A Stepwise Integrated Approach to Personalized Risk Predictions in Stage III Colorectal Cancer.

Manuela Salvucci  
*Royal College of Surgeons in Ireland, manuelasalvucci@rcsi.ie*

Maximilian L. Würstle  
*Royal College of Surgeons in Ireland, maximilianwurstle@rcsi.ie*

Clare Morgan  
*Royal College of Surgeons in Ireland, claremorgan@rcsi.ie*

Sarah Curry  
*Royal College of Surgeons in Ireland, sarahcurry@rcsi.ie*

Mattia Cremona  
*Royal College of Surgeons in Ireland, mattiacremona@rcsi.ie*

See next page for additional authors

Citation
A stepwise integrated approach to personalized risk predictions in stage III colorectal cancer

Manuela Salvucci1,2, Maximilian L. Würstle1,2, Clare Morgan1,3, Sarah Curry1,4, Mattia Cremona1,3, Andreas U. Lindner1,2, Orna Bacon1,5, Alexa J. Resler1,2, Aine C. Murphy1,2, Robert O’Byrne1,2, Lorna Flanagan1,2, Sonali Dasgupta6, Nadege Rice7, Camila Pilati7, Elisabeth Zink1,2, Lisa M. Schöller1,2, Sinead Toomey3, Mark Lawler6, Patrick G. Johnston6, Richard Wilson6, Sophie Camilleri-Broët8, Manuel Salto-Téllez6, Deborah A. McNamara5, Elaine W. Kay1,4, Pierre Laurent-Puig8, Sandra Van Schaeybroeck6, Bryan T. Hennessy1,3, Daniel B. Longley6, Markus Rehm1,2,9*, and Jochen H.M. Prehn1,2,*

1Centre for Systems Medicine, 2Department of Physiology & Medical Physics, 3Department of Medical Oncology, Royal College of Surgeons in Ireland, Dublin 2, Ireland; 4Department of Pathology, 5Department of Surgery, Beaumont Hospital, Dublin 9, Ireland; 6Centre for Cancer Research and Cell Biology, School of Medicine, Dentistry and Biomedical Science, Queen's University Belfast, Belfast, United Kingdom; 7Faculté de Médecine, 8UMR-S775 INSERM Laboratory, Université Paris Descartes, Paris, France. 9Institute of Cell Biology and Immunology, University of Stuttgart, Germany. *Shared senior authorship.

Corresponding author: Prof. Jochen H. M. Prehn, Department of Physiology and Medical Physics, Royal College of Surgeons in Ireland, 123 St. Stephen’s Green, Dublin 2, Ireland. Tel.: +353-1-402-2255; Fax: +353-1-402-2447; E-mail: prehn@rcsi.ie

Running Title: Apoptosis-based prognostic signatures in colorectal cancer

Character count: 53

Word Count: 4823 (excluding Title Page, Abstract, Translation Relevance Statement, References and Figure Captions).

Keywords: apoptosis / machine learning / mathematical modelling / prognostic biomarker / systems biology.
ABSTRACT

Purpose: Apoptosis is essential for chemotherapy responses. In this discovery and validation study, we evaluated the suitability of a mathematical model of apoptosis execution (APOPTO-CELL) as a stand-alone signature and as a constituent of further refined prognostic stratification tools.

Experimental Design: Apoptosis competency of primary tumor samples from n=120 stage III colorectal cancer patients was calculated by APOPTO-CELL from measured protein concentrations of Procaspase-3, Procaspase-9, SMAC and XIAP. An enriched APOPTO-CELL signature (APOPTO-CELL-PC3) was synthesized to capture apoptosome-independent effects of Caspase-3. Furthermore, a machine learning Random Forest approach was applied to APOPTO-CELL-PC3 and available molecular and clinicopathological data to identify a further enhanced signature. Association of the signature with prognosis was evaluated in an independent colon adenocarcinoma cohort (TCGA COAD, n=136).

Results: We identified three prognostic biomarkers (p=0.04, p=0.006 and p=0.0004 for APOPTO-CELL, APOPTO-CELL-PC3 and Random Forest signatures, respectively) with increasing stratification accuracy for stage III colorectal cancer patients. The APOPTO-CELL-PC3 signature ranked highest among all features. The prognostic value of the signatures was independently validated in stage III TCGA COAD patients (p=0.01, p=0.04 and p=0.02 for APOPTO-CELL, APOPTO-CELL-PC3 and Random Forest signatures, respectively). The signatures provided further stratification for patients of CMS1-3 molecular subtype.

Conclusions: The integration of a systems-biology-based biomarker for apoptosis competency with machine learning approaches is an appealing and innovative strategy towards refined patient stratification. The prognostic value of apoptosis competency is independent of other available clinicopathological and molecular factors, with tangible potential of being introduced in the clinical management of stage III colorectal patients.

Word count: 249 (max 250)
Despite extensive research efforts, powerful biomarkers capable of identifying stage III CRC patients at high-risk of chemotherapy-resistance, and thus relapse, still remain an unmet clinical demand. Lack of response to chemotherapeutics has been ascribed to defective apoptosis. Using an integrative stepwise approach, we investigated the potential clinical utility of a mathematical model of caspase activation for identifying patient-specific risk of unfavorable outcomes.

Among all clinicopathological, demographic, and molecular predictors analyzed, enriched-apoptosis systems modeling delivered the highest ranking independent prognostic biomarker. Furthermore, apoptosis modelling can be combined with molecular tumor subtyping to further refine risk predictions.

We report the clinical validation of diagnostic tools for stage III CRC patients that could deliver superior quality of care by personalizing cancer treatment.

