

1 **Exploring novel inflammation-related genetic and hematological predictors of response to**  
2 **neoadjuvant chemoradiotherapy in locally advanced rectal cancer**

3 Mladen Marinkovic<sup>1</sup>, Suzana Stojanovic-Rundic<sup>1,2</sup>, Aleksandra Stanojevic<sup>3</sup>, Marija Ostojic<sup>3</sup>,  
4 Dusica Gavrilovic<sup>4</sup>, Radmila Jankovic<sup>3</sup>, Natasa Maksimovic<sup>5</sup>, Rafael Strogilos<sup>6</sup>, Jerome  
5 Zoidakis<sup>6,7</sup>, Sergi Castellví-Bel<sup>8</sup>, Remond J.A. Fijneman<sup>9</sup>, Milena Cavic<sup>3\*</sup>

6  
7 <sup>1</sup>Clinic for Radiation Oncology and Diagnostics, Department of Radiation Oncology, Institute for  
8 Oncology and Radiology of Serbia, Belgrade, Serbia

9 <sup>2</sup>Faculty of Medicine, University of Belgrade, Belgrade, Serbia

10 <sup>3</sup>Department of Experimental Oncology, Institute for Oncology and Radiology of Serbia,  
11 Belgrade, Serbia

12 <sup>4</sup>Data Center, Institute for Oncology and Radiology of Serbia, Belgrade, Serbia

13 <sup>5</sup>Institute of Epidemiology, University of Belgrade, Belgrade, Serbia

14 <sup>6</sup>Department of Biotechnology, Biomedical Research Foundation, Academy of Athens, Athens,  
15 Greece

16 <sup>7</sup>Department of Biology, National and Kapodistrian University of Athens, Athens, Greece

17 <sup>8</sup>Gastroenterology Department, Fundació Recerca Clínic Barcelona-Institut d'Investigacions  
18 Biomèdiques August Pi i Sunyer, Centro de Investigación Biomédica en Red de Enfermedades  
19 Hepáticas y Digestivas, Clínic Barcelona, University of Barcelona, Barcelona, Spain

20 <sup>9</sup>Department of Pathology, The Netherlands Cancer Institute, Amsterdam, the Netherlands

21  
22 *\*Corresponding author:*

23 Dr Milena Cavic

24 Institute for Oncology and Radiology of Serbia

25 Department of Experimental Oncology

26 Pasterova 14, 11000 Belgrade, Serbia

27 Phone number: + 381 11 2067210

28 Fax number: + 381 11 2067294

29 E-mail: milena.cavic@ncrc.ac.rs

## 30 Abstract

31 **Background:** The standard initial treatment for locally advanced rectal cancer (LARC) is  
32 neoadjuvant chemoradiotherapy (nCRT). In order to select patients who would benefit the most  
33 from nCRT, there is a strong need for predictive biomarkers. The aim of this study was to evaluate  
34 the role of clinical, pathological, radiological, inflammation-related genetic, and hematological  
35 parameters in the prediction of response after nCRT.

36 **Methods:** *In silico* analysis of published transcriptomics datasets was conducted to identify  
37 the best candidate genes, whose expression will be measured using quantitative Real Time PCR  
38 (qRT-PCR) in pretreatment formaline-fixed paraffin-embedded (FFPE) samples. In this study, 75  
39 patients with LARC, between June 2020 and January 2022, were prospectively included. Patients  
40 were assessed for tumor response in the 8th week after nCRT completion with pelvic MRI scan and  
41 rigid proctoscopy. For patients with a clinical complete response (cCR) and initially distant located  
42 tumor no immediate surgery was suggested (“watch and wait” approach). The response after surgery  
43 was assessed using histopathological tumor regression grading (TRG) categories from postoperative  
44 specimens by Mandard. Responders (R) were defined as patients with cCR without operative  
45 treatment, and those with TRG 1 and TRG 2 postoperative categories. Non-responders (NR) were  
46 patients classified as TRG 3-5.

47 **Results:** Responders group comprised 35 patients (46.6%) and NR group included 53.4%  
48 of patients. Analysis of published transcriptomics data identified genes that could predict response  
49 to treatment and their significance was assessed in our cohort by qRT-PCR. When comparison was  
50 made in the subgroup of patients who were operated (TRG1 vs. TRG4), the expression of IDO1 was  
51 significantly deregulated ( $p<0.05$ ). Among hematological parameters between R and NR a  
52 significant difference in the response was detected for neutrophil-to-monocyte ratio (NMR), initial  
53 basophil, eosinophil and monocyte counts ( $p<0.01$ ). According to MRI findings, non-responders  
54 were more often presented with extramural vascular invasion ( $p<0.05$ ).

55 **Conclusion:** Based on logistic regression model, factors associated with favorable response  
56 to nCRT were found to be tumor morphology as well as hematological parameters which can be  
57 easily and routinely derived from initial laboratory results (NMR, eosinophil, basophil and  
58 monocyte counts) in a minimally invasive manner. Using various metrics, an aggregated score of  
59 the initial eosinophil, basophil, and monocyte counts demonstrated the best predictive performance.

60 **Keywords:** inflammation, locally advanced rectal cancer, neoadjuvant chemoradiotherapy,  
61 predictive biomarkers.

## 62 Introduction

63 In 2020, colorectal cancer (CRC) was the third most common malignant disease with 1.9 million  
64 new cases worldwide <sup>1</sup>. With 0.9 million deaths, it held the second place of cancer-related mortality  
65 causes in 2020 <sup>1</sup>. In Serbia, in 2020, there were 2,956 new cases and a total of 1,493 deaths related  
66 to rectal cancer, which placed Serbia in the group of countries with a high incidence and mortality  
67 rate for this disease <sup>1</sup>. In the majority of cases, it is diagnosed in advanced stages, when treatment  
68 options are limited. In this regard, in the past we have profiled the diagnostic, prognostic and  
69 predictive factors for cancers of the digestive system, leading to improved research strategies for  
70 patient management and care <sup>2-7</sup>. However, there is a need for better primary prevention, more  
71 effective screening program, diagnosis at an earlier stage of the disease and improvement of existing  
72 treatment modalities in our country and on a global level.

73 The standard treatment for locally advanced rectal cancer (LARC) is neoadjuvant  
74 chemoradiotherapy (nCRT) followed by total mesorectal excision with or without adjuvant  
75 chemotherapy. The pathologic complete response (pCR) after nCRT is achieved in 10-30% of cases  
76 <sup>8</sup>. It has been reported that pCR, independent of the initial clinical T and N stage of the disease, was  
77 associated with better local and distant disease control, as well as longer disease-free and overall  
78 survival <sup>9</sup>. Other reports showed that radical surgical treatment was related to significant morbidity,  
79 including postoperative complications <sup>10</sup>. Further investigations were directed towards less invasive  
80 surgical treatment or avoiding surgery (“watch and wait” approach) in patients with favorable  
81 response to nCRT, in order to improve the quality of life. Since 2004, a group of researchers led by  
82 Angelita Habr-Gama have contributed greatly in this area by pointing out the effectiveness and  
83 safety of this approach <sup>11</sup>. The current management of LARC uses the clinical complete response  
84 (cCR) as the point of reference for identifying patients for whom a non-operative approach may be  
85 a viable option <sup>12</sup>. However, the clinical response poorly correlates with the pathologic response <sup>13</sup>.

86 Other research trends in this field were dedicated to prolonging the period between completion of  
87 neoadjuvant treatment and surgery, changing the type and regimen of chemotherapy, as well as  
88 increasing the radiotherapy doses. These approaches aimed to achieve a higher percentage of good  
89 response to the initial treatment. As not all patients will benefit from these treatment modifications,  
90 there is a need to categorize them initially before treatment. In order to select patients who would  
91 benefit the most from a neoadjuvant treatment, there is a strong demand to discover and characterize  
92 predictive biomarkers. Despite numerous studies in this field, until now no molecular marker has  
93 been implemented as a diagnostic or predictive parameter in routine clinical practice of LARC. This  
94 is stressed by the fact that there was not enough matching regarding results of published studies in  
95 this area and only two genes (MMP4 and FLNA) were shown to be significant in more than one  
96 study <sup>14</sup>. Limitations of previous studies included a small number of patients, the absence of  
97 reproducibility of measurements, the use of different methodologies, the retrospective nature of the

98 studies, the heterogeneity of the studied groups and applied treatment modalities, as well as the lack  
99 of verification of the findings. Further research was aimed at examining the cumulative effect of  
100 molecular markers in combination with radiological and clinical data. An example of such  
101 successful research is the examination of the correlation between the expression of three protein  
102 molecular markers (c-MYC, PCNA and TIMP1) and magnetic resonance imaging (MRI)  
103 parameters<sup>15</sup>.

104 The association between inflammatory bowel disease and the higher risk of developing colorectal  
105 cancer is well known<sup>16,17</sup>. Also, there is evidence of the role of inflammation in sporadic colorectal  
106 cancer<sup>18,19</sup>. Chronic inflammation in the tumor microenvironment has also been shown to favorize  
107 tumor growth and invasiveness and stimulate synthesis of epithelial to mesenchymal transition  
108 promoting transcription factors<sup>20</sup>. Yet, no inflammation-related genetic or circulating biomarkers  
109 have been investigated in detail or established as predictive parameters in the LARC setting so far.

110 The aim of this study was to evaluate the role of clinical, pathological, radiological, inflammation-  
111 related genetic and hematological parameters in prediction of response after nCRT in patients with  
112 LARC.

## 113 **Methods**

114 *In silico* analysis of published transcriptomics datasets

115 *In silico* analysis of published transcriptomics datasets was conducted to identify the best candidate  
116 genes, whose expression will be measured using quantitative Real Time PCR (qRT-PCR) in  
117 pretreatment formaline-fixed paraffin-embedded (FFPE) samples.

118 Gene expression patterns were analyzed using publicly available datasets. By searching the public  
119 database the National Center for Biotechnology Information Gene Expression Omnibus (NCBI  
120 GEO) using key words rectal cancer, chemoradiotherapy and response to treatment, five adequate  
121 sets of data that analyzed pretreatment samples were identified: GSE45404, GSE68204,  
122 GSE139255, GSE46862 and GSE3493<sup>21-25</sup>. Three datasets were selected where inflammatory  
123 response significantly correlated with treatment outcome to nCRT. Gene expression profiles of  
124 GSE46862, GSE139255, and GSE45404\_570 were obtained from GEO database. The total number  
125 of patients of each dataset was 69, 156, and 42 respectively. In all selected studies, the response to  
126 treatment was classified according to pathohistological tumor regression grading (TRG) categories  
127 from the postoperative specimen using Mandard scoring system. Patients were subdivided into  
128 responders (TRG 1-2) and non-responders (TRG 3-5). Microarray data were processed and  
129 normalized with the Robust Multichip Average method<sup>26</sup>. Analysis for statistically significant  
130 differences between the two groups was conducted using the standard moderated t-test from the  
131 limma package<sup>27,28</sup>. Gene set enrichment analysis (GSEA) was performed on selected datasets, and  
132 Hallmark, Kyoto encyclopedia of genes and genomes (KEGG), and Reactome gene sets were used

133 to identify pathway alterations in patients who responded well to the therapy (TRG 1-2) versus those  
134 who did not (TRG 3-5)<sup>29,30</sup>.

135 Next, the top 100 genes from selected datasets, ranked by the default Signal2Noise metric used in  
136 previously described GSEA analysis, were extracted and overlapped using Venn diagram software.  
137 Cytoscape (version 3.10.0) was applied as bioinformatics software to evaluate the potential  
138 correlation between finally selected genes<sup>31,32</sup>.

### 139 *Patient characteristics, treatment and follow-up*

140 In this study 75 patients with LARC treated at the Institute for Oncology and Radiology of Serbia,  
141 between June 2020 and January 2022, were prospectively included. The inclusion criteria were  
142 histopathologically verified adenocarcinoma of the rectum, with a distant margin up to 12 cm from  
143 the anal verge by rigid proctoscopy. LARC was defined as T3-T4N0 or any T stage N positive.  
144 Pretreatment evaluation included an abdominal and pelvic MRI scan and a computed tomography  
145 (CT) scan or X ray of the chest. All patients were treated with long-course nCRT. Radiotherapy  
146 (RT) was delivered using volumetric modulated arc therapy-simultaneous integrated boost  
147 technique (VMAT-SIB). The dose to mesorectum and pelvic lymph nodes was 45 Gy (1.8  
148 Gy/fraction). A SIB was delivered on macroscopic disease region expanded with 2 cm margin with  
149 a total dose of 54 Gy (2.16 Gy/fraction). Concomitant chemotherapy started on the first day of RT  
150 and was administered during the first and the fifth week of RT. The chemotherapy regimen included:  
151 5-FU (350 mg/m<sup>2</sup> on the first day of the first and fifth week of RT) and Leucovorine (25 mg/m<sup>2</sup>  
152 daily, 5 days of the first and fifth week of RT).

