

Optimizing Genetic Testing Strategy for Suspected Attenuated Adenomatous Polyposis: Effective Solutions in Public Health Systems

Natalia García-Simón, Fátima Valentín, Ana Royuela, Beatriz Hidalgo-Calero, Ricardo Blázquez-Martín, Montserrat de-Miguel-Reyes, José M Sánchez-Zapardiel, Luisa Adán-Merino, Alejandro Rodríguez-Festa, Patricia Gallego-Gil, Pilar Mediavilla-Medel, Laura Quiñonero-Moreno, Lourdes Gutiérrez, Alberto Herreros-de-Tejada, Antonio Sánchez, Mariano Provencio, Atocha Romero.

Natalia García-Simón, Alejandro Rodríguez-Festa, Patricia Gallego-Gil, Pilar Mediavilla-Medel, Atocha Romero, Hereditary Cancer Unit, Medical Oncology Department, Puerta de Hierro University Hospital, Majadahonda, Madrid, 28222, Spain.

Fátima Valentín, Alberto Herreros-de-Tejada, Gastroenterology Department, Biomedical Research Institute (IDIPHISA), Puerta de Hierro University Hospital, Majadahonda, Madrid, 28222, Spain.

Ana Royuela, Biostatistics Unit, Puerta de Hierro Biomedical Research Institute (IDIPHISA), CIBERESP, ISCIII. Majadahonda, Madrid, 28222, Spain.

Beatriz Hidalgo-Calero, Ricardo Blázquez-Martín, Montserrat de-Miguel-Reyes, José M Sánchez-Zapardiel, Hereditary Cancer Laboratory, 12 de Octubre University Hospital, Madrid, 28041, Spain

Beatriz Hidalgo-Calero, Health Research Institute 12 de Octubre Hospital (imas12), Madrid, 28041, Spain

Luisa Adán-Merino, Gastroenterology Department, Infanta Leonor University Hospital, Madrid, 28031, Spain.

Laura Quiñonero-Moreno, Lourdes Gutiérrez, Antonio Sánchez, Mariano Provencio, Medical Oncology Department, Puerta de Hierro University Hospital, Majadahonda, Madrid, 28222, Spain.

ORCID: 0000-0002-1634-7397

Supported by Rio Hortega Fellowship grant no. CM22/00213 from the Instituto de Salud Carlos III (ISCIII). P.M-M is supported by PFIS fellowship grant no. FI22/00321, ISCIII.

Corresponding author: Atocha Romero. Hereditary Cancer Unit, Medical Oncology Department, Puerta de Hierro University Hospital, Majadahonda, Madrid, 28222, Spain. atocha10@hotmail.com

ABSTRACT

BACKGROUND

Attenuated familial adenomatous polyposis (AFAP) and MUTYH-associated polyposis (MAP) represent key hereditary attenuated adenomatous polyposis syndromes, with *APC* and *MUTYH* being the associated genes. While guidelines recommend genetic testing based on polyp count for diagnosing these syndromes, age is often overlooked despite its influence on polyp prevalence.

AIM

To enhance genetic testing strategies for suspected attenuated adenomatous polyposis by combining polyp count and age in a probability calculator.

METHODS

Retrospective study of adult patients referred to NGS genetic testing for suspected attenuated adenomatous polyposis (accumulated history of <100 adenomas) (discovery cohort, N=138). Data collected included age, adenoma count at the time of genetic testing, and test results. Multivariable logistic regression model was developed to link a positive genetic test result with age and adenoma count. The model was externally validated with populations from two tertiary hospitals in our region (validation cohort, N=259).

RESULTS

In the discovery cohort, 13 (9.4%) patients tested positive for pathogenic mutations. Positive cases were younger (odds ratio (OR): 0.91, 95% confidence interval (95%CI): 0.86-0.96) and developed more adenomas (OR: 1.08, 95%CI: 1.04-1.13) compared to negative cases. The logistic regression model combining age and polyp count demonstrated an AUC of 0.92. Using a cutoff probability of 3.5%, the model achieved 100% sensitivity and 58% specificity in identifying positive cases. In the external validation, the model accurately predicted 14 out of 16 positive cases (88%). The remaining two positive cases were a patient with an *AXIN2* mutation in heterozygosis, and a patient with a *NTHL1* mutation in homozygosis. Performance evaluation of both hospitals yielded AUC values of 0.77 and 0.90.

CONCLUSIONS

Older individuals, particularly those with few polyps, may have diminished chances of detecting hereditary syndromes. Incorporating age as a criterion for genetic testing referral has the potential to refine patient selection and improve cost-effectiveness.

Key Words: Hereditary attenuated adenomatous polyposis; Age; Adenomas; Genetic testing; *APC*; *MUTYH*

Core tip: This retrospective multi-center observational study aims to investigate the weight of age and polyp count on patients diagnosed with hereditary attenuated adenomatous polyposis syndrome compared to those with sporadic polyposis. Based on this, a probability calculator has been developed to enhance the efficiency and cost-effectiveness of genetic testing protocols for attenuated adenomatous polyposis syndromes within public health systems.

INTRODUCTION

Hereditary polyposis syndromes are known to be accountable for about 2-3% of all cases of colorectal cancer (CRC)^[1-3]. The most common polyposis syndromes are familial adenomatous polyposis (FAP) (OMIM #175100), attenuated FAP (AFAP) (OMIM #175100), and *MUTYH*-associated polyposis (MAP) (OMIM #608456), while other syndromes such as hamartomatous polyposis are less frequent^{4,5}. The main genes associated with hereditary adenomatous polyposis syndromes are *APC* (OMIM #611731) gene, for FAP and AFAP, and *MUTYH* (OMIM #604933) gene, for MAP.

For suspected patients, guidelines recommend offering genetic testing based on the number of polyps, with a threshold of more than 100 adenomatous polyps for FAP, and more than 10 or 20 adenomatous polyps (depending on the guideline) for AFAP and MAP^[6-10]. However, since polyps are not only caused by mutations in polyposis genes but are also intrinsic to age, the older the patient is, the more likely it is to detect polyps, lowering the probability of being a case of hereditary syndrome, especially when the polyp burden is low. Therefore, despite the selection of patients, germline multigene testing continues to have a high demand in laboratories, which decreases the rate of mutation detection, making these studies not cost-effective. Stanich *et al.*^[11] demonstrated that, on the one hand, the prevalence of mutations in adenomatous polyposis syndromes genes (*APC* and *MUTYH*) increases with the number of polyps developed, and on the other hand, older populations have a lower prevalence of finding significant mutations in these genes. More specifically, in those patients with a low number of polyps (<20), this prevalence is reduced below 2% from 50 years and onwards.

Consequently, age should also be included as a criterion for referring to genetic testing, helping the selection of patients, although very few guidelines include it. In this paper, we aim to improve genetic testing performance in suspected attenuated adenomatous

polyposis by establishing a probability calculator based on the number of polyps and age upon which recommend referring to genetic testing.