**Word count:** 119.
Colorectal cancer (CRC) is a leading contributor to morbidity and mortality worldwide (1). CRC incidence is expected to rise due to longer life expectancy and lifestyle changes, such as poor diet and decreased physical activity (1). The current standard-of-care for stage III CRC patients prescribes surgical resection followed by 5-fluorouracil (5-FU) based adjuvant chemotherapy (2). Response to treatment varies greatly, and numerous studies have attempted to reconcile differential response to chemotherapy with the underlying molecular characteristics of the primary tumor (3). The mechanism of action of chemotherapeutic agents such as 5-FU is centered on their DNA-damaging effects and ability to induce cell cycle arrest and apoptosis (4). Alterations in apoptotic signaling pathways foster tumor progression and contribute to insufficient response to treatment (5,6). The major pathway of apoptosis activation by 5-FU and other genotoxic drugs is mitochondrial outer membrane permeabilisation (MOMP), which triggers a complex signaling network of pro- and anti-apoptotic proteins controlling the activation of effector caspases (7). Cytochrome c and SMAC are released into the cytosol following MOMP and activate Apaf-1 and Procaspase-9 or antagonize the major caspase inhibitor XIAP. Caspase-9 can activate Caspase-3, which drives multiple positive feedback loops that ensure an efficient execution of apoptotic cell death (7).

Systems analysis of apoptosis signaling has the potential to deliver superior prognostic markers by accounting for the relative abundance of key proteins, network connectivity, and regulations of the pathways involved (8,9). In previous work, our group developed a mathematical model, APOPTOCELL, based on ordinary differential equations that reliably simulates signaling kinetics within the caspase activation network downstream of MOMP (10,11). As input, APOPTO-CELL requires protein concentrations of the key downstream apoptosis regulators Apaf-1, Procaspase-9, XIAP, SMAC and Procaspase-3 to calculate cellular apoptosis competency. Apoptosis competency is calculated as Caspase-3’s ability to cleave sufficient amounts of its substrates to ensure efficient execution of apoptosis. The mathematical model has been extensively validated in-house using single-cell- and population-based approaches in cervical, colorectal and glioblastoma cell line models.
In initial proof-of-concept work, high apoptosis competency was associated with better outcome among chemotherapy-treated CRC and glioblastoma patients (13–15).

Here, we performed a discovery and validation study aimed at developing stratification signatures to predict personalized risk in stage III CRC (Fig. 1). We investigated the potential of APOPTO-CELL as a stand-alone tool. Moreover, we synthetized a systems-biology signature by linking APOPTO-CELL with protein expression data and harnessed a machine learning approach to examine the prognostic relevance of ODE-based modelling, proteomics, and clinicopathological data. We validated our signatures in an independent external cohort and assessed their prognostic value for distinct molecular subtypes of CRC.
MATERIALS AND METHODS

Patient cohorts

We developed prognostic signatures for stage III colorectal cancer patients from three distinct and independent collections: discovery, expansion and validation cohorts. The discovery cohort was an in-house multi-center study of n=120 stage III CRC patients. The expansion dataset included n=157 stage III colon cancer participants with known MSI and BRAF status (GEO repository under accession 39582, (16)). The validation cohort was composed of n=136 stage I-IV colon cancer patients (COAD) from The Cancer Genome Atlas (TCGA) project. A detailed description of the patients data handling and inclusion criteria for downstream analyses is presented in Suppl. Methods SM1.

Prognostic signatures

We developed 3 prognostic signatures: APOPTO-CELL, APOPTO-CELL-PC3 and RF.

APOPTO-CELL signature

APOPTO-CELL is a mathematical model of caspases activation resulting in apoptosis (10,11). The model is comprised of 53 ordinary differential equations (ODEs), 19 state variables, and 75 kinetic parameters. It calculates apoptosis execution kinetics from the input proteins Apaf-1, Procaspase-9, XIAP, SMAC and Procaspase-3. Previous quantitative studies found that Apaf-1 protein levels were not rate-limiting for apoptosome formation in colon cancer cells (14). Thus, Apaf-1 patient-specific protein levels were replaced by the median expression (0.123 µM) previously determined in stage II/III CRC tumors (14). An amount of >25% substrate cleavage (SC) by active Caspase-3 served as the APOPTO-CELL prediction for apoptosis competency in line with previous single-cell imaging
findings (10). While APOPTO-CELL was designed to use protein concentrations as inputs, surrogates such as RNA transcripts or gene expression can be used when protein expression is not available. Molar protein concentrations were estimated from the normalized signal intensities via a novel pipeline illustrated in Fig. 2A.

First, a kernel distribution object was constructed from the normalized protein intensities (or surrogate transcript amounts) with the MATLAB function \texttt{fitdist} (Statistics Toolbox) with a normal kernel smoothing, unbounded kernel support and default bandwidth (steps 1-2). This step resulted in a smooth non-parametric representation of the distribution of the protein (or surrogate) concentrations (step 3). Second, a kernel probability distribution object was constructed from the reference distribution with the same method as for the protein (or surrogate) distribution, except with positive kernel support (right hand side, steps 1-3). Third, the kernel distribution of the protein (or surrogate) intensities was used to calculate the smoothed cumulative distribution corresponding to the protein (or surrogate) intensity of each patient (left hand side, step 4). Finally, the protein (or surrogate) cumulative distribution for each patient was converted into a corresponding absolute concentration using the inverse cumulative distribution of the reference distribution kernel (MATLAB function \texttt{icdf} from the Statistics Toolbox, step 5). Step 6 shows the transformation function from protein (or surrogate) signal intensities to estimated protein concentration.

We estimated APOPTO-CELL inputs from RPPA, gene expression and RNA transcripts for the discovery, expansion and validation cohorts, respectively (Suppl. Methods SM2).

