original report

A Machine Learning Platform to Optimize the Translation of Personalized Network Models to the Clinic

Manuela Salvucci, PhD[1]; Arman Rahman, MD, MPH, PhD[2]; Alexa J. Resler, PhD[1]; Girish M. Udupi, MS[2]; Deborah A. McNamara, MD[3]; Elaine W. Kay, MD[3]; Pierre Laurent-Puig, MD, PhD[4]; Daniel B. Longley, PhD[5]; Patrick G. Johnston†, MD, PhD[5]; Mark Lawler, PhD[5]; Richard Wilson, MD[5]; Manuel Salto-Tellez, MD[5]; Sandra Van Schaeybroeck, MD, PhD[5]; Mairin Rafferty, PhD, MBA[2]; William M. Gallagher, PhD, MSB[2]; Markus Rehm, PhD[1,6]; and Jochen H.M. Prehn, PhD[1]

abstract

PURPOSE Dynamic network models predict clinical prognosis and inform therapeutic intervention by elucidating disease-driven aberrations at the systems level. However, the personalization of model predictions requires the profiling of multiple model inputs, which hampers clinical translation.

PATIENTS AND METHODS We applied APOPTO-CELL, a prognostic model of apoptosis signaling, to showcase the establishment of computational platforms that require a reduced set of inputs. We designed two distinct and complementary pipelines: a probabilistic approach to exploit a consistent subpanel of inputs across the whole cohort (Ensemble) and a machine learning approach to identify a reduced protein set tailored for individual patients (Tree). Development was performed on a virtual cohort of 3,200,000 patients, with inputs estimated from clinically relevant protein profiles. Validation was carried out in an in-house stage III colorectal cancer cohort, with inputs profiled in surgical resections by reverse phase protein array (n = 120) and/or immunohistochemistry (n = 117).

RESULTS Ensemble and Tree reproduced APOPTO-CELL predictions in the virtual patient cohort with 92% and 99% accuracy while decreasing the number of inputs to a consistent subset of three proteins (40% reduction) or a personalized subset of 2.7 proteins on average (46% reduction), respectively. Ensemble and Tree retained prognostic utility in the in-house colorectal cancer cohort. The association between the Ensemble accuracy and prognostic value (Spearman $\rho = 0.43$; $P = .02$) provided a rationale to optimize the input composition for specific clinical settings. Comparison between profiling by reverse phase protein array (gold standard) and immunohistochemistry (clinical routine) revealed that the latter is a suitable technology to quantify model inputs.

CONCLUSION This study provides a generalizable framework to optimize the development of network-based prognostic assays and, ultimately, to facilitate their integration in the routine clinical workflow.

INTRODUCTION

In recent years, there has been a transition from making clinical decisions on the basis of macroscopic characteristics of the tumor to molecular-based biomarkers.[1-5] Advances in omic technologies have led to the development of prognostic and predictive molecular signatures for the majority of solid tumors.[6-17] A few transcriptomic-based taxonomies have been commercialized, and their introduction in the clinic is being evaluated.[18-21] Dynamic systems models, historically relegated to comprehending cancer biology in basic research, have emerged as valuable biomarkers.[22-27] Systems models, which primarily are based on ordinary or partial differential equations, have the inherent ability to encode pathway properties that can serve both as prognostic indicators and as screening platforms to inform patient treatment. The network dynamics, and thus model predictions, can be personalized by tuning critical model inputs with patient-specific measurements, such as protein concentrations or expression of genes and microRNAs. However, robust quantification of model inputs requires a large tumor volume for each sample, highly specialized equipment, and time-consuming protocols not suitable for a fast-paced environment such as a clinical histopathology department.

We developed a framework to identify the essential model inputs that require de novo patient-specific quantification for a specific clinical setting, which therefore assists with the integration of pathway-based biomarkers in the routine clinical workflow (Fig 1). As a case study, we selected a mathematical model of apoptosis execution, APOPTO-CELL, that our group developed[28] and showed to be prognostic in colorectal cancer (CRC)[22,26] and glioblastoma multiforme.[29] We developed two methods to apply APOPTO-CELL with a minimal set of protein inputs: Ensemble and Tree. Ensemble estimates the APOPTO-CELL signature by

ASSOCIATED CONTENT

Appendix

Data Supplement

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on October 10, 2018 and published at ascopubs.org/journal/cci on April 17, 2019: DOI https://doi.org/10.1200/CCI.18.00056

†Deceased.

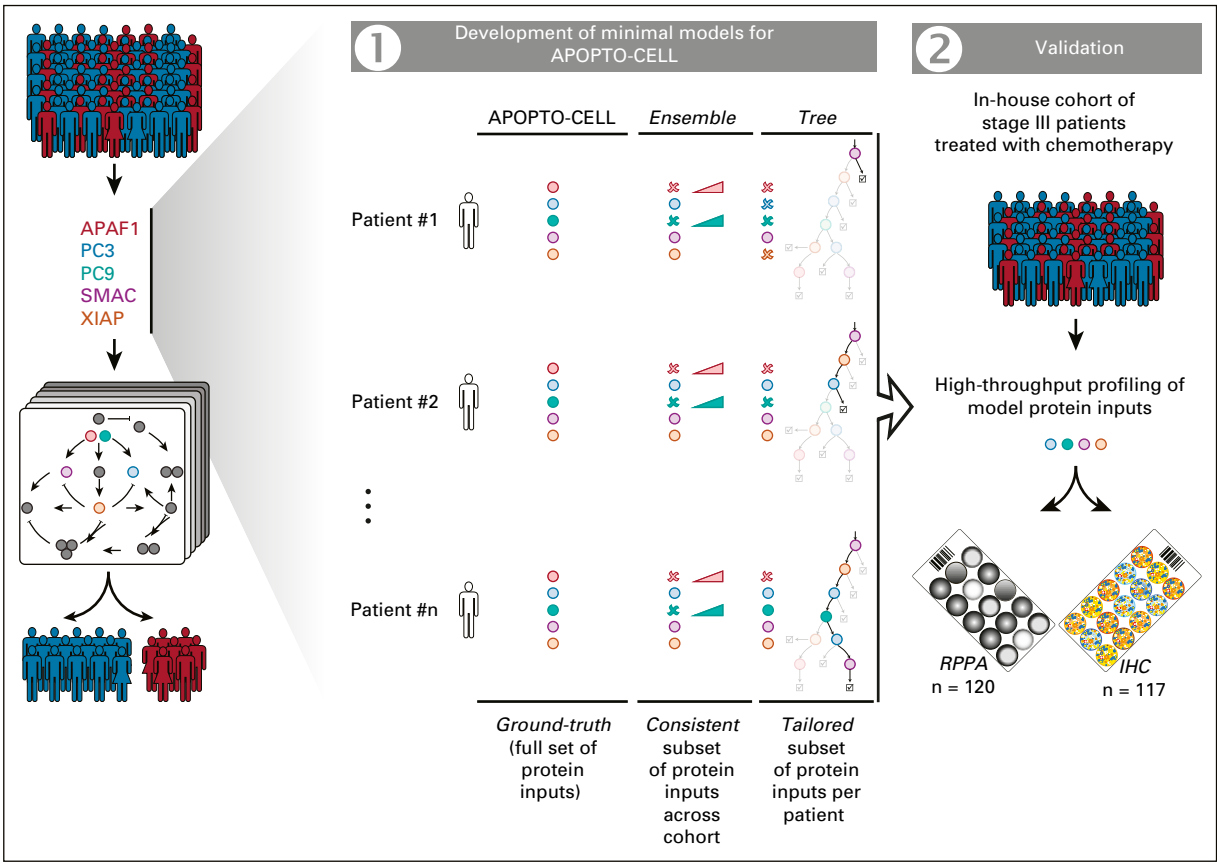


FIG 1. Workflow for the development and validation of two variants of minimal models for APOPTO-CELL: Ensemble and Tree. Ensemble uses a probabilistic approach with clinically grounded distributions in place of individualized measurements for a subset of proteins. In contrast, Tree can be used to personalize the proteins to measure for every patient by traversing the classification decision tree to a node with the predicted APOPTO-CELL signature.

