MULTICENTRIC VALIDATION OF A LOGISTIC MODEL BASED ON PHENOTYPIC FEATURES TO PREDICT MICROSATELLITE INSTABILITY IN COLORECTAL ADENOCARCINOMAS.

R Román[1], M Verdú[1,3], A Vidal[3,5], M Calvo[4], X Sanjuan[5], M Jimeno[6], A Salas[7], J Autonell[8], I Trias[9], N Rodón[1], B García[1], M González[1], X Puig[1,2,3]. [1]BIOPAT. Biopatologia Molecular, SL, Grup Assistència; [2]Hospital de Barcelona, SCIAS, Grup Assistència; [3]Histopat Laboratoris, [4]Statistics Department, Universitat de Barcelona; [5]Hospital Universitari de Bellvitge; [6]Hospital del Mar; [7]Hospital Mútua de Terrassa; [8]Hospital General de Vic; [9]Hospital Plató, Barcelona, Spain.

$$1 - \frac{1}{1 + e^{(x)}}$$

>0.29 MSI-H

a= Location: 0= proximal, 1=distal
b= Growth pattern: 0= expansive, 1=infiltrative
c= Crohn: 0=present, 1=absent
d= Solid %
e= Mucinous %
f= Cribriform %
g= Ki-67%
h= p53 accumulation %

x = (-2.648-(a*1.6)-(b*1.884)-(c*2.401)+(d*0.045)+(e*0.021)-(f*0.118)+(g*0.05)-(h*0.029))

Figure 1: MSI-H prediction model based on phenotypic features

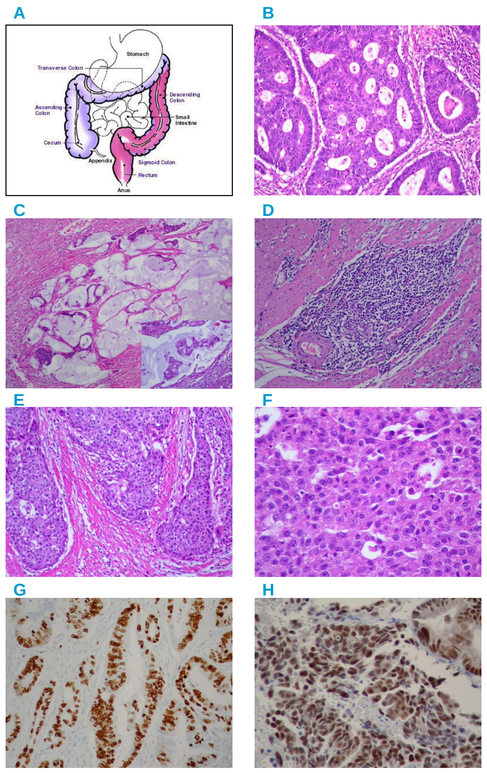

Figure 2: The eight parameters included in our prediction model. A: Proximal location (purple), B: Cribiform pattern, C: Extracellular mucinous pattern, D: "Crohn like" inflammatory response, E: Expansive growth, F: Solid pattern, G: Ki-67 proliferative index, H: p53 expression.

BACKGROUND

High microsatellite instability (MSI-H) allows the identification of a subset of colorectal adenocarcinomas associated with good prognosis and a higher incidence of Lynch syndrome. The objective of this work was to assess the interobserver variability of a MSI-H prediction model based on phenotypic features. This model was previously generated using a series of 204 cases and achieved a sensitivity of 77.8%, a specificity of 96.8% and a negative predictive value of 97.8% (Figure 1).

MATERIALS AND METHODS

Patients and Specimens

For our validation series a total of 265 unselected colorectal cancer cases were prospectively collected from five different hospitals in the area of Catalonia, Spain. Cases were evaluated by the corresponding pathologist at each participating centre, and the eight histopathologic parameters included in our prediction model, namely tumor location, expansive growth, solid, mucinous and cribriform patterns, presence of peritumoral Crohn-like inflammatory response, Ki-67 proliferative index and p53 immunophenotype recorded (Figure 2). Representative formalin-fixed paraffin-embedded blocks of paired tumor and normal tissue were sent to our centre for MSI-H prediction using our statistic model (Figure 1) and MSI-H molecular analysis. Simultaneously a prospective series of 148 cases was collected and identically assessed in our institution; this added interobserver variability and also serve to increase the global validation series.
No special training regarding the evaluation of the parameters included in our logistic model was given to any of the pathologists in order to test its robustness despite interobserver heterogeneity. Table 1 summarizes the results of the eight parameters included in the prediction model for the validation series and our own prospective series.

