Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer

R. Dienstmann¹,²*, G. Villacampa¹, A. Sveen³,⁴, M. J. Mason², D. Niedziecki⁵, A. Nesbakken⁴,⁶, V. Moreno⁷, R. S. Warren⁸, R. A. Lothe³,⁴ & J. Guinney²

¹Oncology Data Science Group, Vall d’Hebron Institute of Oncology, Barcelona, Spain; ²Computational Oncology Group, Sage Bionetworks, Seattle, USA; ³Department of Molecular Oncology, Institute for Cancer Research and K.G. Jebsen Colorectal Cancer Research Centre, Oslo University Hospital, Oslo, Norway; ⁴Faculty of Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway; ⁵Department of Bioinformatics and Biostatistics, Duke University, Durham, USA; ⁶Department of Gastrointestinal Surgery, K.G. Jebsen Colorectal Cancer Research Centre, Oslo University Hospital, Oslo, Norway; ⁷Unit of Biomarkers and Susceptibility, Cancer Prevention and Control Program, Catalan Institute of Oncology, Oncobell Program of IDIBELL, CIBERESP, University of Barcelona, Barcelona, Spain; ⁸Department of Surgery, The Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, USA

*Correspondence to: Mr Rodrigo Dienstmann, Vall d’Hebron Institute of Oncology, Pg. Vall d’Hebron 119-129, Barcelona 08035, Spain. Tel: +34-932746077; Fax: +34-932746059; E-mail: rdienstmann@vhio.net

Background: It remains unknown to what extent consensus molecular subtype (CMS) groups and immune-stromal infiltration patterns improve our ability to predict outcomes over tumor–node–metastasis (TNM) staging and microsatellite instability (MSI) status in early-stage colorectal cancer (CRC).

Patients and methods: We carried out a comprehensive retrospective biomarker analysis of prognostic markers in adjuvant chemotherapy-untreated (N = 1656) and treated (N = 980), stage II (N = 1799) and III (N = 837) CRCs. We defined CMS scores and estimated CD8⁺ cytotoxic lymphocytes (CytoLym) and cancer-associated fibroblasts (CAF) infiltration scores from bulk tumor tissue transcriptomes (CMSclassifier and MCPcounter R packages); constructed a stratified multivariable Cox model for disease-free survival (DFS); and calculated the relative proportion of explained variation by each marker (clinicopathological [ClinPath], genomics [Gen: MSI, BRAF and KRAS mutations], CMS scores [CMS] and microenvironment cells [MicroCells: CytoLym+CAF]).

Results: In multivariable models, only ClinPath and MicroCells remained significant prognostic factors, with both CytoLym and CAF infiltration scores improving survival prediction beyond other markers. The explained variation for DFS models of ClinPath, MicroCells, Gen markers and CMS4 scores was 77%, 14%, 5.3% and 3.7%, respectively, in stage II; and 55.9%, 35.1%, 4.1% and 0.9%, respectively, in stage III. Patients whose tumors were CytoLym high/CAF low had better DFS than other strata [HR=0.71 (0.6–0.9); P = 0.004]. Microsatellite stable tumors had the strongest signal for improved outcomes with CytoLym high scores (interaction P = 0.04) and the poor prognosis linked to high CAF scores was limited to stage III disease (interaction P = 0.04).

Conclusions: Our results confirm that tumor microenvironment infiltration patterns represent potent determinants of the risk for distant dissemination in early-stage CRC. Multivariable models suggest that the prognostic value of MSI and CMS groups is largely explained by CytoLym and CAF infiltration patterns.

Key words: colorectal cancer, prognosis, microsatellite instability, CMS, immune, stromal
Introduction

Management of locoregional colorectal cancer (CRC) is still largely dictated by tumor–node–metastasis (TNM) status at diagnosis, despite the multitude of prognostic biomarker research over the recent years. Depth of tumor wall invasion (pT1–pT4) and lymph node (LN) involvement (pN0–pN2) are strongly associated with 5-year disease-free survival (DFS), which ranges from 30% to 90% [1, 2]. The decision to offer adjuvant chemotherapy, particularly in stage II relies on clinicopathological features such as bowel obstruction or perforation at presentation, number of LNs examined, lymphovascular or perineural invasion and tumor grade [3]. The only molecular marker with clinical utility in early-stage CRC is microsatellite instability (MSI), which is universally recommended for diagnosis of Lynch Syndrome, but also defines a stage II population (irrespective of germline or sporadic background) with significantly improved outcome when treated with surgery alone [4]. There are no validated molecular markers to help identify patient populations with increased benefit from adjuvant chemotherapy in locoregional CRC.