aggregating the results from a collection of APOPTO-CELL simulations performed with a range of concentrations for the proteins that were not quantified. In contrast, Tree uses a decision tree to estimate the APOPTO-CELL signature from a smaller number of protein inputs. The set of proteins that need to be measured is dictated by the decision tree tailored for the individual patient. We validated our framework in an in-house CRC cohort, with protein inputs quantified by high-throughput techniques, such as reverse phase protein array (RPPA) and immunohistochemistry (IHC).

PATIENTS AND METHODS

Development and Validation of Minimal Models for APOPTO-CELL With Virtual and Real-World CRC Cohorts

APOPTO-CELL is a mathematical model that takes as input tumor protein expression of APAF1, procaspase-3 (PC3), procaspase-9 (PC9), SMAC, and XIAP and predicts apoptosis sensitivity. The development of the Ensemble and Tree variants of a minimal model for the APOPTO-CELL signature with a virtual patient cohort is described in the Data Supplement.

We validated our minimal models in a real-world cohort from a retrospective collection of patients with stage II to IV CRC, as previously reported.[26] Clinical follow-up of patients is standardized, and every included patient was reviewed every 6 months for 3 years followed by 6-month or annual visits for years 4 and 5. We used disease-free survival (DFS) and overall survival (OS) as clinical end points. DFS was defined as the absence of a suspicious lesion on surveillance computed tomography of thorax, abdomen, and pelvis performed according to guidelines in place at the treating hospital (every 6 months for 2 years followed annually for 3 years or annually for 3 years). OS was defined as the most recent clinical contact with a patient. We focused our study on patients with stage III CRC treated with fluorouracil-based chemotherapy with clean resection margins and follow-up of at least 1 month after surgery; with sufficient bulk tumor material available for quantification of PC3, PC9, SMAC, and XIAP by IHC; and previous measurements of those proteins by RPPA.[26] The workflow that illustrates data handling and patient inclusion is shown in Appendix Figure A1, and clinical baseline characteristics are listed in the Data Supplement Table.

Statistical Analysis

We assessed differences among DFS and OS curves using Kaplan-Meier estimates and evaluated statistical significance by log-rank tests. We estimated the relative risk of relapse and death associated with the signatures by unadjusted and two multivariate Cox proportional hazards regression models (Data Supplement). In multivariate model 1, we controlled for clinical covariates previously found to be associated with survival in univariate analyses,[26] namely T stage (T2 to T3 *v* T4) and lymphovascular invasion (invasion *v* no invasion). In multivariate model 2, we additionally adjusted for age (continuous, linear), sex (male *v* female), and tumor location (distal, proximal *v* rectal). We reported hazard ratios, 95% CIs, and likelihood ratio test *P* values. We assessed model performance with the concordance index.

Development and validation of Ensemble and Tree variants of the reduced model were performed in MATLAB with Parallel Computing and Statistics Toolboxes release 2014b (The MathWorks, Cambridge, United Kingdom). Simulations and construction of decision trees were performed with batch jobs run at the Irish Centre for High-End Computing (MATLAB with Parallel Computing Toolbox release 2016b). Batch corrections and survival analyses were performed in R version 3.4.3[30] using the packages sva version 3.26.0[31] and survival version 2.41-3.[32,33] Decision trees were visualized using Graphviz version 2.38.0.[34]

Data and Software Availability

Data sets and analysis source code are available at Zenodo (https://doi.org/10.5281/zenodo.1162683).

RESULTS

Investigating Minimal Models for APOPTO-CELL

We synthetized a virtual patient cohort to develop and evaluate the performance of two variants of a minimal model for APOPTO-CELL: Ensemble and Tree. We previously profiled all five proteins required by APOPTO-CELL in a cohort of patients with stage II/III CRC by quantitative Western blotting (qWB)[22] (Fig 2A, 1). We used these protein profiles to build clinically relevant distributions of each protein input discretized into 20 percentile bins (ie, 20 concentrations; Fig 2A, 2). We built a virtual CRC cohort with one patient for each permutation of the five protein inputs and the 20 concentrations considered, which resulted in approximately 3,200,000 patients (five-dimensional cube with 20 concentrations$^{\text{five proteins}}$; Fig 2A, 3). We ran one simulation for each virtual patient and computed the corresponding APOPTO-CELL signature using 25% substrate cleavage (SC) as the cutoff (SC less than or equal to 25% *v* SC greater than 25% for apoptosis-resistant and apoptosis-susceptible predictions, respectively).[22,26,28]

Next, we examined systematically the predictions landscape obtained by running APOPTO-CELL with an arbitrary subset of the proteins. For each measurement status (available *v* not available) and for each of five protein inputs, we calculated the APOPTO-CELL signature. These 32 combinations (two-measurement status$^{\text{five proteins}}$) span the complete set from having all five protein measurements available to having no quantifications. For each virtual patient, we computed both the ground-truth APOPTO-CELL and the Ensemble signatures. The ground-truth APOPTO-CELL signature was obtained by running APOPTO-CELL using all five proteins as input. The Ensemble signature was determined by running multiple simulations (20 concentrations$^{\text{No. unavailable proteins}}$) using the available inputs and permutations of possible values for the remaining unavailable protein concentrations. For each virtual patient, we enumerated the simulations classified as apoptosis sensitive or apoptosis resistant and used the majority vote as the overall prediction (Fig 2A, 4). For each combination of proteins, we assessed the fractions of virtual patients for whom the Ensemble and ground-truth APOPTO-CELL predictions matched (Fig 2B). Unavailability of APAF1 expression affected only a very limited number of simulations (less than 1%), which rendered PC3 + PC9 + SMAC + XIAP the optimal quartet of proteins. Among the trios and duos of proteins, PC3 + SMAC + XIAP and SMAC + XIAP ranked highest (fourth and eighth out of 32), which matched ground-truth APOPTO-CELL predictions in 92% and 87% of the simulations, respectively. We deemed the Ensemble signatures conclusive when , we computed the same predictions for more than 75% of the simulations (supermajority vote; Fig 2A, 4). The PC3 + SMAC + XIAP and SMAC + XIAP protein combinations conclusively and correctly categorized 84% and 74% of the virtual patients, respectively. However, we conclusively made incorrect predictions for 3% and 5% of patients (ie, more than 75% of the simulations did not match with ground-truth APOPTO-CELL; Fig 2B). SMAC closely followed by XIAP ranked highest as single proteins and achieved 75% and 73% accuracy, respectively. For all other protein combinations, the ability to make conclusive predictions and to identify apoptosis-resistant virtual patients decreased drastically.

