Um modelo de predição para seleccionar para co-gestão doentes de cirurgia colo-rectal.

A multivariable prediction model to select colorectal surgical patients for co-management.

Alexandra Bayão Hortaa,b, Carlos Geraldesa,c, Cátia Salgadod, Susana Vieirad, Miguel Xaviera , Ana Luísa Papoilaa,c

a NOVA Medical School|Faculdade de Ciências Médicas - Campo Mártires da Pátria, 130
1169-056 Lisboa, Portugal. abhorta257@gmail.com; ana.papoila@nms.unl.pt; carlos.geraldes@nms.unl.pt; miguel.xavier@nms.unl.pt

b Hospital da Luz-Lisboa – Av. Lusíada, 100, 1600-650 Lisboa, Portugal

c Centro de Estatística e Aplicações, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

d IDMEC, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal. catia.salgado@tecnico.ulisboa.pt; susana.vieira@tecnico.ulisboa.pt

**CORRESPONDING AUTHOR:**

Alexandra Bayão Horta

ORCID ID: orcid.org/0000-0002-8696-6089

E-MAIL: abhorta257@gmail.com

Carta de Resposta aos Revisores do manuscrito

Os autores agradecem ao editor e ao revisor B pela valiosa revisão que em muito contribuiu para aumentar a qualidade do artigo.

**# REVISOR B**

**All points raised were attended to except this:**

**“RevisorB – Comentário 1:**

*Some alternative methodologies should be used, at least as a robustness check. The Linear Probability Model, which is easy to implement, in some circumstances LPM beats logit and probit model in terms of forecasting. This should be addressed. In the limit, a combination of predictions - using (weighted) averages - obtained from these three models shall be tested.*

**Response:**

Thank you very much for your comments and suggestions. In Health Sciences it is common to use Logistic Regression Models whenever we have binary outcomes. In fact, medical doctors are familiarized with these models and easily interpret the results. Additionally, we didn’t choose to use the Linear Probability Model because it can produce probability estimates above 1 or below 0, and assumptions like normal distribution of errors and homoscedasticity might be violated. Regarding Probit Model, results are usually similar to the Logit Model. In summary, you are completely right, we could have used one of these three statistical approaches to model our data. It was a matter of choice.”

**RevisorB – Novo Comentário:**

I still believe that the linear probability model should be considered as well, at least as a forecast benchmark.

It is true that the LPM can produce negative estimates or higher than one. However, the estimated probabilities can always be truncated, to zero or to one.

On the other hand, the non-normality of the error and the heteroskedasticity have no impact on the estimates. They only impact standard errors and inference, but in any case, robust inference can be made.

Hence, the LPM must be considered because:

1) if it predicts better than Logit then these predictions should be taken into account.

2) if it is worse than Logit, we have a benchmark.

However, if the authors claim on not considering this model, the article may already be accepted for publication.

**Response:**

Thank you very much for your comments and suggestions. Accordingly, we used LPM to fit the data and very similar results were obtained regarding both discriminative and predictive performances. As consequence of these results, we decided, for the reasons already presented in the previous revision, and because you do not object as well, to maintain the multivariable logistic regression approach.