Previously, we have estimated the value of clinicopathological features, MSI and additional molecular markers, namely *BRAF*V600E and *KRAS* exon 2 mutations, in prognostic models of stage II/III (early-stage) CRC [4]. We found that incorporation of MSI and driver gene mutation status to overall survival models with TNM staging does improve the prognostic discriminatory power, but only modestly increases prediction accuracy in multivariable models that include detailed clinicopathological annotation, particularly in chemotherapy-treated patients. Importantly, there was only one genomically defined subgroup of CRC with consistently higher risk of death across multiple cohorts, namely patients whose tumors were microsatellite stable (MSS) and *BRAF*V600E mutated, which corresponded to only 6% of the stage II/III population. Discovery of additional clinically relevant subgroups with potentially targetable pathway dependencies and/or cancer microenvironment features has the potential to guide adjuvant treatment strategies in broader patient populations.

Tumor microenvironment markers have demonstrated independent prognostic value in stage II/III CRC, with a high density of CD8+ cytotoxic T lymphocyte (CytoLym) infiltration being consistently associated with prolonged survival [5–8]. Indeed, an immunohistochemistry-based scoring system has been developed (termed Immunoscore®) to quantify cytotoxic and memory T cells in the tumor center/invasive margin, which proved to be a strong prognostic index in early-stage CRC [8]. Several authors have suggested that the dense CytoLym infiltrate could explain the better prognosis of tumors displaying MSI, compared with MSS CRC [9, 10]. In addition, a subset of MSS tumors harbors prominent expression of immune cytotoxic markers (up to 30% of the stage II/III population), and multivariable models revealed that Immunoscore® was superior to MSI status in predicting disease-specific recurrence and patient survival [10]. Furthermore, gene expression signatures that reflect epithelial–mesenchymal transition and tumor infiltration with cancer-associated fibroblasts (CAFs) also constitute independent prognostic indicators in early stages, with dismal outcomes for patients with an invasive stromal-rich microenvironment [11, 12]. These markers are not as thoroughly validated as Immunoscore®), but they are reflected in the transcriptomic-based consensus molecular subtypes (CMS) of CRC, which have also been linked to patient outcome in early-stage CRC. High CAFs content (CMS4 Mesenchymal) and CytoLym infiltration (CMS1 MSI Immune) are important features of the tumor microenvironment in early-stage CRC, with poor and good prognosis, respectively. The dismal DFS rates for CMS4 tumors is independent of clinicopathological markers, *KRAS* or *BRAF*V600E mutations and MSI status [13, 14]. Interestingly, CMS4 tumors have high expression of genes specific for both CAFs and CytoLym, counterbalanced by immunosuppressive cells, such as Treg cells, myeloid-derived suppressor cells, monocyte derived cells and T N17 cells [15], previously linked to chemotherapy resistance [16, 17].

Given the known interactions between MSI, immune-stromal markers and mesenchymal activation states, only an unbiased multivariable prognostic model that incorporates TNM staging, genomic markers, transcriptomic subtypes and microenvironment features can identify the most critical drivers of disease recurrence in CRC. It is unknown to what extent the combined analysis of immune and stromal infiltration patterns improves prediction of DFS over traditional clinicopathological/molecular markers in stage II/III CRC and whether the potential prognostic effect of microenvironment cells is modulated by adjuvant chemotherapies. Here we describe the results of a comprehensive retrospective biomarker analysis of prognostic markers in chemotherapy-untreated and treated stage II/III CRCs from a large aggregated cohort of clinical studies with molecular data. Our hypothesis was that microenvironment features constitute stronger determinants of disease recurrence in early-stage CRC than genomic or transcriptomic subtypes. We aimed to perform a holistic assessment of the prognostic value of well-known intrinsic biological features of CRC.