We next sought to identify concentration ranges where apoptosis predictions are driven largely by a single protein (Fig 2C). For each concentration of the protein of interest, we aggregated by majority vote the APOPTO-CELL signature across 160,000 simulations (20 concentrations$^{\text{four available proteins}}$). For extreme SMAC (greater than or equal to the 95th percentile) or XIAP (less than the 5th percentile) concentrations, APOPTO-CELL predicted apoptosis sensitivity almost independently of any other protein. Conversely, high XIAP (greater than or equal to the 95th percentile) produced predominantly predictions of apoptosis impairment (86%). The fraction of simulations that predicted apoptosis resistance decreased with increasing concentrations of PC3 and, albeit more moderately, PC9. In contrast, APAF1 expression influenced

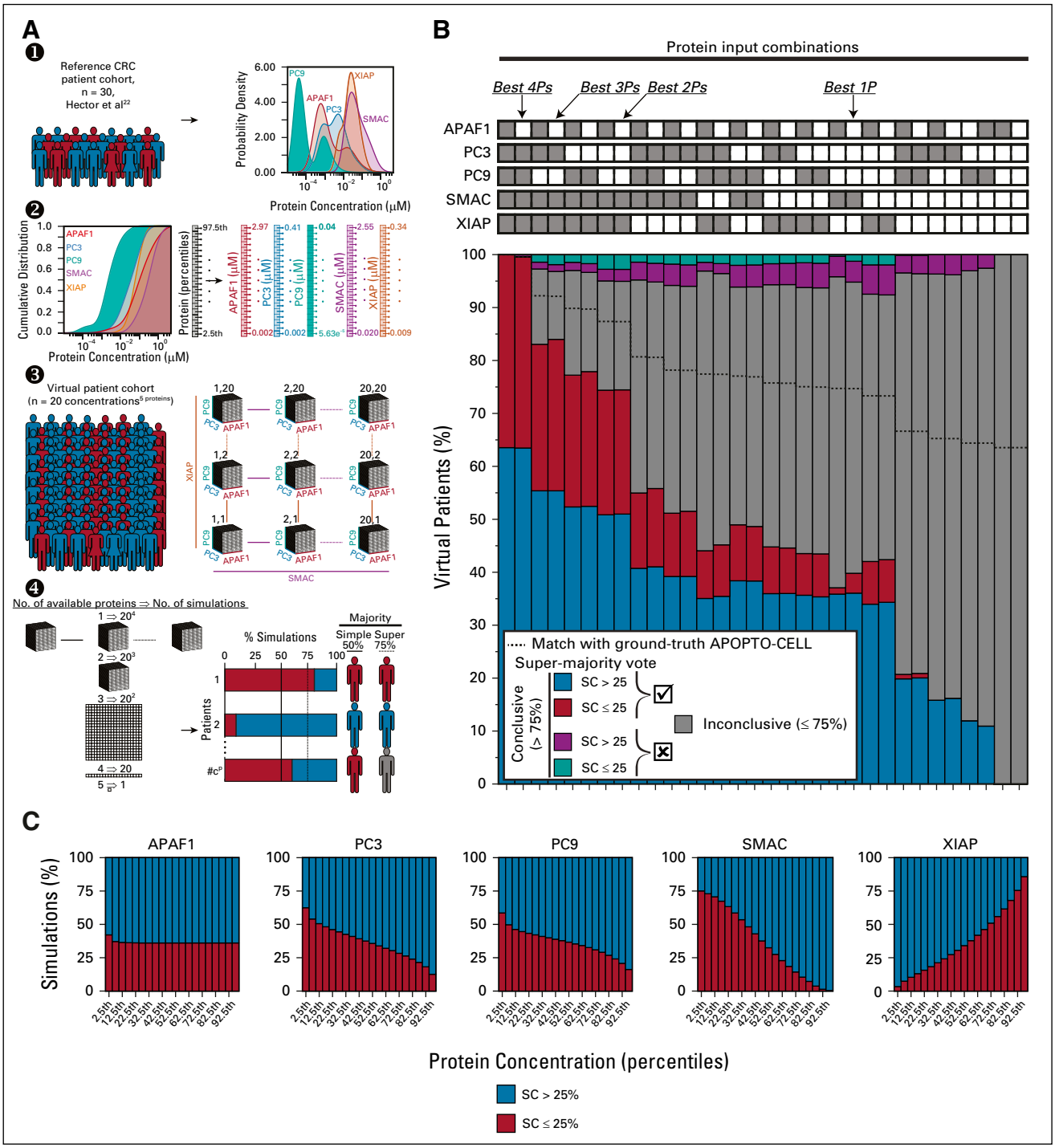


FIG 2. Exploration of alternative strategies to develop a minimal model of the APOPTO-CELL signature on the basis of simulation ensembles (Ensemble). (A) Schematic representation of the pipeline to evaluate systematically how APOPTO-CELL model predictions are affected when computed with a reduced number of inputs. For each protein, $P_i (i \in [1, 5])$ was estimated from a reference cohort[22] (1), and its cumulative distribution was discretized into 20 bins of 5-percentile increments (2). The concentrations corresponding to the centers of each bin were used as inputs to APOPTO-CELL for a full factorial design (five-dimensional cube with $\#c^{\#p}$ simulations where $\#c$ and $\#p$ indicate the number of concentrations and number of proteins, respectively (3). For each minimal model, there are multiple simulations $\#c^{\#p - \#\text{available } p}$, and the overall APOPTO-CELL signature (substrate cleavage [SC] ≤ 25% *v* SC > 25%) was computed using a majority vote (simple- and super-majority vote for > 50% and > 75% of simulations, respectively; 4). (B) and (C) Dependency of the APOPTO-CELL signature by all combinations of measurement status (available *v* not available) for the inputs and by single proteins. In (B), correct/incorrect conclusive predictions by simple- and super-majority vote are indicated by blue/purple, and red/teal for apoptosis susceptibility (SC > 25%) and resistance (SC ≤ 25%), respectively.

APOPTO-CELL predictions only at low concentrations (predominantly less than the 5th percentile).[22,26,28]

Development of Classification Decision Trees for Optimal Selection of Model Inputs

Results from Ensemble revealed high prediction accuracy for subpopulations of patients from only a single or a reduced set of inputs. Hence, we examined whether we could optimally select which proteins to measure, personalized to each patient, and quantify additional inputs only when essential.

We trained a binary classification decision tree (Tree) on the 3,200,000 virtual patients using protein concentrations as input features and the ground-truth APOPTO-CELL

