SUPPLEMENTARY MATERIAL

Supplement to: Multicentre prospective evaluation of real-time optical diagnosis of T1 colorectal cancer in large non-pedunculated colorectal polyps using narrow band imaging (the OPTICAL-study)

Table of Contents

Supplementary Methods

Development of the risk score chart

We aimed to develop a score chart with the estimated risk for T1 CRC within LNPCPs with use of endoscopic features. Therefore, this analysis was restricted to non-invasive polyps and T1 CRCs (prevalence T1 CRC: 10.9%; 36/329). All endoscopic features were considered candidate predictors. However, for the robustness of our results, features with low variation were omitted beforehand (resulting in the exclusion of presence of excavation in this analysis). Moreover, in view of the low number of LNPCPs with Hiroshima type A, we binned type A and B for this analysis, in line with previous literature.[1, 2] This resulted in a final set of the following parameters selected for the model: location (proximal vs distal location), surface structure (non-granular, homogeneous granular, granular with a large nodule or granular with a non-granular erythematous area), depressed area (presence vs absence), spontaneous bleeding (presence vs absence), and the Hiroshima classification (type A-B vs type C1 vs type C2 vs type C3).

For the development of the chart, we conducted a penalized logistic regression analysis, named L1 penalized least absolute shrinkage and selection (LASSO) logistic regression, augmented with 5-fold cross-validation for internal validation.[3] LASSO is a statistical approach that allows robust multivariable prediction modelling, including predictor selection, in cohorts such as ours with a relative high number of predictors (i.e. endoscopic features) in comparison with cases (i.e. T1 CRCs).[3] LASSO shrinks the size of the coefficients of the endoscopic features based on the value of the penalty lambda (λ). With larger penalties, the estimates of weaker endoscopic features shrink towards zero, so that only the strongest predictors remain. We conducted mixed effects binary logistic LASSO models with random endoscopist and patient intercepts to account for clustering. We modelled the endoscopic features as fixed effects, and only used these fixed effect coefficients for our risk score. We did not include the random effects in our risk score (these cannot be used in new patients and other centers), and updated the intercept in a final step to obtain accurate calibration in the large. Fifty λ's were evaluated, and models were fit in each cross-validation set for each λ (Figure S1). The optimal λ was defined as the λ with the maximum cross-validation area under the curve (cvAUC). A coefficient profile plot was drawn in which the size of the coefficients of the endoscopic features shrinks based on the value of the λ, with the features selected for the prediction rule based on the optimal λ (Figure S2). The area under the receiver-operating curve (ROC) was used to assess the

performance of the model to discriminate between T1 CRC and non-invasive polyps, and were plotted for each cross-validation set for the optimal λ (Figure S3). To check the calibration of the multivariate model, a calibration plot was made presenting the predicted probability against the observed risk for T1 CRC (Figure S4). The absolute risk of T1 CRC was calculated with $P = (1/[1 + \exp(-1 * (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k))]) * 100$, in which β_0 is the intercept, and β_k is the regression coefficient for predictor X_k (the different endoscopic features that contributed to the differentiation of T1 CRCs from non-invasive polyps in LASSO regression-analysis as presented in Table S4). The accuracy (i.e. sensitivity, specificity, negative and positive predictive values (NPV and PPV) at different probability thresholds (i.e. absolute risk of T1 CRC) were plotted.

Patients age and sex were not included as candidate predictors in the analysis. However, a secondary analysis was performed to assess whether addition of these factors could improve the risk score (i.e., the cvAUC).

We included bootstrap-based 95% confidence intervals (CI) surrounding the regression coefficients and predicted probabilities. Within each bootstrap sample, we fitted the LASSO model at the optimal penalty derived in a preceding step from cross-validation. We did not include a reassessment of this optimal penalty in each bootstrap sample. The resulting 95%CI (based on the percentile method) cannot be interpreted in the classical sense of a 95%CI; they merely represent an indication of the level of uncertainty surrounding the reported estimates, due to the relatively small dataset and complex estimation procedures.