**APOPTO-CELL-PC3 signature**

Procaspase-3 exerts pivotal roles in multiple aspects of apoptosis both via apoptosome-dependent and –independent pathways. Thus, we generated a new signature (APOPTO-CELL-PC3) by combining the APOPTO-CELL model output with the expression of Procaspase-3 (Results and Fig. 4A). As with the APOPTO-CELL signature, Procaspase-3 concentration was estimated from RPPA, gene expression and RNA transcripts for the discovery, expansion and validation cohorts, respectively.
The random forest (RF) machine learning approach was conducted in MATLAB (built-in `TreeBagger` class (Statistics Toolbox) with 100,001 trees, surrogate decision split flags ‘on’; other options were set to ‘default’) and was used to calculate the RF classifier as a predictor for disease recurrence (17). We selected 36 months as a cut-off since most recurrence events (87.9%) occurred within this time. For training purposes, we only considered patients known to have a recurrence or to be recurrence-free at 36 months (only 12.5% of patients were lost to follow-up). The training set consisted of two-thirds of the patients, randomly sampled for each decision tree in the RF. For validation, recurrence predictions averaged from RF decision trees were analyzed for accuracy on patients that did not contribute to defining these trees (‘out of bag’ predictions). This allowed an unbiased validation without a separate validation cohort. The importance of each variable as a predictor was assessed by calculating the Permutated Variable Delta Error, a measurement for the increase in the prediction error upon random perturbation of the predictor values. The predictors that contributed positively to the accuracy of the RF (permutated variable delta >0.05) were used to define the reduced RF classifier. Classifier performance was assessed by the area under the receiver operating characteristic (ROC) curve computed from the validation results.

We generated a synthetic cohort to further understand how the RF classifier makes predictions. This synthetic cohort had one entry for each unique hypothetical patient that each of the decision trees included in the RF ensemble was learned from. The complete set of hypothetical patients was compiled from all permutations of the unique values of each predictor. For each categorical predictor, values were given by each level (for example: low-, medium- and high- risk for the APOPTO-CELL-PC3 signature). For discrete numerical predictors (age and nodal count), the full set of values was taken from the complete list of unique cut-off points extracted from all the trees included in the RF ensemble.

We used the distribution of the patient-specific recurrence probabilities determined by the RF classifier in the discovery cohort to define risk groups (Fig. 5E). Survival analysis was performed using all n=120 patients, including patients lost to follow-up that did not contribute to the RF
classifier identification. The probability of recurrence, and thus the risk group, was computed based on the out-of-bag predictions for patients included in the training set and de novo predictions for patients lost to follow-up within the first 36 months.

Statistical Analysis

Kaplan-Meier estimates were used to compare disease-free (DFS) and overall survival (OS) curves between groups and statistically significant differences were determined by log-rank tests. We used univariate and multivariate Cox proportional hazard models to estimate the relative risks of outcome associated with the signatures (APOPTO-CELL, APOPTO-CELL-PC3, and RF) and clinical factors. We computed hazard ratios (HR) and 95% confidence intervals (CI) and we evaluated statistical significance with likelihood ratio tests. An in-depth account of the survival analyses is provided in Suppl. Methods SM3. Association between APOPTO-CELL and APOPTO-CELL-PC3 signatures with T stage and lymphovascular invasion was assessed by Chi-square ($\chi^2$) test.

Data processing and statistical analysis were performed in MATLAB (MATLAB and Statistics Toolbox Release 2014b, The MathWorks, Inc., Cambridge, UK) unless stated otherwise. Median follow-up time among censored patients, log-rank tests, Cox regression modeling, and assessment of proportional hazards assumptions were conducted in R (version 3.3.0) employing the functions Surv, coxph and cox.zph from the package “survival” (version 2.39-2) and Anova from the package “car” (version 2.1-2). All tests were 2-sided and p-values <0.05 were considered statistically significant.
RESULTS

Quantitative profiling of apoptosis execution proteins in a discovery cohort of CRC patients.

We explored whether mathematical modelling of effector caspase activation in the mitochondrial apoptosis pathway could identify high-risk patients in a multi-center cohort of n=120 stage III CRC patients (Fig. 1). Suppl. Table ST1 presents clinicopathological and demographic characteristics.

Protein amounts of Procaspase-9, XIAP, SMAC and Procaspase-3 were quantified in primary tumor samples by reverse phase protein arrays (RPPAs), (Suppl. Methods SM2).

We established a workflow to determine the molar concentrations required by APOPTO-CELL as input from normalized RPPA data (Fig.2A and Materials and Methods). For each protein, absolute concentrations (Fig. 2B iv) were estimated from normalized RPPA signal intensities (Fig. 2B i) via a transformation function (Fig. 2B iii) determined from a reference distribution (Fig. 2B ii). We previously determined these reference distributions in fresh tumor tissues from a CRC cohort, with analogous demographic and clinicopathological characteristics, using quantitative Western blotting (14). Fig.2B i-iv presents the normalized RPPA signal intensities, reference distributions, results from the alignments, and final estimated concentrations for all proteins from patient tumor samples included in the discovery cohort.

APOPTO-CELL delivers a stratification signature for stage III CRC patients of the discovery cohort treated with 5-FU-based adjuvant chemotherapy.

Tumor concentrations of Procaspase-9, XIAP, SMAC and Procaspase-3 were input into the APOPTO-CELL mathematical model. We calculated apoptosis execution kinetics for a duration of 300 min (14) and computed substrate cleavage (SC) over time as a surrogate for the efficiency of apoptosis execution. SC profiles demonstrated that tumor apoptosis competency was variable between patients (Fig. 2C). We observed two major sub-populations displaying ≥90% or ≤10% SC (58% and
32% of patient tumors, Fig. 2D), reflecting very high and low apoptosis competency, respectively. Thus, apoptosis execution functions as a biological all-or-nothing switch between apoptosis competency and resistance, emanating from the systems-level interplay of key proteins (18,19).

We investigated whether SC was associated with clinical outcome among stage III CRC patients treated with adjuvant chemotherapy. As 25% SC serves as an optimal decision threshold for apoptotic cell death in cancer cell lines (10,12) and CRC patient tumors (14), patients were dichotomized as those with SC> or ≤25% at 300 min (Fig. 2C-D). We observed statistically significant differences between overall (OS) and disease-free survival (DFS) curves when comparing these groups (log-rank p<0.05), (Fig. 2E-F). Patients with SC≤25% had increased risks of relapse (HR 2.05, 95% CI 1.03–4.05, p=0.04) and death (HR 3.78, 95% CI 1.42–10.08, p=0.006) compared to those with SC>25%.

