

Clinical Diagnosis Nomogram of an Early Phase of the Sporadic Carcinoma of Colon (I and Ila Dukes)

Juan Luis Méndez Mora^{1*}, Villadiego-Sánchez JM², Praena-Fernández JM³, Vilches-Arenas A⁴, Santos-Lozano JM⁵ and Ortega-Calvo M⁶

¹Specialist in General Surgery and Digestive Diseases, Andalusian Health Service, Spain

²Specialist in Family and Community Medicine, Research Coordinator EPES-061, Seville, Spain

³Statistic, Andalusian Public Foundation for Health Research Management, Seville (FISEVI), Statistics, Methodology and Research Evaluation Unit, Spain

⁴Professor, Department of Preventive Medicine, Sevilla University, Spain

⁵Family Medicine Specialist, San Pablo University Health Center, CIBER of the Pathophysiology of Obesity and Nutrition, ISCIII, Associate Professor of the Department of Medicine, Sevilla University, Spain

⁶Family Medicine Specialist, Las Palmeritas Health Center, Head of Research DSAP Sevilla, CIBER of the Pathophysiology of Obesity and Nutrition, ISCIII, Spain

***Corresponding Author:** Juan Luis Méndez Mora, Specialist in General Surgery and Digestive Diseases, Andalusian Health Service, Spain.

Received: January 23, 2020; **Published:** March 30, 2020

Abstract

Introduction: sporadic colon carcinoma is one of the most frequently diagnosed malignancies, reaching almost 10% of all cancer cases. Its early diagnosis should be a health objective at both primary health care as hospital services. Nomography is a graphical tool to provide scientific support to healthcare decisions.

Methodology: On a case-control design we obtained a diagnostic predictive model of sporadic colon neoplasia in stages I and Ila Dukes. Using binary logistic regression, we constructed a nomogram using the R program.

Results: It contains three predicting variables (age, total cholesterol and CA 19.9). Its calibration graph shows a 0.176 Brier index, Brier scaled index of 0.158 and an area under the curve of 0.749.

Discussion: The values obtained from the calibration of the nomogram reported in this paper estimate that gives a predictive level and an acceptable discriminative power.

Keywords: Colonic Neoplasm; Logistic Models; CA 19.9; Early Diagnosis; Early Detection of Cancer; Calibration; Evidence-Based -Medicine

Introduction

Sporadic colon carcinoma is one of the most frequently diagnosed malignancies in the human species, reaching almost 10% of all cancer cases [1]. From a population point of view there seems to be a diverse spectrum of risk [2]. In 2008 it caused 600,000 deaths world-

wide. Its early diagnosis seems to be a health objective both at the level of primary care and general hospital services. Early diagnosis of sporadic colon neoplasia does not lead to greater and better survival according to some researchers [3] and according to others [4]. On the other hand, when screening programs are carried out, cases are often misled at an early stage (Dukes A and B) [5].

Nomography is increasingly used in the clinic because it seems to improve communication between the doctor and the patient, giving the former a useful tool for answering [6] clinical questions. It is an important phase of the development of predictive models. Although there are also voices against its excessive use [7]. In any case, it is a graphic instrument to give scientific support to health decisions [8].

Objective of the Study

Our objective in this article has been the development and construction of a diagnostic nomogram of clinical utility with the data of a predictive model of sporadic colorectal carcinoma (SCC) in the early clinical phase [9,10].

Materials and Methods

The construction of the nomogram has been carried out based on the data obtained in a case-control study carried out on patients with sporadic colorectal malignancies (cases) and healthy or affected patients with other non-oncological diseases studied both at the hospital level and in care. primary (controls). All cases had a diagnostic biopsy and were in stages I and IIa of Dukes in the presurgical phase. The inclusion criteria for the cases were an endoscopy with positive biopsy. The exclusion criteria were the existence of metastases, severe dyslipidemia, coexistence with other neoplasia, hereditary polyposis syndrome, hereditary non-polypoid colorectal cancer, inflammatory bowel disease, non-epithelial neoplasms and immunodeficiencies.