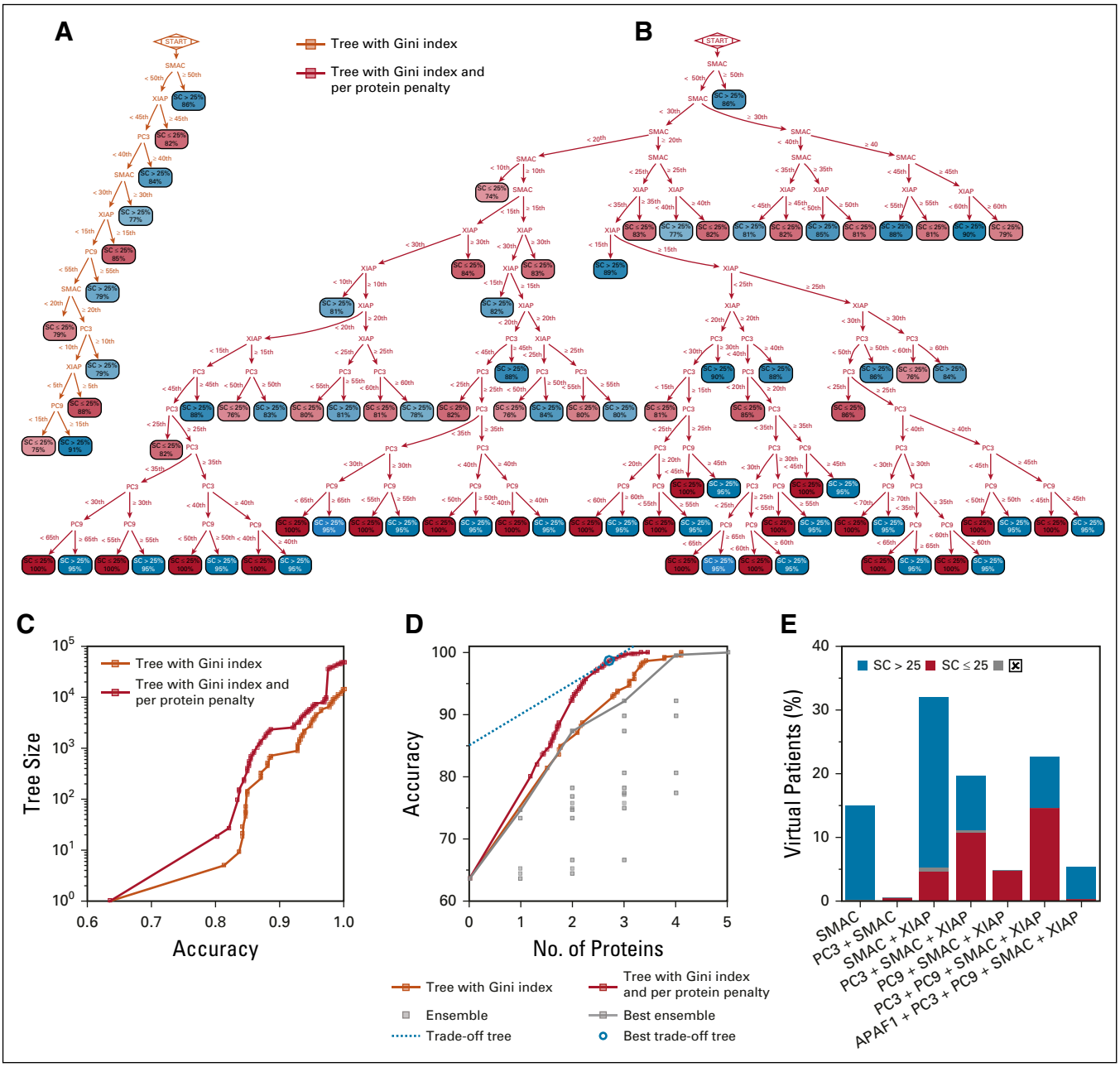


FIG 3. Custom-made classification decision trees (Tree) allow for personalization of which protein inputs to measure without compromising APOPTO-CELL prediction accuracy. (A) and (B) Tree built using Gini index and a modified Gini index including a penalty for any new protein to measure as cost functions. Trees in (A) and (B) were generated for a confidence threshold of 0.74 for illustrative purposes. (C) Tree size as a function of the accuracy for the Tree-based approach. (D) Accuracy (ie, match with the ground-truth APOPTO-CELL signature) as a function of available proteins determined by the Ensemble (gray) and Tree (light and dark red) variants for minimal models. Best trade-off tree for an illustrative example (dotted blue line) is circled in blue. (E) Breakdown of protein requirements and corresponding APOPTO-CELL predictions for the best trade-off tree (blue circle in panel D) for the virtual colorectal cancer cohort. Virtual patients where predictions by Tree and ground-truth APOPTO-CELL signature do not match are shown in gray. See also the Data Supplement.

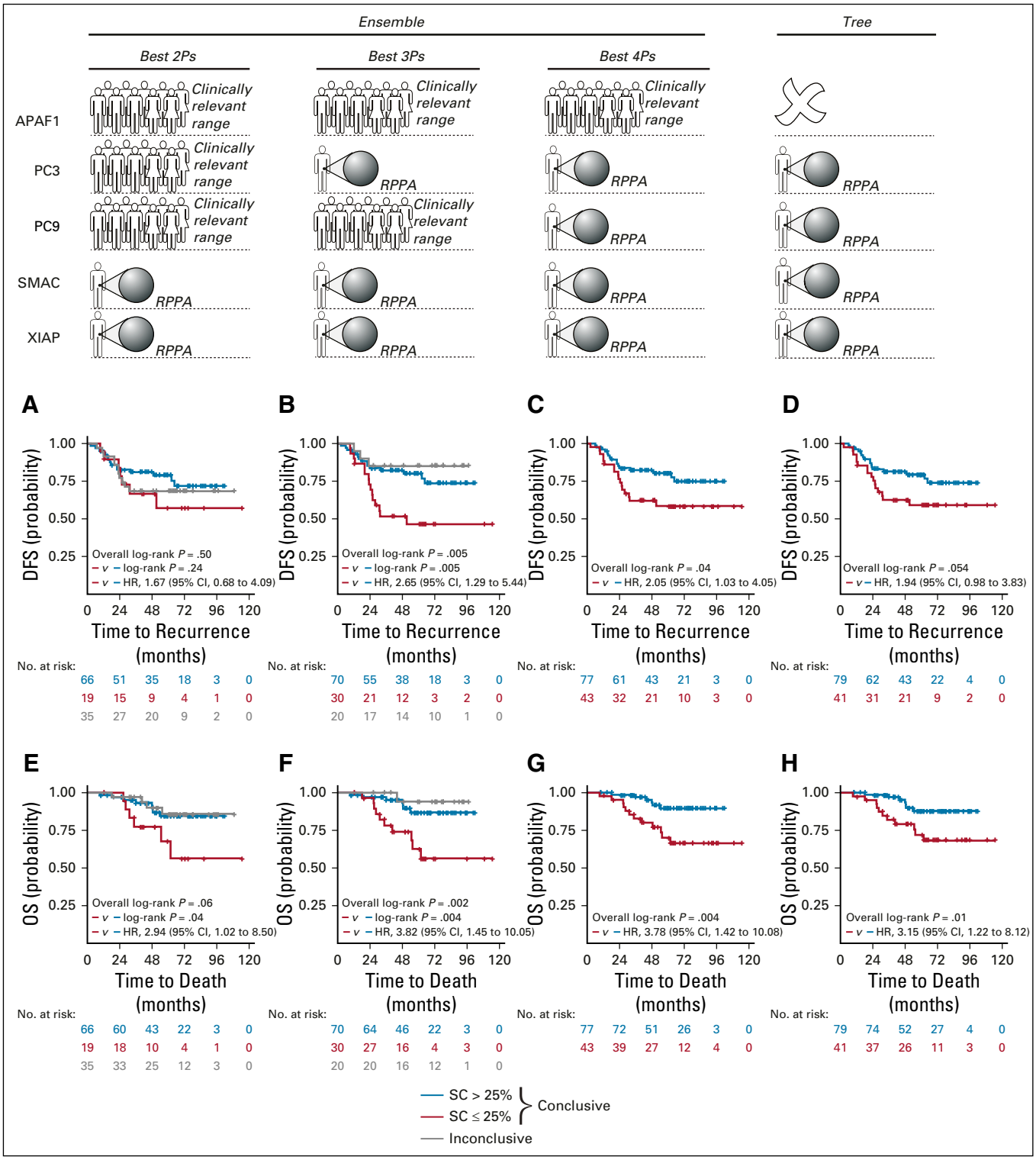


FIG 4. Validation of minimal models for APOPTO-CELL with reverse phase protein array–based measurements in an in-house cohort of patients with stage III colorectal cancer (CRC). (A) to (H) Kaplan-Meier survival estimates for disease-free survival (DFS; top row) and overall survival (OS; bottom row) for patients with stage III CRC. Patients were grouped by the Ensemble signature computed by supermajority vote from patient-specific inputs for the best (A) and (E) duo (SMAC + XIAP, 20^3 simulations per patient), (B) and (F) trio (procaspase-3 [PC3] + SMAC + XIAP, 20^2 simulations per patient), and (C) and (G) quadruple (PC3 + procaspase-9 [PC9] + SMAC + XIAP, 20 simulations per patient) combination of proteins and by (D) and (H) the Tree signature (best trade-off tree for PC3 + PC9 + SMAC + XIAP determined as in Fig 3D; Tree-4P).