METHODS

Subjects

We conducted a retrospective analysis of all patients aged 18 years and older who were referred for genetic testing at Puerta de Hierro Hospital due to suspicion of attenuated adenomatous polyposis (AFAP or MAP) between 2015 and 2023 (N=138). The suspicion of attenuated polyposis was based on an accumulated history of 10 to 100 adenomatous polyps, according to the Community of Madrid (CAM) guidelines followed by our center¹². Patients with two or more hamartomatous polyps were excluded from the study, regardless of the number of adenomas. The presence of hamartomatous polyps suggests hamartomatous polyposis and excludes adenomatous polyposis⁶. The study received approval from the ethics committee of Puerta de Hierro Hospital (internal code: PI_48/24). Pre-test genetic counseling was conducted, and clinical consent for genetic testing was obtained.

Only pathogenic (P) (class 5) and likely pathogenic (LP) (class 4) variants in genes related to hereditary adenomatous polyposis (*APC* and *MUTYH*) were considered positive cases. Being a recessive gene, *MUTYH* variants were classified as positive cases only if found in homozygosity or compound heterozygosity. No variants detected, benign (class 1) and probably benign (class 2) variants, or monoallelic variants in *MUTYH* gene, were all considered negative cases.

Genetic testing

At Puerta de Hierro Hospital, germline DNA was extracted from peripheral blood using the Maxwell RSC whole blood DNA kit (Promega). Genetic testing was performed by massive sequencing (NGS) on a (MiSeq sequencer (Illumina) using the Hereditary Cancer Solution (HCS) kit (Sophia Genetics) and following the manufacturer's instructions. The panel included *APC* and *MUTYH* as relevant genes associated with adenomatous polyposis. Bioinformatic analysis was performed using the Sophia DDM-V4 (Sophia Genetics) data analysis platform. Relevant SNPs and indels were confirmed by Sanger sequencing. The reference sequences used to name variants were NM_001128425.2 for *MUTYH* and NM_000038.6 for *APC*.

Age and number of polyps

Age data is referred to as years of age at the moment of the genetic testing.

Number of polyps refers to the total accumulative polyps until genetic testing. Information related to colonoscopy examinations was collected with ENDOBSE® (Olympus Corporation, Tokyo, Japan). The program reflected all the data from the colonic examination: quality of the preparation, type of sedation, number of polyps, and location, along with other patient data relevant to the examination such as reason for endoscopic study and medication.

Histologically, polyps were classified into the following groups: adenomatous polyps, sub-classified in tubular, tubulovillous and villous, and non-adenomatous polyps, sub-classified in hyperplastic and serrated polyps. There were some reports that describe resected polyps just as “adenomatous” with no sub-classification, so they are here reported as “not classified” adenomatous polyps and were only considered in the adenomatous vs non-adenomatous polyps comparison and not in the subtype comparison.

External validation

In order to validate the findings, we utilized two independent cohort datasets (N=259). The datasets were obtained from 12 de Octubre University Hospital (n=162) and Infanta Leonor University Hospital (n=97).

As for methodology, the 12 de Octubre University Hospital, extracted DNA from whole blood using the Maxwell RSC Whole Blood kit (Promega). The Custom Hereditary Cancer Solution (CHCS) kit (Sophia Genetics) was employed for genetic testing, and software analysis was conducted using Sophia DDM-V4 (Sophia Genetics). Genes included in the sequencing kit were *APC*, *MUTYH*, *POLE*, *POLD1*, *AXN2* and *NTHL1*. Any pathogenic or likely pathogenic variants identified through massive sequencing were subsequently validated via Sanger sequencing.

Infanta Leonor Hospital utilized the QIAamp Blood DNA kit (QIAcube) for the extraction and purification of DNA from peripheral blood. Genetic testing was conducted by NGS on a MiSeq (Illumina) using the SureSelect QXT Target Enrichment (Agilent) kit for the coding region and flanking zones of the analyzed genes (*APC*, *MUTYH*, *POLE*, *POLD1*, *NTHL1*, *MSH3*). The bioinformatic analysis was carried out using custom-designed analysis pipelines, assisted by the SureCall and Alissa Interpreter software (Agilent). Sanger sequencing was employed to confirm relevant SNPs.

Statistics

The Shapiro-Wilk test was employed to assess normality. Non-normally distributed quantitative variables were presented as median along with the 25th (P25) and 75th

(P75) percentiles. For nonparametric comparisons, the Chi square test and Mann-Whitney test were used for categorical and quantitative variables respectively. Multivariable logistic regression (logit) was used to establish the association between having a positive result in the genetic test (dependent variable) and the age and number of adenomatous polyps. The model was internally validated through the `bsvalidation` command in Stata^[13]. This command performs an internal validation through calibration and discrimination. Resampling techniques were performed by bootstrapping, with 500 replications. To evaluate calibration, a calibration plot was generated, in which the quintiles of the observed and expected probabilities of having the event were graphically confronted. If calibration is perfect, the line between the two risks will lie on the main diagonal of the plot. The expected/observed (E/O) ratio will equal 1, the calibration in the large (CITL) will be 0 and the slope equal to 1. Discrimination is measured by the C-statistic, which is an analog of the AUC, with values ranging from 0.5 for no discrimination to 1.0 for perfect discrimination. The Brier scale (range 0-100) was also calculated as an overall performance measure, with high values indicating predictions closer to the actual outcome. It was obtained from the Brier score: $\text{Brier scaled} = 1 - \text{Brier score} / \text{Brier max}$.

For the external validation, the calibration plot assessed the calibration and the C-statistic, the discrimination. The Brier score is also shown, with a range between 0 and 1, being the lower value, the more accurate prediction.

From the model predicted probability, we pursued an optimal cutoff point with the maximal sensitivity and developed an online calculator available at <https://investigacionpuertadehierro.com/laboratorio-biopsia-liquida/>.

P value <0.05 was considered statistically significant.

Statistical analysis was carried out using MedCalc Statistical Software version 11.4.2.0 program (MedCalc Software bvba, Ostend, Belgium; <http://www.medcalc.org>; 2018), Stata v18 (StataCorp. 2023. *Stata Statistical Software: Release 18*. College Station, TX: StataCorp LLC.).

RESULTS

Of the total of patients included in the Puerta de Hierro cohort (N=138), 13 patients (9.4%) had a positive result in the genetic testing. From these, 11 patients had a P/LP variant in *MUTYH* gene, of which three were in homozygosis and eight in compound heterozygosis; two patients had a P variant in *APC* gene in heterozygosis (Supplementary figure 1). The most prevalent *MUTYH* mutation was c.1187G>A p.(Gly396Asp) (commonly known as G396D), present in the three homozygous cases

and half of compound heterozygous cases, followed by c.536A>G p.(Tyr179Cys) (commonly known as Y179C) mutation found in the other half of compound heterozygous cases. Other mutations identified in this biallelic group were c.1012C>T p.(Gln338Ter) in 3/8 cases, c.1227_1228dup p.(Glu410fs) in 2/8 cases, and c.933+3A>C, c.736G>T p.(Val246Phe) and c.1101dup in 1/8 each one. Table 1 shows detailed information on the positive cases.

