (100)	27 (38)
Female (%)	4 (17)	7 (44)	1 (8)	19 (53)	1 (10)	30 (59)	0 (0)	44 (62)
Age (years)								
Median	60	63	57	60	57	57	56	57
Minimum	50	38	52	35	52	35	55	35
Maximum	82	80	82	82	82	84	74	84
Dukes								
A	5	NA	3	NA	2	NA	1	NA
B	8	NA	4	NA	4	NA	1	NA
C	7	NA	5	NA	3	NA	2	NA
D	3	NA	1	NA	1	NA	0	NA
(b) Characteristics of misclassified cases and controls in the validation set (total number of cases and controls 50 and 82)								
Total misclassified (%)	18 (36)	4 (4.9)	12 (24)	7 (8.5)	9 (18)	11 (13)	2 (4.0)	21 (26)
Sensitivity	0.640		0.760		0.820		0.960	
Specificity	0.951		0.914		0.866		0.780	
Gender								
Male (%)	9 (50)	2 (50)	5 (42)	2 (29)	4 (44)	2 (18)	0 (0)	6 (29)
Female (%)	9 (50)	2 (50)	7 (58)	5 (71)	5 (56)	9 (82)	2 (100)	15 (71)
Age (years)								
Median	68	34	71	33	67	35	62	41
Minimum	39	26	39	23	39	23	57	23
Maximum	86	51	86	52	82	52	67	69
Dukes								
A	7	NA	6	NA	5	NA	1	NA
B	2	NA	1	NA	1	NA	0	NA
C	6	NA	4	NA	2	NA	0	NA
D	3	NA	1	NA	1	NA	1	NA

NA, not applicable.

DISCUSSION

In this study we have evaluated mass spectrometry-based peptide and protein signatures for improved early cancer detection, motivated by that fact that the success rate of currently available tests for early diagnosis of CRC is rather low (11). These signatures, or profiles, were used successfully to distinguish CRC patients from healthy controls with a high discriminative power. It was found that the applied technology is a potential candidate for screening and early detection of CRC.

Despite the high discriminative power larger studies are essential (and on-going) to investigate the "tumour-specificity" of the obtained discriminating signatures. Survival

is relatively good when CRC is diagnosed in an early stage (3). Early detection will identify cancer when it is still localized and curable, not only preventing mortality, but also reducing morbidity and costs. Detection of symptomless CRC or precursor lesions through population screening allows for more effective treatment, which would likely improve long-term survival (3;4).

Full automation of the preparation and analysis process ensures standardization and robustness together with high discriminative power, supporting the potential of this test for screening programs. Encouraging results of well above 90% were obtained with regard to specificity- and sensitivity values. Cut-off levels can be chosen depending on the defined strategy for patient management and availability of colonoscopy facilities. Implementing a test for population screening requires consideration of factors such as compliance and costs. Enhanced sensitivity is an essential goal in the development of a screening test; however the fine-tuning of the ideal cut off value also depends on the organisation of the healthcare environment. False positive results are associated with patient anxiety and unnecessary colonoscopy (20). Zorzi and co-workers (21) evaluated five (large) population screening programs using iFOBT, that reported a total of 267,769 screened individuals of which 13,388 (5.0%) had a positive iFOBT test. From this group 90.3% (12,089 persons) were followed-up with colonoscopy, of which 748 individuals (6.2%) had a screen detected cancer. Thus, more than 90% of the persons with a positive iFOBT resulted in a negative colonoscopy (21).

Colonoscopy is an invasive procedure with a complication rate of 0.8-2% (22;23), which often requires sedation, which includes monitoring, extra nursing support and risk. Furthermore colonoscopy is time consuming and not really suited for screening. Both colonoscopy and CTC require bowel preparation and have a low participation rate of respectively 22% and 34% (1). The serum proteomics test is based on the analysis of one tube of peripheral blood, which is more convenient for the patient.

In conclusion, serum protomics analysis is easy to apply, cheap and patient-friendly with good sensitivity and specificity. The next step is to compare the test performance to current screening methods such as iFOBT. To this end, population screening will be evaluated, comparing serum proteomics analysis with iFOBT using colonoscopy as the gold standard. This may ultimately result in new guidelines for CRC screening in the Netherlands.