Methods—patient population and molecular data

To perform this analysis, we aggregated data from multiple public cohorts and collaborated with different academic groups to have access to private data from prospective series and one clinical trial (Alliance CALGB9581). This project was approved by the Vall d’Hebron Institute of Oncology Ethics Committee. Patients signed informed consent for exploratory biomarker research on samples prospectively collected in accordance with the guidelines of Institutional Review Boards from each organization/clinical trial. Table 1 summarizes the final study population that included 2636 patients diagnosed with stage II/III CRC, untreated (N = 1656) or treated (N = 980) with adjuvant chemotherapy, with clinicopathological and molecular annotation for variables of interest. Transcriptomic data were normalized following standard bioinformatics procedures for CMSclassifier and MCPcounter application independently in each cohort (see supplementary Table S1, available at Annals of Oncology online for details on gene expression platform and tissue source). We obtained CMS1, CMS2, CMS3 and CMS4 Random Forest posterior probabilities as a continuous value (each ranging from 0–1) and final CMS labels using CMSclassifier R-package [13].

Likewise, the abundance of immune- and nonimmune-stromal cell populations was estimated from gene expression data using **MCPcounter** R-package [18]. Given the fact that **MCPcounter** scores are affected by gene expression platform and tissue source, microenvironment cell infiltration scores were scaled (from 0 to 1) first within three subgroups [Affymetrix in fresh frozen samples; Agilent in fresh frozen samples; and Almac-Affymetrix in formalin-fixed paraffin-embedded (FFPE) samples] and then rescaled after data aggregation to facilitate cross-study comparisons. Multiple imputation of random missing values was carried out via the **mice** R package in the aggregated cohort (supplementary Table S2, available at *Annals of Oncology* online).

### Table 1. Patient’s and tumor characteristics

<table>
<thead>
<tr>
<th>Dataset public</th>
<th>Variable</th>
<th>All (N = 2636)</th>
<th>No adjuvant chemotherapy (N = 1656)</th>
<th>Adjuvant chemotherapy (N = 980)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>E-MTAB-864</td>
<td>144</td>
<td>48.37%</td>
<td>144</td>
<td>49.15%</td>
</tr>
<tr>
<td>GSE14333</td>
<td>183</td>
<td>69.32%</td>
<td>138</td>
<td>69.19%</td>
</tr>
<tr>
<td>GSE17536</td>
<td>109</td>
<td>39.9%</td>
<td>109</td>
<td>39.9%</td>
</tr>
<tr>
<td>GSE24550</td>
<td>76</td>
<td>28.6%</td>
<td>76</td>
<td>28.6%</td>
</tr>
<tr>
<td>GSE31595</td>
<td>37</td>
<td>14.0%</td>
<td>37</td>
<td>14.0%</td>
</tr>
<tr>
<td>GSE33113</td>
<td>87</td>
<td>32.2%</td>
<td>87</td>
<td>32.2%</td>
</tr>
<tr>
<td>GSE37892</td>
<td>128</td>
<td>48.4%</td>
<td>128</td>
<td>48.4%</td>
</tr>
<tr>
<td>GSE38832</td>
<td>72</td>
<td>27.5%</td>
<td>72</td>
<td>27.5%</td>
</tr>
<tr>
<td>GSE39582</td>
<td>439</td>
<td>163.8%</td>
<td>439</td>
<td>163.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dataset private</th>
<th>Variable</th>
<th>All (N = 2636)</th>
<th>No adjuvant chemotherapy (N = 1656)</th>
<th>Adjuvant chemotherapy (N = 980)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRCSC</td>
<td>589</td>
<td>51.63%</td>
<td>110</td>
<td>50.85%</td>
</tr>
<tr>
<td>Oslo</td>
<td>281</td>
<td>108.0%</td>
<td>241</td>
<td>106.9%</td>
</tr>
<tr>
<td>CALGB9581</td>
<td>393</td>
<td>152.4%</td>
<td>393</td>
<td>152.4%</td>
</tr>
<tr>
<td>Colonomics</td>
<td>98</td>
<td>37.3%</td>
<td>98</td>
<td>37.3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Median (min—max)</th>
<th>69 (22 – 97)</th>
<th>70 (24–97)</th>
<th>65 (22–97)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Male</td>
<td>1382 (52.43%)</td>
<td>867 (52.36%)</td>
<td>515 (52.55%)</td>
</tr>
<tr>
<td>Stage</td>
<td>T1-2N0</td>
<td>5 (0.19%)</td>
<td>5 (0.30%)</td>
<td>0 (0%</td>
</tr>
<tr>
<td>Primary site</td>
<td>Right colon</td>
<td>1292 (49.01%)</td>
<td>860 (51.94%)</td>
<td>432 (46.08%)</td>
</tr>
<tr>
<td>Microsatellite status</td>
<td>Stable (MSS)</td>
<td>2184 (82.85%)</td>
<td>1343 (81.10%)</td>
<td>841 (85.82%)</td>
</tr>
<tr>
<td>BRAF&lt;sup&gt;V600E&lt;/sup&gt; status</td>
<td>Wt</td>
<td>2306 (87.48%)</td>
<td>1414 (85.39%)</td>
<td>892 (91.02%)</td>
</tr>
<tr>
<td>KRAS codons 12/13 status</td>
<td>Wt</td>
<td>1684 (63.88%)</td>
<td>1057 (63.83%)</td>
<td>627 (63.98%)</td>
</tr>
<tr>
<td>CMS label</td>
<td>CMS1</td>
<td>507 (19.23%)</td>
<td>339 (20.47%)</td>
<td>168 (17.14%)</td>
</tr>
<tr>
<td>CMS4 score</td>
<td>Median (IQR)</td>
<td>0.10 (0.02–0.38)</td>
<td>0.10 (0.02–0.35)</td>
<td>0.11 (0.02–0.43)</td>
</tr>
<tr>
<td>Cytotoxic lymphocyte (CytoLym) infiltration score</td>
<td>Median (IQR)</td>
<td>0.43 (0.38–0.49)</td>
<td>0.43 (0.37–0.5)</td>
<td>0.43 (0.38–0.49)</td>
</tr>
<tr>
<td>Cancer-associated fibroblast (CAF) infiltration score</td>
<td>Median (IQR)</td>
<td>0.59 (0.49–0.67)</td>
<td>0.59 (0.48–0.67)</td>
<td>0.59 (0.51–0.67)</td>
</tr>
</tbody>
</table>
**Methods—multivariable survival modeling**