Patient characteristics of positive and negative groups are shown in Table 2. According to sex, there were no statistical differences (odds ratio (OR): 0.91, 95% confidence interval (95%CI): 0.86 to 0.96, $P = 0.012$). Proportionally, development of CRC was similar between the two groups, with 4 (24.8%) CRC patients in the positive group and 31 (30.8%) in the negative group (OR: 1.35, 95%CI: 0.39 to 4.69, $P = 0.64$). Regarding family history (FH) of CRC, 43.2% of patients in the negative group had at least one family member with CRC while for the positive group, the percentage rose to 61.5%, although the difference did not reach statistical significance (OR: 2.04, 95%CI: 0.63 to 6.57, $P = 0.23$).

Parameters that did show significant differences between the negative and positive groups were age (OR: 0.91, 95%CI: 0.86 to 0.96, $P = 0.012$), adenoma number (OR: 1.08, 95%CI: 1.04 to 1.13, $P < 0.001$) and smoking (OR: 8.17, 95%CI: 1.97 to 33.8, $P = 0.004$).

Age comparison study

The youngest positive case was 28 years old, and the oldest one was 79 years old. For the negative group, ages ranged from 39 to 83 years old. The median age was 51 years in the positive group (percentile25-P75: 44 to 66), whereas the median age was 67 years (P25-P75: 61 to 72) in negative cases (OR: 0.91, 95%CI: 0.86 to 0.96; Figures 1 and 2). Among all negative cases, 74% were aged over 60, whereas the positive group had only five cases above that age.

Polyp comparison study

As previously mentioned, comparison data revealed a relation between the genetic test result and the number of adenomas. Both groups developed more adenomatous polyps than non-adenomatous polyps, but overall, the positive group developed significantly more adenomas (median: 35, P25-P75: 32 to 74) than the negative group (median: 22, P25-P75: 16 to 28) (OR: 1.08, 95%CI: 1.04 to 1.13) (Table 3). At the time of genetic testing, the majority of positive cases (85%) accumulated more than 30 adenomas, while only 24% of negative cases reached that threshold (Figure 3).

There were no significant differences according to the subtypes of polyps. For adenomatous subtypes, the most common one was tubular in the positive group as well as in the negative group (91% vs 89.9% respectively), followed far behind by tubulovillous (5.2% vs 6.1%) and villous polyp subtypes (0.4% vs 0.2%) ($P = 0.42$). For non-adenomatous polyps, the hyperplastic subtype was the most prevalent (83% vs 81%) and serrated polyps were less frequent in both groups (17% vs 19%). The graphic distribution of polyps is represented in Supplementary figure 2.

Calculator

Model development

Once all data was gathered, we aimed to establish a mathematical relation between the combination of age and number of adenomatous polyps, and the genetic test result. Using multivariable logistic regression, we estimated the probability of a patient having a positive genetic result based on their age and number of adenomas at genetic testing. The regression equation was: ***logit (genetic test (+)/1- genetic test (+)) = 0,3822 + (-0,0814*age in years) + (0,0731*number of adenomas)*** © IDIPHIM, (2024), All rights reserved. Overall model performance was ranked by a Brier score of 32.6%. Calibration scores were 1 for E/O ratio, 0 (95%CI: -0.73 to 0.73) for CITL, and 1 (95%CI: 0.56 to 1.44) for slope. Discrimination was assessed by an AUC of 0.924 (95%CI: 0.85 to 0.99; $P < 0.01$) (Figure 4).

The next step was to establish a cut-off point from the predicted probability model upon which to decide whether to refer patients to genetic testing or not. The requirement set to select this point was having a 100% sensibility with maximum specificity, so the false negative rate would be 0% but minimizing the number of false positives. These criteria were fulfilled at a probability of 3.5%, with a sensibility of 100% and a specificity of 58%. Applying the model retrospectively, it was found that 74 cases meeting the polyposis criteria of the CAM had a probability of a positive genetic test below 3.5%. This could have potentially led to savings amounting to 50,000€.

Internal validation

The internal validation was made by bootstrap resampling. The Brier score was 24.3% for overall model performance. In terms of calibration, the results obtained were an E/O ratio of 0.97 (95% CI: 0.57 to 1.38), CITL of 0.07 (95%CI: -0.8 to 1.01), and a calibration slope of 0.89 (95%CI: 0.39 to 1.51). C-statistic for discrimination was 0.9 (95%CI: 0.78 to 1). After the adjustment of the model by bootstrapping, the OR for age was 0.93 (95%CI: 0.88 to 0.98), and the OR for number of polyps was 1.07 (95%CI: 1.03 to 1.1).

External validation

The final validation was made by using data from other centers (N=259), located in the same geographical area. We gathered data on the number of polyps and age at the time of genetic testing, and the results of such test, classifying patients between “positive” (when genetic results revealed a mutation in genes related to polyposis) and “negative” (when no mutation related to polyposis was found). The general patient characteristics from external centers closely resembled those of our own, with a higher proportion of men than women and comparable median age and number of polyps (Table 4).

At 12 de Octubre Hospital (n=162), 11 patients were reported as positive. Of these, 4 cases presented *MUTYH* mutations (3 homozygous and 1 compound heterozygous), 4 cases carried *APC* mutations in heterozygosis and 1 case presented a heterozygous *POLD1* mutation. The model correctly predicted the positive result in 9 out of these 11 cases. The remaining two positive cases were predicted as negative. One patient, in their late 60s, had 19 adenomatous polyps and a heterozygous *AXIN2* gene mutation c.1994dup p.(Asn666fs). The second patient, in their early 70s, had 20 adenomatous polyps a homozygous *NTHL1* gene mutation c.268C>T p.(Gln90Ter). Calibration performance yielded a Hosmer-Lemeshow p-value of 0.45 (Table 5), while the discrimination study resulted in an AUC of 0.77 (95%CI: 0.61 to 0.93) (Figure 5).

At Infanta Leonor Hospital (n=97), 92 cases were negative, and 5 cases were positive. Mutations identified in these patients included three in *MUTYH* (1 homozygous and 2 compound heterozygous), one in *APC*, and one in *PTEN*. All 5 mutated patients were accurately predicted as positive. Performance evaluation indicated a Hosmer-Lemeshow p-value of 0.38 (Table 5) and an AUC of 0.90 (95%CI: 0.78 to 1) (Figure 5).