Sensitivity analyses demonstrate high robustness of the APOPTO-CELL signature. We examined whether APOPTO-CELL predictions of apoptosis susceptibility would notably change when accounting for noise in the protein concentrations (Fig. 3A). For each patient, we built a normal distribution centered on the reference (unperturbed) value with a given standard deviation for each of the input proteins (including Apaf-1). Next, for each patient we ran 1000 simulations with values randomly drawn from these distributions (bootstrapping, (20)) and computed the predictions for apoptosis competency. We then computed a robustness index (RI) as the fraction of simulations matching the observed substrate cleavage (SC>25% vs. SC≤25%) in the absence of noise. We investigated the RI for up to 30% variation (upper limit for intercellular heterogeneity observed in (21,22)) and defined patient-specific predictions as robust when RI≥90% (gray line in Fig. 3B). When increasing the standard deviation from 10% to 30%, we observed a marginal decrease (87% to 71%) in the number of patients consistently classified as low- vs. high-risk. We previously found the relative standard variation in intra-tumor protein amounts in a set of CRC patients to be approximately 11% (14). Thus, in this study we selected 10% as the coefficient of variation when assessing robustness in downstream analyses. Visualizing the distribution of SC simulations revealed that despite adding variation, apoptosis competency remained stable in regions of very low or high SC for
most patients (Fig. 3C). Next, we sought to examine the clinical implications of the robustness analyses. We categorized patients into three sub-populations (‘bootstrap robustness group’): i) robust low-risk if at least 90% of the simulations classified the patient as low-risk, ii) robust high-risk if at least 90% of the simulations classified the patients as high-risk and iii) non-robust if otherwise. We observed statistically significant differences in DFS ($p=0.01$) and OS ($p=0.0004$) curves when comparing the robust low- vs. high-risk groups (Fig. 3D-E). Only 16 patients (13%) were not classified as low- or high-risk for at least 90% of the simulations (Fig. 3F). These results suggest that while a degree of uncertainty may be present in the protein measurements (both in terms of technical measurement error and true biological variability), APOPTO-CELL still retained its prognostic value. This robustness is a distinctive advantage over classical threshold-based stratification by protein biomarkers, for which continuous readouts and cut-off values are strongly influenced by measurement noise.

Combining the APOPTO-CELL signature with Procaspase-3 expression (APOPTO-CELL-PC3) resulted in enhanced stratification in stage III patients of the discovery cohort.

Apart from their defined roles in the mitochondrial apoptosis pathway, Procaspase-9, XIAP, SMAC and Procaspase-3 may play roles in MOMP- and apoptosome-independent cell death pathways. They may also regulate other cellular processes such as proliferation, autophagy, immune response, and differentiation (23–29). As these processes could affect patient response to therapy, we examined whether these proteins were independently associated with clinical outcome. Only Procaspase-3 (PC3) showed potential as an independent biomarker (Suppl. Fig. S1). We therefore included Procaspase-3 levels as an additional prognostic factor in our analysis. This approach compensates for apoptosome-independent Caspase-3 activation. To explore whether PC3 could further improve prognostic accuracy, we combined PC3 quantifications with APOPTO-CELL model outputs (APOPTO-CELL-PC3), (Fig. 4A). We categorized patients into three groups: i) high-risk, patients with low PC3 expression ($\leq$median) and apoptosis resistance (SC$\leq$25%); ii) low-risk, patients with high PC3 expression ($>$.median) and apoptosis competency (SC$>$.25%), and iii) medium-risk, all other
patients. We observed statistically significant differences between OS and DFS curves when comparing these groups (log-rank p<0.01), (Fig. 4B-C), which was driven by differences between the high- and low-risk groups (pairwise log-rank p<0.01). Patients categorized as high-risk by the APOPTO-CELL-PC3 signature had significantly increased risks of relapse (HR 3.90, 95% CI 1.59–9.57, p=0.008) and death (HR 9.30, 95% CI 2.06–41.98, p=0.002) compared to patients categorized as low-risk. Harrell's concordance index suggested that the APOPTO-CELL-PC3 signature had superior prognostic discrimination compared to the APOPTO-CELL signature alone for both OS (0.73 vs. 0.67) and DFS (0.64 vs. 0.59).

The APOPTO-CELL-PC3 signature is an independent prognostic marker for stage III CRC patients of the discovery cohort.

We used Cox proportional hazards regression models to examine the prognostic value of APOPTO-CELL-PC3 and established clinical risk factors (Fig. 4D). T stage and lymphovascular invasion were significantly associated with the risks of relapse or death in univariate analyses. Patients categorized as high-risk by the APOPTO-CELL or APOPTO-CELL-PC3 signature were significantly over-represented in stage T4 (chi-square p=0.03 and p=0.006, respectively). No association was detected between lymphovascular invasion and the APOPTO-CELL (chi-square p=0.58) or APOPTO-CELL-PC3 signatures (chi-square p=0.65). When adjusting for T stage and lympho-vascular invasion, patients categorized as high-risk by the APOPTO-CELL-PC3 signature had increased risks of relapse (HR 3.26, 95% CI 1.27–8.35, p=0.04) and death (HR 11.10, 95% CI 2.35–52.33, p=0.001) compared to patients categorized as low-risk.

We explored whether staging, primary tumor location, and lymphovascular invasion further aided in patient stratification (Suppl. Fig. S2). We observed significant differences between OS and DFS curves by lymphovascular invasion (log-rank p=0.04 for both, Suppl. Fig. S2C i-ii), though not for staging or tumor location (Suppl. Fig. S2A-B i-ii). When stratifying by stage and location, we observed significant differences between OS and DFS curves (log-rank p<0.05) by the APOPTO-CELL-PC3 signature among stage III-A/B and proximal tumors (Suppl. Fig. S2A-B iii-iv). There
were differences between DFS curves by APOPTO-CELL-PC3 among patients without lymphovascular invasion (Suppl. Fig. S2C iii) and between OS curves among patients with lymphovascular invasion (Suppl. Fig. S2C iv).