Validation of the risk score chart

We validated the risk score chart using data from 100 consecutive LNPCPs from 96 patients (61.5% male, median age 64.5 years; IQR 60.3 – 72.8) assessed in one of the participating centres (UMC Utrecht) by two endoscopists (LMGM and PD) after the study period (i.e., temporal validation). Similar to the development cohort, obvious advanced cancers (≥T2) or LNPCPs that were known to harbour cancer due to previous biopsies were not included. The validation set contained 15 T1 CRCs (superficial invasion: N= 6; deep invasion: N=6, invasion depth unable to determine: N=3) and 85 non-invasive LNPCPs (adenomas: N=75, traditional serrated adenomas: N=4, sessile serrated adenomas: N=6). Assessment of the endoscopic features was performed blinded for the histological diagnosis during colonoscopy in the same manner as the development set.

We used the developed formula ((1/[1+ exp (-1 * (-1.90 + 0.90 * spontaneous bleeding + 0.08 * non-granular surface + -0.34 * homogeneous granular + -0.02 * granular with large nodule + 0.41 * distal location + -0.02 * depressed area + -1.72 * Hiroshima type A-B + 1.57 * Hiroshima type C2 + 2.82 * Hiroshima type C3))]) * 100) to calculate the T1 CRC risk for each LNPCP, and quantified the discriminative ability using the AUC. To check the calibration, a calibration plot was made presenting the predicted probability against the observed risk for T1 CRC (Figure S4). As the validation set was derived from a tertiary referral centre in which relatively more T1 CRCs are encountered than in general hospitals, we expected a slight underestimation of the overall risk of T1 CRC (Figure S4C). We therefore also present a calibration plot with an updated intercept (Figure S4D), in which the intercept was used from the same tertiary hospital as derived in the development set (-1.20 instead of -1.90). By updating the intercept the average predicted risks of the updated model became more equal to the observed risk in the validation set.

Comparison of AUCs

Finally, we aimed to evaluate the value of NBI in addition to white-light features. For this analysis, we repeated the LASSO analysis two times: once with location and white-light features as endoscopic predictors and once with location, white-light features and NBI features as endoscopic predictors. For both models, we plotted the apparent AUC, and compared both AUCs using the DeLongs test in the pROC package of R. Similarly, the AUC of the score chart was compared with the AUC of the endoscopists' optical diagnosis.

Table S1. List of endoscopes used

Type of scope	Narrow-band imaging	High definition / Quality
CF-Q180AL/I	Yes	Quality
CF-H180AL/I	Yes	High definition
CF-H180DL	Yes	High definition
CF-H190L	Yes	High definition
CF-HQ190I	Yes	High definition
CF-Q180AL	Yes	Quality
GIF-1TH190	Yes	High definition
GIF-1TQ160	Yes*	Quality
GIF-H180	Yes	High definition
GIF-H190	Yes	High definition
GIF-HQ190	Yes	High definition
PCF-H180AL	Yes	High definition
PCF-H190DL	Yes	High definition
PCF-H190L	Yes	High definition

* After modification

Table S2. Hiroshima classification [1, 2]

Type	Definition
A	Microvessels are not visible or extremely opaque
B	• Regular surface pattern • Regular microvessel pattern
C1	• Irregular surface pattern • Thickness and distribution of vessels are homogeneous
C2	• More irregular surface pattern • Thickness and distribution of vessels are heterogeneous
C3	• Unclear surface pattern • Thickness and distribution of vessels are heterogeneous • Avascular area and scattered microvessel fragments

Supplementary results

Table S3. Diagnostic accuracy of optical diagnosis of T1 CRC and T1 CRC with deep submucosal invasion when excluding ≥T2 CRC or LNPCPs in which biopsies had been taken before optical assessment