DISCUSSION

The diagnosis of Hereditary Polyposis Syndromes is of utmost importance due to the implications for not only patients but also their family relatives. This diagnosis begins with an oligopoliposis phenotype, and it is confirmed by genetic testing when a mutation in one of the implicated genes is found. Ensuring an accurate selection of patients for genetic testing is indispensable for the efficiency of public health systems. Since polyps are not only caused by genetic mutations but are also caused by aging, there is a phenotype overlap between hereditary polyposis and sporadic polyposis. The CAM implemented the *Prevecolon* program in 2017 for the early diagnosis of CRC through the fecal occult blood test (FOBT)^[14]. The campaign involves mass screening of asymptomatic individuals aged 50 to 69, who undergo the FOBT, and if positive, are referred for colonoscopy. This campaign has resulted in an increase in the detection of

asymptomatic polyposis, leading to a greater number of patients being referred for genetic consultation and consequently resulting in higher expenditure of both material and human resources^[15-18]. Therefore, this translates into a high rate of negative results, and thus, a low diagnostic yield. All of this emphasizes the need to implement new tools for better patient selection. To improve it, guidelines include other features such as personal or family history of CRC to help a better distinction between genetic and sporadic polyposis^[19-21]. Nowadays age is beginning to be included too, although very few guidelines do it and there is no consensus about the cut-off limit^[6,9,22]. Consequently, we sought to identify differences between hereditary and sporadic polyposis in order to improve the selection process. Out of that, we created a calculator based on the number of adenomas and age to help the health professionals make a better selection of patients and optimize the diagnostic yield of genetic testing.

In the Puerta de Hierro cohort, the prevalence of biallelic *MUTYH* mutations was 8% (11/138), and 1.5% (2/138) for *APC* mutations, in line with previous studies which ranged prevalence of these mutations in patients with oligopolyposis from 3% to 15% for *MUTYH* and from 2% to 9% for *APC*^[23-27].

Among the *MUTYH* pathogenic variants found, the most represented ones were G396D and Y179C. This is consistent with what has been previously found since most patients belonged to European population, in which these two variants are considered founder mutations^[24,28-31]. In the case of *APC* mutations, it has been reported that phenotype arising from alterations in this gene, varies depending on the location of the mutation. The ones found in this study fall into the regions of 5' end (codons 1 to 233) and exon 9 (codons 311 to 412), which have been associated with AFAP^[7,32-35].

Confrontation of other features between positive and negative group, showed no differences in sex, as described in other studies^[36-38]. Personal history of CRC and family history of CRC did not reach statistical significance between the two groups, demonstrating that the CRC risk for mutated patients in this study has been lowered due to the early diagnosis and prophylactic surgical strategies carried out (polypectomies and colectomies) that prevented developing CRC^[39-43].

In terms of tobacco consumption, the negative group exhibited higher rates of smoking and former smoking compared to the positive group (OR: 8.17, 95%IC: 1.97 to 33.8). Tobacco is a known carcinogen and has been linked to the development of polyposis^[44-46]. Our data imply that smoking was a significant contributing factor in sporadic polyposis cases.

Among the different polyps that can be present in general in polyposis, adenomatous polyps represent about two-thirds of all colonic polyps (being tubular adenomas the major representative of this group) followed by hyperplastic polyps as the second most common polyps. Tubulovillous and villous adenomas, as well as serrated polyps, can also be normally found but are less frequent^[38,47]. In this study, mutated patients developed in proportion more adenomas and much less hyperplastic polyps than not mutated cases. Interestingly, if the comparison was made separately among adenomatous subtypes on one side, and non-adenomatous subtypes on the other, there were no differences in the distribution. For both positive and negative groups, tubular polyp histology was the most common one among the adenomatous polyps, followed by tubulovillous polyps and villous polyps, while 80% of non-adenomatous polyps had hyperplastic histology. This demonstrates that altered genomics has more penetrance than sporadic predisposing factors (such as age) for the development of adenomatous polyps. However, mutations in polyposis-related genes do not influence the distribution of subtypes.

For age comparison, the median age for patients with an *APC* or biallelic *MUTYH* mutation was 51 years, significantly younger than those patients in the negative group (OR: 0.91, 95%CI: 0.86 to 0.96). Our results complement other studies that reached the same conclusion^[11,32,48-50]. This consolidates the use of age as a complementary criterion for referring to genetic test. In sum, our data stated that hereditary polyposis patients develop polyps at an earlier age and with a higher proportion of adenomas.

Using data on the number of adenomas, age, and genetic test results, we constructed a model to estimate the likelihood of detecting a polyposis mutation based on adenoma count and age. After evaluating various thresholds, we established the decision point at a 3.5% probability, ensuring 100% sensitivity and nearly 60% specificity. Had this model been applied to the patients of the Puerta de Hierro Hospital cohort, more than half of the genetic tests (53.6%) could have been saved, avoiding any missed positive cases and resulting in savings of 50,000€. External validation was conducted using data from two different hospitals within the same geographical area. This approach aimed to minimize potential confounding factors associated with variations in patient characteristics. Performance evaluation at both centers reported a Hosmer-Lemeshow p-value above 0.05, indicating no significant differences between observed and model-predicted values. In terms of discrimination, the area under the curve (AUC) was satisfactory for both centers. However, 12 de Octubre Hospital exhibited slightly poorer performance due to two positive cases being incorrectly predicted as negative by the model. Both cases involved elderly patients (over 65 years) with a low number of polyps

(19 and 20 adenomas). Guidelines^{6,8,9,10} are gradually shifting the adenoma count threshold for recommending genetic testing from 10 to 20, and those that include age, criteria are more restricted for patients over 60 years old. These two cases fell into a grey area, as one was below the 20-adenoma threshold and the other one was just in the limit with an advanced age. Consequently, depending on the guidelines applied, these two patients might not have met the requested criteria for genetic testing referral.

Our study has some limitations. The model is constructed on the basis of solely common adenomatous polyposis genes, *APC* and *MUTYH*, as positive cases. However, in recent years, new genes such as *POLE*, *POLD1*, *AXIN2*, and *NTHL1* have been associated with adenomatous polyposis. These genes are being incorporated into genetic panels but were not available at the time of the testing in our center, and therefore, were not included in the model. However, given that the phenotype is very common among the different genes, this limitation may not be significant. The model correctly predicted a *POLD1* carrier, while the *AXIN2* and *NTHL1* cases that were wrongly predicted were on the limit of the cutoff point and could potentially have been correctly predicted with data from the next colonoscopies.

CONCLUSION

In conclusion, hereditary polyposis syndromes present themselves at an early age and with a higher burden of adenomas than sporadic polyposis. Both features should be taken into consideration for selecting patients to refer to genetic testing. To ease the process, we developed a calculator that provides the probability of obtaining an informative genetic result based on these two characteristics. This will aid in deciding whether to proceed with genetic testing.