In the controls, the inclusion criterion was the absence of colorectal neoplasia. The exclusion criteria were in turn: any type of neoplastic disease, existence of pre-neoplastic colorectal lesions, severe dyslipidemia and immunodeficiency. Controls were not performed colonoscopy or opaque enema. There was a telephone check two years after the selection to rule out neoplastic disease at any level.

The mathematical model is published in the year 2009 [10] and it included three predictive variables (patient age, total plasma cholesterol level and CA 19.9 level) and an interaction (age x CA 19.9). By criteria of clinical utility we have constructed the nomogram with the model that did not include the interaction.

The same data package from the original investigation was used, but this time the statistical program used was R [11]. The nomogram was built with the “rms” package and the “val.prob.ci” function belonging to the R [12] library. We build the ROC curve of the model with the R Commander package (RcmdrPlugin.ROC).

Results

The origin of the controls is shown in table 1. The nomogram is shown in figure 1 and the calibration chart made with the “rms” package is in figure 2.

Figure 3 shows the ROC curve together with the specific confidence intervals of the binary logistic regression model.

	Primary Care Controls	Hospital Controls	Cases of CCE
Number	128	147	126
Percentage	31.92	36.66	31.42

Table 1: Total number of cases and distribution of controls.

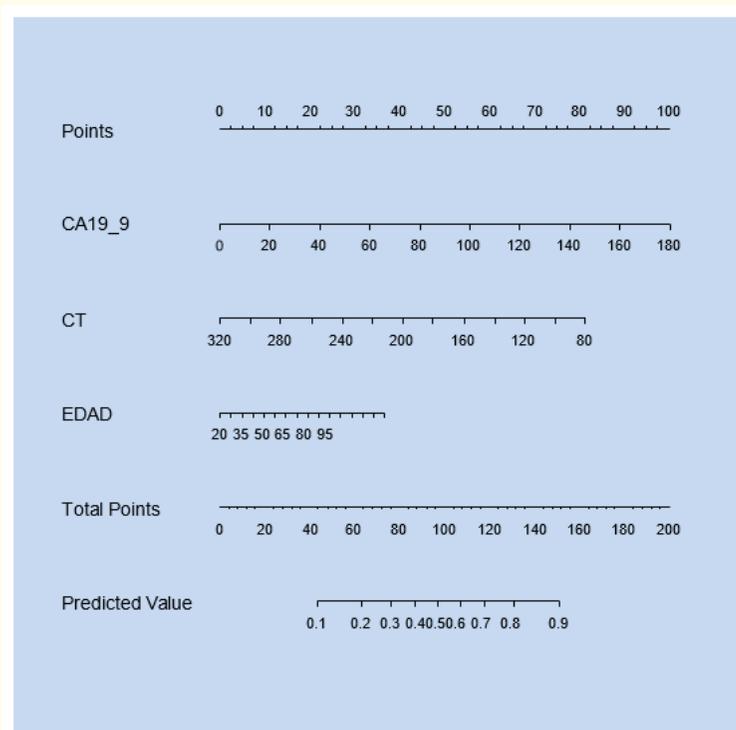


Figure 1: Nomogram for the clinical diagnosis of CEC in phases I and IIa of Dukes.