153 Patients were assessed for tumor response in the 8th week after nCRT completion with pelvic MRI  
154 scan, rigid proctoscopy and digital rectal examination. For patients with cCR and initially distant  
155 located tumor no immediate radical surgery was suggested and they were enrolled in a strict follow-  
156 up program (“watch and wait” approach). Patients with cCR where sphincter preservation surgery  
157 treatment can be delivered, were referred to surgical resection between weeks 8 and 12 from nCRT  
158 completion. For patients with partial response (PR), surgery was delayed until week 12-15,  
159 approximately. The pathohistological response after surgery was assessed according to  
160 classification by Mandard. The response to treatment was classified according to pathohistological  
161 TRG categories from the postoperative specimen. Responders were defined as patients with cCR  
162 without operative treatment, and those with TRG 1 and TRG 2 postoperative categories. Non-  
163 responders were patients classified as TRG 3-5.

164 Formalin-fixed paraffin-embedded (FFPE) samples taken at the time of disease diagnosis were  
165 collected. The project was approved by the Ethics Committee of the Institute for Oncology and  
166 Radiology of Serbia (Approval No. 2211-01 from 11.06.2020.) and Ethics Committee of the Faculty  
167 of Medicine, University of Belgrade (Approval No. 1322/XII-17 from 03.12.2020.). All patients  
168 signed an informed consent.

169 Before initiation of treatment, ethylenediaminetetraacetic acid (EDTA) peripheral blood was drawn  
170 by venipuncture and hematological parameters were derived from the absolute differential counts  
171 of a complete blood count (CBC). The neutrophil-to-lymphocyte ratio (NLR) was calculated as a  
172 ratio of circulating neutrophil and lymphocyte counts, and the platelet-to-lymphocyte ratio (PLR)  
173 was defined as the absolute count of platelets divided by the absolute lymphocyte count. The derived  
174 neutrophil-to-lymphocyte ratio (dNLR) was calculated as absolute neutrophil count divided by  
175 absolute leukocyte minus absolute neutrophil count. The lymphocyte-to-monocyte ratio (LMR),  
176 platelet-to-monocyte ratio (PMR), and neutrophil-to-monocyte ratio (NMR) were also analyzed.  
177 Patients' pre-treatment hemoglobin levels were obtained. The staging of the tumor was assessed  
178 according to the eighth edition of the Union for International Cancer Control (UICC) TNM staging  
179 system for rectal cancer<sup>33</sup>. The general condition of the patients was classified using the Eastern  
180 cooperative oncology group (ECOG) Scale of Performance Status<sup>34</sup>.

### 181 *RNA isolation and cDNA synthesis*

182 Total RNA was isolated from 2-5 10 µm thick FFPE tissue sections using RNeasy FFPE Kit  
183 (Qiagen, Manchester, UK). RNA quality and concentration were determined spectrophotometrically  
184 using BioSpec-nano (Shimadzu Scientific Instruments, Kyoto, Japan). The complementary DNA  
185 (cDNA) was accessed from 1 µg total RNA using random primers and MultiScribe™ Reverse  
186 Transcriptase (50 U/µL) from the High-Capacity cDNA Reverse Transcription kit (Applied  
187 Biosystems, Foster City, CA, USA). The reaction was performed in 20 µL, using the following  
188 program: 25°C for 10 min, 37°C for 120 min, and inactivation at 85°C for 5 min.

### 189 *Quantitative Real Time PCR (qRT-PCR)*

190 The mRNA levels of IL6 (RefSeq. NM\_000600.5), CXCL9 (RefSeq. NM\_002416.3), IDO1  
191 (RefSeq. NM\_002164.6) and CYBB (RefSeq. NM\_000397.4) were detected by quantitative real-  
192 time PCR (qRT-PCR) using oligonucleotides primers (Integrated DNA Technology, Coralville,  
193 Iowa, USA) previously designed using NCBI Primer Blast and SybrGreen Gene Expression Master  
194 Mix (Applied Biosystems), on ABI Prism 7500 Sequence Detection System (Applied Biosystems).  
195 The thermal cycling conditions consisted of an UDG activation at 50°C, initial denaturation step at  
196 95°C for 2 min followed by 40 cycles of denaturation (15 sec at 95°C) and annealing/extension (1  
197 min at 60°C). All experiments were performed in duplicate, including non-template controls in each  
198 amplification. Gene expression data were normalized to glyceraldehyde-3-phosphate  
199 dehydrogenase (GAPDH, RefSeq. NM\_002046.5). Data was analyzed using the classical delta-  
200 delta-Ct method, and results expressed in relative units.

### 201 *Statistical analysis*

202 For normal distribution data testing, the Kolmogorov-Smirnov and Shapiro-Wilk tests were used.  
203 Descriptive methods (frequencies, percentage, mean, median, standard deviation (SD) and range)  
204 were used to summarize the data. The statistical significance level was set at  $p < 0.05$ . For comparison  
205 of disease and treatment characteristics among different subgroups the Wilcoxon rank sum, Pearson

206 chi-square and Fisher exact tests were used. Also, for evaluating potential predictors of the response,  
207 univariate and multivariate logistic regression was used (odds ratio with 95% CI for description,  
208 Likelihood Ratio and Wild test), and the responders versus non-responders was set as a dependent  
209 variable. We evaluated the sensitivity, specificity, positive predictive value, negative predictive  
210 value, and predictive accuracy for clinical assessment of disease presence in comparison with  
211 pathohistological response as a gold standard in group of patients where operative treatment was  
212 conducted <sup>35</sup>. The Receiver Operating Characteristics curve (ROC) methods were applied to  
213 investigate the discriminative potential of NLR, PLR, dNLR, LMR, PMR, NMR, initial basophil,  
214 eosinophil and monocyte counts, for the good response to treatment (AUC ROC-Area Under the  
215 ROC curve according DeLong's method; Likelihood ratio test for AUC ROC; the best cut-off value  
216 was set as value with maximum sensitivity and specificity). The statistical analysis was performed  
217 using the program R (version 3.3.2 (2016-10-31) --"Sincere Pumpkin Patch"; Copyright (C) 2016  
218 The R Foundation for Statistical Computing; Platform: x86\_64-w64-mingw32/x 64 (64-bit);  
219 downloaded: January 21, 2021). In the search for a measure that outperforms the individual  
220 variables, numerical variables that remained significant in the multivariate analysis were utilized to  
221 create the composite scores. The predictive power of these scores was then tested using the AUC,  
222 the Area Under Precision-Recall Curve (AUCRP), the Root Mean Square Error (RMSE) as a metric  
223 (using the ROCR package) and a random forest classifier (using the randomForest package, with  
224 the MeanDecreaseAccuracy metric) <sup>36</sup>.

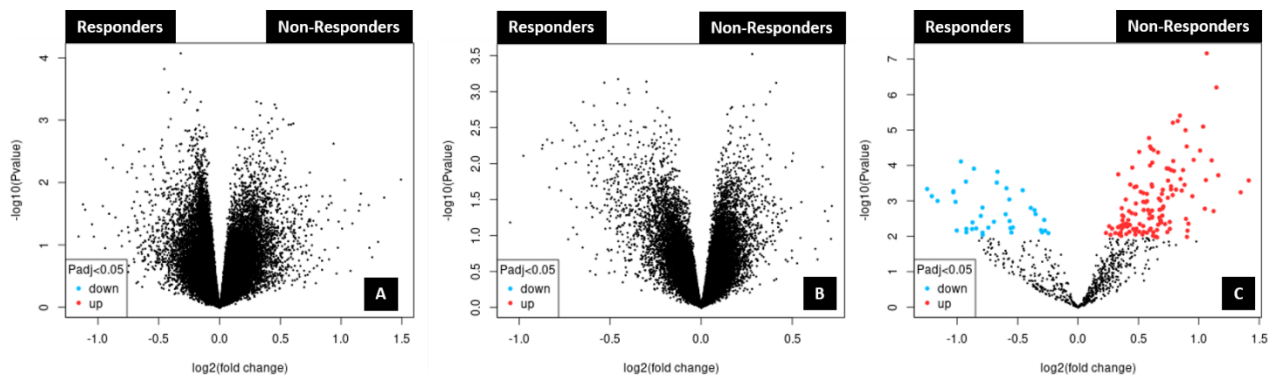
## 225 **Results**

226 After extensive search of the GEO database according to the key words rectal cancer,  
227 chemoradiotherapy and response to treatment, three gene expression datasets were finally obtained.  
228 Volcano plots were employed to identify genes that exhibited statistically significant differential  
229 expression between responders and non-responders, as determined by the adjusted p-value, among  
230 selected datasets (Figure 1). Significant alteration defined as those with adjusted p-value lower than  
231 0.05 was obtained only for GSE 139255 dataset. The results of the differential expression analysis  
232 within GSE 139255 are reported in Supplementary Material 1. While KEGG and Reactome GSEA  
233 analysis yielded no overlap among the selected datasets, the Hallmark analysis exhibited consistent  
234 and significant parameters across two datasets. Results of GSEA Hallmark analysis presented in  
235 Table 1 showed parameters which reached significant levels within GSE46862 and GSE139255  
236 datasets in relation to Hallmark inflammatory response pathway (NOM p-value < 0.05, FDR q-  
237 value < 0.25). Enrichment plots were used to present the expression of genes in selected datasets in  
238 Figure 2.

239

240

241



242

243 **Figure 1. Volcano plots for selected datasets: GSE45404\_570 (A); GSE46862 (B);**

244

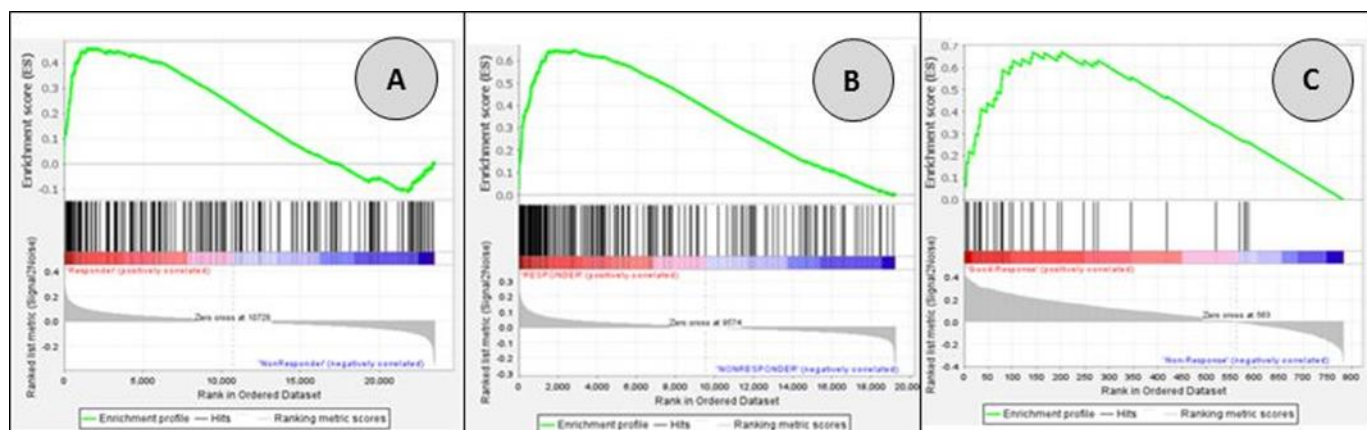
**GSE139255 (C).**

245 **Table 1. GSEA Hallmark analysis in relation to Hallmark inflammatory response pathway.**

Values	Datasets		
	GSE45404_570	GSE46862	GSE139255
NES <sup>1</sup>	1.09	1.63	1.85
NOM p-value <sup>2</sup>	0.373	0.024	0.002
FDR q-value <sup>3</sup>	0.661	0.056	0.054

246 <sup>1</sup>NES - Normalized Enrichment Score; <sup>2</sup>NOM p-value – Nominal p-value; <sup>3</sup>FDR q-value – False Discovery Rate q-

247 value;



248

249 **Figure 2. GSEA enrichment plots for genes included in Hallmark inflammatory response**

250

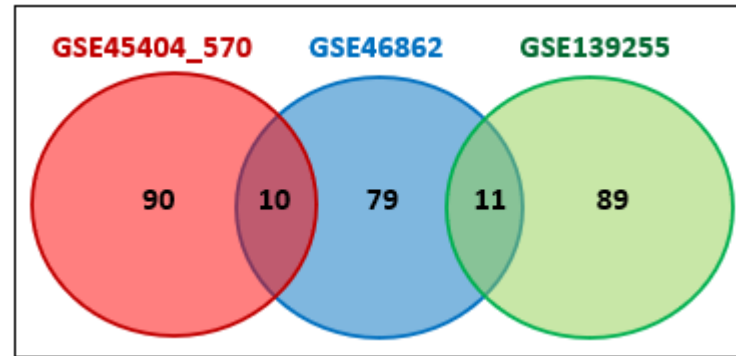
**pathway: GSE45404\_570 (A); GSE46862 (B); GSE139255 (C).**

251 The top 100 genes from each database (Supplementary Material 2) were chosen, and overlapped

252 among these three datasets using Venn diagram. The results are presented in Figure 3.