Our primary end point was DFS, measured from time of cancer diagnosis until disease relapse or death from any cause. Survival was censored at 72 months based on median follow-up of patients alive. The covariates considered for inclusion in the prognostic models were as follows:

(i) Clinicopathological features (ClinPath): AJCC version 7 pathological tumor stage (pT-stage; pT1, pT2, pT3 and pT4) and pathological nodal stage (pN-stage; pN0, pN1 and pN2), age (continuous), sex (male versus female), primary tumor location [right (caecum to transverse colon) versus left (splenic flexure to sigmoid) or rectum].

(ii) Genomic markers (Gen): MSI status (MSI high versus MSS or MSI low), mutations in KRAS codons 12/13 or \(BRAF^{V600E}\) (versus wild-type).

(iii) Transcriptomic markers (CMS groups): CMS labels either as categorical variables (CMS1, CMS2, CMS3 or CMS4) or as independent continuous scores (CMS1 score, CMS2 score, CMS3 score, CMS4 score) from CMSclassifier R package [13].

(iv) Microenvironment infiltration markers (MicroCells): MCPcounter R package scores predict the abundance of 10 cell populations from transcriptomic profiles (CD3+ T cells, CD8+ T cells [CytoLym], CTLs [cytotoxic lymphocytes], NK [Natural Killer] cells, B lymphocytes, mononuclear lineage cells, myeloid dendritic cells, neutrophils, endothelial cells and CAFs) [18] as a continuous variable.

Study methodology is summarized in supplementary Figure S1, available at *Annals of Oncology* online. First, following data aggregation and imputation, we assessed proportional hazards assumption using survival R package. Sex did not hold the assumption of proportional hazards (\(P < 0.05\)) and was included in the survival models as a stratification variable, together with gene expression profiling subgroups as defined above (Affymetrix fresh-frozen; Agilent fresh-frozen; Affymetrix-Almac FFPE) and adjuvant chemotherapy status. In order to select variables within the CMS and MicroCells categories with the highest prognostic impact on DFS estimation, we carried out forward and backward stepwise regression using the Bayesian information criterion. Then, multivariable Cox proportional hazards models were formulated using all factors that demonstrated statistical significance for DFS in univariate models (with \(P < 0.1\) according to log-rank test). We investigated significant interactions among genomic, transcriptomic, microenvironment infiltration markers and clinicopathological features with impact on patient outcomes (\(P < 0.05\) according to ANOVA test). Next, the following multivariable models were compared: (i) ClinPath + Gen; (ii) ClinPath + Gen + CMS; (iii) ClinPath + Gen + MicroCells; and (iv) ClinPath + Gen + CMS + MicroCells. We then calculated, using multiple permutations, the relative proportion of explained variation in DFS that was accounted for by the different categories of predictor covariates using survMise R package [19]. For illustration purposes, continuous scores (CMS and micro-environment cell infiltration) were dichotomized based on the maximization of the log-rank statistic to generate Kaplan–Meier DFS curves using survminer R package. All analyses were carried out using R statistical software version 3.2.5 [20].