APOPTO-CELL and APOPTO-CELL-PC3 are independent prognostic markers of recurrence in the expansion cohort

The prognostic relevance of BRAF and MSI status has been comprehensively examined, and testing of MSI status has been introduced in clinical practice (2). We used an expansion cohort (GSE39582 dataset, (16)) to explore whether MSI and BRAF status potentially confound associations between our prognostic signatures and recurrence risk (Fig. 1). We analyzed stage III patients with known status for microsatellites and BRAF mutation (n=157), (Suppl. Table ST1).

APOPTO-CELL and APOPTO-CELL-PC3 signatures, estimated from gene expression (Suppl. Methods SM2), were associated with DFS in univariate and multivariate analyses (Suppl. Table ST2). In unadjusted analyses, patients categorized as high- vs. low-risk by APOPTO-CELL (SC≤ vs. >25%) had an increased risk of relapse (HR 1.94, 95% CI 1.17-3.24, p=0.01). Patients categorized as high- and medium-risk by APOPTO-CELL-PC3 had an increased risk of relapse compared to those classified as low-risk (HR 2.39, 95% CI 1.20-4.78 and HR 2.66, 95% CI 1.32-5.34, p=0.008). We fit two multivariate models i) stratifying by MSI (dMMR vs. pMMR) status and adjusting for BRAF (wild-type vs. mutant) as well as treatment received (chemotherapy vs. no chemotherapy), and ii) additionally adjusting for age (continuous linear), sex (male vs. female), tumor location (distal vs. proximal), and KRAS mutational status (wild-type vs. mutant). Relative risk estimates were similar between unadjusted and adjusted models, suggesting that these factors did not likely confound associations between our signatures and DFS. Further, both signatures remained independent prognostic markers for DFS (Suppl. Table ST2).
A machine learning Random Forest signature for recurrence identifies the APOPTO-CELL-PC3 signature as the most salient predictor.

We built a Random Forest (RF) using the APOPTO-CELL-PC3 signature, the remaining single proteins (dichotomized as > or ≤median), and the clinical variables analyzed in Fig. 4D (‘RF-all predictors’). A reduced RF classifier was constructed using the most influential predictors of relapse (‘RF-reduced predictors set’). The reduced classifier predicted disease recurrence without any loss in accuracy compared to the full RF classifier (McNemara p>0.05). Thus, we performed our downstream analysis with the reduced classifier, which included the APOPTO-CELL-PC3 signature, T stage, sex, N stage, age, and nodal count (Fig. 5B). The AUC (see Materials and Methods) suggested this model satisfactorily predicted relapse ((0.73, 95% CI 0.62-0.74), Fig. 5C).

RF classifiers are considered ‘black box’ machine learning algorithms that have limited interpretability. To investigate how the recurrence probability predicted by the RF classifier depended on each feature and their interactions (Suppl. Fig. S3), we generated a synthetic cohort where each patient was characterized by a unique combination of input features (see Materials and Methods). The RF classifier predicted a reduced likelihood of recurrence for patients categorized as low-risk by the APOPTO-CELL-PC3 signature, aged >40, with a nodal count between 10-50, males, and those with less advanced T and N stage (Suppl. Fig. S3A-F).

We examined how the RF features interact to deploy the final recurrence predictions. For each combination of predictors, we visualized recurrence risk relative to the overall average recurrence risk (the reference baseline) (Fig. 5D). The relative risks associated with age and nodal count were aggregated into three levels by averaging over all risks in the corresponding group. Suppl. Fig. S3G presents this visualization without aggregation of these predictors. Overall, patients categorized as low-risk by the APOPTO-CELL-PC3 signature had a reduced recurrence risk. This effect was lost in advanced cancers (T4 and N2), which exhibited an overall higher probability of recurrence. Females also had a higher risk of relapse than males.

We categorized patients into three risk groups according to the probability of recurrence predicted by the reduced RF classifier: <20%, 20-50%, and >50%. There were significant differences between DFS
curves by these groups (log-rank p<0.001), (Fig. 5F). Patients with >50% risk had an approximately 3-fold increased recurrence risk (HR 5.13, 95% CI 2.12–12.41, p=0.002) compared to patients with <20% risk. Results were similar after adjusting for T stage and lymphovascular invasion. Harrell’s concordance index suggested that the RF signature was better at discriminating recurrence than the APOPTO-CELL-PC3 signature (0.67 vs. 0.64).

Validation of the prognostic signatures in an independent external cohort.

We evaluated the prognostic value of our signatures using an independent publically available cohort: TCGA COAD (Fig. 1). Suppl. Table ST1 presents clinicopathological and demographic characteristics of patients who met our inclusion criteria (Suppl. Methods SM1).

The three signatures identified in the discovery cohort were tested on stage III patients in the validation cohort. In this dataset, protein expression of SMAC and XIAP were available. For the amounts of Procaspase-3 and Procaspase-9, transcript abundance was used as surrogate for protein expression (Suppl. Methods SM2). Despite the limited sample size and follow-up time, we nevertheless observed statistically significant differences between DFS curves when comparing risk groups for the APOPTO-CELL, APOPTO-CELL-PC3, and RF signatures (log-rank p<0.05), (Fig. 6A-C).

Evaluation of the prognostic value of the signatures in the context of the CRC Consensus Molecular Subtypes (CMS).

A recent study identified four distinct consensus molecular subtypes (CMS) for CRC (30), with CMS4 (‘mesenchymal’) vs. CMS1-3 subtypes being associated with poorer outcome. Among our validation cohort, there were not differences between DFS curves by CMS4 vs. 1-3 subtypes among stage III (log-rank p=0.90) or stage I-IV patients (log-rank p=0.54). Nevertheless, SC retained its prognostic value among CMS1-3 (log-rank p=0.02), though not among CMS4 (log-rank p=0.11) (Fig.
Analogously, the APOPTO-CELL-PC3 and RF signatures had prognostic value among CMS1-3, but not CMS4 (Fig. 6E-F).
We evaluated APOPTO-CELL, a mathematical model that calculates apoptosis competency of patient tumors, as a signature of stage III CRC patient outcome to 5FU-based chemotherapy. Moreover, we present a novel systems medicine workflow to optimize model predictions. This study is the first large-scale demonstration that determining apoptosis competency by systems modeling adds independent prognostic value to current clinicopathological markers.