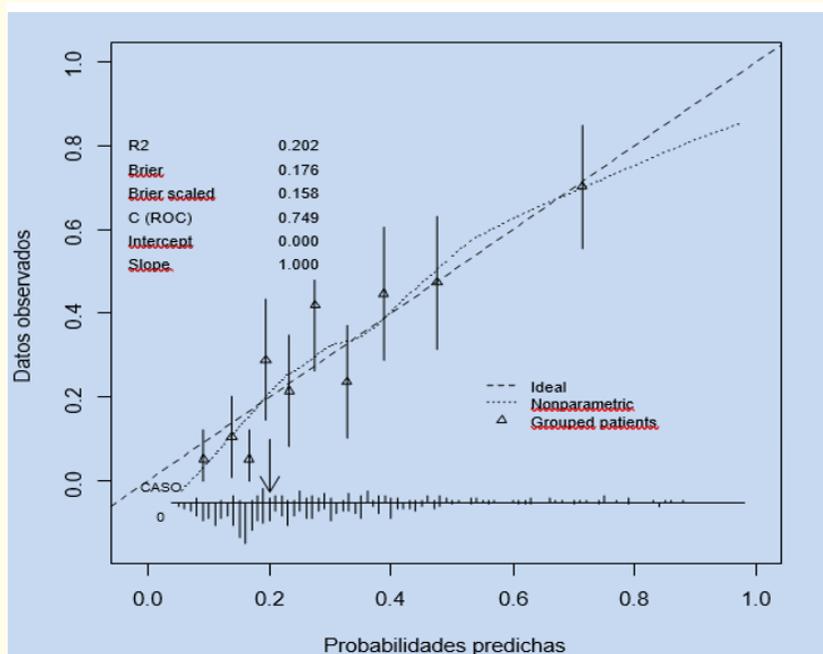


Figure 2: Calibration graph of the nomogram of figure 1.
 R2: R2 from Nagelkerke; Brier: Brier index; Brier scaled: Brier index scaled; C (ROC): Concordance index/area under the ROC curve; Intercept: Constant; Slope: Pending.

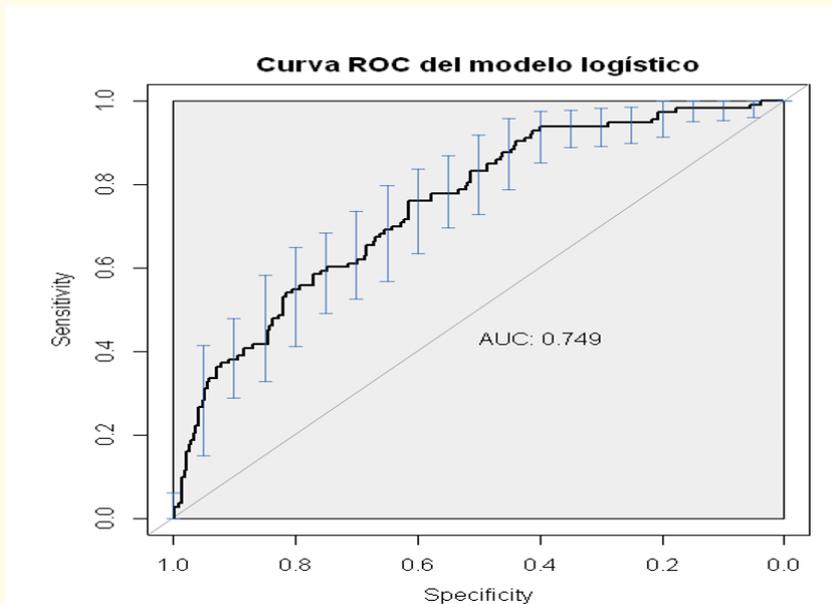


Figure 3: ROC curve of the binary logistic model of clinical diagnosis of CCE 10. AUC: Area under the ROC curve.

Discussion and Conclusion

Without entering into the methodological question of whether nomograms are a setback or an advance in statistical and epidemiological research, we have prepared this because we understand that it can be a useful tool to manage the uncertainty of the diagnosis of sporadic colon neoplasia in early stages. Given that a diverse risk spectrum exists from a population point of view [2], we understand that this nomogram can be useful in both primary and specialized care.

In the initial stages of our investigation [9] the variable CEA (pre-surgical carcinoembryonic antigen) was not statistically significant and we decided not to continue collecting it. The carcinoembryonic antigen has more diagnostic utility in the postoperative follow-up of patients than in the diagnosis of the early stages of the disease as in this case [14]. We could have imputed values to the records of those we did not own, but we gave up on it and analyzed only the values actually observed.

On the other hand, CA 19.9, a carbohydrate antigen also called Lewis sialized antigen, has value in any diagnostic phase of sporadic colon neoplasia and is the one we have exploited in this work. The highest values have been described in a pancreatic tail neoplasm with liver metastases [15].