REFERENCES

1. Stoop EM, de Haan MC, de Wijkerslooth TR, Bossuyt PM, van BM, Nio CY, et al. Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial. Lancet Oncol 2012 Jan;13(1):55-64.
2. Lieberman D. Progress and challenges in colorectal cancer screening and surveillance. Gastroenterology 2010 Jun;138(6):2115-26.
3. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, et al. Cancer statistics, 2008. CA Cancer J Clin 2008 Mar;58(2):71-96.
4. Davies RJ, Miller R, Coleman N. Colorectal cancer screening: prospects for molecular stool analysis. Nat Rev Cancer 2005 Mar;5(3):199-209.
5. Miller S, Steele S. Novel molecular screening approaches in colorectal cancer. J Surg Oncol 2012 Apr 1;105(5):459-67.
6. Levi Z, Rozen P, Hazazi R, Vilkin A, Waked A, Maoz E, et al. A quantitative immunochemical fecal occult blood test for colorectal neoplasia. Ann Intern Med 2007 Feb 20;146(4):244-55.
7. Smith A, Young GP, Cole SR, Bampton P. Comparison of a brush-sampling fecal immunochemical test for hemoglobin with a sensitive guaiac-based fecal occult blood test in detection of colorectal neoplasia. Cancer 2006 Nov 1;107(9):2152-9.
8. Johnson CD, Chen MH, Toledano AY, Heiken JP, Dachman A, Kuo MD, et al. Accuracy of CT colonography for detection of large adenomas and cancers. N Engl J Med 2008 Sep 18;359(12):1207-17.
9. Duffy MJ, van DA, Haglund C, Hansson L, Holinski-Feder E, Klapdor R, et al. Tumour markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines for clinical use. Eur J Cancer 2007 Jun;43(9):1348-60.
10. Winawer S, Classen M, Lambert R et al. World Gastroenterology Organization / International Digestive Cancer Alliance Practice Guidelines for Colorectal Cancer Screening 2007.
11. Smith RA, Cokkinides V, Brooks D, Saslow D, Shah M, Brawley OW. Cancer screening in the United States, 2011: A review of current American Cancer Society guidelines and issues in cancer screening. CA Cancer J Clin 2011 Jan;61(1):8-30.
12. Vogelstein B, Kinzler KW. Winning the war: science parkour. Sci Transl Med 2012 Mar 28;4(127):127ed2.
13. Hanash S. Disease proteomics. Nature 2003 Mar 13;422(6928):226-32.
14. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011 Mar 4;144(5):646-74.
15. Huijbers A, Velstra B, Dekker TJ, Mesker WE, van der Burgt YE, Mertens BJ, et al. Proteomic serum biomarkers and their potential application in cancer screening programs. Int J Mol Sci 2010;11(11):4175-93.
16. Bladergroen MR, Derks RJ, Nicolardi S, de VB, van BS, van der Burgt YE, et al. Standardized and automated solid-phase extraction procedures for high-throughput proteomics of body fluids. J Proteomics 2012 Dec 21;77:144-53.
17. Jimenez CR, El FZ, Knol JC, Hoekman K, Kruyt FA, Giaccone G, et al. Automated serum peptide profiling using novel magnetic C18 beads off-line coupled to MALDI-TOF-MS. Proteomics Clin Appl 2007 Jun;1(6):598-604.
18. Callesen AK, Madsen JS, Vach W, Kruse TA, Mogensen O, Jensen ON. Serum protein profiling by solid phase extraction and mass spectrometry: a future diagnostics tool? Proteomics 2009 Mar;9(6):1428-41.

19. Velstra B, van der Burgt YE, Mertens BJ, Mesker WE, Deelder AM, Tollenaar RA. Improved classification of breast cancer peptide and protein profiles by combining two serum workup procedures. J Cancer Res Clin Oncol 2012 Dec;138(12):1983-92.
20. Taylor KL, Shelby R, Gelmann E, McGuire C. Quality of life and trial adherence among participants in the prostate, lung, colorectal, and ovarian cancer screening trial. J Natl Cancer Inst 2004 Jul 21;96(14):1083-94.
21. Zorzi M, Fedato C, Grazzini G, Stocco FC, Banovich F, Bortoli A, et al. High sensitivity of five colorectal screening programmes with faecal immunochemical test in the Veneto Region, Italy. Gut 2011 Jul;60(7):944-9.
22. Arora G, Mannalithara A, Singh G, Gerson LB, Triadafilopoulos G. Risk of perforation from a colonoscopy in adults: a large population-based study. Gastrointest Endosc 2009 Mar;69(3 Pt 2):654-64.
23. Ko CW, Riffle S, Michaels L, Morris C, Holub J, Shapiro JA, et al. Serious complications within 30 days of screening and surveillance colonoscopy are uncommon. Clin Gastroenterol Hepatol 2010 Feb;8(2):166-73.

APPENDIX S1. DISCRIMINATION BETWEEN PATIENTS AND CONTROLS

Both the WCX and RPC18 data were calibrated separately with a discriminant rule using logistic ridge regression based on the training- or test set. Joint calibration of the classifiers and unbiased estimation of the class probabilities on the training set was achieved using double cross-validation, as described previously (26). The two sets of double cross-validated class probabilities which were thus obtained on the training data were then used as inputs for the estimation of an ordinary logistic regression model, which combines the predictions from the WCX and RPC18 training data. To evaluate this combination classifier, first the logistic ridge regression models were refit on the WCX and RPC18 data separately using the optimum penalty term identified in the previous double cross-validatory analysis. Then, for each validation sample these two logistic regression models were applied to obtain class predictions on the WCX and RPC18 sets separately. Finally, these two predictions were combined using the above described ordinary logistic regression combination model, which gives a single output class probability for each individual in the validation data. The double cross-validated results yielded a total recognition rate with an AUC of 0.95. For the validation set the AUC was 0.93 (see Figure 2). Different cut-off values were evaluated to match the most optimal test performance in a given population with respect to the colonoscopy capacities/facilities, as is summarized in Table 3. As an example, at a cut-off value of 0.5 the sensitivity and specificity numbers in the test set are 82% and 94% (validation: 64% and 95%), whereas at a cut-off of 0.2 the sensitivity and specificity are 92% and 82% (validation: 82% and 87%). A low cut-off value results in optimal sensitivity at the cost of specificity. Whereas when a higher specificity is preferred a higher cut-off value can be chosen. In Tables 3A and 3B an overview is given of the patient characteristics, which were misclassified in this study design at an associated chosen cut-off value.