ACKNOWLEDGEMENTS

Patients and their caregivers

REFERENCES

- 1 **Kidambi TD**, Kohli DR, Samadder NJ, Singh A. Hereditary Polyposis Syndromes. *Curr Treat Options Gastroenterol* 2019; **17**: 650-665 [PMID: 31705372 DOI: 10.1007/s11938-019-00251-4]
- 2 **Byrne RM**, Tsikitis VL. Colorectal polyposis and inherited colorectal cancer syndromes. *Ann Gastroenterol* 2018; **31**: 24-34 [PMID: 29333064 DOI: 10.20524/aog.2017.0218]

- 3 **Talseth-Palmer BA**. The genetic basis of colonic adenomatous polyposis syndromes. *Hered Cancer Clin Pract* 2017; **15**: 5 [PMID: 28331556 DOI: 10.1186/s13053-017-0065-x]
- 4 **Chen L**, Ye L, Hu B. Hereditary Colorectal Cancer Syndromes: Molecular Genetics and Precision Medicine. *Biomedicines* 2022; **10**: 3207 [PMID: 36551963 DOI: 10.3390/biomedicines10123207]
- 5 **Nielsen M**, Hes FJ, Nagengast FM, Weiss MM, Mathus-Vliegen EM, Morreau H, Breuning MH, Wijnen JT, Tops CM, Vasen HF. Germline mutations in APC and MUTYH are responsible for the majority of families with attenuated familial adenomatous polyposis. *Clin Genet* 2007; **71**: 427-433 [PMID: 17489848 DOI: 10.1111/j.1399-0004.2007.00766.x]
- 6 **National Comprehensive Cancer Network**. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment: Colorectal (Version 2.2023). 2023. [cited January 16, 2024] Available from: http://www.nccn.org/professionals/physician_gls/pdf/bone.pdf
- 7 **Syngal S**, Brand RE, Church JM, Giardiello FM, Hampel HL, Burt RW; American College of Gastroenterology. ACG clinical guideline: Genetic testing and management of hereditary gastrointestinal cancer syndromes. *Am J Gastroenterol* 2015; **110**: 223-262 [PMID: 25645574 DOI: 10.1038/ajg.2014.435]
- 8 **Yang J**, Gurudu SR, Koptiuch C, Agrawal D, Buxbaum JL, Abbas Fehmi SM, Fishman DS, Khashab MA, Jamil LH, Jue TL, Law JK, Lee JK, Naveed M, Qumseya BJ, Sawhney MS, Thosani N, Wani SB, Samadder NJ. American Society for Gastrointestinal Endoscopy guideline on the role of endoscopy in familial adenomatous polyposis syndromes. *Gastrointest Endosc* 2020; **91**: 963-982.e2 [PMID: 32169282 DOI: 10.1016/j.gie.2020.01.028]
- 9 **Monahan KJ**, Bradshaw N, Dolwani S, Desouza B, Dunlop MG, East JE, Ilyas M, Kaur A, Laloo F, Latchford A, Rutter MD, Tomlinson I, Thomas HJW, Hill J; Hereditary CRC guidelines eDelphi consensus group. Guidelines for the management of hereditary colorectal cancer from the British Society of Gastroenterology (BSG)/Association of Coloproctology of Great Britain and Ireland (ACPGBI)/United Kingdom Cancer Genetics Group (UKCGG). *Gut* 2020; **69**: 411-444 [PMID: 31780574 DOI: 10.1136/gutjnl-2019-319915]
- 10 **Sociedad Española de Oncología Médica (SEOM)**. Cáncer Hereditario. 3ª Edición. 2019. [cited January 22, 2024] Available from: https://seom.org/images/Libro_Cancer_hereditario_2019.pdf
- 11 **Stanich PP**, Pearlman R, Hinton A, Gutierrez S, LaDuca H, Hampel H, Jasperson K. Prevalence of Germline Mutations in Polyposis and Colorectal Cancer-Associated

Genes in Patients With Multiple Colorectal Polyps. *Clin Gastroenterol Hepatol* 2019; **17**: 2008-2015.e3 [PMID: 30557735 DOI: 10.1016/j.cgh.2018.12.008]

12 Servicio Madrileño de Salud. Plan de Cáncer Familiar. Comunidad de Madrid. Madrid, 2015. [cited March 26, 2024] Available from: <https://www.comunidad.madrid/servicios/salud/programa-cancer-familiar#panel-43292>

13 **Fernandez-Felix BM**, García-Esquinas E, Muriel A, Royuela A, Zamora J. Bootstrap internal validation command for predictive logistic regression models. *The Stata Journal* 2021; **21**: 498-509 [DOI: 10.1177/1536867X211025836]

14 **Consejería de Sanidad Servicio Madrileño de Salud**. Prevecolon: prevención del cáncer de colon y recto [cited April 02, 2024] Available from: <https://www.comunidad.madrid/servicios/salud/prevecolon-prevencion-cancer-colon-recto>

15 **Oficina regional de coordinación oncológica**. Dirección general de coordinación de la asistencia sanitaria. Consejería de Sanidad. Memoria del programa de cribado de cáncer de colon y recto PREVECOLON. 2017. [cited April 02, 2024] Available from: https://www.comunidad.madrid/sites/default/files/doc/sanidad/asis/memoria_prevecolon_2017.pdf

16 **Consejería de Sanidad Servicio Madrileño de Salud**. Observatorio de resultados del Servicio Madrileño de Salud. Datos generales - Detección precoz de cáncer [cited April 15, 2024]. Available from: <http://observatorioresultados.sanidadmadrid.org/HospitalesDatosGeneralesTabla.aspx?ID=106>

17 **Levi Z**, Birkenfeld S, Vilkin A, Bar-Chana M, Lifshitz I, Chared M, Maoz E, Niv Y. A higher detection rate for colorectal cancer and advanced adenomatous polyp for screening with immunochemical fecal occult blood test than guaiac fecal occult blood test, despite lower compliance rate. A prospective, controlled, feasibility study. *Int J Cancer* 2011; **128**: 2415-2424 [PMID: 20658527 DOI: 10.1002/ijc.25574]

18 **Sanford KW**, McPherson RA. Fecal occult blood testing. *Clin Lab Med* 2009; **29**: 523-541 [PMID: 19840685 DOI: 10.1016/j.cll.2009.06.008]

19 **Kastrinos F**, Samadder NJ, Burt RW. Use of Family History and Genetic Testing to Determine Risk of Colorectal Cancer. *Gastroenterology* 2020; **158**: 389-403 [PMID: 31759928 DOI: 10.1053/j.gastro.2019.11.029]

20 **Armelaio F**, de Pretis G. Familial colorectal cancer: a review. *World J Gastroenterol* 2014; **20**: 9292-9298 [PMID: 25071323 DOI: 10.3748/wjg.v20.i28.9292]

21 **Basso G**, Bianchi P, Malesci A, Laghi L. Hereditary or sporadic polyposis syndromes. *Best Pract Res Clin Gastroenterol* 2017; **31**: 409-417 [PMID: 28842050 DOI: 10.1016/j.bpg.2017.05.011]