signature (SC greater than 25% *v* SC less than or equal to 25%) as target class (Data Supplement). The root and internal nodes include a predictor (protein) and a split point (concentration). The root node contains the predictor and concentration, which yielded the best split. The leaf node encodes the APOPTO-CELL signature and its confidence level. Trees that meet a required prediction confidence can be built by tuning the splitting routine. To build a tree, we tested all possible predictors/proteins and split points/concentrations and selected recursively the optimal split (protein and concentration) with the Gini index as evaluation metric (Fig 3A). To build trees that would favor re-use of proteins previously evaluated in the internal nodes rather than require proteins not yet in use, we introduced a per-protein penalty in the Gini index metric (penalized Gini index; Fig 3B). By varying the confidence threshold and the per-protein penalty in a parameter scan, we observed an increase in size (Fig 3C) and accuracy (Fig 3D) when comparing trees built with the default and penalized Gini index metrics (light and dark red solid lines, respectively). Tree outperformed the best (solid line) and individual (single data points) Ensemble approaches (Fig 3D, in gray) previously outlined in Fig 2C.

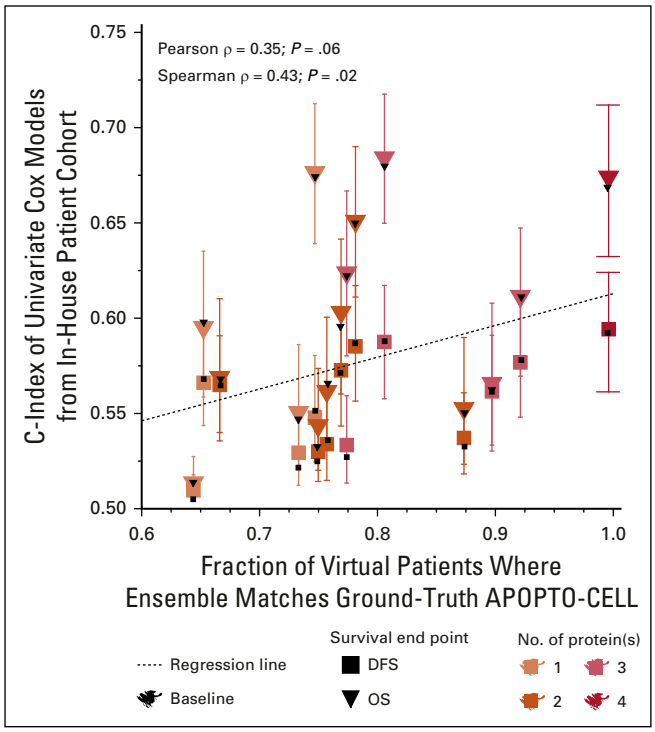


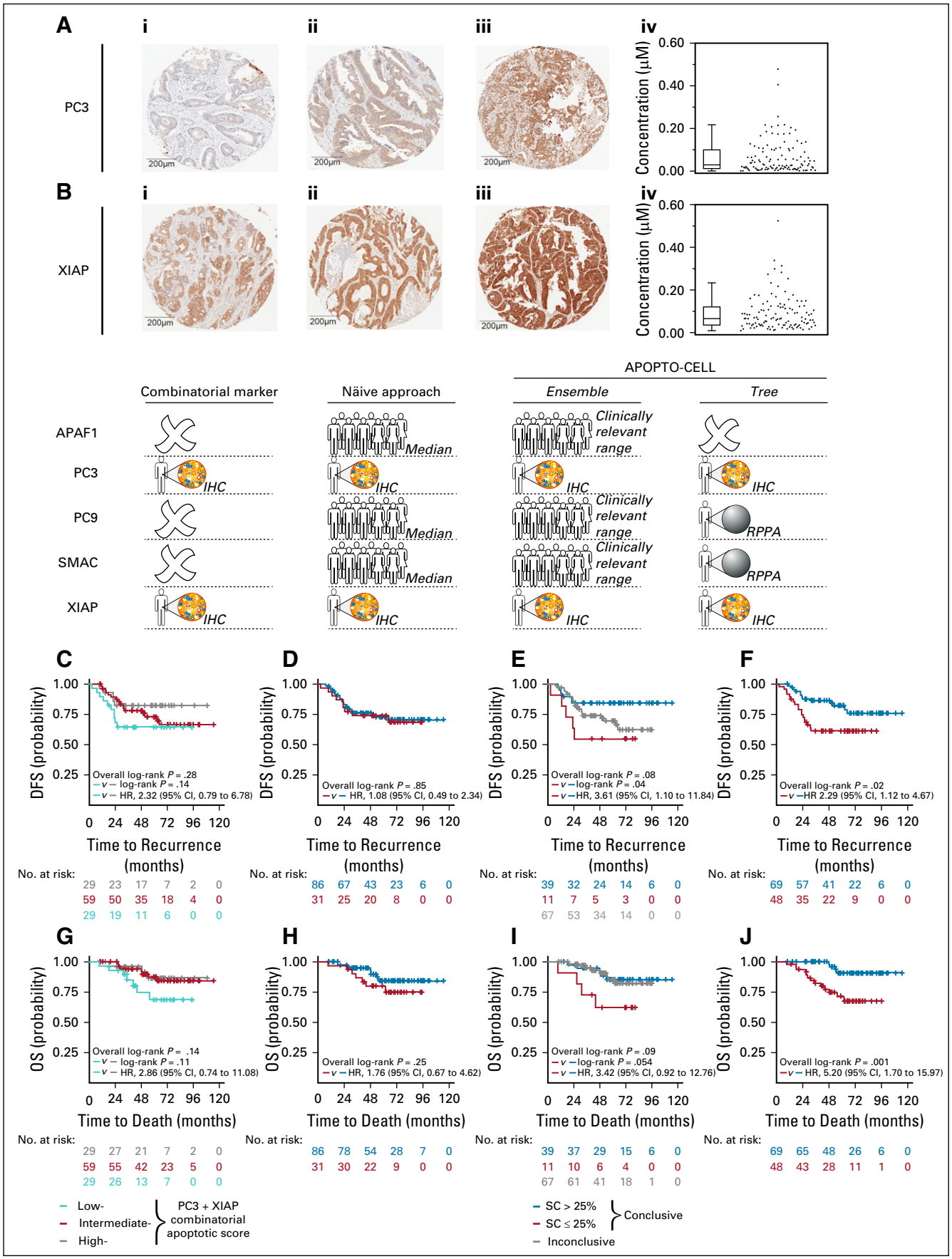
FIG 5. Systematic evaluation of the association between accuracy and prognostic value for Ensemble. Association between accuracy (ie, agreement with ground-truth APOPTO-CELL; Fig 2B) and concordance indices (c-indices) from corresponding univariate Cox proportional hazards regression models for Ensemble simulation groups. Ensemble simulations were run for each combination of input status (available v not available) for each of the four proteins quantified by reverse phase protein array (procaspase3 [PC3], procaspase9 [PC9], SMAC, XIAP) where at least one protein was available, which resulted in 15 groups of simulations ($2^{4\ \text{proteins}}$-1). For performance reasons, 10 concentrations per protein were tested instead of the 20 used in all other analyses presented, which totaled 5,569,200 simulations for 120 patients. Each patient was classified as apoptosis sensitive (substrate cleavage [SC] > 25%) or resistant (SC ≤ 25%) for each set of simulations on the basis of simple-majority vote. For each Ensemble simulation group, univariate Cox proportional hazards regression models were fitted for disease-free survival (DFS) and overall survival (OS), and the corresponding c-indices were computed for the baseline (unperturbed patients [black]) and 1,000 bootstrapped (resampled patients) sets. Medians and interquartile ranges are shown for c-indices computed from bootstrapped sets.