253





254

255 **Figure 3. Venn diagram showing overlapping of inflammation-related genes between three**  
256 **analyzed datasets.**

257 Our focus was on the genes included in the inflammatory response. As a result of overlapping three  
258 datasets, there were 11 genes present in two of them (PLAU, TGFB2, HGF, IL6, CXCL10, CXCL9,  
259 IDO1, INHBA, PDGFRB, CYBB, IL24). Statistical significance of these genes among responders  
260 and non-responders in all three datasets was examined and the results are presented in Table 2.

261

**Table 2. Comparison between responders and non-responders within analyzed datasets in relation to expression of selected genes.**

Gene	Data Sets									
	GSE45404_570			GSE46862			GSE139255			
	R	NonR	Test*	R	NonR	Test*	R	NonR	Test*	
<i>N (%)</i>	19/42 (45.2%)	23/42 (54.8%)	-	28/69 (40.6%)	41/69 (59.4%)	-	89/156 (57.1%)	67/156 (42.9%)	-	
<i>PLAU</i>	Mean (SD)	7.6 (0.6)	7.5 (0.9)	ns	8.1 (0.9)	7.6 (0.9)	<i>p</i> <0.05	1967.0 (2152.2)	766.5 (637.0)	<i>p</i> <0.01
	Median (Range)	7.7 (6.2-8.5)	7.4 (6.0-9.5)		8.3 (5.8-9.8)	7.8 (5.6-9.3)		1351.0 (75.0-9876.0)	604.7 (88.4-3027.0)	
<i>TGFB2</i>	Mean (SD)	5.0 (0.3)	5.2 (0.3)	ns	5.5 (0.7)	5.1 (0.6)	<i>p</i> <0.05	64.9 (50.9)	35.9 (25.9)	<i>p</i> <0.01
	Median (Range)	5.1 (4.4-5.7)	5.1 (4.5-5.7)		5.5 (4.3-6.5)	5.1 (3.7-6.7)		51.6 (1.7-261.9)	30.8 (3.3-114.8)	
<i>HGF</i>	Mean (SD)	3.8 (0.3)	3.8 (0.3)	ns	5.6 (0.9)	5.2 (0.8)	ns	332.6 (371.5)	166.0 (104.4)	<i>p</i> <0.01
	Median (Range)	3.8 (3.5-4.4)	3.8 (3.4-4.4)		5.6 (4.0-7.6)	5.0 (3.5-6.7)		224.6 (22.2-2733)	160.0 (30.7-540.1)	
<i>IL6</i>	Mean (SD)	6.0 (1.3)	5.6 (1.1)	ns	5.7 (1.4)	4.9 (1.1)	<i>p</i> <0.01	432.9 (996.6)	104.3 (200.9)	<i>p</i> <0.01
	Median (Range)	5.8 (4.0-8.5)	5.4 (4.2-8.8)		5.7 (4.0-9.7)	4.5 (3.8-8.8)		90.7 (1.7-5968.0)	40.7 (1.1-1207.0)	
<i>CXCL10</i>	Mean (SD)	7.9 (1.4)	6.7 (1.6)	<i>p</i> <0.05	6.4 (1.3)	5.8 (1.1)	ns	-	-	NA
	Median (Range)	7.8 (5.5-10.3)	6.6 (3.6-10.4)		6.2 (4.1-9.5)	5.6 (4.0-9.5)		-	-	
<i>CXCL9</i>	Mean (SD)	6.8 (1.6)	5.6 (1.6)	<i>p</i> <0.05	5.5 (1.0)	5.0 (0.8)	<i>p</i> <0.05	-	-	NA
	Median (Range)	6.7 (3.4-9.2)	5.6 (3.2-9.8)		5.1 (4.3-8.6)	4.8 (4.1-8.5)		-	-	
<i>IDO1</i>	Mean (SD)	6.5 (1.5)	5.8 (1.2)	ns	5.9 (1.0)	5.4 (1.2)	<i>p</i> <0.05	-	-	NA
	Median (Range)	6.2 (4.5-10.6)	5.6 (4.2-8.8)		5.7 (3.7-8.2)	5.2 (3.6-8.5)		-	-	
<i>INHBA</i>	Mean (SD)	8.0 (0.8)	7.9 (1.4)	ns	6.8 (1.1)	6.2 (0.9)	ns	539.0 (768.1)	240.6 (217.8)	<i>p</i> <0.01
	Median (Range)	7.9 (6.7-9.4)	8.1 (5.4-10.7)		6.6 (5.2-9.2)	6.4 (4.3-8.3)		295.1 (25.2-4582.0)	172.2 (17.7-1138.0)	
<i>PDGFRB</i>	Mean (SD)	6.9 (0.6)	6.8 (1.0)	ns	7.6 (0.8)	7.1 (0.7)	<i>p</i> <0.01	1064.0 (894.3)	629.1 (614.6)	<i>p</i> <0.01
	Median (Range)	6.7 (6.1-8.2)	6.9 (4.7-8.9)		7.5 (6.5-9.2)	7.1 (5.6-8.7)		781.7 (43.7-3629.0)	470.9 (59.3-3896.0)	
<i>CYBB</i>	Mean (SD)	5.3 (0.6)	4.7 (0.5)	<i>p</i> <0.01	8.1 (0.9)	7.5 (1.1)	<i>p</i> <0.01	-	-	NA
	Median (Range)	5.0 (4.5-6.4)	4.6 (3.7-5.5)		8.4 (5.5-9.9)	7.7 (4.6-9.6)		-	-	
<i>IL24</i>	Mean (SD)	5.3 (1.2)	5.5 (1.4)	ns	4.8 (0.9)	4.4 (0.8)	ns	150.9 (241.4)	59.7 (66.8)	ns
	Median (Range)	5.3 (3.3-7.9)	5.4 (3.4-8.8)		4.6 (3.7-7.1)	4.1 (3.5-7.1)		56.5 (1.0-1321.0)	32.7 (2.2-391.9)	

263 *ns* - not statistically significant; *R* - responder; *NonR* - non-responder; \*Wilcoxon rank sum test; NA – not available (without data in dataset); *PLAU* - Plasminogen Activator,  
264 *Urokinase*; *TGFB2* - Transforming Growth Factor Beta 2; *HGF* - hepatocyte growth factor; *IL6* - Interleukin-6; *CXCL10* - C-X-C Motif Chemokine Ligand 10; *CXCL9* - C-X-C Motif  
265 Chemokine Ligand 9; *IDO1* - Indoleamine 2,3-dioxygenase-1; *INHBA* - Inhibin Subunit Beta A; *PDGFRB* - Platelet derived growth factor receptor beta; *CYBB* - Cytochrome B-245  
266 Beta Chain; *IL24* - Interleukin 24;

267 None of the selected genes were found to be statistically significant in all three datasets. In order to  
 268 validate the potential of gene expression to predict treatment outcome, genes included in the  
 269 Hallmark inflammatory response pathway (IL6, CXCL9, CYBB) were chosen, along with a gene  
 270 which had promising potential according to literature search (IDO1)<sup>37,38</sup>. After the connection  
 271 among selected genes was checked using Cytoscape network, it was found that three of them (IL6,  
 272 CXCL9, CYBB) were part of a pathway related to avoiding immune detection.

273 In order to explore the significance of *in silico* obtained results, the expression of candidate genes  
 274 was analyzed in the cohort of LARC patients from the Institute for Oncology and Radiology of  
 275 Serbia. Patients, disease, treatment and outcomes characteristics are presented in Table 3. The  
 276 majority of patients had T3 stadium and N positive disease. One third of patients were female. All  
 277 patients completed the planned nCRT. Operative treatment was conducted in 63 patients, and the  
 278 pathohistological complete response rate was 20.6 % (Table 3). Twelve patients with distally located  
 279 tumor and complete clinical response were involved in a “watch and wait” approach. One patient  
 280 had to be excluded from the hematological ratios analysis, because of his chronic lymphocytic  
 281 leukemia and its influence to the parameters of this analysis.

282 **Table 3. Patients', disease, treatment and outcomes characteristics.**

Characteristics	N (%)	Characteristics	N (%)
<b>Age (years)</b>		<b>LMR<sup>6</sup></b>	
Mean (SD)	60.8 (10.6)	N (%)	74/75 (98.7%)
Median (Range)	62.0 (33.0-81.0)	Mean (SD)	3.5 (1.3)
<b>Gender</b>		Median (Range)	3.5 (1.3-6.6)
Female	25 (33.3%)	<b>PMR<sup>7</sup></b>	
Male	50 (66.7%)	N (%)	74/75 (98.7%)
<b>Performance status (PS)<sup>1</sup></b>		Mean (SD)	555.9 (237.6)
ECOG 0	52 (69.3%)	Median (Range)	512.5(231.0-1795)
ECOG 1	23 (30.7%)	<b>NMR<sup>8</sup></b>	
<b>T in clinical TNM</b>		N (%)	74/75 (98.7%)
T2	2 (2.7%)	Mean (SD)	8.8 (5.3)
T3	64 (85.3%)	Median (Range)	7.7 (2.0-46.0)
T4	9 (12.0%)	<b>Tumor location<sup>9</sup>(cm)</b>	
<b>N in clinical TNM</b>		Lower (anal verge≤ 8)	60 (80.0%)
N0	1 (1.3%)	Middle (8<anal verge≤12)	15 (20%)
N1	22 (29.4%)	<b>Tumor morphology</b>	
N2	52 (69.3%)	Polypoid	2 (2.6%)
<b>UICC<sup>2</sup> staging</b>		Semi-annular	17 (22.7%)
IIA	1 (1.3%)	Annular	56 (74.7%)
IIIA	1 (1.3%)	<b>Extramural vascular invasion</b>	
IIIB	40 (53.4%)	Yes	25 (33.3%)
IIIC	33 (44.0%)	No	50 (66.7%)
<b>Tumor differentiation</b>		<b>Stenotic character</b>	
well	39 (52.0%)	Yes	23 (30.7%)
moderate	30 (40.0%)	No	52 (69.3%)
poor	6 (8.0%)	<b>IL6</b>	
<b>Mucinous histological type</b>		N (%)	57/75 (76.0%)
Yes	13 (17.3)	Mean (SD)	3.5 (6.6)
No	62 (85.7)	Median (Range)	0.68 (0.01-36.1)
<b>Initial hemoglobin level (g/L)</b>		<b>CXCL</b>	
Mean (SD)	130.9 (20.6)	N (%)	60/75 (80.0%)
Median (Range)	134.4 (79.0-163.8)	Mean (SD)	1.6 (2.9)
<b>Absolute basophil count (10<sup>9</sup>/L)</b>		Median (Range)	1.0 (0.1-21.1)
Mean (SD)	0.1 (0.1)	<b>CYBB</b>	
Median (Range)	0.0 (0.0-1.0)	N (%)	59/75 (78.7%)

<b>Absolute eosinophil count (<math>10^9/L</math>)</b>		Mean (SD)	2.2 (3.2)
Mean (SD)	0.2 (0.2)	Median (Range)	1.0 (0.03-15.7)
Median (Range)	0.2 (0.0-1.0)	<b>IDO1</b>	
<b>Absolute monocyte count (<math>10^9/L</math>)</b>		N (%)	57/75 (76.0%)
Mean (SD)	0.6 (0.2)	Mean (SD)	1.7 (2.3)
Median (Range)	0.5 (0.2-1.4)	Median (Range)	1.3 (0.2-17.0)
<b>NLR<sup>3</sup></b>		<b>Operative treatment</b>	
N (%)	74/75 (98.7%)	No (cCR <sup>10</sup> )	12 (16.0%)
Mean (SD)	2.9 (2.7)	Yes	63 (84.0%)
Median (Range)	2.4 (0.9-23.0)	<b>TRG<sup>11</sup> score (operated patients)</b>	
<b>PLR<sup>4</sup></b>		TRG1	13/63 (20.6%)
N (%)	74/75 (98.7%)	TRG2	10/63 (15.9%)
Mean (SD)	184.4 (131.6)	TRG3	30/63 (47.6%)
Median (Range)	148.9 (66.3-897.5)	TRG4	10/63 (15.9%)
<b>dNLR<sup>5</sup></b>		<b>Response to the treatment</b>	
N (%)	74/75 (98.7%)	R <sup>12</sup> (cCR+TRG1+TRG2)	35/75 (46.7%)
Mean (SD)	1.9 (1.5)	NonR <sup>13</sup> (TRG3+TRG4)	40/75 (53.3%)
Median (Range)	1.6 (0.1-13.1)	<b>Total</b>	<b>75 (100%)</b>