		Excluding ≥T2 CRC	Excluding LNPCPs in which biopsies had been taken before optical assessment
Optical diagnosis of T1 CRC *(i.e., non-invasive vs. invasive LNPCPs)*	**Nr. of LNPCPs (disease positives)**	329 (36)	308 (42)
	Agreement (95%CI)	92.1 (89.9 – 95.1)	92.5 (89.5 – 95.3)
	Sensitivity (95%CI)	77.8 (60.9 – 89.9)	78.6 (63.2 – 89.7)
	Specificity (95%CI)	94.2 (90.9 – 96.6)	94.7 (91.3 – 97.1)
	PPV (95%CI)	62.2 (50.1 – 73.0)	70.2 (58.0 – 80.1)
	NPV (95%CI)	97.2 (94.9 – 98.5)	96.6 (94.0 – 98.0)
Optical diagnosis of T1 CRC with deep invasion *(i.e., endoscopically resectable vs endoscopically unresectable LNPCPs)*	**Nr. of LNPCPs (disease positives)**	323 (19)	303 (27)
	Agreement (95%CI)	96.6 (94.4 – 98.5)	96.0 (93.9 – 98.0)
	Sensitivity (95%CI)	57.9 (33.5 – 79.8)	66.7 (46.0 – 83.5)
	Specificity (95%CI)	99.0 (97.1 – 99.8)	98.9 (96.9 – 99.8)
	PPV (95%CI)	78.6 (52.7 – 92.3)	85.7 (65.4 – 95.0)
	NPV (95%CI)	97.4 (95.7 – 98.5)	96.8 (94.7 – 98.1)

Abbreviations: CI: confidence interval; CRC: colorectal cancer; LNPCP: large non-pedunculated colorectal polyp; NPV: negative predictive value; PPV: positive predictive value

Figure S1. Plot of the cross-validation area under the curve according to the penalty.

The green lines indicate the standard error. The red line indicates the optimal penalty (i.e. the penalty with the maximum cross-validation AUC).