Tree can be customized for a given clinical application. As an illustrative example, we deemed it worth measuring an additional protein only if it would result in an increase of at least 5% in accuracy (best trade-off tree; Fig 3D, blue dotted line). The tree meeting this requirement (Tree-5P; Fig 3D, blue circle; Data Supplement) yielded results that matched ground-truth APOPTO-CELL in 99% of the virtual patients by measuring, on average, 2.7 proteins rather than five. Tree-5P demonstrated that measurement of the full protein set is required only for 5% of the virtual patients, whereas for approximately one half (48%), one to two proteins are sufficient (Fig 3E).

Validation of Minimal Models in a Cohort of Patients With Stage III CRC, With Protein Inputs Quantified by RPPA

We validated the minimal APOPTO-CELL models in a real-world cohort of patients with stage III CRC treated with fluorouracil-based chemotherapy for whom the expression of PC3 + PC9 + SMAC + XIAP had been previously determined by RPPA.[26] We evaluated the performances of Ensemble for the best double, triple, and quadruple combination of proteins identified in Fig 2C. For each patient, we tested 20 concentration$^{\text{No. unavailable proteins}}$ scenarios, which resulted in 20^3, 20^2, and 20^1 simulations for the best double, triple, and quadruple protein combinations, respectively, which totaled approximately 1,010,400 evaluations for 120 patients. We observed differences in Kaplan-Meier estimates for OS (P = .04) but not for DFS (P = .24) when comparing patients categorized as apoptosis sensitive or resistant by Ensemble for the best duo combination (Figs 4A and 4E; Data Supplement Table). When testing the best trio and quartet, we found that patients predicted as apoptosis impaired by Ensemble showed reduced DFS and OS compared with those predicted as apoptosis sensitive (P < .05; Figs 4B, 4C, 4F, and 4G; Data Supplement Table). Furthermore, when testing systematically all combinations of proteins, we found a statistically significant association (Spearman ρ = 0.43; P = .02) between Ensemble accuracy and prognostic value (Fig 5).

We observed differences in DFS (P = .054) and OS (P = .01) probabilities when comparing patients categorized by Tree-4P (best trade-off tree built with only the four proteins measured by RPPA and selected as in Fig 3D [Figs 4D and 4H; Data Supplement Table]). These results



closely resembled those obtained by using the best quartet of proteins (0.0% and 3.0% difference in concordance index computed for DFS and OS, respectively) while requiring measurement of, on average, only 2.8 rather than four proteins (Figs 4D and 4H; Data Supplement Table).

Minimal Models Provide Insights Into Patient Risk, Even With a Suboptimal Protein Set Quantified by IHC

We tested the performances of Ensemble and Tree on protein inputs quantified by IHC, a technique used in the clinical pathologic routine. When comparing IHC- and RPPA-based measurements (n = 152; Appendix Fig A1), we found an acceptable quantitative agreement for PC3 (Spearman $\rho = 0.24$; $P = .003$) and XIAP (Spearman $\rho = 0.28$; $P < .001$), whereas we observed no association for PC9 (Spearman $\rho = 0.12$; $P = .15$) and SMAC (Spearman $\rho = 0.09$; $P = .29$). RPPA is the gold standard for quantitative proteomics[35-37]; thus, for subsequent analyses, we retained only proteins with significant associations between RPPA- and IHC-based measurements. Figs 6A and 6B show representative images for PC3 and XIAP of weak, moderate, and strong IHC staining with validated antibodies (Appendix Fig A2; Data Supplement) and corresponding batch-corrected histoscore distributions (Appendix Fig A3). We did not find statistically significant differences for DFS and OS ($P > .05$) when examining PC3 and XIAP as a single (Appendix Fig A4; Data Supplement Table) or combinatorial (Figs 6C and 6G; Data Supplement Table) biomarker.

Next, we tested various strategies to tackle the missing data of IHC-based protein measurements for PC9 + SMAC + XIAP as inputs to APOPTO-CELL. We found no association between APOPTO-CELL predictions and survival ($P > .05$) when using a naïve approach (individual expression for PC3 + XIAP and median concentrations for other inputs; Figs 6D and 6H; Data Supplement Table) or Tree (customized tree with only PC3 + XIAP, Tree-2P; Appendix Fig A5; Data Supplement Table). In contrast, application of Ensemble (personalized quantifications for PC3 + XIAP coupled with 20^3 simulations for APAF1 + PC9 + SMAC) revealed differences in DFS ($P = .04$) and OS ($P = .054$; Figs 6E and 6I; Data Supplement Table).

In exploratory analyses, we investigated the use of protein fingerprinting by IHC for PC3 + XIAP and by RPPA for PC9 + SMAC as inputs for Tree-4P (Figs 6F and 6J; Data Supplement Table), ground-truth APOPTO-CELL (Appendix Figs A6A and A6B; Data Supplement Table), and an enriched apoptosis signature developed to account for PC3-mediated apoptosome-independent signaling (APOPTO-CELL-PC3[26]; Appendix Figs A6E and A6F; Data Supplement Table). All three classifiers retained prognostic value for DFS and OS ($P < .05$), which corroborates previous results on the basis of RPPA-only quantifications[26] (Appendix Fig A6C). Moreover, permutation analyses (Data Supplement) indicated that PC3 + XIAP expression by IHC critically contributed to the prognostic value of the ground-truth APOPTO-CELL signature (Appendix Fig A6D).

DISCUSSION

The appeal of dynamic pathway models as biomarkers stems from gaining a mechanistic understanding of the network (dys)regulation and potential therapeutic interventions.[38,39] Network models deliver personalized predictions by tailoring the model skeleton with individualized inputs. However, input quantifications represent a major bottleneck in translating mathematical models into the clinic. Inputs should be measurable from minimal amounts of sample with high-throughput, robust, cost-effective, and clinically amenable techniques equally suited for fresh frozen/formalin-fixed paraffin-embedded tissues.