283 <sup>1</sup>ECOG PS - The Eastern Cooperative Oncology Group Performance Status; <sup>2</sup>UICC - Union for International Cancer  
 284 Control; <sup>3</sup>NLR - Neutrophil-to-lymphocyte ratio; <sup>4</sup>PLR - Platelet-to-lymphocyte ratio; <sup>5</sup>dNLR - Derived neutrophil-to-  
 285 lymphocyte ratio; <sup>6</sup>LMR - Lymphocyte-monocyte ratio; <sup>7</sup>PLR - Platelet-lymphocyte ratio; <sup>8</sup>NMR - Neutrophil-to-  
 286 monocyte ratio; <sup>9</sup>Tumor location - distance from anal verge; <sup>10</sup>cCR - Patients without operative treatment due to  
 287 complete clinical response; <sup>11</sup>TRG - Tumor regression grade; <sup>12</sup>R - responders; <sup>13</sup>NonR - non-responders  
 288 Correlation of clinical evaluation and pathological examination as a gold standard in a group of  
 289 patients where operative treatment was conducted is presented in Table 4. Using disease prevalence  
 290 of 79.4% the sensitivity, specificity, positive predictive value, negative predictive value and  
 291 predictive accuracy were calculated (Table 5).

292 **Table 4. Correlation of clinical and pathological CR within a group of patients where**  
 293 **operative treatment was conducted.**

		Pathological assessment of disease presence		
		Yes	No	Total
Clinical assessment of disease presence	Yes	46 (73.02%)	3 (4.76%)	49 (77.78%)
	No	4 (6.35%)	10 (15.87%)	14 (22.22%)
	Total	50 (79.37%)	13 (20.63%)	63 (100%)

295 **Table 5. Sensitivity, specificity, positive predictive value, negative predictive value and**  
 296 **predictive accuracy of clinical evaluation for prediction of disease status using pathological**  
 297 **examination as a gold standard.**

Characteristics	Clinical evaluation
<i>Sensitivity (95% CI)</i>	92.0% (80.8-97.8%)
<i>Specificity (95% CI)</i>	76.9% (46.2-95.0%)
<i>PPV (95% CI)</i>	93.9% (85.0-97.6%)
<i>NPV (95% CI)</i>	71.4% (48.2-87.0%)
<i>Predictive accuracy (95% CI)</i>	88.9% (78.4-95.4%)

298 *CI – confidence interval; NPV – negative predictive value; PPV -positive predictive value*  
 299

300 Research interest was the comparisons between responders (comprised 35 patients) and non-  
 301 responders (included 40 patients) (Table 6). Initial T and N stadium of disease were not significantly  
 302 different between these two groups. Patients with poorly differentiated tumors and those with  
 303 mucinous histological type responded to treatment significantly worse than patients with well or  
 304 moderate tumor differentiation and those without mucinous type ( $p < 0.05$  and  $p < 0.01$ , respectively).  
 305 According to MRI findings, non-responders presented more often with extramural vascular invasion  
 306 (EMVI) ( $p < 0.05$ ). Among hematological parameters, significance was found for absolute basophil,  
 307 eosinophil and monocyte counts, dNLR and NMR.

308 In the whole patient group, there was no significant correlation between *in silico* selected genes  
 309 (IL6, CYBB, CXCL9, IDO1) and response to treatment. On the other hand, when comparison  
 310 between patients where pCR (TRG1) was detected and those who responded the worst (TRG4),  
 311 statistical significance was found based on IDO1 expression (Wilcoxon rank sum test,  $p = 0.036$ )  
 312 (Supplementary Table 1).

313 **Table 6. Comparison of characteristics of responders and non-responders to neoadjuvant**  
 314 **chemoradiotherapy.**

Characteristic	The response to treatment		
	Responders	Non-responders	Wilcoxon rank sum test
<b>Age (years)</b>			
Mean (SD)	61.5 (10.7)	60.3 (10.6)	
Median (Range)	63.0 (38.0-81.0)	62.0 (33.0-76.0)	<i>ns</i>
<b>Gender</b>			
Male	25 (71.4%)	25 (62.5%)	
Female	10 (28.6%)	15 (37.5%)	<i>ns*</i>
<b>T in clinical TNM</b>			
T2	2 (5.7%)	0 (0%)	
T3	29 (82.9%)	35 (87.5%)	<i>ns#</i>
T4	4 (11.4%)	5 (12.5%)	
<b>N in clinical TNM</b>			
N0	1 (2.9%)	0 (0%)	
N1	13 (37.1%)	9 (22.5%)	<i>ns#</i>
N2	21 (60.0%)	31 (77.5%)	
<b>UICC staging</b>			
IIA+ IIIA+ IIIB	22 (62.9%)	20 (50.0%)	<i>ns*</i>

IIC	13 (37.1%)	20 (50.0%)	
<b>Tumor differentiation</b>			
Well and moderate	35 (100%)	34 (85%)	<i>p</i> <0.05 <sup>#</sup>
Poor	0 (0%)	6 (15%)	
<b>Mucinous histological type</b>			
No	35 (100%)	27 (67.5%)	<i>p</i> <0.01 <sup>#</sup>
Yes	0 (0%)	13 (32.5%)	
<b>Extramural vascular invasion</b>			
No	28 (80%)	22 (55%)	<i>p</i> <0.05 <sup>*</sup>
Yes	7 (20%)	18 (45%)	
<b>Tumor morphology</b>			
Polypoid and semi-annular	14 (40.0%)	5 (12.5%)	<i>p</i> <0.01 <sup>*</sup>
Annular	21 (60.0%)	35 (87.5%)	
<b>Stenotic character</b>			
No	29 (82.9%)	23 (57.5%)	<i>p</i> <0.05 <sup>*</sup>
Yes	6 (17.1%)	17 (42.5%)	
<b>Absolute basophil count</b>			
Mean (SD)	0.03 (0.04)	0.08 (0.16)	<i>p</i> <0.01
Median (Range)	0.02 (0-0.1)	0.1 (0-1.0)	
<b>Absolute eosinophil count</b>			
Mean (SD)	0.17 (0.12)	0.27 (0.22)	<i>p</i> <0.05
Median (Range)	0.1 (0-0.53)	0.2 (0-1.0)	
<b>Absolute monocyte count</b>			
Mean (SD)	0.52 (0.18)	0.63 (0.21)	<i>p</i> <0.01
Median (Range)	0.5 (0.2-1.1)	0.6 (0.4-1.4)	
<b>Neutrophil-to-lymphocyte ratio</b>			
Mean (SD)	3.38 (3.62)	2.56 (1.5)	<i>ns</i>
Median (Range)	2.5 (1.17-23.0)	2.27 (0.93-7.46)	
<b>Platelet-to-lymphocyte ratio</b>			
Mean (SD)	184.8 (141.5)	184.0 (123.9)	<i>ns</i>
Median (Range)	144.4 (72.3-897.5)	150.0 (66.3-681.1)	
<b>dNLR<sup>§</sup></b>			
Mean (SD)	2.21 (2.02)	1.61 (0.73)	<i>p</i> <0.05
Median (Range)	1.81 (0.84-13.14)	1.49 (0.09-3.37)	
<b>Neutrophil-to-monocyte ratio</b>			
Mean (SD)	10.3 (7.01)	7.44 (2.48)	<i>p</i> <0.01
Median (Range)	9.18 (4.56-46.0)	7.0 (2.0-15.5)	
<b>IL6</b>			
N (%)	28/35 (80.0%)	29/40 (72.5%)	<i>ns</i>
Mean (SD)	4.3 (8.6)	2.8 (4.0)	
Median (Range)	0.4 (0.01-36.1)	1.4 (0.05-16.8)	
<b>CYBB</b>			
N (%)	29/35 (82.8%)	30/40 (75.0%)	<i>ns</i>
Mean (SD)	2.1 (2.8)	2.2 (3.6)	
Median (Range)	1.0 (0.03-10.3)	1.1 (0.05-15.7)	
<b>CXCL9</b>			
N (%)	29/35 (82.8%)	31/40 (77.5%)	<i>ns</i>
Mean (SD)	1.1 (0.7)	2.1 (3.9)	
Median (Range)	0.9 (0.3-2.4)	1.2 (0.1-21.1)	
<b>IDO1</b>			
N (%)	28/35 (80.0%)	29/40 (72.5%)	<i>ns</i>
Mean (SD)	1.5 (1.1)	1.9 (3.1)	
Median (Range)	1.4 (0.2-3.6)	1.3 (0.2-17.0)	
<b>Total</b>	<b>35 (100%)</b>	<b>40 (100%)</b>	<b>-</b>

315 \*Pearson  $\chi^2$  Test; <sup>#</sup>Fisher Exact Test; *ns* - not statistically significant; <sup>§</sup>dNLR-Derived neutrophil-to-lymphocyte ratio  
316

317 Next, ROC analysis was performed and it revealed the optimal cut-off values for absolute basophil,  
318 eosinophil and monocyte counts and NMR, above/below which the possibility of achieving  
319 favorable response after nCRT increased significantly (Table 7, Figure 4). The optimal cut-off value,  
320 which might distinguish patients with and without good response was not found only for dNLR.

321

322

323 **Table 7. Results of the ROC analysis for NMR, dNLR, absolute basophil, eosinophil and**  
 324 **monocyte counts, and relevant events.**

Characteristics	Absolute count			dNLR	NMR
	Basophil	Eosinophil	Monocyte		
<b>AUC ROC<sup>a</sup></b>	68.2%	66.1%	67.8%	62.9%	68.2%
<b>(95% CI)</b>	(56.5-79.9%)	(53.7-78.6%)	(55.5-80.1%)	(50.0-75.8%)	(55.6-80.8%)
<b>Likelihood ratio test<sup>b</sup></b>	<i>p</i> <0.01	<i>p</i> <0.05	<i>p</i> <0.05	<i>ns</i>	<i>p</i> <0.01
<b>ROC-cut-off value<sup>c</sup></b>	0.05	0.15	0.57	-	8.12
<b>Sensitivity</b>	53.8%	74.4%	61.5%	-	71.8%
<b>(95% CI)</b>	(38.5-69.2%)	(59.0-87.2%)	(46.2-76.9%)	-	(56.4-84.6%)
<b>Specificity</b>	80.0%	57.1%	71.4%	-	65.7%
<b>(95% CI)</b>	(65.7-91.4%)	(40.0-71.4%)	(57.1-85.7%)	-	(48.6-80.0%)