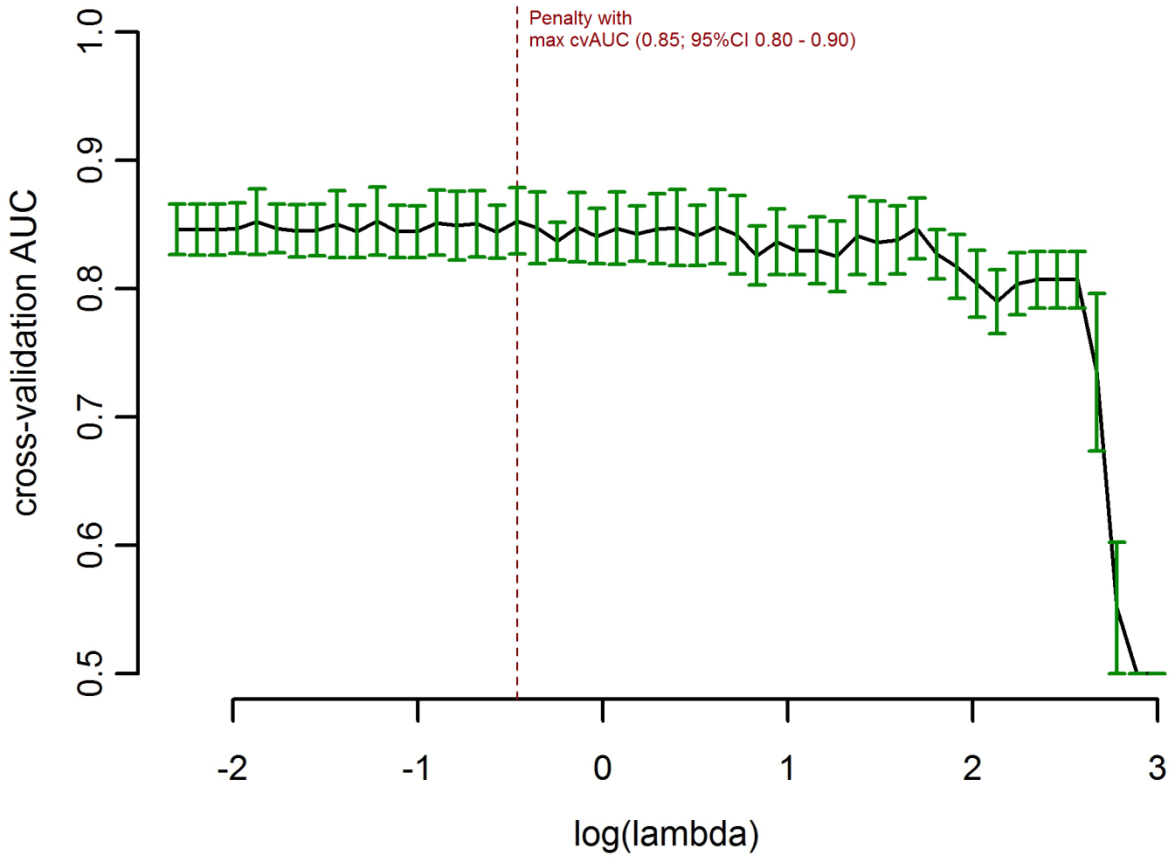


Figure S2. Coefficient profile plot showing how the size of the coefficients of the endoscopic features shrink with increasing value of the λ penalty, with the features and their coefficients selected for the risk score chart based on the optimal λ (i.e. the penalty with the maximum cross-validation AUC). Abbreviations: cvAUC: cross-validation area under the curve

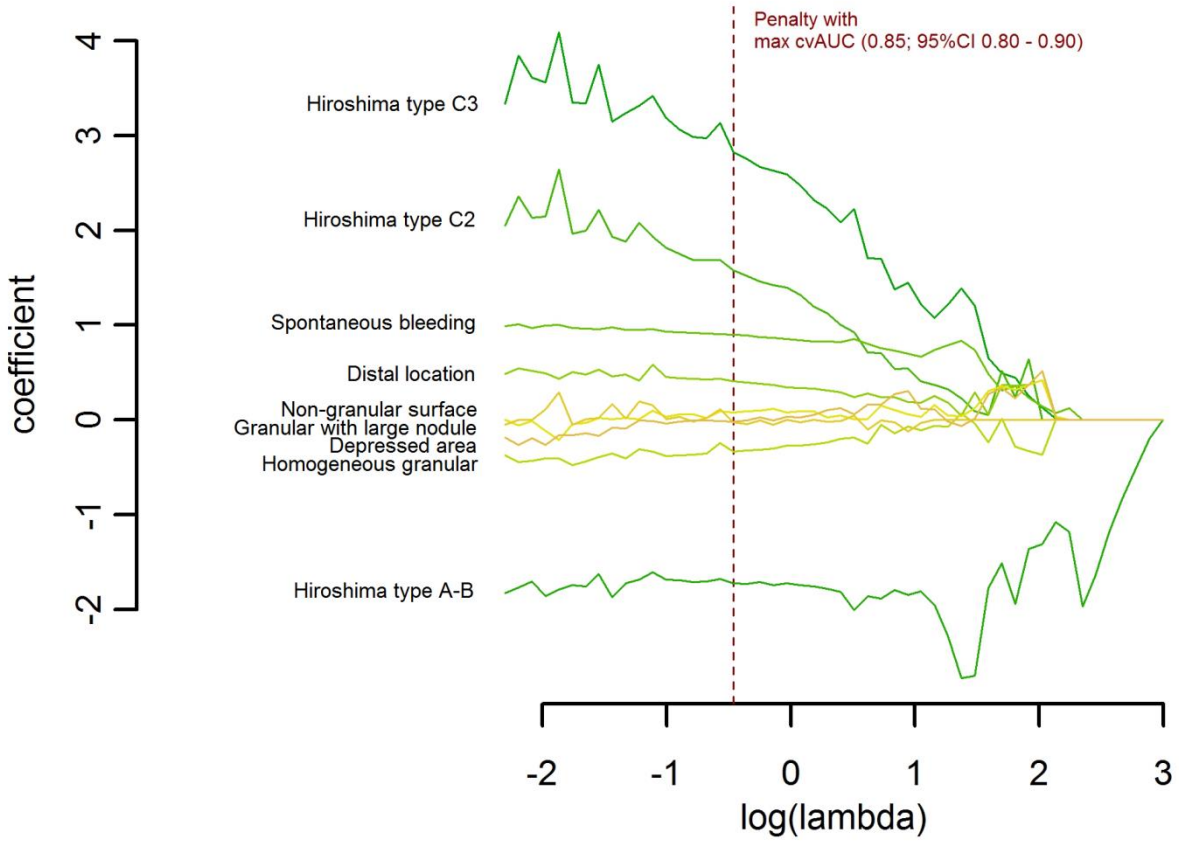


Figure S3. Receiver-operating curve (ROC) to assess the performance of the model in terms of accuracy of correct prediction of T1 CRC, plotted for each cross-validation set.