- 22 **Cubiella J**, Marzo-Castillejo M, Mascort-Roca JJ, Amador-Romero FJ, Bellas-Beceiro B, Clofent-Vilaplana J, Carballal S, Ferrándiz-Santos J, Gimeno-García AZ, Jover R, Mangas-Sanjuán C, Moreira L, Pellisè M, Quintero E, Rodríguez-Camacho E, Vega-Villaamil P; Sociedad Española de Medicina de Familia y Comunitaria y Asociación Española de Gastroenterología. Clinical practice guideline. Diagnosis and prevention of colorectal cancer. 2018 Update. *Gastroenterol Hepatol* 2018; **41**: 585-596 [PMID: 30245076 DOI: 10.1016/j.gastrohep.2018.07.012]
- 23 **Grover S**, Kastrinos F, Steyerberg EW, Cook EF, Dewanwala A, Burbidge LA, Wenstrup RJ, Syngal S. Prevalence and phenotypes of APC and MUTYH mutations in patients with multiple colorectal adenomas. *JAMA* 2012; **308**: 485-492 [PMID: 22851115 DOI: 10.1001/jama.2012.8780]
- 24 **Guarinos C**, Juárez M, Egoavil C, Rodríguez-Soler M, Pérez-Carbonell L, Salas R, Cubiella J, Rodríguez-Moranta F, de-Castro L, Bujanda L, Serradesanferm A, Nicolás-Pérez D, Herráiz M, Fernández-Bañares F, Herreros-de-Tejada A, Aguirre E, Balmaña J, Rincón ML, Pizarro A, Polo-Ortiz F, Castillejo A, Alenda C, Payá A, Soto JL, Jover R. Prevalence and characteristics of MUTYH-associated polyposis in patients with multiple adenomatous and serrated polyps. *Clin Cancer Res* 2014; **20**: 1158-1168 [PMID: 24470512 DOI: 10.1158/1078-0432.CCR-13-149]
- 25 **Morak M**, Laner A, Bacher U, Keiling C, Holinski-Feder E. MUTYH-associated polyposis - variability of the clinical phenotype in patients with biallelic and monoallelic MUTYH mutations and report on novel mutations. *Clin Genet* 2010; **78**: 353-363 [PMID: 20618354 DOI: 10.1111/j.1399-0004.2010.01478.x]
- 26 **Landon M**, Ceulemans S, Saraiya DS, Strike B, Arnell C, Burbidge LA, Moyes K, Theisen A, Fernandes PH, Ji JQ, Abbott B, Kaldate RR, Roa B. Analysis of current testing practices for biallelic MUTYH mutations in MUTYH-associated polyposis. *Clin Genet* 2015; **87**: 368-372 [PMID: 24620956 DOI: 10.1111/cge.12375]
- 27 **Mak S**, Alexander JL, Clark SK, Hawkins M, Cuthill V, Latchford A, Monahan KJ. The Diagnostic Yield of Genetic Testing in Patients With Multiple Colorectal Adenomas: A Specialist Center Cohort Study. *Clin Transl Gastroenterol* 2024; **15**:e00645 [PMID: 37856205 DOI: 10.14309/ctg.0000000000000645]
- 28 **Aretz S**, Tricarico R, Papi L, Spier I, Pin E, Horpaopan S, Cordisco EL, Pedroni M, Stienen D, Gentile A, Panza A, Piepoli A, de Leon MP, Friedl W, Viel A, Genuardi M. MUTYH-associated polyposis (MAP): evidence for the origin of the common European mutations p.Tyr179Cys and p.Gly396Asp by founder events. *Eur J Hum Genet* 2014; **22**: 923-929 [PMID: 23361220 DOI: 10.1038/ejhg.2012.309]

- 29 **Barreiro RAS**, Sabbaga J, Rossi BM, Achatz MIW, Bettoni F, Camargo AA, Asprino PF, A F Galante P. Monoallelic deleterious MUTYH germline variants as a driver for tumorigenesis. *J Pathol* 2022; **256**: 214-222 [PMID: 34816434 DOI: 10.1002/path.5829]
- 30 **Paller CJ**, Tukachinsky H, Maertens A, Decker B, Sampson JR, Cheadle JP, Antonarakis ES. Pan-Cancer Interrogation of MUTYH Variants Reveals Biallelic Inactivation and Defective Base Excision Repair Across a Spectrum of Solid Tumors. *JCO Precis Oncol* 2024; **8**: e2300251 [PMID: 38394468 DOI: 10.1200/PO.23.00251]
- 31 **Sampson JR**, Jones N. MUTYH-associated polyposis. *Best Pract Res Clin Gastroenterol* 2009; **23**: 209-218 [PMID: 19414147 DOI: 10.1016/j.bpg.2009.03.006]
- 32 **PDQ Cancer Genetics Editorial Board**. Genetics of Colorectal Cancer (PDQ®): Health Professional Version. 2024 Feb 2. In: PDQ Cancer Information Summaries. Bethesda (MD): National Cancer Institute (US); 2002 [PMID: 26389505]
- 33 **Ibrahim A**, Barnes DR, Dunlop J, Barrowdale D, Antoniou AC, Berg JN. Attenuated familial adenomatous polyposis manifests as autosomal dominant late-onset colorectal cancer. *Eur J Hum Genet* 2014; **22**: 1330-1333 [PMID: 24549056 DOI: 10.1038/ejhg.2014.20]
- 34 **Nieuwenhuis MH**, Vasen HF. Correlations between mutation site in APC and phenotype of familial adenomatous polyposis (FAP): a review of the literature. *Crit Rev Oncol Hematol* 2007; **61**: 153-161 [PMID: 17064931 DOI: 10.1016/j.critrevonc.2006.07.004]
- 35 **Knudsen AL**, Bisgaard ML, Bülow S. Attenuated familial adenomatous polyposis (AFAP). A review of the literature. *Fam Cancer* 2003; **2**: 43-55 [PMID: 14574166 DOI: 10.1023/a:1023286520725]
- 36 **Qumseya BJ**, Coe S, Wallace MB. The effect of polyp location and patient gender on the presence of dysplasia in colonic polyps. *Clin Transl Gastroenterol* 2012; **3**: e20 [PMID: 23238292 DOI: 10.1038/ctg.2012.14]
- 37 **McCashland TM**, Brand R, Lyden E, de Garmo P; CORI Research Project. Gender differences in colorectal polyps and tumors. *Am J Gastroenterol* 2001; **96**: 882-886 [PMID: 11280569 DOI: 10.1111/j.1572-0241.2001.3638_a.x]
- 38 **Kazem-Shahmoradi M**, Soleimaninejad M, Sharifian M. Evaluation of colonoscopy data for colorectal polyps and associated histopathological findings. *Ann Med Surg (Lond)* 2020; **57**: 7-1 [PMID: 32685144 DOI: 10.1016/j.amsu.2020.07.010]
- 39 **Kalady MF**, Church JM. Prophylactic colectomy: Rationale, indications, and approach. *J Surg Oncol* 2015; **111**: 112-117 [PMID: 25418116 DOI: 10.1002/jso.23820]
- 40 **Vogelsang HE**. Prophylactic Surgery and Extended Oncologic Radicality in Gastric and Colorectal Hereditary Cancer Syndromes. *Visc Med* 2019; **35**: 231-239 [PMID: 31602384 DOI: 10.1159/000501919]