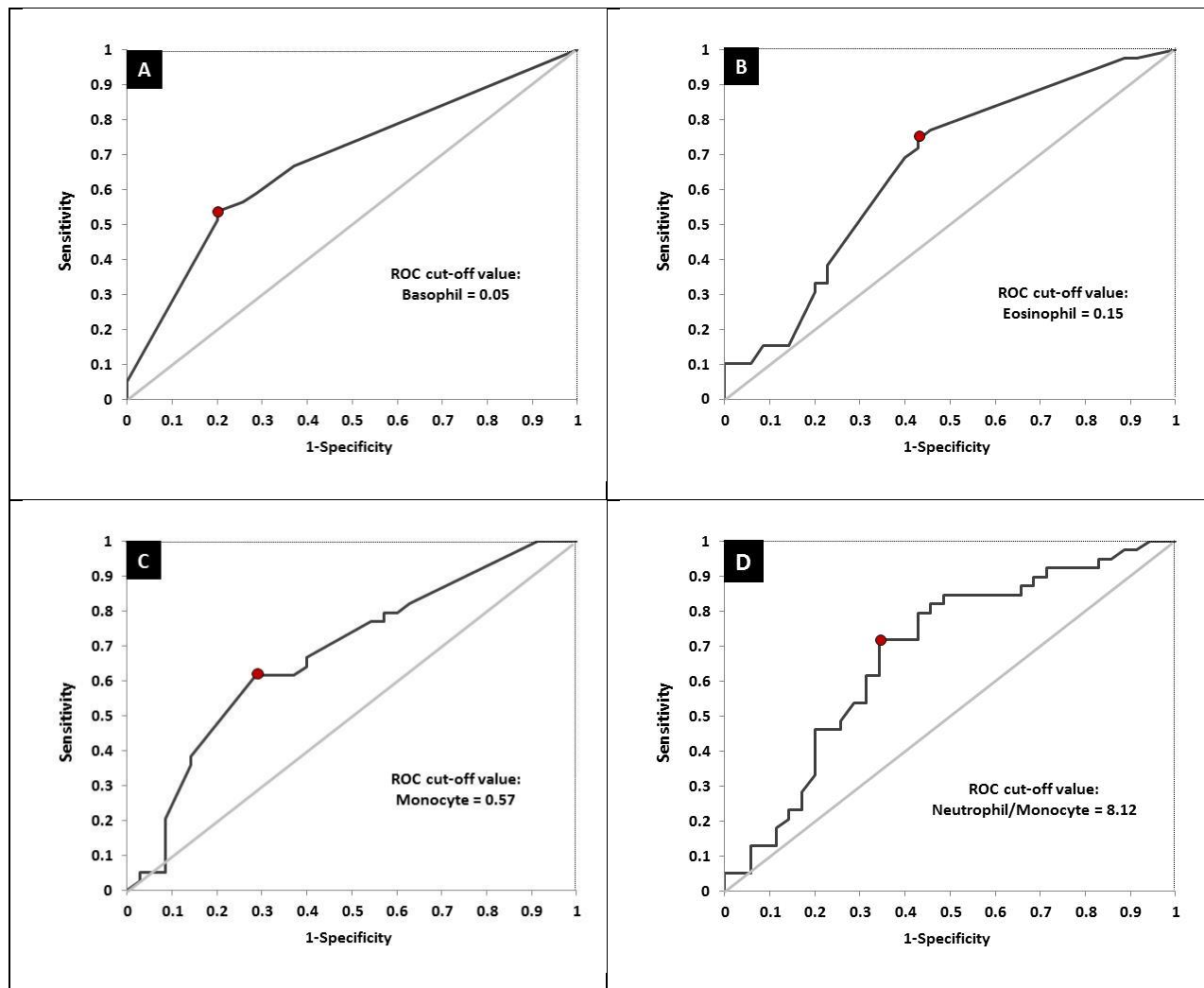
325 <sup>a</sup>Area Under the ROC curve (DeLong's method); <sup>b</sup>Likelihood ratio test for AUC ROC; <sup>c</sup>Value with maximum sensitivity  
 326 and specificity; *ns* - not statistically significant

327

328

329

330



331 **Figure 4. ROC curves for the absolute basophil count (A), absolute eosinophil count (B),**  
332 **absolute monocyte count (C) and NMR (D) in relation to response to treatment.**

333 Afterwards, differences between responders and non-responders according to the cut-off values  
334 obtained by ROC analysis were examined (Table 8). According to the achieved cut-off values a  
335 statistically significant difference in the response was confirmed for the initial basophil, eosinophil  
336 and monocyte counts ( $p < 0.01$  for all variables). Initial higher level of these parameters (greater  
337 than 0.05, 0.15, 0.57 respectively) were associated with unfavourable responses.

338

339



340 **Table 8. The value of NMR, dNLR, absolute basophil, eosinophil and monocyte counts in**  
 341 **prediction the response to nCRT.**

<b>Parameters (ROC cut-off value)</b>	<b>Response to nCRT</b>		
	<b>responders</b>	<b>non-responders</b>	<b>Pearson <math>\chi^2</math> test</b>
<b><i>Absolute basophil count</i></b>			
$\leq 0.05$	28 (80.0%)	18 (45.0%)	$p < 0.01$
$> 0.05$	7 (20.0%)	22 (55.0%)	
<b><i>Absolute eosinophil count</i></b>			
$\leq 0.15$	20 (57.1%)	10 (25.0%)	$p < 0.01$
$> 0.15$	15 (42.9%)	30 (75.0%)	
<b><i>Absolute monocyte count</i></b>			
$\leq 0.57$	25 (71.4%)	15 (37.5%)	$p < 0.01$
$> 0.57$	10 (28.6%)	25 (62.5%)	
<b><i>Neutrophil-to-monocyte ratio</i></b>			
$\leq 8.12$	12 (34.3%)	28 (70.0%)	$p < 0.05$
$> 8.12$	32 (65.7%)	12 (30.0%)	
<b><i>Total</i></b>	<b>35 (46.7%)</b>	<b>40 (53.3%)</b>	-

342

343 Significant variables from the analyses were then used for the construction of a logistic regression  
 344 model. The UICC staging was included as parameter which unit T and N stadium of disease and has  
 345 high clinical importance. Finally, the model comprised ten variables: UICC staging, tumor  
 346 differentiation, mucinous histological type, tumor morphology, stenotic character, extramural  
 347 vascular invasion, as well as NMR, absolute basophil, eosinophil and monocyte counts (Table 9).  
 348 After univariate analyses were conducted, the extremely high OR values were observed for tumor  
 349 differentiation and mucinous histological type categories. These values were in correlation with the  
 350 fact that all patients with mucinous histological type and/or poorly differentiated tumor had achieved  
 351 bad response. Previously mentioned parameters as well as UICC staging were excluded after  
 352 univariate analyses. The final model included tumor morphology, NMR, absolute basophil,  
 353 eosinophil, and monocyte counts.

**Table 9. Logistic regression analysis of the response to nCRT.**

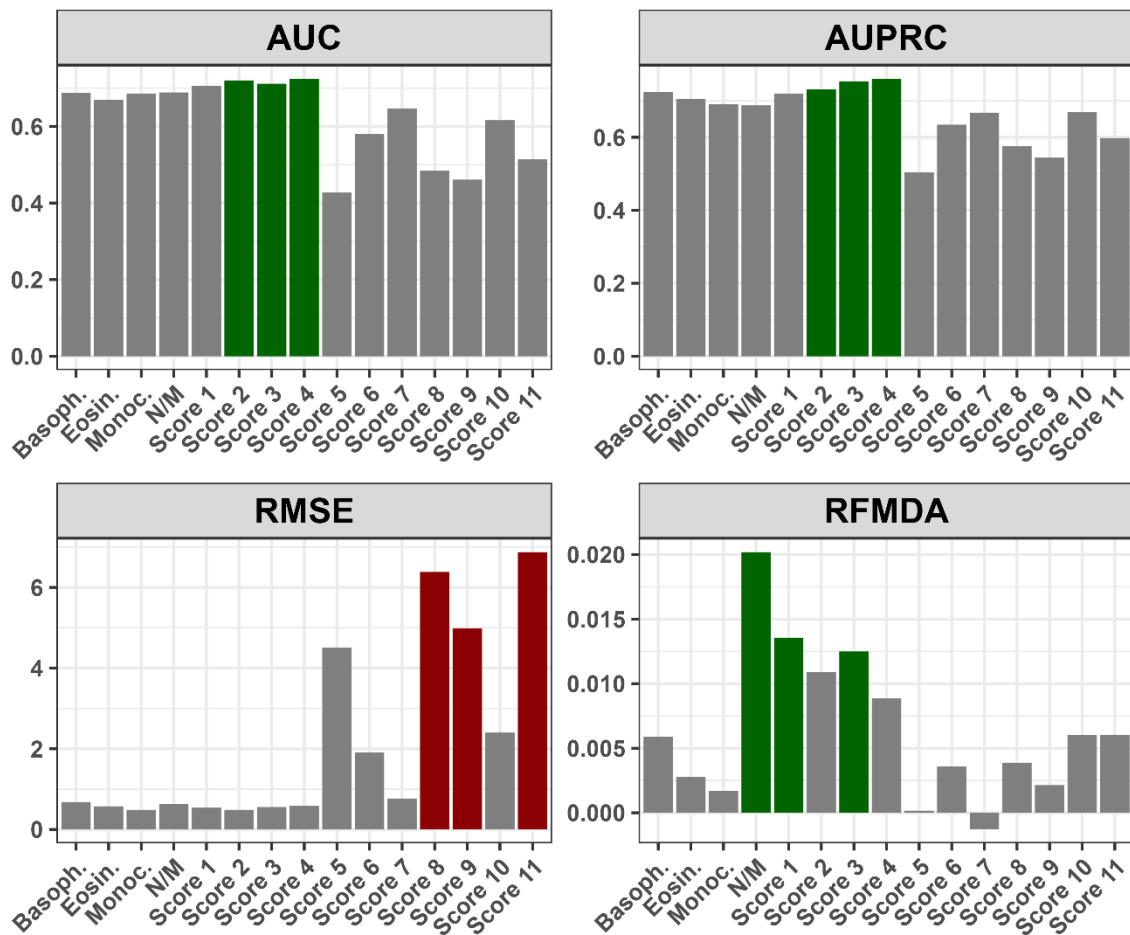
Characteristic	Logistic regression			
	Univariate		Multivariate	
	Odds Ratio (95%CI)	Wald test	Odds Ratio (95%CI)	Likelihood Ratio test
<b><i>UICC staging<sup>s</sup></i></b>				
IIIC vs. IIA+ IIIA+ IIIB	1.69 (0.67-4.26)	<i>p</i> =0.265	-	-
<b><i>Tumor differentiation</i></b>				
Poor vs. Well and moderate	43.8*10 <sup>6</sup> (0-∞)	<i>p</i> =0.991	-	-
<b><i>Mucinous histological type</i></b>				
Yes vs. No	14.99*10 <sup>7</sup> (0-∞)	<i>p</i> =0.992	-	-
<b><i>Tumor morphology</i></b>				
Annular vs. Polypoid and semi-annular	4.67 (1.47-14.82)	<i>p</i> =0.009	10.11 (1.81-56.39)	<i>p</i> <sup>#</sup> =0.008
<b><i>Stenotic character</i></b>				
Yes vs. No	3.57 (1.21-10.52)	<i>p</i> =0.021	-	<i>p</i> <sup>#</sup> =0.230
<b><i>Extramural vascular invasion</i></b>				
Yes vs. No	3.27 (1.16- 9.23)	<i>p</i> =0.025	-	<i>p</i> <sup>#</sup> =0.131
<b><i>Absolute basophil count</i></b>				
> 0.05 vs. ≤ 0.05	4.89 (1.73-13.78)	<i>p</i> =0.003	4.55 (1.21-17.13)	<i>p</i> <sup>#</sup> =0.025
<b><i>Absolute eosinophil count</i></b>				
> 0.15 vs. ≤ 0.15	4.00 (1.50-10.66)	<i>p</i> =0.005	3.86 (1.09-13.71)	<i>p</i> <sup>#</sup> =0.037
<b><i>Absolute monocyte count</i></b>				
> 0.57 vs. ≤ 0.57	4.17 (1.57-11.03)	<i>p</i> =0.004	3.46 (1.01-11.89)	<i>p</i> <sup>#</sup> =0.049
<b><i>NMR<sup>e</sup></i></b>				
≤ 8.12 vs. > 8.12	4.47 (1.69-11.82)	<i>p</i> =0.003	6.38 (1.74-23.39)	<i>p</i> <sup>#</sup> =0.005

355 <sup>s</sup>UICC - Union for International Cancer Control; <sup>e</sup>NMR - Neutrophil-to-monocyte ratio; <sup>#</sup>Wild test.

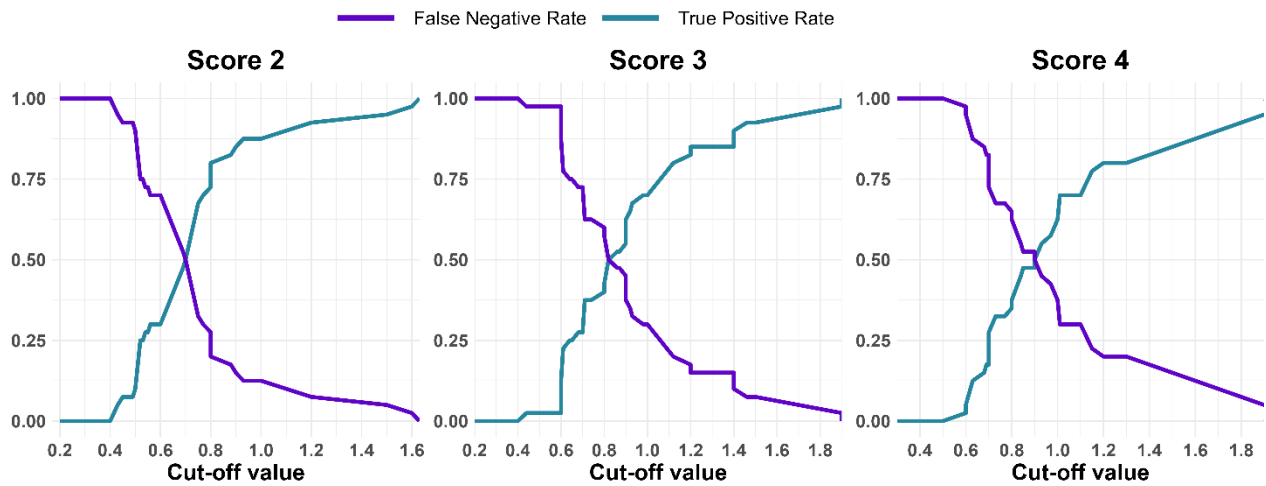
356

357 The numerical variables that remained significant in the multivariate analysis were utilized to create  
 358 eleven different composite scores. These scores were calculated using various combinations of the  
 359 significant variables (Supplementary Table 2). The best predictive power was observed when the  
 360 initial eosinophil, basophil, and monocyte counts were combined (Figure 5). The changes in the  
 361 false negative and true positive rates for the top three composite scores with respect to different cut-  
 362 off values of these three scores are shown in Figure 6.