Black line is cross-validation AUC; grey lines are the AUC as achieved in the 5 cross-validation sets; dashed line is the reference line.

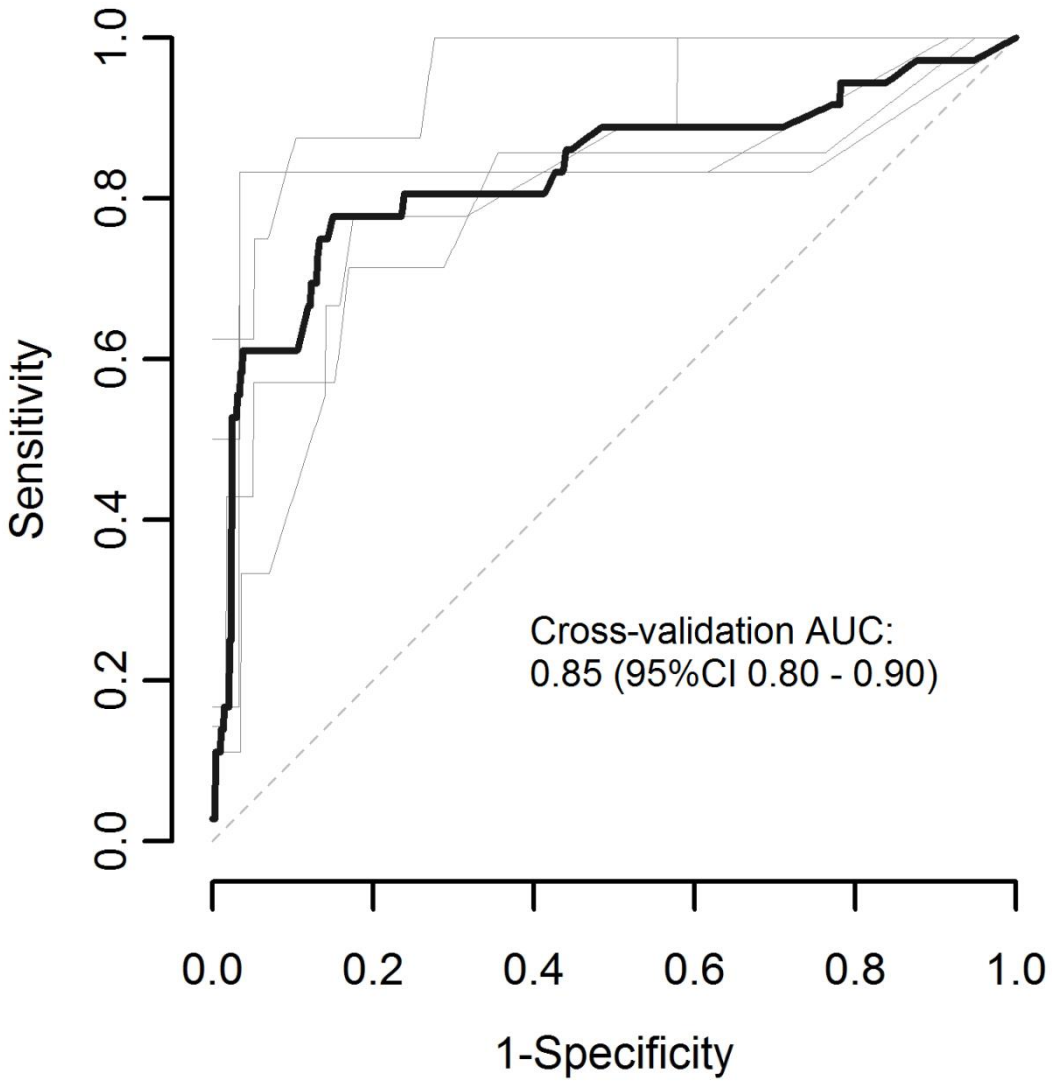


Figure S4. Calibration plot presenting the predicted probability against the observed risk for T1 CRC. A. Cross-validation calibration curve development set, B. Apparent calibration curve development set, C. Calibration curve validation set without updated intercept (intercept: -1.90), D. Calibration curve validation set with updated intercept (intercept: -1.20). Solid grey line shows perfect calibration (i.e. perfect agreement between predicted and observed T1 CRC risk); solid black line shows the smoothed calibration curve; the dotted line is the top 5% of the predicted probabilities; histograms depict predicted probability distribution of patients with T1 CRC (top) or with non-invasive polyps (bottom).