- 41 **Cross AJ**, Robbins EC, Pack K, Stenson I, Patel B, Rutter MD, Veitch AM, Saunders BP, Duffy SW, Wooldrage K. Colorectal cancer risk following polypectomy in a multicentre, retrospective, cohort study: an evaluation of the 2020 UK post-polypectomy surveillance guidelines. *Gut* 2021; **70**: 2307-2320 [PMID: 33674342 DOI: 10.1136/gutjnl-2020-323411]
- 42 **Stoffel EM**, Boland CR. Genetics and Genetic Testing in Hereditary Colorectal Cancer. *Gastroenterology* 2015; **149**: 1191-1203.e2 [PMID: 26226567 DOI: 10.1053/j.gastro.2015.07.021]
- 43 **Grady WM**. Genetic testing for high-risk colon cancer patients. *Gastroenterology* 2003; **124**: 1574-1594 [PMID: 12761718 DOI: 10.1016/s0016-5085(03)00376-7]
- 44 **Yoshida N**, Ishikawa H, Eguchi H, Okazaki Y, Hirose R, Inoue K, Dohi O, Itoh Y, Mutoh M, Ishiguro S, Ishida H. Promotion Effects of Smoking in Polyp Development in Monozygotic Twins with Atypical Colorectal Polyposis. *Case Rep Gastroenterol* 2022; **16**: 375-381 [PMID: 35949244 DOI: 10.1159/000524944]
- 45 **Botteri E**, Iodice S, Raimondi S, Maisonneuve P, Lowenfels AB. Cigarette smoking and adenomatous polyps: a meta-analysis. *Gastroenterology* 2008; **134**: 388-395 [PMID: 18242207 DOI: 10.1053/j.gastro.2007.11.007]
- 46 **Ji BT**, Weissfeld JL, Chow WH, Huang WY, Schoen RE, Hayes RB. Tobacco smoking and colorectal hyperplastic and adenomatous polyps. *Cancer Epidemiol Biomarkers Prev*. 2006; **15**: 897-901 [PMID: 16702367 DOI: 10.1158/1055-9965.EPI-05-0883]
- 47 **Eshghi MJ**, Fatemi R, Hashemy A, Aldulaimi D, Khodadoostan M. A retrospective study of patients with colorectal polyps. *Gastroenterol Hepatol Bed Bench* 2011; **4**: 17-22 [PMID: 24834150]
- 48 **Terlouw D**, Suerink M, Singh SS, Gille HJJP, Hes FJ, Langers AMJ, Morreau H, Vasen HFA, Vos YJ, van Wezel T, Tops CM, Ten Broeke SW, Nielsen M. Declining detection rates for APC and biallelic MUTYH variants in polyposis patients, implications for DNA testing policy. *Eur J Hum Genet* 2020; **28**: 222-230 [PMID: 31527860 DOI: 10.1038/s41431-019-0509-z]
- 49 **Valle L**, Vilar E, Tavtigian SV, Stoffel EM. Genetic predisposition to colorectal cancer: syndromes, genes, classification of genetic variants and implications for precision medicine. *J Pathol* 2019; **247**: 574-588 [PMID: 30584801 DOI: 10.1002/path.5229]
- 50 **Aelvoet AS**, Buttitta F, Ricciardiello L, Dekker E. Management of familial adenomatous polyposis and MUTYH-associated polyposis; new insights. *Best Pract Res Clin Gastroenterol* 2022; **58-59**: 101793 [PMID: 35988966 DOI: 10.1016/j.bpg.2022.101793]

Footnotes

Institutional review board statement: The study received approval from the ethics committee of Puerta de Hierro University Hospital, internal code: PI_48/24.

Informed consent statement: written informed consent for data publication was obtained from patients.

Conflict-of-interest statement

The authors have no conflicts of interest to declare.

Figures and Tables

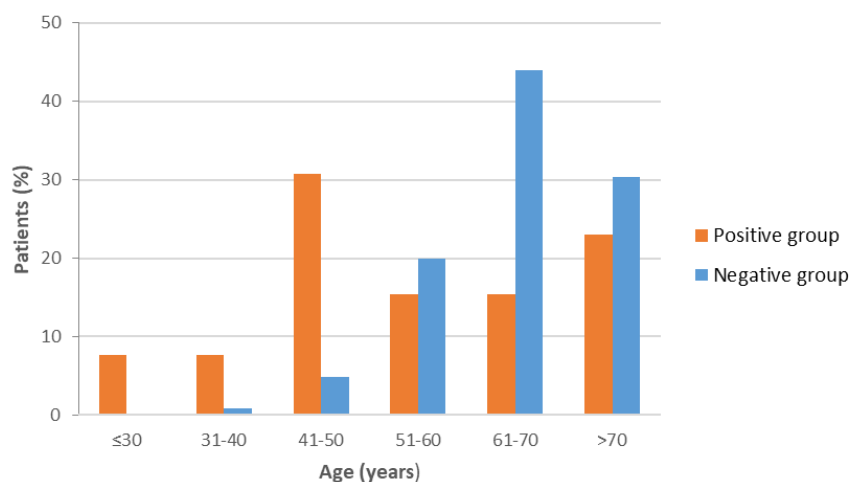


Figure 1. Distribution of ages (by decades of years) at which genetic testing was performed.

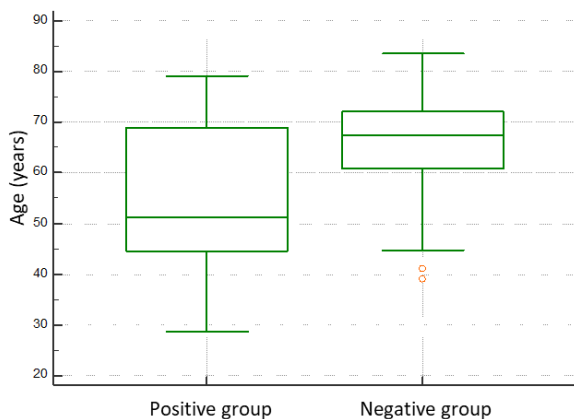


Figure 2. Box and whiskers plot of age at genetic testing for negative and positive

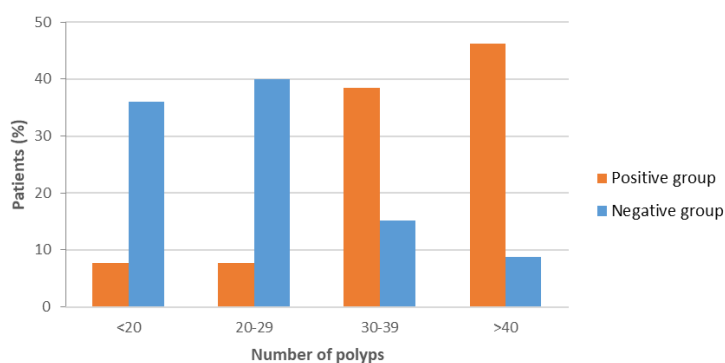


Figure 3. Distribution of number of adenomatous polyps by group.

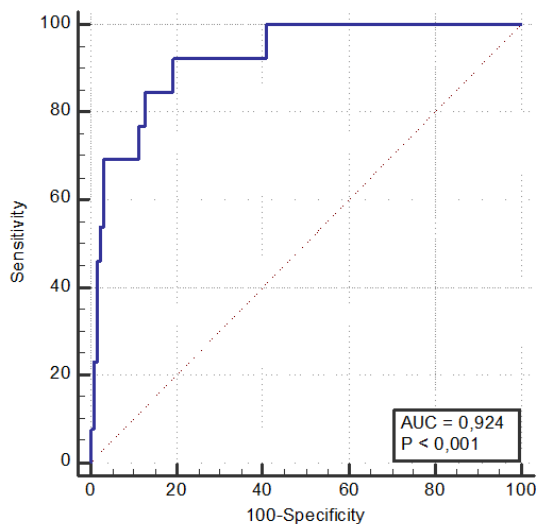


Figure 4. AUC for probability of having a positive genetic test.