363



364  
 365 **Figure 5. Performance of the composite scores with respect to various metrics: AUC - Area**  
 366 **Under Curve; AUPRC - Area Under Precision-Recall Curve; RMSE - Root Mean Square**  
 367 **Error; RFMDA - Random Forest Mean Decrease in Accuracy; Basoph. - Absolute basophil**  
 368 **count; Eosin. - Absolute eosinophil count; Monoc. - Absolute monocyte count; N/M -**  
 369 **Neutrophil-to-monocyte ratio; Score 1 - Absolute basophil + eosinophil count; Score 2 -**  
 370 **Absolute basophil + monocyte count; Score 3 - Absolute eosinophil + monocyte count; Score**  
 371 **4 - Absolute basophil + eosinophil + monocyte count; Score 5 - Neutrophil-to-monocyte ratio**  
 372 **+ Absolute monocyte count; Score 6 - Neutrophil-to-monocyte ratio + Absolute eosinophil**  
 373 **count; Score 7 - Neutrophil-to-monocyte ratio + Absolute basophil count; Score 8 -**  
 374 **Neutrophil-to-monocyte ratio + Absolute monocyte + eosinophil count; Score 9 - Neutrophil-**  
 375 **to-monocyte ratio + Absolute monocyte + Absolute basophil count; Score 10 - Neutrophil-**  
 376 **to-monocyte ratio + Absolute eosinophil + basophil count; Score 11 - Neutrophil-to-monocyte**  
 377 **ratio + Absolute monocyte + eosinophil count + basophil count;**



378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

**Figure 6. Relationship between False Negative and True Positive Rates for Top Three Composite Scores at Different Cut-off Values: Score 2 - Absolute basophil + monocyte count; Score 3 - Absolute eosinophil + monocyte count; Score 4 - Absolute basophil + eosinophil + monocyte count;**

## 400 Discussion

401 The optimal time for assessment of tumor response after nCRT, time for surgery, and how to profile  
402 the best candidates for the „watch and wait” approach is still unknown. In this study, we aimed to  
403 select patients who would benefit the most from an increase of RT dose and waiting periods longer  
404 than 6 weeks after nCRT completion according to initial clinical, pathological, radiological, and  
405 hematological parameters, as well as inflammation-related genetic biomarkers chosen by *in silico*  
406 analysis. The identification of these predictive clinical and molecular markers would enable the  
407 intensification of treatment in selected groups of patients. Better selection of patients with a higher  
408 probability of a favorable response to neoadjuvant treatment would contribute to the reduction of  
409 morbidity, while improving survival and local control of the disease. On the other hand, patients  
410 where a good response to neoadjuvant treatment is not expected would be candidates for other  
411 treatment modalities in the initial approach, such as induction polychemotherapy, application of  
412 target therapy (e.g. epidermal growth factor receptor inhibitors) or surgery without delay after  
413 completion of neoadjuvant treatment.

414 In some cases, pelvic MRI scan performed at 8th week after nCRT completion cannot clearly  
415 distinguish residual tumor due to post treatment changes and still probably did not achieve the  
416 maximum response. The sensitivity of clinical evaluation, according to our results, was 92%. It  
417 refers to high probability that disease evaluation will indicate an incomplete response when viable  
418 tumor cells are present, which is confirmed with pathohistological examination as a gold standard.  
419 Therefore, the combination of MRI scan and proctoscopy examination is beneficial when it comes  
420 to a group of patients who still have residual disease after nCRT, at which point operative treatment  
421 is indicated. On the other hand, lower specificity and negative predictive value (NPV) (76.92%,  
422 71.39%, respectively) suggest that this kind of evaluation is not selective enough for patients who  
423 are candidates for the “watch and wait” approach. This method is particularly important in the case  
424 of distally located rectal cancer when abdominoperineal resection is the only option. In our study,  
425 the majority of patients had distant located tumor (80%). The only way to confirm CR after nCRT  
426 is strict follow-up with reevaluation every 2-3 months in the first 2 years after treatment completion,  
427 followed by continuation of the protocols<sup>39</sup>. Evidence suggest that in the case of local regrowth,  
428 salvage surgery can be done in 95% of patients, which indicate the safety of this approach<sup>10</sup>.  
429 However, when near CR is found at the first assessment, the protocols are not well established yet.  
430 It is well known that prolongation of period after nCRT completion is associated with higher pCR  
431 rate<sup>40</sup>. In the case when primary response 6-8 weeks after treatment completion is close to cCR, it  
432 is beneficial for patients who are not candidates for sphincter preservation surgery to delay surgery  
433 with one more clinical assessment after 8-12 weeks in order to achieve the maximum response.  
434 Simpson at al. reported local regrowth on repeated assessment for 37% of patients whose response  
435 was defined as near CR<sup>41</sup>. Another article which investigated the role of prolongation of period

436 after nCRT in order to achieve the maximum response, found that 90% of patients with initial near  
437 CR at the first assessment were found to be cCR at the reassessment after 6-12 weeks<sup>42</sup>. On the  
438 other hand, delaying surgery in order to achieve better response is associated with a higher  
439 probability of distant metastases<sup>43</sup>. This fact can be related to local regrowth, but it has not been  
440 proved yet.

441 These circumstances stress the necessity of additional parameters which can guide the selection of  
442 patients who can be expected to achieve a complete response. Molecular markers in combination  
443 with good MRI and rigid proctoscopy examination may allow longer delays in surgery and one more  
444 pelvic MRI scan after 8-12 weeks. In this study, we aimed to investigate some genetic factors that  
445 were found to be promising candidates using *in silico* methods of previously published datasets.  
446 However, statistical significance between responders and non-responders in relation to expression  
447 of selected genes (*IL6*, *CYBB*, *CXCL9*, *IDO1*) was not reached. When comparisons were made in  
448 the subgroup of patients who were operated, a significantly higher expression of *IDO1* ( $p<0.05$ ) was  
449 found for TRG1 compared to TRG4. *IDO1* is critical for tryptophan metabolism, and is regarded to  
450 have a significant effect on the modulation of T-cell behavior and differentiation of regulatory T-  
451 cells. In a previous study which explored *IDO1* expression using immunohistochemistry in  
452 postoperative specimens, the relation to pathological response was not found ( $p=0.44$ ). The same  
453 study showed that higher expression of *IDO1* was associated with worse prognosis<sup>38</sup>. However,  
454 another study exploring nodal-positive LARC revealed that high *IDO1* expression in specimens  
455 after nCRT completion was associated with improved overall survival (OS)<sup>44</sup>. In our study, all but  
456 one patient were nodal-positive and our analyses were conducted on the initial specimens, which  
457 enabled us to analyze potential predictive biomarkers.

458 Concerning liquid biopsy parameters, periodic measurement of markers during patient follow up  
459 may also be crucial to prove the absence of the disease as well as for early detection of disease  
460 progression. This kind of approach has been investigated in metastatic colorectal cancer and it was  
461 shown that periodic sampling of liquid biopsy accompanied with ctDNA levels measurements can  
462 be valid for monitoring status of the disease and profile the response to treatment<sup>45</sup>.

463 The importance of EMVI as a prognostic factor in LARC setting is well established. By comparison  
464 of disease-free survival (DFS) between II and III stadium of disease, it was shown that independent  
465 from disease stadium, the presence of EMVI results in the worse prognosis<sup>46</sup>. The predictive role  
466 of EMVI has not been defined yet. Sun et al. found that EMVI status was the only factor by  
467 multivariate analysis which influences the response to treatment. The focus of this research was the  
468 role of initial MRI characteristics on treatment outcome of T3 LARC patients. Patients with ypT0-  
469 2N0 postoperative category were previously defined as good responders<sup>47</sup>. In our cohort, 33.3% of  
470 patients were EMVI positive, and it was shown that they were more likely to have poor response  
471 ( $p<0.05$ ). Worse response in EMVI positive group of patients can be connected with tumour hypoxia

472 and consequent radioresistancy, due to the fact that primary mechanism of radiotherapy  
473 effectiveness is formation of reactive oxygen radicals. Hypoxia in solid tumors is a well known  
474 problem because of insufficient vascularisation of rapid tumor growth. In order to resolve this in  
475 our study, we tried to increase the administrated dose per fraction on the gross disease region (2.16  
476 Gy/fraction). By doing this, we attempt to cause cell death related to direct DNA damage caused by  
477 radiation, and to overcome lower level of oxygen in some parts of the tumor. By combining pCR  
478 rate in group of patients where operative treatment was conducted and patients who were enrolled  
479 in “watch and wait“ program, we achieved 33.3% complete response rate. On the other hand, in  
480 EMVI positive group of patients, the complete response was achieved for only 16% of them. The  
481 option for this group of patients might be further dose escalation using adaptive MRI-guided  
482 radiotherapy which had shown potential for higher cCR rates and wider implementation of organ  
483 preservation approaches <sup>48</sup>.

484 The role of initial basophile count has been previously investigated as a prognostic factor in  
485 colorectal cancer. The association between lower basophile level and worse survival as well as  
486 aggressive tumor potential has been shown <sup>49</sup>. To the best of our knowledge, this is the first study  
487 to find the predictive role of basophile counts in the rectal cancer settings. Patients with an initial  
488 basophile count lower than 0.05 are more likely to achieve good response ( $p<0.01$ ). Similar results  
489 were found in advanced gastric cancer, where worse response to programmed death 1 inhibitor (anti-  
490 PD-1 inhibitor) plus chemotherapy was in correlation with a higher level of peripheral basophils <sup>50</sup>.

491 Comparing literature data on the predictive role of initial eosinophil counts, it has been proposed as  
492 a potential predictive marker for immunotherapy in lung cancer, with a higher levels detected in  
493 patients with better treatment outcome <sup>51</sup>. It was also found that higher initial level of eosinophil is  
494 connected to more effective outcomes when immunotherapy is administrated together with  
495 chemotherapy in advanced melanoma <sup>52</sup>. In our study, a higher initial eosinophil level is associated  
496 with worse response, which might be explained by different treatment modalities and the addition  
497 of the radiotherapy component.

498 Analyzing initial monocyte counts, a predictive role was previously reported in the CRC settings,  
499 with higher levels detected in patients with poor outcome <sup>53</sup>. The same was found in our research  
500 where the absolute monocyte levels a higher than 0.57 were related to worse response. The NMR  
501 has been investigated in low-risk differentiated thyroid carcinoma as a prognostic factor, and it was  
502 found that lower initial level is related to a worse prognosis, which is in relation to our findings <sup>54</sup>.  
503 Our group has previously shown that hematological parameters easily derived and routinely  
504 determined by low-cost and minimally invasive methods might be useful in predicting the response  
505 to chemoradiotherapy in patients with anal cancer <sup>7</sup>. Also, we successfully evaluated the role of  
506 hematological parameters in predicting the survival and toxicity to specific treatment in the lung  
507 cancer setting <sup>55</sup>.

508 According to our results, mucinous tumor differentiation was significantly associated with poor  
509 response ( $p < 0.01$ ). The study conducted by Simha et al. also found that presence of mucin is  
510 associated with larger residual disease and worse prognosis<sup>56</sup>. Previously it has also been described  
511 that mucinous rectal carcinoma is associated with a unique genetic pattern, including more frequent  
512 presence of microsatellite instability (MSI), which is caused by a defect in DNA mismatch repair  
513<sup>57</sup>. The connection of MSI in rectal carcinoma and poorer prognosis has also been reported<sup>58</sup>.  
514 Bearing it in mind, recently presented preliminary results with focus on usefulness of introduction  
515 of the anti-PD-1 inhibitor dostarlimab in patients with mismatch repair-deficient (dMMR) LARC  
516 patients can be promising to individualise treatment in this group of patients<sup>59</sup>.