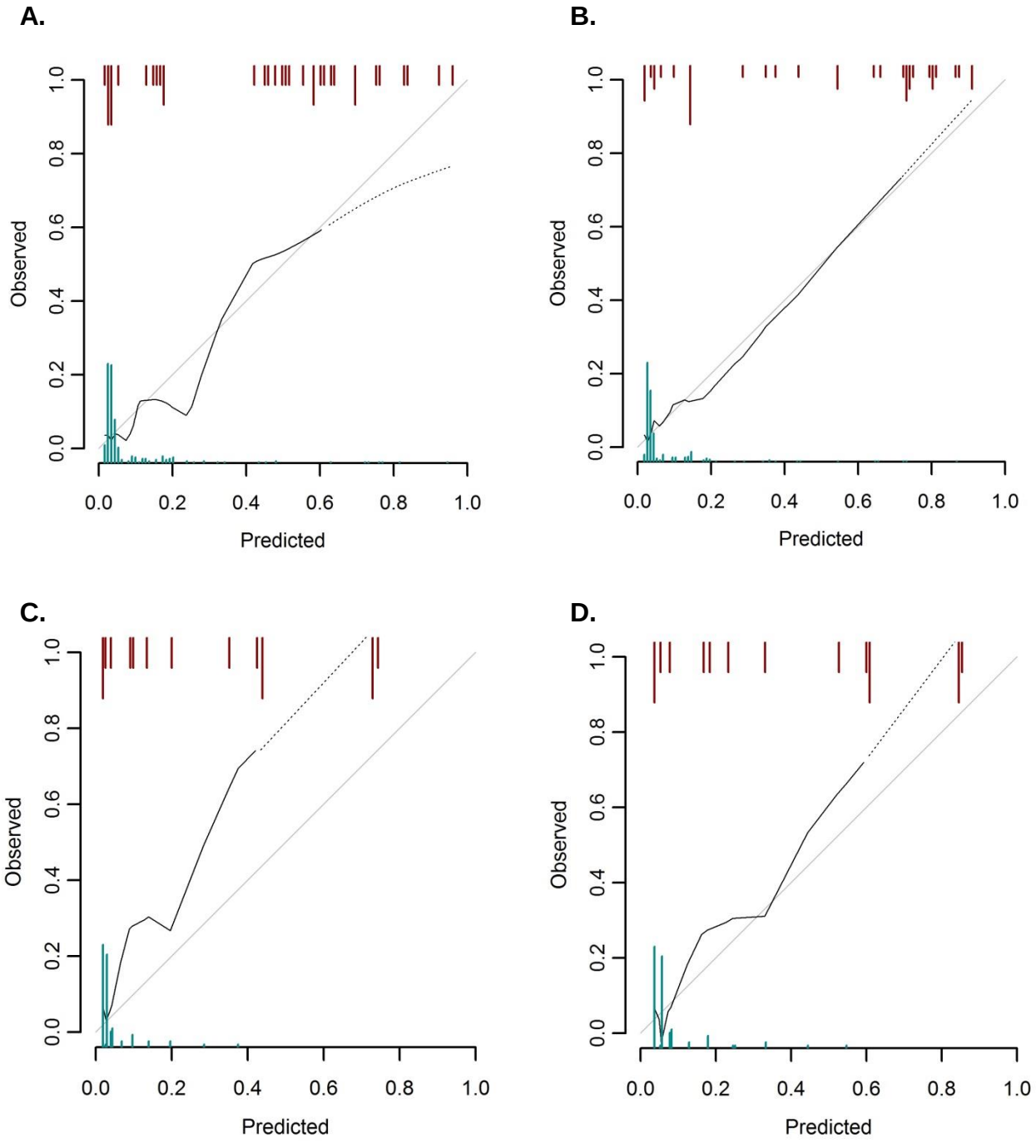


Table S4. LASSO-derived regression coefficients and multivariate odds ratios of endoscopic predictors of T1 CRC

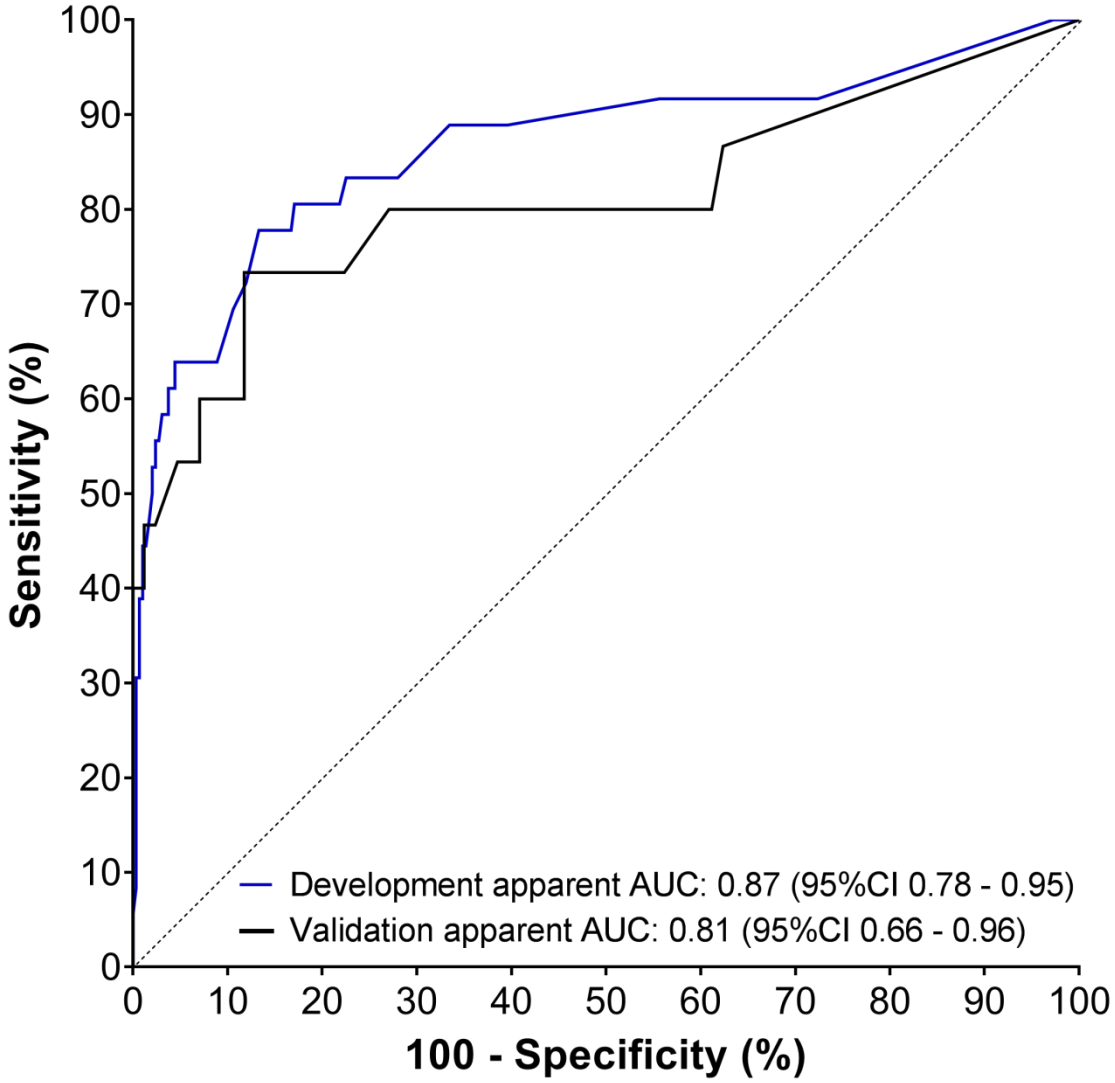
	Regression coefficient (95%CI)[a,b]	Multivariate odds ratio (95%CI)[b]
Location		
- Proximal location (proximal to splenic curve)	Reference	Reference
- Distal location (distal to splenic curve)	0.41 (-0.21 to 1.44)	1.51 (0.81 to 4.23)
Spontaneous bleeding		
- Absent	Reference	Reference
- Present	0.90 (0.02 to 1.98)	2.45 (1.02 to 7.22)
Depressed area		
- Absent	Reference	Reference
- Present	-0.02 (-0.92 to 0.78)	0.98 (0.40 to 2.19)
Granularity		
- Homogeneous granular	-0.34 (-1.50 to 0.37)	0.72 (0.22 to 1.45)
- Granular with large nodule	-0.02 (-0.88 to 0.85)	0.98 (0.41 to 2.34)
- Granular with NG erythematous area	Reference	Reference
- Non-granular	0.08 (-0.66 to 0.90)	1.08 (0.52 to 2.46)
Hiroshima classification		
- Hiroshima type A-B	-1.72 (-3.15 to -0.64)	0.18 (0.04 to 0.53)
- Hiroshima type C1	Reference	Reference
- Hiroshima type C2	1.57 (0.25 to 3.30)	4.83 (1.28 to 27.05)
- Hiroshima type C3	2.82 (1.40 to 5.15)	16.84 (4.06 to 172.24)
Intercept	-1.90	
Cross validation area under the curve	0.85 (95%CI 0.80 to 0.90)	