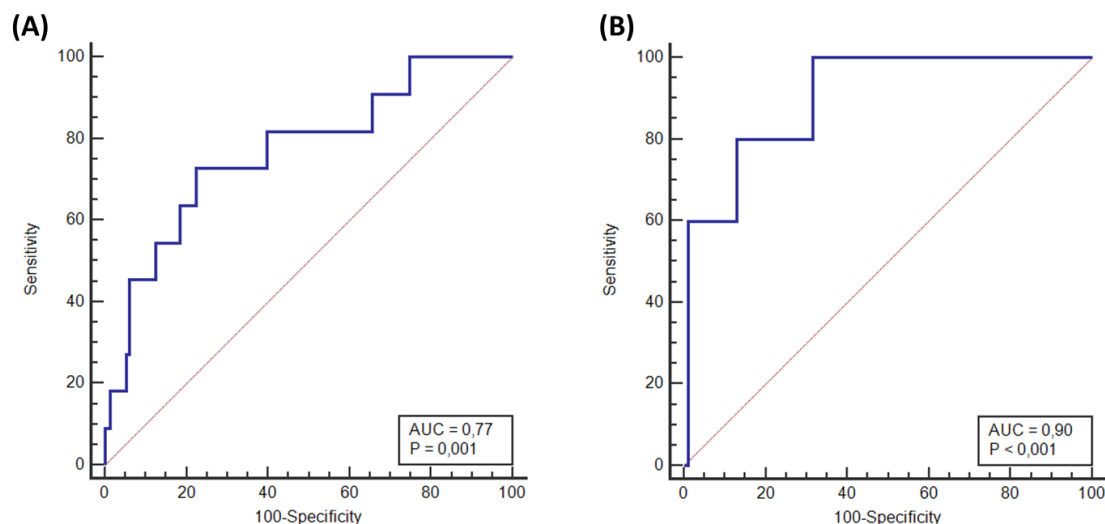


Figure 5. AUC resulting from applying the model to external validation databases. (A) AUC for 12 de Octubre Hospital, and (B) AUC for Infanta Leonor Hospital

Patient	Range age at GT (years)*	Gene	Variant	Heterozygosity	CRC (Range age diagnosis)*	FH CRC
Case 1	70-75	<i>MUTYH</i>	c.1187G>A	Homozygous	YES (55-60)	NO
Case 2	25-30	<i>MUTYH</i>	c.536A>G + c.933+3A>C	Compound heterozygous	NO	YES
Case 3	45-50	<i>MUTYH</i>	c.1012C>T + c.536A>G	Compound heterozygous	NO	YES
Case 4	40-45	<i>MUTYH</i>	c.1012C>T + c.536A>G	Compound heterozygous	NO	YES
Case 5	45-50	<i>MUTYH</i>	c.1012C>T + c.536A>G	Compound heterozygous	NO	YES
Case 6	60-65	<i>MUTYH</i>	c.1187G>A + c.736G>T	Compound heterozygous	YES (55-60)	YES
Case 7	65-70	<i>MUTYH</i>	c.1187G>A	Homozygous	YES (55-60)	YES
Case 8	75-80	<i>MUTYH</i>	c.1187G>A	Homozygous	NO	YES
Case 9	50-55	<i>MUTYH</i>	c.1187G >A + c.1227_1228dup	Compound heterozygous	NO	NO
Case 10	50-55	<i>MUTYH</i>	c.1187G >A + c.1101dup	Compound heterozygous	NO	NO

Case 11	40-45	<i>MUTYH</i>	c.1187G >A + c.1227_1228dup	Compound heterozygous	YES (40-45)	NO
Case 12	35-40	<i>APC</i>	c.697C>T	Heterozygous	NO	YES
Case 13	75-80	<i>APC</i>	c.423G >C	Heterozygous	NO	NO

Table 1. Characteristics of positive group patients. GT: genetic testing; CRC: colorectal cancer; FH: family history. * The exact age data is not shown due to editorial request; it will be available upon journal publication.

Patients' characteristics	Positive group n=13	Negative group n=125	P value
Age (years) median (P25-P75)	51 (44 - 66)	67 (61 - 72)	0.012
Sex			0.51
Women, n (%)	6 (46.2)	46 (36.8)	
Men, n (%)	7 (53.8)	79 (63.2)	
Polyp type	701 (100)	3685 (100)	<0.001
Adenomatous, n (%)	677 (96.6)	2916 (79.1)	
Non-adenomatous, n (%)	24 (3.4)	769 (20.9)	
CRC, n (%)	4 (30.8)	31 (24.8)	0.64
FH CRC, n (%)	8 (61.5)	55 (44)	0.23
Smoking, n (%)			0.004
Yes	2 (16.6)	42 (33.3)	
Former	1 (5.6)	49 (39.7)	
No	7 (55.6)	26 (20.6)	
ND	3 (22.2)	8 (6.4)	

Table 2. Patients' characteristic for positive and negative group. CRC: colorectal cancer; FH CRC: family history of colorectal cancer; ND: no data.

Polyp subtype	Positive group n=701	Negative group n=3685	P value
Adenomatous, n (%)	677 (96.6)	2916 (79.1)	0.42
Tubular	616 (91)	2620 (89.9)	
Tubulovillous	35 (5.2)	179 (6.1)	
Villous	3 (0.4)	7 (0.2)	
Not sub- classified	23 (3.4)	110 (3.8)	
Non- adenomatous, n (%)	24 (3.4)	769 (20.9)	0.81
Hyperplastic	20 (83)	626 (81)	
Serrated	4 (17)	143 (19)	

Table 3. Distribution of adenomatous and non-adenomatous polyps by subtype.

P value was calculated comparing only tubular, tubulovillous and villous polyps, leaving out the not sub-classified polyps.

Patients characteristics	Puerta de Hierro (N=138)	12 de Octubre (n=162)	Infanta Leonor (n=97)
Sex			
Women, n (%)	52 (37.7)	43 (26.5)	23 (23.7)
Men, n (%)	86 (62.3)	119 (73.5)	74 (76.3)
Age, median (P25-P75)	67 (60-72)	66 (59-71)	67 (60-73)
Nº of polyps, median (P25-P75)	23 (16-31)	21 (15-29)	24 (19.8-32)

Table 4. Patients' characteristics for study center and external validation centers.

Hospital	Patients			Model prediction		Performance measures	
	All n	Positive n, (%)	Negative n, (%)	TP n, (%)	TN n, (%)	Hosmer-Lemeshow p-value	AUC
12 de Octubre	162	11 (6.8)	151 (93.2)	9 (81.8)	75 (49.7)	0.45	0.77
Infanta Leonor	97	5 (5.2)	92 (94.8)	5 (100)	49 (53.3)	0.38	0.90

Table 5. Perform results for external validation data