Given the epidemiological utility of the product variable (CA 19.9 x Age) defined in our article [10], we did not want to miss the opportunity to use it by name. But the results we obtained in a first analysis were not admissible because the expected probability line ("predicted value") was very small and not manageable for the clinician. We decided to construct the nomogram with the natural logarithms of the variables that made up the model with interaction and we also did not obtain very adequate results from a visual point of view (the predicted probability line was still small and of little use, so we finally stopped including the interaction variable).

The variable total cholesterol is very interesting in this predictive model because its reduction is in line with the existence of a general cancerous syndrome at an early stage because it decreases as the diagnostic probability of SCC increases. Although from the princeps publication of Rose, *et al.* [16] there is disparate evidence on this issue [17].

We believe that this nomogram is of clinical utility both in specialty consultations (Digestive, Medical Oncology) and family medicine (Primary Care) because the controls were collected at both levels thus giving representativeness to the sample. Also, with a size of around 400 elements, we estimate that the model has adequate internal functioning, by including nothing more than three significant predictors [18].

Calibration graphs are necessary when creating a nomogram [19,20] although in our case it is not a recurrence calibration but a diagnostic calibration. The graphic functions of the R program provide a very suitable solution for the validation of a model like this. The fact that the smoothing line in figure 2 is very close to the diagonal graphically exposes the acceptable calibration level of this model. The area under the ROC curve of our model (~ 0.75) is a very interesting value and supports the clinical utility of the nomogram (acceptable discriminative level). In predictive models with dichotomous outcome variables, the concordance statistic C is equivalent to the area under the ROC curve.

Nagelkerke R [2] is similar to the coefficient of determination in linear regression. Quantify the percentage of variability explained by the predictive variables of the adjusted model. The Brier index measures the difference between what is observed and what is predicted on a quadratic scale. The closer to 0 the better predictive ability the model has. At lower value, better predictions. The Brier index of our model is 0.17.

The Brier index began to be used in the analysis of weather forecasts. Nowadays it is used more and more frequently in the health literature [21].

It depends on the prevalence of the disease (the a priori probability) in the population on which the predictive model is built. It is therefore suitable for population comparisons. There is a standardized form that is the scaled Brier index [22] that ranges from 0 to 1 and that is independent of prevalence.

The scaled Brier index ranges from 0 to 1. The scaled Brier of our model, 0.158, we believe that it also gives clinical utility [23,24] in predictive terms. In any case, this nomogram is believed to be useful in the case-finding technique and not in the screening technique. Our intention has not been to create an overdiagnosis instrument, but to develop a method to manage clinical uncertainty [25].

For example, a primary care physician who performs a diagnostic evaluation of a 76-year-old male (non-predictive variable) (predictor variable) with a total plasma cholesterol of 137 (predictor variable) and a CA 19.9 of 43 (predictor variable), controlling for the rest of the variables, you would get a score of: Age = ~ 27 points, Plasma cholesterol = ~ 62 points and Ca 19.9 = ~ 24 points.

Adding all these scores on the bottom line of the nomogram we obtain a total score of 113 points. The perpendicular of that score corresponds approximately to a predicted probability of 0.67 (67%) (Figure 1). The biological profile described corresponds to our registry No. 43, which was a case of colon adenocarcinoma.

In summary, we developed and described a diagnostic nomogram of early Dukes CCE (I and IIb) based on a case-control study conducted previously. The Brier index and the area under the ROC curve obtained make it useful, both at a predictive and discriminative level, in clinical consultations of primary and specialized care.

Acknowledgement

The researchers thank Prof. Emilio Sánchez Catalejo for his teaching at the Andalusian School of Public Health in Granada (EASP).

Author's Contribution

JLMM, JMVS and MOC conceived and designed the study that was part of the Doctoral Thesis of the first two. JLMM, JMVS, AVA, JMPF, JMSL and MOC collected the data. All authors analyzed and interpreted the results. All authors participated in the drafting, revision and approval of the submitted manuscript.