Abbreviations: CI: confidence interval

a. The regression coefficient is the log of the odds ratio. For regression coefficient X, the multivariate odds ratio was calculated with exp(X). The absolute risk for T1 CRC within a LNPCP was calculated using the formula $P = (1/[1 + \exp(-1 * (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k))]) * 100$, in which β_0 is the intercept, and β_k is the regression coefficient for predictor X_k (the different endoscopic features that contributed to the differentiation of T1 CRCs from non-invasive polyps). For example, the risk of T1 CRC in a non-granular LNPCP (regression coefficient: 0.08) located in the distal colon (regression coefficient: 0.41) without spontaneous bleeding or depression, and on NBI-assessment Hiroshima type B (regression coefficient: -1.72) is (1/[1 + exp(-1 * (-1.90 + (0.08 + 0.41 - 1.72)]) * 100 = 4.2%. All other risk estimates are provided in the risk chart (see Table 5).

b. Bootstrap-based 95%CI are provided (based on the percentile method). These cannot be interpreted in the classical sense of a 95%CI; they merely represent an indication of the level of uncertainty surrounding the reported estimates, due to the relatively small dataset and complex estimation procedures.

Figure S5. Apparent receiver operating characteristic curves for optical diagnosis of T1 CRC using white-light and NBI features in the development set (blue line) and validation set (black line)



References

1 Kanao H, Tanaka S, Oka S*, et al.* Narrow-band imaging magnification predicts the histology and invasion depth of colorectal tumors. Gastrointest Endosc 2009;**69**:631-6.

2 Jang HW, Park SJ, Cheon JH*, et al.* Does magnifying narrow-band imaging or magnifying chromoendoscopy help experienced endoscopists assess invasion depth of large sessile and flat polyps? Dig Dis Sci 2014;**59**:1520-8.

3 Tibshirani R. Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society Series B-Methodological 1996;**58**:267-88.