We used a mathematical model of caspase-driven apoptosis in CRC (APOPTO-CELL[28]) with proven prognostic value[22,26] and treatment stratification capabilities[40] as an illustrative example. We built and validated a computational platform to capitalize on the systems-level understanding provided by the network dynamic while optimizing the inputs required. This approach not only reduces the cost for molecular characterization but also limits the exhaustion of clinical material. We developed two computational methods (Ensemble and Tree) with distinct and complementary applications in clinical settings.

We designed Ensemble to identify a consistent inputs subset to quantify for all patients enrolled in a study, whereas Tree provides a personalized input prioritization system. Ensemble could be applied when designing studies with different cohort characteristics, such as different CRC stages and cancer types. Ensemble also could be applied in clinical trials when planning the development

FIG 6. Assessment of minimal models for APOPTO-CELL with a subset of the inputs quantified by immunohistochemistry (IHC) for the in-house cohort of patients with stage III colorectal cancer (CRC). (A) and (B) Representative images of tumor cores with weak (i), moderate (ii), and strong (iii) staining and corresponding protein distributions (iv) for procaspase-3 (PC3) and XIAP. (C) to (J) Kaplan-Meier survival estimates for disease-free survival (DFS; top row) and overall survival (OS; bottom row). (C) and (G) Patients were grouped by the PC3 + XIAP combinatorial apoptotic score (low, PC3 ≤ median + XIAP > median; high, PC3 > median + XIAP ≤ median; intermediate, otherwise). Patients were categorized by the APOPTO-CELL signature computed on the basis of various inputs and computational strategies: (D) and (H) patient-specific quantifications for PC3 + XIAP and population median values for APAF1 + procaspase-9 (PC9) + SMAC (naïve approach for APOPTO-CELL), (E) and (I) patient-specific quantifications for PC3 + XIAP and clinically relevant range for APAF1 + PC9 + SMAC (20^3 simulations per patient, Ensemble), and (F) and (J) best trade-off tree for PC3 + PC9 + SMAC + XIAP (Tree-4P). All patient-specific quantifications were IHC based, except in (F) and (J), where measurements for PC9 and SMAC were reverse phase protein array based. See also Appendix Figures A1 to A6.

of companion assays for clinical testing of inhibitor of apoptosis protein antagonists. Furthermore, it enables the integration of historical data sets. Of note, Ensemble also would provide estimates of loss of accuracy and prognostic power as rationale for optimizing the cost-effectiveness of the model reduction.

In contrast, Tree is the optimal approach for individual patients who come to the clinic, preserving 99% accuracy while requiring 46% fewer quantifications. However, this approach has the overhead necessary to manage allocation of patients through the various quantification paths. Thus, for this method to be valuable and easily adopted in busy and fast-paced clinical settings, a robust handling system would need to accompany the deployment of tree-based approaches.

For some applications, not all model inputs may be detected accurately and reliably by technologies that integrate seamlessly in the clinic. Techniques such as qWB and quantitative polymerase chain reaction may be optimal for protein and transcript profiling in the development, validation, and refinement phases of a dynamic system model cycle. However, correlation of mRNA and protein levels often is poor in tumor tissue,[41,42] and qWB is not a routine biochemistry/histopathology technique in the clinic. We previously showed how APOPTO-CELL inputs could be quantified reliably not only by qWB but also by RPPA.[26] However, integration of an RPPA-based workflow into the clinical environment is also not feasible because RPPA is an exploratory tool used for the collective high-throughput analysis of larger patient cohorts. In contrast, we demonstrated that at least for a subset of the inputs, IHC, a more clinically amenable protein profiling alternative, represents a valid substitute. Ensemble predictions using patient-specific quantifications for PC3 + XIAP outperformed the prognostic value of the same two proteins used as a combinatorial biomarker. These results highlight how Ensemble makes it possible to exploit system-level properties, even with a suboptimal inputs set, instead of falling back to combinatorial static biomarkers. Of note, Ensemble provides an estimate of confidence in the model predictions and thus allows the flagging of patients who require further investigations.

Although our study has the limitation that the proteins under investigation currently are not assessed routinely, the methods developed can be applied to multiple diagnostic assays that investigate oncogenic pathways and proteins that are indeed currently routinely assessed in the clinic. Although many challenges are at play when attempting to bridge the gap between bench and bedside, this work presents actionable methods to assist in the application of systems medicine approaches in the clinic.

AFFILIATIONS

[1]Royal College of Surgeons in Ireland, Dublin, Ireland
[2]OncoMark, Dublin, Ireland
[3]Beaumont Hospital, Dublin, Ireland
[4]Université Paris Descartes, Paris, France
[5]Queen's University Belfast, Belfast, United Kingdom
[6]University of Stuttgart, Stuttgart, Germany

CORRESPONDING AUTHOR

Jochen H.M. Prehn, PhD, Department of Physiology and Medical Physics, Royal College of Surgeons in Ireland, 123 St Stephen's Green, Dublin 2, Ireland; e-mail: prehn@rcsi.ie.

PRIOR PRESENTATION

Presented at the Annual Meeting of the Irish Association for Cancer Research, Dublin, Ireland, February 21-23, 2018, and Annual Research Day, Dublin, Ireland, March 7, 2018.

SUPPORT

Supported by the European Union Framework Programme 7 (FP7 APO-DECIDE, contract No. 306021). D.B.L., M.Re., and J.H.M.P were supported by a Science Foundation Ireland/Department of Enterprise and Learning Partnership Award (14/IA/2582). J.H.M.P. was supported by a Science Foundation Ireland Investigator Award (13/IA/1881) and by the Irish Cancer Society Collaborative Cancer Research Centre BREAST-PREDICT (CCRC13GAL).

AUTHOR CONTRIBUTIONS

Conception and design: Manuela Salvucci, Pierre Laurent-Puig, Mairin Rafferty, William M. Gallagher, Markus Rehm, Jochen H.M. Prehn

Administrative support: Manuela Salvucci, Mairin Rafferty, Jochen H.M. Prehn

Provision of study material or patients: Deborah A. McNamara, Elaine W. Kay, Sandra Van Schaeybroeck

Collection and assembly of data: Manuela Salvucci, Arman Rahman, Elaine W. Kay, Pierre Laurent-Puig, Patrick G. Johnston, Richard Wilson, Manuel Salto-Tellez, Sandra Van Schaeybroeck, Jochen H.M. Prehn

Data analysis and interpretation: Manuela Salvucci, Arman Rahman, Alexa J. Resler, Girish M. Udupi, Deborah A. McNamara, Elaine W. Kay, Pierre Laurent-Puig, Daniel B. Longley, Mark Lawler, William M. Gallagher, Markus Rehm, Jochen H.M. Prehn

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/site/ifc.