Previous studies demonstrated that systems biological approaches towards apoptosis signaling may have great potential as novel avenues towards discovery of biomarkers (31) and therapeutic viable targets (for example, SMAC-mimetics (32)) resulting in an educated stratification of patients for clinical trials (33). Further development and exploitation of such new strategies is warranted, since despite extensive efforts to identify prognostic and predictive biomarkers for CRC TNM staging remains the most reliable prognostic marker that still drives treatment decisions. Other clinicopathological tumor characteristics, such as tumor location, differentiation and lymphovascular invasion, have been associated with prognosis and response to treatment (2). Microsatellite instability and wild-type status for KRAS, BRAF, PIK3CA and TP53 have been associated with better outcome (2,34,35). Conversely, wild-type APC patients showed adverse survival (35). Recent studies have highlighted the intertwined relationship of MSI and somatic mutations status (35–39). MSI/BRAF\textsuperscript{wild-type} and MSS/BRAF\textsuperscript{mutant} patients exhibited the most favorable and unfavorable outcome, respectively (36–39). Remarkably though, MSI/BRAF\textsuperscript{mutant} showed superior prognosis compared to MSS/BRAF\textsuperscript{wild-type}, suggesting that microsatellite status prevails over BRAF mutation in steering patients survival in these sub-groups (36–39). Mutations in TP53, KRAS and APC have been more frequently observed in MSS patients with the most adverse outcome being observed in the MSS triple-mutant sub-population (35). Remarkable advances in ‘-omics’ technologies have brought about a shift in prognostic marker discovery, moving from conventional clinicopathological risk factors towards their integration with molecular characteristics. Multi-gene-based classification algorithms have been shown to aid in stratifying patient outcome, but these classifiers so far provide only small
improvements in stratification compared to the use of clinical markers alone (40). Reviewing and re-
assessing promising multi-gene and multi-protein panels using systems biological approaches may
allow us to further improve the prognostic value of such signatures.

We demonstrated that patient-specific predictions by the APOPTO-CELL signature were highly
robust against noise in measurements and parameterization. This robustness emanates from the multi-
protein interplay during the apoptosis execution phase, which includes several positive feedback
loops. If sufficiently stimulated, these bring about an irreversible cell death decision with a
pronounced binary/switch-like character (9,10,18). As our discovery cohort was composed of patients
recruited at three different centers in three European countries, this high robustness may be
fundamental for successful model-based patient stratification. Our approach of apoptosis modeling
was successful in the expansion and validation cohort, which made use of FF materials and transcript
abundance in CRC tumor samples collected at various centers across the France and US, respectively.

We introduced the APOPTO-CELL-PC3 signature which capitalizes on both the prognostic
information embedded in the apoptosome-dependent cell death pathways modelled by APOPTO-
CELL and Caspase-3 paracrine signaling to further improve prognostic accuracy. Procaspase-3 exerts
important roles outside of MOMP-dependent apoptosis execution including mediating tissue
regeneration and immune responses (23–29). Upon Caspase-3 activation, apoptotic cells release
growth signals such as prostaglandin E2 that ultimately induce tissue regeneration and wound healing
in the neighbor cells via the “Phoenix Rising” pathway (25). Further, anti-tumor immune responses
trigger apoptosis cascades through Caspase-8 activation or granzyme B activation, that can drive
Procaspase-3 activation independent of the mitochondrial pathway (23–25,29).

There is extensive genomic, epigenomic, and molecular inter-patient heterogeneity in CRC (41). A
key step towards personalized precision oncology is to develop panels of combinatorial biomarkers,
as opposed to single biomarkers, that can better capture disease complexity and provide
individualized predictions on cancer progression and treatment response. To develop optimal
stratification tools, it is paramount to integrate clinical patient-specific characteristics with molecular
phenotypes or, as in our case, patient-specific systems-level characteristics. By applying RF machine
learning, we investigated how the APOPTO-CELL-PC3 signature could be combined with conventional clinical variables currently used for CRC prognosis. The RF approach identified APOPTO-CELL-PC3 as the most impactful predictor, and a reduced RF classifier allowed us to optimally stratify patients into risk groups from a limited number of variables. Reducing the number of clinicopathological variables and molecular markers required for optimal and individualized prognostication minimizes the economic burden of otherwise possibly excessively complex assays and data acquisition. We generated a valuable visualization of how the predictors contribute to shifting recurrence risk, which could serve as a nomogram to aid medical practitioners in planning disease management.

Extensive RNA profiling work (16,42–45) recently resulted in the definition of four major subtypes for CRC (CMS 1-4, (30)). Compared to patients with CMS1-3, CMS4 patients had higher risks of relapse and mortality. Analysis of the interplay between apoptosis competency and CMS revealed a potential differential role of apoptosis signaling in CMS1-3 compared to CMS4, with apoptosis competency associated with good prognosis in CMS1-3. The salient features of the CMS4 subtype (‘mesenchymal’) include CIN/MSI alterations, NOTCH3 overexpression, TGF-β and VEGF activation (30). TGF-β and VEGF signaling has been reported to suppress Caspase-3-dependent apoptosis execution in various cell types (46,47). Conversely, caspases have been shown to be enriched in the CMS 1 (‘immune’) subtype.