**Results**

Demographics, tumor-related characteristics and molecular markers of the numerous cohorts of patients with stage II or stage III CRC included in survival models are described in Table 1. Patients were recruited in the different studies between 1990 and 2005. Clinicopathological features of our population are in line with other prospective biomarker series (non-clinical trial cohorts), including a relatively elderly population, with stage II representing 85% and 35% of the cases in the untreated and chemotherapy-treated cohorts, respectively. Information on type of adjuvant chemotherapy is missing in most public cohorts, but given the standards of treatment at time of study recruitment, we estimate that <10% of the combined public–private treated populations received oxaliplatin in addition to 5-fluorouracil. The prevalence of molecular markers also mirrors published literature, with 17% MSI tumors, 36% KRAS codons 12/13 mutated, 13% \(BRAF^{V600E}\) mutated and 21% CMS4 tumors.

In the variable selection process, CMS4 as a continuous score was a better predictor of DFS than CMS 1/2/3 scores or CMS 1/2/3/4 labels. Also, from 10 MicroCells populations, CytoLym and CAFs infiltration scores were the strongest predictors of DFS. We found no significant inter-study heterogeneity when assessing CytoLym and CAFs scores as dichotomous variables in DFS models (supplementary Figure S2, available at *Annals of Oncology* online). On average, CytoLym infiltration scores were highest in CMS1 samples, while CAF scores were highest in CMS4 samples (supplementary Figure S3, available at *Annals of Oncology* online). Indeed, 66% of CMS1 tumors fall in the CytoLym high categories, ~65% of CMS2 and CMS3 samples are classified as CytoLym low/CAF low and 49% of CMS4 samples are CytoLym low/CAF high (supplementary Figure S3, available at *Annals of Oncology* online). We also found high association between CytoLym infiltration scores and microsatellite status, with 68% of MSI tumors in the CytoLym high categories (supplementary Figure S4, available at *Annals of Oncology* online).

Clinicopathological and genomic markers, together with continuous CMS4 scores and CytoLym/CAF infiltration scores, were assessed in univariate and multivariable models detailed in Table 2. Importantly, only age, pT stage, pN stage, primary tumor location and immune-stromal infiltration scores were independent prognostic factors in multivariable models. Patients with right-sided tumors had better DFS outcomes than left-sided primaries, both in stage II and III CRC. MSI status and CMS4 scores did not significantly improve prognosis prediction over ClinPath features when CytoLym and CAFs infiltration scores were considered. Analysis of deviance showed that the addition of CytoLym and CAF infiltration scores to ClinPath + Gen +/- CMS models provides significant prognostic information (ANOVA \(P < 0.05\)), as detailed in supplementary Table S3, available at *Annals of Oncology* online. We found a significant interaction between MSI status and CytoLym scores on DFS models (interaction \(P = 0.04\)), with MSS tumors having the strongest signal for improved outcomes when displaying CytoLym high infiltration scores.