Bibliography

1. Win AK, *et al.* "Risk prediction models for colorectal cancer: a review". *Cancer Epidemiology, Biomarkers and Prevention* 21.3 (2012): 398-410.
2. Jemal A, *et al.* "Global patterns of cancer incidence and mortality rates and trends". *Cancer Epidemiology, Biomarkers and Prevention* 19.8 (2010): 1893-907.
3. Porta M, *et al.* "Influence of "diagnostic delay" upon cancer survival: an analysis of five tumour sites". *Journal of Epidemiology and Community Health* 45.3 (1991): 225-230.
4. Tørring ML, *et al.* "Diagnostic interval and mortality in colorectal cancer: U-shaped association demonstrated for three different datasets". *Journal of Clinical Epidemiology* 65.6 (2012): 669-678.
5. Logan RF, *et al.* "Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests". *Gut* 61.10 (2012): 1439-1446.
6. Kattan MW and Marasco J. "What is a real nomogram?" *Seminars in Oncology* 37.1 (2010): 23-26.
7. Grimes D A. "The nomogram epidemic: resurgence of a medical relic". *Annals of Internal Medicine* 149.4 (2008): 273-275.
8. Grimes DA and Schulz KF. "Refining clinical diagnosis with likelihood ratios". *Lancet* 365.9469 (2005): 1500-1505.
9. Méndez-Mora JL, *et al.* "Comportamiento del CA 19.9 y de la fracción HDL - colesterol en una muestra de pacientes con carcinoma de colon esporádico". *Anales de Medicina Interna* 21 (2004): 227-230.
10. Villadiego-Sánchez JM, *et al.* "Multivariate explanatory model for sporadic carcinoma of the colon in Dukes' stages I and IIa". *International Journal of Medical Sciences* 6.1 (2009): 43-50.
11. Ihaka R and Gentleman RR. "A language for data analysis and graphics". *Journal of Computational and Graphical Statistics* 5 (1996): 299-314.
12. Harrel Jr FJ. "Package "rms"". Versión 4.1-0 (2014).
13. Duffy MJ. "Tumor markers in clinical practice: a review focusing on common solid cancers". *Medical Principles and Practice* 22.1 (2013): 4-11.
14. Menéndez-Sánchez P, *et al.* "Marcadores tumorales en cáncer colorectal". *Cirugía y Cirujanos* 81 (2013): 169-175.
15. Al-Khallaf H, *et al.* "The Highest Ever Reported Level of Carbohydrate Antigen 19-9: A Case Report". *Journal of Medical Cases* 2 (2011): 71-75.
16. Rose G, *et al.* "Colon cancer and blood cholesterol". *Lancet* 1.7850 (1974): 181-183.
17. Wulaningsih W, *et al.* "Serum Lipids and the Risk of Gastrointestinal Malignancies in the Swedish AMORIS Study". *Journal of Cancer Epidemiology* (2012): 792034.

18. Ortega Calvo M and Cayuela Domínguez A. "Regresión Logística no condicionada y tamaño de muestra: una revisión bibliográfica". *Revista Española de Salud Pública* 76.2 (2002): 85-93.
19. Weiser MR, *et al.* "Individualized prediction of colon cancer recurrence using a nomogram". *Journal of Clinical Oncology* 26.3 (2008): 380-385.
20. Chen D, *et al.* "Validation of the Memorial Sloan-Kettering cancer center nomogram to predict disease-specific survival after r0 resection in a chinese gastric cancer population". *PLoS One* 8.10 (2013): e76041.
21. Wu YC and Lee WC. "Alternative performance measures for prediction models". *PLoS One* 9.3 (2014): e91249.
22. Steyerberg EW, *et al.* "Assessing the performance of prediction models: a framework for traditional and novel measures". *Epidemiology* 21 (2010): 128-138.
23. Royston P, *et al.* "Prognosis and prognostic research: developing a prognostic model". *BMJ* 338 (2009): b604.
24. Mallett S, *et al.* "Reporting performance of prognostic models in cancer: a review". *BMC Medicine* 8 (2010): 21.
25. Morgan S and Coleman J. "We live in testing times - Teaching rational test ordering in general practice". *Australian Family Physician* 43.5 (2014): 273-276.

Volume 7 Issue 4 April 2020

©All rights reserved by Juan Luis Méndez Mora, *et al.*