Previous proof-of-concept studies on apoptosis modeling in CRC have been limited to the use of FF materials for the generation of tissue protein extracts (14). Our study shows that FFPE-derived protein extracts are suitable for calculating apoptosis competency, representing a key advancement towards aligning with routine surgical and pathological workflows. It also supports independent retrospective validation in larger bio-banked collections of stage III CRC cohorts. While RPPA-based protein quantification pipelines may not be amenable for integration into standard clinical workflows, the development of calibrated protein quantification assays, such as multiplex ELISAs or novel quantitative digital histopathology techniques, may address current limitations in routine quantitative clinical proteomics (48).
Recent research has pinpointed the pivotal role played by the tumor microenvironment in regulating response to chemotherapeutics (49). Factoring in relative abundance of key apoptotic proteins both in the tumor and the surrounding tissue (stroma, immune cells, etc.) will be further fine-tune risk scores.

Emerging technologies, such as multiplexed fluorescence microscopy (MxIF, (50)) allow now simultaneous co-staining of multiple-proteins and reliable quantitative measurements within different compartments which will be instrumental in future studies to tackle these challenges.
**Author contributions:** MS, AUL, AJR, EZ, LMS, MR and JHMP analyzed and interpreted the data. OB, SD, NR and CP gathered the clinical data. MLW, ACM, SVS, MR and JHMP coordinated the exchange of samples and data across collaborators. MS and MLW curated the clinical data. CM, SC, MC, ROB, LF and ST performed and analyzed experiments. ML, PGJ, RW, SCB, MST, DAMN, EWK, BRH, PLP, SVS, MR and JHMP provided reagents, tools, clinical samples and methodological advice. MR and JHMP reviewed the results and supervised the project. MS, DL, MR and JHMP wrote the manuscript. All authors read and approved the manuscript.

**Conflict of Interests:** J.H.M.P. and M.R. hold a patent on “A computer-implemented system and method for the prediction of cancer response to genotoxic chemotherapy and personalised neoadjuvant treatments (pccp)” (Derwent primary accession number: 2013-A25393 [52]).

**Funding:** The authors acknowledge support for their work by the European Union Framework Programme 7 (FP7 APO-DECIDE) under contract No. 306021. J.H.M.P. is supported by a Science Foundation Ireland Investigator Award (13/IA/1881) and by the Irish Cancer Society BreastPredict Collaborative Research Centre (CCRC13GAL). J.H.M.P., M.R. and D.L. are supported by a Science Foundation Ireland/Department of Enterprise and Learning Partnership Award (14/IA/2582). E.Z. was supported by a Health Research Board Translational Research Supplementary Award to J.H.M.P. (TRA/2007/26). L.M.S. was supported by the European Community Action Scheme for the Mobility of University Students (ERASMUS).

**Abbreviations**

Colorectal Cancer (CRC), Reverse Phase Protein Array (RPPA), The Cancer Genome Atlas Colon Adenocarcinoma (TCGA COAD), Ordinary Differential Equation (ODE), Hazard Ratio (HR), Confidence Interval (CI), Substrate Cleavage (SC), Procaspace-3 (PC3), Enriched APOPTO-CELL Signature (APOPTO-CELL-PC3), Machine Learning Random Forest Signature (RF signature), Consensus Molecular Subtyping (CMS), Fresh Frozen (FF), Formalin-Fixed Paraffin-Embedded (FFPE).
REFERENCES


Kolata G. The Art of Learning from Experience: Statistician Bradley Efron tells what his field is about and how a new method, the bootstrap, exploits the power of large-scale computing. Science. 1984;225:156–8.


Figure 1. Workflow for the stepwise development and validation of personalized risk signatures in stage III CRC patients.

Three signatures with increasing prognostic accuracy were developed in a stepwise integrative approach based on ODEs-modelling, proteomics and clinicopathological features in the discovery cohort. The prognostic value of the apoptosis-based signatures was assessed in the context of microsatellite instability and BRAF mutational status using an expansion cohort. Validation was performed in an external cohort, including the evaluation of the prognostic value for CRC molecular subtypes.
Figure 2. The APOPTO-CELL signature derived from ODE-based computations of absolute protein concentrations in tumor resections is a prognostic marker for stage III CRC patients.

A. Workflow of the pipeline developed to estimate absolute protein concentrations from the normalized signal intensities determined by RPPA (see Materials and Methods).

B. Normalized RPPA signal intensities (i), reference distribution (ii), transformation function (iii) and estimated concentrations (iv) for Procaspase-9, XIAP, SMAC and Procaspase-3. Data are color-coded based on data density. Further details on these procedures are provided in the Materials and Methods section.

C. APOPTO-CELL-based simulations for the cleavage of effector caspase substrates are shown for n=120 stage III CRC patients. Time zero represents the event of mitochondrial outer membrane permeabilisation (MOMP). The dashed line at 25% substrate cleavage represents the decision threshold beyond which tumor cells are committed to die.

D. Simulation results from (C) were binned using 10 percentage points increments in substrate cleavage. The majority of patient tumors presented with very low or very high apoptosis competency.

E-F. Kaplan-Meier plots of disease-free survival (E) and overall survival (F) are shown. Patients were stratified according to the APOPTO-CELL signature (substrate cleavage (SC) reached at 300 min in C-D being above (n=77) or below (n=43) the threshold of 25%). P-values were determined by log-rank tests.
Figure 3. APOPTO-CELL predictions are robust against perturbations in protein concentrations.

A. Schematic representation of the workflow to test the robustness of the APOPTO-CELL signature by bootstrapping analysis. For each patient, perturbations (± a given s. d.) were applied to all protein concentrations and 1000 simulations were run by randomly sampling with replacement from these parameterizations.

B. Robustness of the apoptosis competency predictions as a function of the percentage of the standard deviation in the APOPTO-CELL proteins. Patient predictions were defined as robust if the robustness index RI (percentage of the simulations matching the observed substrate cleavage values) ≥90% (gray line).

C. Heatmap showing the distribution of the substrate cleavage values reached at 300 min for each of 1000 simulations run with ±10% s. d. for each of the 120 patients of the discovery cohort (black, robust SC; shades of yellow, fragile SC). Hierarchical clustering (on the left) is based on correlations (computed with average as linkage method) between rows of patients to highlight groups of patients who behaved similarly.