517 This study has some limitations. The sample size is relatively low, but has met the criteria of a  
518 minimum number of LARC samples taking into account its incidence and population size in Serbia  
519 (95% confidence level)<sup>60</sup>. The evaluation of potential prognostic parameters has not been included,  
520 as the enrolled patients are currently under follow-up for long-term outcomes. The predictive model  
521 constructed in our study is currently being validated in an independent prospective cohort of patients  
522 with LARC treated with nCRT.

## 523 **Conclusions**

524 Based on the logistic regression model, important factors associated with favorable response to  
525 nCRT were tumor morphology and hematological parameters which can be easily and routinely  
526 derived from initial laboratory results (NMR, eosinophile, basophil and monocyte counts) in a  
527 minimally invasive manner. Here, we present evidence that a combined score derived by summing  
528 the initial absolute counts of basophils, eosinophils, and monocytes holds the highest predictive  
529 value and potential clinical utility. Further studies involving larger cohorts are necessary to validate  
530 these initial observations.

531 **Acknowledgements:** This study was funded by the Horizon Europe Project STEPUPIORS  
532 (Agreement No. 101079217) and the Ministry of Education and Science of the Republic of Serbia  
533 (Agreement No. 451-03-47/2023-01/200043).

534 **Disclosures:** The authors declare no conflict of interest.

535 **Study approval:** This study has been approved by the Ethics Committee of the Institute for  
536 Oncology and Radiology of Serbia and the Ethics Committee of the Faculty of Medicine, University  
537 of Belgrade, Serbia. All patients signed an informed consent. All experiments have been performed  
538 in accordance with the Helsinki Declaration of 1975, as revised in 2013.

539 **Data availability statement:** The data that support the findings of this study are available upon  
540 reasonable request from the corresponding author. The data are not publicly available due to ethics  
541 restrictions as their containing information could compromise the privacy of patients.



## 543 References

- 544 1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates  
545 of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.*  
546 2021;71(3):209-249. doi:<https://doi.org/10.3322/caac.21660>
- 547 2. Nikolic N, Radosavljevic D, Gavrilovic D, et al. Prognostic Factors for Post-Recurrence  
548 Survival in Stage II and III Colorectal Carcinoma Patients. *Medicina (Kaunas).*  
549 2021;57(10). doi:10.3390/medicina57101108
- 550 3. Cavic M, Krivokuca A, Boljevic I, et al. Pharmacogenetics in cancer therapy - 8 years of  
551 experience at the Institute for Oncology and Radiology of Serbia. *J BUON.* 2016;21:1287-  
552 1295.
- 553 4. Brotto K, Malisic E, Cavic M, Krivokuca A, Jankovic R. The Usability of Allele-Specific  
554 PCR and Reverse-Hybridization Assays for KRAS Genotyping in Serbian Colorectal  
555 Cancer Patients. *Dig Dis Sci.* 2012;58. doi:10.1007/s10620-012-2469-9
- 556 5. Jakovljevic K, Malisic E, Cavic M, Krivokuca A, Dobricic J, Jankovic R. KRAS and BRAF  
557 mutations in Serbian patients with colorectal cancer. *J BUON.* 2012;17(3):575-580.  
558 <http://www.ncbi.nlm.nih.gov/pubmed/23033302>
- 559 6. Stanojevic A, Samiotaki M, Lygirou V, et al. Data independent acquisition mass  
560 spectrometry (DIA-MS) analysis of FFPE rectal cancer samples offers in depth proteomics  
561 characterization of response to neoadjuvant chemoradiotherapy. *medRxiv*. Published online  
562 January 1, 2023:2023.05.12.23289671. doi:10.1101/2023.05.12.23289671
- 563 7. Stojanovic-Rundic S, Marinkovic M, Cavic M, et al. The role of haematological parameters  
564 in predicting the response to radical chemoradiotherapy in patients with anal squamous cell  
565 cancer. *Radiol Oncol.* 2021;55(4):449-458. doi:10.2478/raon-2021-0039
- 566 8. Zorcolo L, Rosman AS, Restivo A, et al. Complete pathologic response after combined  
567 modality treatment for rectal cancer and long-term survival: a meta-analysis. *Ann Surg*  
568 *Oncol.* 2012;19(9):2822-2832. doi:10.1245/s10434-011-2209-y
- 569 9. Maas M, Nelemans PJ, Valentini V, et al. Long-term outcome in patients with a  
570 pathological complete response after chemoradiation for rectal cancer: a pooled analysis of  
571 individual patient data. *Lancet Oncol.* 2010;11(9):835-844. doi:10.1016/S1470-  
572 2045(10)70172-8
- 573 10. Dossa F, Chesney TR, Acuna SA, Baxter NN. A watch-and-wait approach for locally  
574 advanced rectal cancer after a clinical complete response following neoadjuvant  
575 chemoradiation: a systematic review and meta-analysis. *lancet Gastroenterol Hepatol.*

- 576 2017;2(7):501-513. doi:10.1016/S2468-1253(17)30074-2
- 577 11. Habr-Gama A, Gama-Rodrigues J, São Julião GP, et al. Local recurrence after complete  
578 clinical response and watch and wait in rectal cancer after neoadjuvant chemoradiation:  
579 impact of salvage therapy on local disease control. *Int J Radiat Oncol Biol Phys*.  
580 2014;88(4):822-828. doi:10.1016/j.ijrobp.2013.12.012
- 581 12. Ferrari L, Fichera A. Neoadjuvant chemoradiation therapy and pathological complete  
582 response in rectal cancer. *Gastroenterol Rep*. 2015;3(4):277-288.  
583 doi:10.1093/gastro/gov039
- 584 13. Maas M, Beets-Tan RGH, Lambregts DMJ, et al. Wait-and-see policy for clinical complete  
585 responders after chemoradiation for rectal cancer. *J Clin Oncol Off J Am Soc Clin Oncol*.  
586 2011;29(35):4633-4640. doi:10.1200/JCO.2011.37.7176
- 587 14. Conde-Muñoz R, Cuadros M, Zambudio N, Segura-Jiménez I, Cano C, Palma P. Predictive  
588 Biomarkers to Chemoradiation in Locally Advanced Rectal Cancer. *Biomed Res Int*.  
589 2015;2015:921435. doi:10.1155/2015/921435
- 590 15. Li X-F, Jiang Z, Gao Y, Li C-X, Shen B-Z. Combination of three-gene  
591 immunohistochemical panel and magnetic resonance imaging-detected extramural vascular  
592 invasion to assess prognosis in non-advanced rectal cancer patients. *World J Gastroenterol*.  
593 2016;22(38):8576-8583. doi:10.3748/wjg.v22.i38.8576
- 594 16. Eaden JA, Abrams KR, Mayberry JF. The risk of colorectal cancer in ulcerative colitis: a  
595 meta-analysis. *Gut*. 2001;48(4):526-535. doi:10.1136/gut.48.4.526
- 596 17. Bajpai M, Seril DN, Van Gorp J, et al. Effect of Long-Term Mesalamine Therapy on  
597 Cancer-Associated Gene Expression in Colonic Mucosa of Patients with Ulcerative Colitis.  
598 *Dig Dis Sci*. 2019;64(3):740-750. doi:10.1007/s10620-018-5378-8
- 599 18. Schmitt M, Greten FR. The inflammatory pathogenesis of colorectal cancer. *Nat Rev*  
600 *Immunol*. 2021;21(10):653-667. doi:10.1038/s41577-021-00534-x
- 601 19. Long AG, Lundsmith ET, Hamilton KE. Inflammation and Colorectal Cancer. *Curr*  
602 *Colorectal Cancer Rep*. 2017;13(4):341-351. doi:10.1007/s11888-017-0373-6
- 603 20. Vuletić A, Mirjačić Martinović K, Tišma Miletić N, Zoidakis J, Castellvi-Bel S, Čavić M.  
604 Cross-Talk Between Tumor Cells Undergoing Epithelial to Mesenchymal Transition and  
605 Natural Killer Cells in Tumor Microenvironment in Colorectal Cancer. *Front cell Dev Biol*.  
606 2021;9:750022. doi:10.3389/fcell.2021.750022
- 607 21. Agostini M, Janssen K-P, Kim I-J, et al. An integrative approach for the identification of  
608 prognostic and predictive biomarkers in rectal cancer. *Oncotarget*. 2015;6(32):32561-

- 609 32574. doi:10.18632/oncotarget.4935
- 610 22. Millino C, Maretto I, Pacchioni B, et al. Gene and MicroRNA Expression Are Predictive of  
611 Tumor Response in Rectal Adenocarcinoma Patients Treated With Preoperative  
612 Chemoradiotherapy. *J Cell Physiol.* 2017;232(2):426-435. doi:10.1002/jcp.25441
- 613 23. Park IJ, Yu YS, Mustafa B, et al. A Nine-Gene Signature for Predicting the Response to  
614 Preoperative Chemoradiotherapy in Patients with Locally Advanced Rectal Cancer.  
615 *Cancers (Basel).* 2020;12(4). doi:10.3390/cancers12040800
- 616 24. Gim J, Cho YB, Hong HK, et al. Predicting multi-class responses to preoperative  
617 chemoradiotherapy in rectal cancer patients. *Radiat Oncol.* 2016;11:50.  
618 doi:10.1186/s13014-016-0623-9
- 619 25. Watanabe T, Komuro Y, Kiyomatsu T, et al. Prediction of sensitivity of rectal cancer cells  
620 in response to preoperative radiotherapy by DNA microarray analysis of gene expression  
621 profiles. *Cancer Res.* 2006;66(7):3370-3374. doi:10.1158/0008-5472.CAN-05-3834
- 622 26. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of  
623 Affymetrix GeneChip probe level data. *Nucleic Acids Res.* 2003;31(4):e15.  
624 doi:10.1093/nar/gng015
- 625 27. Smyth GK. Linear models and empirical bayes methods for assessing differential  
626 expression in microarray experiments. *Stat Appl Genet Mol Biol.* 2004;3:Article3.  
627 doi:10.2202/1544-6115.1027
- 628 28. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for  
629 RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.  
630 doi:10.1093/nar/gkv007
- 631 29. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-  
632 based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci.*  
633 2005;102(43):15545-15550. doi:10.1073/pnas.0506580102
- 634 30. Mootha VK, Lindgren CM, Eriksson K-F, et al. PGC-1 $\alpha$ -responsive genes involved in  
635 oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet.*  
636 2003;34(3):267-273. doi:10.1038/ng1180
- 637 31. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated  
638 models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498-2504.  
639 doi:10.1101/gr.1239303
- 640 32. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: new features for  
641 data integration and network visualization. *Bioinformatics.* 2011;27(3):431-432.

- 642 doi:10.1093/bioinformatics/btq675
- 643 33. O’Sullivan B, Brierley J, Byrd D, et al. The TNM classification of malignant tumours-  
644 towards common understanding and reasonable expectations. *Lancet Oncol.*  
645 2017;18(7):849-851. doi:10.1016/S1470-2045(17)30438-2
- 646 34. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern  
647 Cooperative Oncology Group. *Am J Clin Oncol.* 1982;5(6):649-655.
- 648 35. MedCalc Software Ltd. Diagnostic test evaluation calculator.  
649 [https://www.medcalc.org/calc/diagnostic\\_test.php](https://www.medcalc.org/calc/diagnostic_test.php)
- 650 36. Liaw A, Wiener M. Classification and Regression by RandomForest. *Forest.* 2001;23.
- 651 37. Chen B, Alvarado DM, Iticovici M, et al. Interferon-Induced IDO1 Mediates Radiation  
652 Resistance and Is a Therapeutic Target in Colorectal Cancer. *Cancer Immunol Res.*  
653 2020;8(4):451-464. doi:10.1158/2326-6066.CIR-19-0282
- 654 38. Takasu C, Nishi M, Yoshikawa K, et al. Role of IDO expression in patients with locally  
655 advanced rectal cancer treated with preoperative chemoradiotherapy. *BMC Cancer.*  
656 2022;22(1):1263. doi:10.1186/s12885-022-10357-1
- 657 39. Maas M, Lambregts DMJ, Nelemans PJ, et al. Assessment of Clinical Complete Response  
658 After Chemoradiation for Rectal Cancer with Digital Rectal Examination, Endoscopy, and  
659 MRI: Selection for Organ-Saving Treatment. *Ann Surg Oncol.* 2015;22(12):3873-3880.  
660 doi:10.1245/s10434-015-4687-9
- 661 40. Macchia G, Gambacorta MA, Masciocchi C, et al. Time to surgery and pathologic complete  
662 response after neoadjuvant chemoradiation in rectal cancer: A population study on 2094  
663 patients. *Clin Transl Radiat Oncol.* 2017;4:8-14. doi:10.1016/j.ctro.2017.04.004
- 664 41. Simpson G, Hopley P, Wilson J, et al. Long-term outcomes of real world “watch and wait”  
665 data for rectal cancer after neoadjuvant chemoradiotherapy. *Color Dis Off J Assoc*  
666 *Coloproctology Gt Britain Irel.* 2020;22(11):1568-1576. doi:10.1111/codi.15177
- 667 42. Hupkens BJP, Maas M, Martens MH, et al. Organ Preservation in Rectal Cancer After  
668 Chemoradiation: Should We Extend the Observation Period in Patients with a Clinical  
669 Near-Complete Response? *Ann Surg Oncol.* 2018;25(1):197-203. doi:10.1245/s10434-017-  
670 6213-8
- 671 43. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical  
672 complete responders after neoadjuvant treatment for rectal cancer in the International  
673 Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*  
674 *(London, England).* 2018;391(10139):2537-2545. doi:10.1016/S0140-6736(18)31078-X