Analysis of Microsatellite Instability

MSI status was evaluated using the five microsatellites from the NCI panel (BAT25, BAT26, D5S346, D2S123 and D17S250) in a multiplex PCR reaction. Instability was assigned to a marker if its fragment pattern displayed either additional peaks or the appearance of separated novel fragments when the profiles of normal and tumor tissue were compared with each other. In addition instability observed using the six microsatellites markers aimed at elucidating the LOH status of chromosomes 18q or 17p (TP53 locus) was also recorded and a separate global MSI status was assigned with the 11 microsatellite markers. According to the consensus definitions of the US NCI, tumors were classified as exhibiting high microsatellite instability (MSI-H) when 30% or more of the tested loci resulted unstable and non-MSI-H when they were less than 30%.

Statistical Analyses

Statistical analysis was carried out by Fisher's exact tests (for categorical variables) or Kruskal-Wallis test (for continuous variables). Probability values (p-values) were considered significant when less than 0.05. Data were analyzed using the R package v.2.7.0 (©2008, R Development Core Team) supplied with the coin package. "

RESULTS

Homogeneity assessment revealed significant differences between the five participating hospitals in the estimation of the expansive growth, presence of Crohn-like inflammatory response, percentage of cribriform pattern and Ki-67 expression (Table 2). Despite this observation, our model was globally able to predict MSI-H with a negative predictive value of 97.0%. The results obtained with our own prospective series were equivalent, achieving a negative predictive value of 97.8% and did not differ from those obtained in the initial study where a negative predictive value achieved was 97.8% (Table 3a). Crude data on predicted *vs* observed MSI-H cases from the validation series has been included in table 3b.

DISCUSSION

The complexity and importance of the MSI study has triggered the development of models based on pathological features to predict instability status. The model we presented and has been now further validated, focus on a non-selected population of patients with colorectal cancer aiming at identifying not only possible Lynch syndrome candidates, but any carcinomas exhibiting an MSI-H phenotype and thus a better prognosis. The high negative predictive value achieved by our model allows the reduction of the cases to be tested for MSI to approximately 10%, only those cases predicted as MSI-H should be confirmed as such. Furthermore, the easy evaluation of the parameters included in the model and the fact that it is not influenced by interobserver variability, renders it a very useful tool for the routine practice and can reinforce the current clinical protocols to detect the subset of colorectal cancer patients bearing HNPCC risk and/or MSI-H phenotype.

REFERENCES

•Colomer A. *et al*. A novel logistic model based on clinico-pathological features predicts microsatellite instability in colorectal carcinomas. *Diagn Mol Pathol* 2005;14:213–23.
•Jenkins MA. *et al*. Pathology features in Bethesda Guidelines predict colorectal cancer microsatellite instability: a population-based study. *Gastroenterology* 2007;133:48-56.
•Thibodeau SN. *et al*. Microsatellite instability in cancer of the proximal colon. *Science* 1993;53:5849-5852.

CATEGORICAL VARIABLES	VALIDATION SET	OUR PROSPECTIVE SET
Location		
Proximal	104	48
Distal	161	100
Growth pattern		
Infiltrative	114	45
Expansive	151	103
Crohn-like inflammatory response		
Present	113	69
Absent	152	79
NUMERICAL VARIABLES	MEAN SD	MEAN SD
Solid carcinoma (%)	3.1 ± 12	6.4 ± 20
Mucinous carcinoma (%)	9.5 ± 20	11 ± 23
Cribriform structures (%)	16 ± 22	10 ± 18
Ki-67 proliferative index (%)	69 ± 20	64 ± 20
P53 accumulation (%)	46 ± 41	41 ± 38

Table 1: Clinicopathological parameters included in the prediction model for the validation series and our own prospective series.

	Centre 1	Centre 2	Centre 3	Centre 4	Centre 5	p-value
Proximal	6	12	18	42	26	0.33
Distal	4	15	42	58	42	
Infiltrative	6	4	37	44	26	<0.001
Expansive	4	23	23	56	50	
Crohn-like present	1	11	15	69	17	<0.001
Crohn-like absent	9	16	45	31	51	
Solid carcinoma (%)	-	5.7	1.2	1.2	6.5	0.20
Mucinous carcinoma (%)	-	11	10	10	6.5	0.81
Cribriform structures (%)	-	25	17	11	18	<0.001
Ki-67 proliferative index (%)	-	68	69	67	78	0.004
P53 accumulation (%)	-	40	55	43	47	0.32

Table 2: Homogeneity assessment between the five participating hospitals.

A

	Initial study	Validation set	Our prospective set
Accuracy	95.10	91.32	91.41
Sensitivity	66.67	66.67	83.33
Specificity	96.81	93.44	92.41
Positive predictive value	75.00	46.67	57.69
Negative predictive value	97.85	97.02	97.81
Total	204	265	148

B

	Observed MSI-H	Observed MSS	Total
Predicted MSI-H	14	16	30
Predicted MSS	7	228	235
Total	21	244	265

Table 3: A: Statistical parameters achieved in our initial study, validation and prospective series. B: Crude data from the validation series, the green circle indicates those cases predicted as MSI-H whose status would require confirmation; the red circle indicates those MSI-H cases that would have been missed by our prediction model.