Figure 1 illustrates adjusted Kaplan–Meier DFS curves by CytoLym and CAF infiltration scores in chemotherapy-untreated and treated cohorts, as well as stage II and III CRC populations. Overall, CytoLym high/CAF low population represented 16% of
samples, compared with 18% with CytoLym high/CAF high, 22% with CytoLym low/CAF high and 44% harboring CytoLym low/CAF low scores. Patients whose tumors were CytoLym high/CAF low had better DFS than other strata [adjusted HR = 0.71 (0.6–0.9); P = 0.004]. In both untreated and treated cohorts (Figure 1A and B), patients with CytoLym low/CAF high infiltration scores had the highest risk of disease recurrence. In the untreated population, patients with CytoLym high/CAF low tumors had significantly better DFS than CytoLym low/CAF high tumors (Figure 1A). In both untreated and stage II CRC, CytoLym infiltration scores were a critical determinant of prognosis, with significant differences in DFS when comparing CytoLym high versus CytoLym low categories (Figure 1A and C). The same associations were found in both chemotherapy-treated and stage III CRC, but CAF infiltration scores helped segregate the strata even further (Figure 1B and D). Indeed, the poor DFS outcomes linked to high CAF infiltration scores were limited to stage III CRC (interaction P = 0.04). We found no major differences in CytoLym and CAF infiltration scores between stage III low-risk (T1–3, N1) and high-risk (T4 or N2) groups (supplementary Figure S5, available at Annals of Oncology online). However, 31.3% of high-risk stage III tumors are classified as CytoLym low/CAF high, while 24.5% of low-risk stage III patients fall into this category. Individual Kaplan–Meier DFS curves stratified by stage plus adjuvant chemotherapy exposure can be found in supplementary Figure S6, available at Annals of Oncology online.

Figure 2 illustrates the relative contribution of different factors for prognosis prediction in early-stage CRC. Across all cohorts, the explained variation in multivariable DFS models for clinicopathological features, microenvironment infiltration scores, genomic markers and CMS4 scores were 83%, 11.4%, 3.7% and 1.9%, respectively. The relative contribution of clinicopathological features in prognosis prediction was larger in stage II as compared with stage III disease, while the impact of tumor microenvironment features on DFS is not diminished in more advanced stages of locoregional CRC (Figure 2C and D). In untreated and stage II cohorts (Figure 2A and C), in addition to CytoLym infiltration scores, age was an important contributor to DFS estimation, potentially associated with competing death risks. In adjuvant chemotherapy-treated and stage III cohorts (Figure 2B and C), CAF infiltration scores had substantial impact on DFS estimation.

Discussion

In this study, we investigated whether intrinsic gene expression signatures from cancer and microenvironment cells in stage II/III CRC have significant impact on DFS models adjusted for clinicopathological and genomic markers. For the first time, we show that the prognostic value of CMS groups is largely explained by tumor microenvironment infiltration patterns. In fact, CytoLym and CAF infiltration scores obviate the prognostic value of MSI status and CMS4 scores in multivariable models. We found that a high CytoLym infiltrated microenvironment is a critical determinant of improved outcomes and this ‘protective’ effect is strong across clinicopathological and genomic subgroups. Low microenvironment cytotoxic lymphocyte counts associates with higher chances of disease recurrence or death, particularly when tumors have a stromal invasive phenotype infiltrated with CAFs. Our results are in line with extensive literature built on pathology assessments of the tumor microenvironment showing the protective role of high infiltration by CD8+ cytotoxic T cells (approximately one-third of early-stage CRC population) [8] and the unfavorable outcome of patients whose tumors harbor high infiltration with stromal fibroblasts (~40% of early-stage CRC

original_article
population) [21]. Our data also suggest that the worse outcomes repeatedly seen in patients whose tumors display a mesenchymal-like phenotype may be directly linked to a prometastatic immune evasive and stromal-rich microenvironment. In fact, the impact of CAF infiltration scores on DFS estimation was larger in stage III and treated populations, which may be linked to a chemotherapy-resistance phenotype [16, 17]. Still, most CRC tumors have a ‘microenvironment desert’ phenotype, previously linked to epithelial CRC subtypes (CMS2 Canonical and CMS3 Metabolic) [15], known to have intermediate prognosis in early-stage disease. Here, the prognostic effect of cancer cell-intrinsic genomic markers may be more profound, as previously illustrated for KRAS mutations [22].

Although our results confirm the notion that activated immune cytotoxicity represents a potent determinant of the risk of distant dissemination, it is also known that immune and stromal infiltration patterns can be predicted by assessing intrinsic tumor cell epithelial expression profiles in early-stage CRC samples [23]. Moreover, interactions between immune cells, checkpoint expression and MSI status do have a clinically significant impact on prognosis. The expression levels of Immunoscore-like meta-genes confers favorable prognosis in CRC patients with MSS tumors displaying low levels of CytoLym and immune checkpoints [24]. On the other hand, the expression levels of immune checkpoints annuls the prognostic relevance of CytoLym in highly immunogenic colon tumors and predicts a poor outcome in MSI CRC patients. These results are in line with our data showing a stronger signal for improved outcomes with CytoLym high scores mainly in the MSS population. To summarize, these data suggest that microenvironment markers must be analyzed alongside cancer cell genomic and transcriptomic markers for proper risk stratification.