- 675 44. Schollbach J, Kircher S, Wiegering A, et al. The local immune phenotype influences  
676 prognosis in patients with nodal-positive rectal cancer after neoadjuvant chemoradiation.  
677 *Int J Colorectal Dis.* 2020;35(2):365-370. doi:10.1007/s00384-019-03466-0
- 678 45. van 't Erve I, Medina JE, Leal A, et al. Metastatic colorectal cancer treatment response  
679 evaluation by ultra-deep sequencing of cell-free DNA and matched white blood cells. *Clin*  
680 *Cancer Res.* Published online December 19, 2022:CCR-22-2538. doi:10.1158/1078-  
681 0432.CCR-22-2538
- 682 46. Chand M, Bhangu A, Wotherspoon A, et al. EMVI-positive stage II rectal cancer has  
683 similar clinical outcomes as stage III disease following pre-operative chemoradiotherapy.  
684 *Ann Oncol.* 2014;25(4):858-863. doi:<https://doi.org/10.1093/annonc/mdu029>
- 685 47. Sun Y, Li J, Shen L, Wang X, Tong T, Gu Y. Predictive value of MRI-detected extramural  
686 vascular invasion in stage T3 rectal cancer patients before neoadjuvant chemoradiation.  
687 *Diagn Interv Radiol.* 2018;24(3):128-134. doi:10.5152/dir.2018.17286
- 688 48. Kensen CM, Betgen A, Wiersema L, et al. Online Adaptive MRI-Guided Radiotherapy for  
689 Primary Tumor and Lymph Node Boosting in Rectal Cancer. *Cancers (Basel).* 2023;15(4).  
690 doi:10.3390/cancers15041009
- 691 49. Liu Q, Luo D, Cai S, Li Q, Li X. Circulating basophil count as a prognostic marker of  
692 tumor aggressiveness and survival outcomes in colorectal cancer. *Clin Transl Med.*  
693 2020;9(1):6. doi:10.1186/s40169-019-0255-4
- 694 50. Wu C, Qiu Y, Zhang R, et al. Association of peripheral basophils with tumor M2  
695 macrophage infiltration and outcomes of the anti-PD-1 inhibitor plus chemotherapy  
696 combination in advanced gastric cancer. *J Transl Med.* 2022;20(1):386.  
697 doi:10.1186/s12967-022-03598-y
- 698 51. Caliman E, Fancelli S, Ottanelli C, et al. Absolute eosinophil count predicts clinical  
699 outcomes and toxicity in non-small cell lung cancer patients treated with immunotherapy.  
700 *Cancer Treat Res Commun.* 2022;32:100603. doi:10.1016/j.ctarc.2022.100603
- 701 52. Ferrucci PF, Gandini S, Cocorocchio E, et al. Baseline relative eosinophil count as a  
702 predictive biomarker for ipilimumab treatment in advanced melanoma. *Oncotarget.*  
703 2017;8(45):79809-79815. doi:10.18632/oncotarget.19748
- 704 53. Li Z, Xu Z, Huang Y, et al. The predictive value and the correlation of peripheral absolute  
705 monocyte count, tumor-associated macrophage and microvessel density in patients with  
706 colon cancer. *Medicine (Baltimore).* 2018;97(21):e10759.  
707 doi:10.1097/MD.00000000000010759

- 708 54. Offi C, Romano RM, Cangiano A, Candela G, Docimo G. Clinical significance of  
709 neutrophil-to-lymphocyte ratio, lymphocyte-to-monocyte ratio, platelet-to-lymphocyte  
710 ratio and prognostic nutritional index in low-risk differentiated thyroid carcinoma. *Acta*  
711 *Otorhinolaryngol Ital organo Uff della Soc Ital di Otorinolaringol e Chir Cerv-facc.*  
712 2021;41(1):31-38. doi:10.14639/0392-100X-N1089
- 713 55. Jokic V, Savic-Vujovic K, Spasic J, et al. Hematological parameters in EGFR-mutated  
714 advanced NSCLC patients treated with TKIs: predicting survival and toxicity. *Expert Rev*  
715 *Anticancer Ther.* Published online March 2021:1-7. doi:10.1080/14737140.2021.1893694
- 716 56. Simha V, Kapoor R, Gupta R, Bahl A, Nada R. Mucinous adenocarcinoma of the rectum: a  
717 poor candidate for neo-adjuvant chemoradiation? *J Gastrointest Oncol.* 2014;5(4):276-279.  
718 doi:10.3978/j.issn.2078-6891.2014.020
- 719 57. Mekenkamp LJM, Heesterbeek KJ, Koopman M, et al. Mucinous adenocarcinomas: poor  
720 prognosis in metastatic colorectal cancer. *Eur J Cancer.* 2012;48(4):501-509.  
721 doi:10.1016/j.ejca.2011.12.004
- 722 58. Hasan S, Renz P, Wegner RE, et al. Microsatellite Instability (MSI) as an Independent  
723 Predictor of Pathologic Complete Response (PCR) in Locally Advanced Rectal Cancer: A  
724 National Cancer Database (NCDB) Analysis. *Ann Surg.* 2020;271(4):716-723.  
725 doi:10.1097/SLA.0000000000003051
- 726 59. Cercek A, Lumish M, Sinopoli J, et al. PD-1 Blockade in Mismatch Repair–Deficient,  
727 Locally Advanced Rectal Cancer. *N Engl J Med.* 2022;386(25):2363-2376.  
728 doi:10.1056/NEJMoa2201445
- 729 60. Flikkema RM, Toledo-Pereyra LH. Sample size determination in medical and surgical  
730 research. *J Investig Surg Off J Acad Surg Res.* 2012;25(1):3-7.  
731 doi:10.3109/08941939.2011.648868

732 **Table and Figure legends**

733 **Table 1. GSEA Hallmark analysis in relation to Hallmark inflammatory response pathway.**

734 **Table 2. Comparison between responders and non-responders within analyzed datasets in**  
735 **relation to expression of selected genes.**

736 **Table 3. Patients', disease, treatment and outcomes characteristics.**

737 **Table 4. Correlation of clinical and pathological CR within a group of patients where**  
738 **operative treatment was conducted.**

739 **Table 5. Sensitivity, specificity, positive predictive value, negative predictive value and**  
740 **predictive accuracy of clinical evaluation for prediction of disease status using pathological**  
741 **examination as a gold standard.**

742 **Table 6. Comparison of characteristics of responders and non-responders to neoadjuvant**  
743 **chemoradiotherapy.**

744 **Table 7. Results of the ROC analysis for NMR, dNLR, absolute basophil, eosinophil and**  
745 **monocyte counts, and relevant events.**

746 **Table 8. The value of NMR, dNLR, absolute basophil, eosinophil and monocyte counts in**  
747 **prediction the response to nCRT.**

748 **Table 9. Logistic regression analysis of the response to nCRT.**

749 **Supplementary Material 1. The results of the differential expression analysis within GSE**  
750 **139255.**

751 **Supplementary Material 2. The top 100 genes from each database.**

752 **Supplementary Table 1. Comparison of genes expression of patients with pCR and those with**  
753 **TRG4 postoperative category.**

754 **Supplementary Table 2. Performance of the composite scores with respect to various metrics.**

755 **Figure 1. Volcano plots for selected datasets: GSE45404\_570 (A); GSE46862 (B); GSE139255**  
756 **(C).**

757 **Figure 2. GSEA enrichment plots for genes included in Hallmark inflammatory response**  
758 **pathway: GSE45404\_570 (A); GSE46862 (B); GSE139255 (C).**

759 **Figure 3. Venn diagram showing overlapping of inflammation-related genes between three**  
760 **analyzed datasets.**

761 **Figure 4. ROC curves for the absolute basophil count (A), absolute eosinophil count (B),**  
762 **absolute monocyte count (C) and NMR (D) in relation to response to treatment.**

763 **Figure 5. Performance of the composite scores with respect to various metrics: AUC - Area**  
764 **Under Curve; AUPRC - Area Under Precision-Recall Curve; RMSE - Root Mean Square**  
765 **Error; RFMDA - Random Forest Mean Decrease in Accuracy; Basoph. - Absolute basophil**  
766 **count; Eosin. - Absolute eosinophil count; Monoc. - Absolute monocyte count; N/M -**  
767 **Neutrophil-to-monocyte ratio; Score 1 - Absolute basophil + eosinophil count; Score 2 -**  
768 **Absolute basophil + monocyte count; Score 3 - Absolute eosinophil + monocyte count; Score**  
769 **4 - Absolute basophil + eosinophil + monocyte count; Score 5 - Neutrophil-to-monocyte ratio**  
770 **+ Absolute monocyte count; Score 6 - Neutrophil-to-monocyte ratio + Absolute eosinophil**  
771 **count; Score 7 - Neutrophil-to-monocyte ratio + Absolute basophil count; Score 8 -**  
772 **Neutrophil-to-monocyte ratio + Absolute monocyte + eosinophil count; Score 9 - Neutrophil-**  
773 **to-monocyte ratio + Absolute monocyte + Absolute basophil count; Score 10 - Neutrophil-to-**  
774 **monocyte ratio + Absolute eosinophil + basophil count; Score 11 - Neutrophil-to-monocyte**  
775 **ratio + Absolute monocyte + eosinophil count + basophil count;**

776 **Figure 6. Relationship between False Negative and True Positive Rates for Top Three**  
777 **Composite Scores at Different Cut-off Values: Score 2 - Absolute basophil + monocyte count;**  
778 **Score 3 - Absolute eosinophil + monocyte count; Score 4 - Absolute basophil + eosinophil +**  
779 **monocyte count;**

780

781

782

783

784

785

786



787 **List of abbreviations:**

788 CRC - colorectal cancer

789 LARC - locally advanced rectal cancer

790 MRI - magnetic resonance imaging

791 nCRT – neoadjuvant chemoradiotherapy

792 pCR – pathologic complete response

793 cCR - clinical complete response

794 NCBI GEO - National center for biotechnology information gene expression omnibus

795 TRG - tumor regression grading

796 GSEA - gene set enrichment analysis

797 KEGG - Kyoto encyclopedia of genes and genomes

798 CT - computed tomography

799 RT – radiotherapy

800 VMAT-SIB - volumetric modulated arc therapy-simultaneous integrated boost technique

801 PR - partial response

802 FFPE - formaline-fixed paraffin-embedded

803 EDTA - ethylenediaminetetraacetic acid

804 CBC - complete blood count

805 NLR - neutrophil-to-lymphocyte ratio

806 PLR - platelet-to-lymphocyte ratio

807 dNLR - derived neutrophil-to-lymphocyte ratio

808 LMR - lymphocyte-to-monocyte ratio

809 PMR - platelet-to-monocyte ratio

810 NMR - neutrophil-to-monocyte ratio

811 UICC - Union for International Cancer Control

812 ECOG - Eastern cooperative oncology group

813 cDNA - complementary DNA

814 qRT-PCR - quantitative Real Time PCR

- 815 GAPDH - glyceraldehyde-3-phosphate dehydrogenase
- 816 SD - standard deviation
- 817 ROC - receiver operating characteristics curve
- 818 AUC ROC - area under the ROC curve
- 819 AUCRP - Area Under Precision-Recall Curve
- 820 RMSE - Root Mean Square Error
- 821 EMVI - extramural vascular invasion
- 822 RFMDA - Random Forest Mean Decrease in Accuracy
- 823 DFS - disease-free survival
- 824 anti-PD-1 inhibitor - programmed death 1 inhibitor
- 825 MSI - microsatellite instability
- 826 dMMR - mismatch repair-deficient