D-E. Kaplan-Meier estimates and log-rank tests comparing DFS (D) and OS curves (E) by the bootstrap robustness groups suggested that the association between APOPTO-CELL and prognosis is retained despite ±10% noise in protein concentrations.

F. Breakdown of robust vs. non-robust predictions obtained by applying ±10% s. d. for patients categorized as low- vs. high-risk.
Figure 4. An enriched APOPTO-CELL signature (APOPTO-CELL-PC3) is a prognostic biomarker for stage III CRC that is independent of established clinicopathological predictors.

A. Workflow leading to the derivation of the enriched APOPTO-CELL signature based on the combination of the APOPTO-CELL model output with the Procaspase-3 (PC3) expression to define patient subgroups.

B-C. Kaplan-Meier plots comparing patient subgroups identified by the APOPTO-CELL-PC3 signature shown in (A) for disease-free and overall survival. The enriched signature identifies three risk groups, with highly significant separation of high- and low-risk patients. P-values were determined by log-rank tests.

D. Estimated hazard ratios, 95% confidence intervals, and p-values from likelihood ratio tests from univariate (gray shaded) and multivariate (red shaded) Cox proportional hazards models examining associations between clinical factors and the APOPTO-CELL-PC3 signature with the risks of recurrence (left) and death (right). The multivariate analysis (‘APOPTO-CELL-PC3 signature adj.’) was adjusted for clinical variables found to be associated with outcome in the Cox univariate analysis (T stage and lympho-vascular invasion).
Figure 5. A machine learning Random Forest signature for relapse identifies the APOPTO-CELL-PC3 signature as the most important predictor and sheds light on how distinct features contribute to the risk predictions.

A. Workflow for the development of the machine learning Random Forest signature.

B. The identified Random Forest classifier consists of six predictors for recurrence, ranked here by importance (see Materials and Methods).

C. Evaluation of the performance of the Random Forest classifier by ROC analysis. Error bars represent 95% confidence intervals obtained from 1,000 bootstraps. Inset represents area under the curve with 95% confidence interval.

D. Dependence plot highlighting how the predictors (and their interactions) contribute to the relative risk of recurrence predicted by the RF classifier. Size of the dot encodes the magnitude of the relative risk compared to baseline whereas color indicates the direction of change (red and blue for increase and decrease, respectively).

E. Distribution of the probability of recurrence for the patients used to develop the random forest signature highlighting the cut-off thresholds employed to define risk groups.

F. Kaplan-Meier estimates comparing disease-free survival curves for the risk groups defined by the Random Forest classifier in E. P-values were computed by log-rank tests.
Figure 6. Validation of the prognostic value of the signatures identified in the discovery cohort and their relationship to the CRC Consensus Molecular Subtypes (CMS) in an independent external cohort of stage III colon cancer patients.

A-C. Kaplan-Meier plots for disease-free survival stratification of stage III colon cancer patients based on the APOPTO-CELL signature (A), the APOPTO-CELL-PC3 signature (B) and the Random Forest signature (C).

D-F. Kaplan-Meier estimates comparing disease-free survival curves based on the APOPTO-CELL signature (D), the APOPTO-CELL-PC3 signature (E) and the Random forest signature (F) stratified by CMS1-3 vs. CMS4. Results suggest that the effect of the signatures on the risk of recurrence differs between CMS1-3 and CMS4.
Figure 1

Systems biology & machine learning integrative approach to develop prognostic signatures for stage III CRC

APOPTO-CELL signature

Enriched APOPTO-CELL signature

Machine learning
Random Forest signature

Figs. 2-3

Fig. 4 Suppl. Figs. S1-2

Fig. 5 Suppl. Fig. S3

Evaluation of signatures in an expansion cohort

Gene Expression Omnibus
GSE39582

Suppl. Tables ST1-2
Suppl. Met. 1 Fig. SM1.1

Validation of signatures in an external cohort

The Cancer Genome Atlas
Colon Adenocarcinoma

Suppl. Table ST1 Fig. 6
Suppl. Met. 1 Fig. SM1.1
Suppl. Met. 2 Table SM2.1
Suppl. Met. 2 Figs. SM2.2-3

Exploration of the signatures across CRC molecular subtypes

CMS 1-3
CMS 4

Fig. 6
Estimation of absolute concentrations

1. Density probability of protein signal intensities vs. protein concentration
2. Density probability of protein signal intensities vs. protein concentration
3. Cumulative distribution of protein signal intensities vs. protein concentration
4. Cumulative distribution of protein signal intensities vs. protein concentration
5. Cumulative distribution of protein signal intensities vs. protein concentration
6. Protein concentration vs. protein signal intensities

Figure 2

A. APOPTO-CELL simulations

B. i. Procaspe-9 Signal Intensities [a.u.]
   ii. Number of patients
   iii. Prob. dist. of Procaspe-9 [μM]
   iv. Estimated concs. [μM]

B. i. XIAP Signal Intensities [a.u.]
   ii. Number of patients
   iii. Prob. dist. of XIAP [μM]
   iv. Estimated concs. [μM]

B. i. SMAC Signal Intensities [a.u.]
   ii. Number of patients
   iii. Prob. dist. of SMAC [μM]
   iv. Estimated concs. [μM]

B. i. Procaspe-3 Signal Intensities [a.u.]
   ii. Number of patients
   iii. Prob. dist. of Procaspe-3 [μM]
   iv. Estimated concs. [μM]

C. Substrate Cleavage (SC) [%]

D. Fraction of Patients [%]

E. Disease-free survival

F. Overall survival
Figure 3

A. Protein inputs and APOPTO-CELL simulations. Variation 2 x s.d.% and Robustness Index (RI).

B. Fraction of patients and Simulations vs. Variation in simulation inputs [%].


D. Disease-free survival. Disease-free survival vs. Time to recurrence [months].

E. Overall survival. Overall survival vs. Time to death [months].

F. Low-risk and High-risk patients. n=6 and n=10. Robust and non-robust groups.