Our study has limitations related to the retrospective, non-randomized nature of most patient cohorts included in the analysis, with missing values for some genomic markers such as KRASV600E mutations and MSI status in the range of 30%–50% in some cohorts. Therefore, no definite conclusions can be obtained from our study on the role of genomic markers for prognostication in early-stage CRC. Larger and more contemporaneous cohorts are needed to investigate whether the prognostic effect of microenvironment cells is modulated by adjuvant chemotherapies, particularly oxaliplatin-based. On the other hand, our study describes the largest cohort of untreated early-stage CRC patients with clinical, genomic and transcriptomic features, the ideal setting for multivariable survival modeling. In addition, prospective non-clinical trial series included in our analyses are representative of ‘real world’ populations. Indeed, recent insights on clinicopathological prognostic markers, such as primary tumor location, were validated in our aggregated cohort. In a recent SEER database study and a British Columbia cohort, stage II right-sided cancers had better cause-specific survival and relapse-free survival than the left-sided cancers, respectively [25, 26]. In stage III disease, right-sided poorly differentiated mucinous adenocarcinoma showed significantly better survival than left-sided malignancies [25]. These results could be explained by the higher prevalence of MSI in right-sided high-grade tumors.
also found improved DFS for right-sided early-stage CRC when compared with left-sided disease, both in stage II and III disease. Interestingly, in a multivariable model, sidedness remained an independent prognostic factor when adjusting for MSI status. Our findings may be partly explained by unique patient characteristics included in population studies, such as a large proportion of elderly females.

To conclude, our data reinforce the idea that tumor microenvironment features should guide biomarker-drug co-development in the adjuvant setting. With recent advances in targeted immunotherapeutic interventions, we may be able to successfully boost anticancer immune cytotoxicity or inhibit immunosuppressive pathway dependencies that facilitate metastatic spread. The major impact of CAF infiltration scores on DFS of stage III and chemotherapy-treated populations deserves further validation in recent clinical trial cohorts, such as the IDEA consortium. We believe that our results will encourage clinical trial design with novel agents capable of reverting a chemotherapy-resistance phenotype linked to a distinct tumor microenvironment.

When combined with liquid biopsies for detection of minimal residual disease [27], tumor and microenvironment biomarkers will eventually help personalize adjuvant therapies in early-stage CRC.

**Funding**

This study was supported by Merck KGaA, Darmstadt, Germany (Grant for Oncology Innovation 2015, ‘Next generation of clinical trials with matched targeted therapies in colorectal cancer’). Merck KGaA reviewed the manuscript for medical accuracy only before journal submission. The authors are fully responsible for the content of this manuscript, and the views and opinions described in the publication reflect solely those of the authors. This work was also supported by the Norwegian Cancer Society (project numbers 6824048-2016 and 182759-2016) and the Research Council of Norway (project number 250993). VM received fund from the Instituto de Salud Carlos III, co-funded by FEDER funds—a way to build Europe—grant PI14-00613, the

---

**Figure 2.** Relative proportion of explained variation in disease-free survival (DFS) of the full multivariable model in chemotherapy untreated (A), chemotherapy treated (B), stage II (C) and stage III (D) CRC cohorts.
Agency for Management of University and Research Grants (AGAUR) of the Catalan Government grant 2017SGR723 and the Scientific Foundation of the Spanish Association Against Cancer (AECC). RSW was supported by a grant from the NIH (RO184018).

**Disclosure**

RD has received speaker fee from Roche, Symphogen, Ipsen, Amgen, Sanofi, Servier, MSD and had advisory roles for Roche and Novartis over the last 2 years. RD and JG received research grant from Merck from 2015 to 2017. RAL received a Board of Directors member fee from Inven2, an innovation company for research at Oslo University Hospital and the University of Oslo. VM has received research funds from Universal Diagnostics and Bioiberica and had advisory role for Ferrer. All remaining authors have declared no conflicts of interest.

**References**