Pierre Laurent-Puig

Honoraria: Amgen, AstraZeneca, Boehringer Ingelheim, Merck Serono, Merck, Roche, Sanofi

Consulting or Advisory Role: Merck, Bristol-Myers Squibb, Boehringer Ingelheim

Patents, Royalties, Other Intellectual Property: Inventor of mir31-3p licensed to IntegraGen Genomics

Travel, Accommodations, Expenses: Roche, Merck

Daniel B. Longley
Stock and Other Ownership Interests: Fusion Antibodies
Consulting or Advisory Role: Astex Pharmaceuticals
Research Funding: Astex Pharmaceuticals
Patents, Royalties, Other Intellectual Property: Inhibitors of the anti-apoptotic protein FLIP, Patent Cooperation Treaty filed
Travel, Accommodations, Expenses: Astex Pharmaceuticals

Patrick G. Johnston
Stock and Other Ownership Interests: Almac Diagnostics

Mark Lawler
Honoraria: Pfizer

Richard Wilson
Consulting or Advisory Role: Merck Serono, Sirtex Medical, Amgen, Servier, Clovis Oncology, Halozyme Therapeutics, Bristol-Myers Squibb
Research Funding: Almac Group (Inst)
Travel, Accommodations, Expenses: Merck Serono, Amgen

Manuel Salto-Tellez
Consulting or Advisory Role: Philips Healthcare, Visiopharm, Bristol-Myers Squibb, Merck
Patents, Royalties, Other Intellectual Property: Co-inventor of QuPath, an open source system for digital pathology analysis

Mairin Rafferty
Employment: OncoMark, Deciphex
Leadership: Deciphex
Stock and Other Ownership Interests: OncoMark, Deciphex

William M. Gallagher
Employment: OncoMark
Leadership: OncoMark
Stock and Other Ownership Interests: OncoMark
Consulting or Advisory Role: Carrick Therapeutics
Research Funding: Carrick Therapeutics
Patents, Royalties, Other Intellectual Property: Two patents licensed to OncoMark
Travel, Accommodations, Expenses: OncoMark

Markus Rehm
Patents, Royalties, Other Intellectual Property: Patent on a mathematical method to generate predictions on responsiveness to apoptosis-inducing treatments in cancer

Jochen H.M. Prehn
Patents, Royalties, Other Intellectual Property: Inventor on patent application No. WO2013001002: A Computer-Implemented System and Method for the Prediction of Cancer Response to Genotoxic Chemotherapy and Personalised Neoadjuvant Treatments (PCCP) that describes the use of APOPTO-CELL modeling as a prognostic biomarker and stratification tool for solid tumors, including colorectal cancer

No other potential conflicts of interest were reported.

ACKNOWLEDGMENT

We wish to thank the patients who participated in this study. We kindly acknowledge Sonali Dasgupta, Orna Bacon, Sarah Curry, Áine C. Murphy, Lisa M. Schöller, Maximilian L. Würstle, Elisabeth Zink, Sophie Camilleri-Broët, Camilla Pilati, Nadege Rice for their assistance with the procurement of the tumor samples, collection of the clinical data, technical assistance and continuous scientific support. We wish to acknowledge the DJEI/DES/SFI/HEA Irish Centre for High-End Computing (ICHEC) for the provision of computational facilities and support.

REFERENCES

1. Dienstmann R, Rodon J, Tabernero J: Optimal design of trials to demonstrate the utility of genomically-guided therapy: Putting precision cancer medicine to the test. Mol Oncol 9:940-950, 2015
2. Meric-Bernstam F, Brusco L, Shaw K, et al: Feasibility of large-scale genomic testing to facilitate enrollment onto genomically matched clinical trials. J Clin Oncol 33:2753-2762, 2015
3. Le Tourneau C, Delord J-P, Gonçalves A, et al: Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): A multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial. Lancet Oncol 16:1324-1334, 2015
4. Swanton C, Soria J-C, Bardelli A, et al: Consensus on precision medicine for metastatic cancers: A report from the MAP conference. Ann Oncol 27:1443-1448, 2016
5. Stockley TL, Oza AM, Berman HK, et al: Molecular profiling of advanced solid tumors and patient outcomes with genotype-matched clinical trials: The Princess Margaret IMPACT/COMPACT trial. Genome Med 8:109, 2016
6. Guinney J, Dienstmann R, Wang X, et al: The consensus molecular subtypes of colorectal cancer. Nat Med 21:1350-1356, 2015
7. Isella C, Brundu F, Bellomo SE, et al: Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. Nat Commun 8:15107, 2017
8. Hamblin A, Wordsworth S, Fermont JM, et al: Clinical applicability and cost of a 46-gene panel for genomic analysis of solid tumours: Retrospective validation and prospective audit in the UK National Health Service. PLoS Med 14:e1002230, 2017
9. Danielsen HE, Hveem TS, Domingo E, et al: Prognostic markers for colorectal cancer; estimating ploidy and stroma. Ann Oncol 29:616-623, 2018
10. Marisa L, de Reyniès A, Duval A, et al: Gene expression classification of colon cancer into molecular subtypes: Characterization, validation, and prognostic value. PLoS Med 10:e1001453, 2013
11. De Sousa E Melo F, Wang X, Jansen M, et al: Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. Nat Med 19:614-618, 2013
12. Sadanandam A, Lyssiotis CA, Homicsko K, et al: A colorectal cancer classification system that associates cellular phenotype and responses to therapy. Nat Med 19:619-625, 2013
13. Brat DJ, Verhaak RG, Aldape KD, et al: Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. N Engl J Med 372:2481-2498, 2015
14. Bailey P, Chang DK, Nones K, et al: Genomic analyses identify molecular subtypes of pancreatic cancer. Nature 531:47-52, 2016
15. Sørlie T, Perou CM, Tibshirani R, et al: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98:10869-10874, 2001
16. Parker JS, Mullins M, Cheang MCU, et al: Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 27:1160-1167, 2009
17. Garraway LA: Genomics-driven oncology: Framework for an emerging paradigm. J Clin Oncol 31:1806-1814, 2013

18. Trifiletti DM, Sturz VN, Showalter TN, et al: Towards decision-making using individualized risk estimates for personalized medicine: A systematic review of genomic classifiers of solid tumors. PLoS One 12:e0176388, 2017
19. Carlson JJ, Roth JA: The impact of the Oncotype Dx breast cancer assay in clinical practice: A systematic review and meta-analysis. Breast Cancer Res Treat 141:13-22, 2013 [Erratum: Breast Cancer Res Treat 146:233, 2014]
20. Di Narzo AF, Tejpar S, Rossi S, et al: Test of four colon cancer risk-scores in formalin fixed paraffin embedded microarray gene expression data. J Natl Cancer Inst 106:dju247, 2014
21. Goldstein LJ, Gray R, Badve S, et al: Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. J Clin Oncol 26:4063-4071, 2008
22. Hector S, Rehm M, Schmid J, et al: Clinical application of a systems model of apoptosis execution for the prediction of colorectal cancer therapy responses and personalisation of therapy. Gut 61:725-733, 2012
23. Lindner AU, Concannon CG, Boukes GJ, et al: Systems analysis of BCL2 protein family interactions establishes a model to predict responses to chemotherapy. Cancer Res 73:519-528, 2013
24. Flanagan L, Lindner AU, de Chaumont C, et al: BCL2 protein signalling determines acute responses to neoadjuvant chemoradiotherapy in rectal cancer. J Mol Med (Berl) 93:315-326, 2015
25. Lindner AU, Salvucci M, Morgan C, et al: BCL-2 system analysis identifies high-risk colorectal cancer patients. Gut 66:2141-2148, 2017
26. Salvucci M, Würstle ML, Morgan C, et al: A stepwise integrated approach to personalized risk predictions in stage III colorectal cancer. Clin Cancer Res 23:1200-1212, 2017
27. Fey D, Halasz M, Dreidax D, et al: Signaling pathway models as biomarkers: Patient-specific simulations of JNK activity predict the survival of neuroblastoma patients. Sci Signal 8:ra130, 2015
28. Rehm M, Huber HJ, Dussmann H, et al: Systems analysis of effector caspase activation and its control by X-linked inhibitor of apoptosis protein. EMBO J 25:4338-4349, 2006
29. Murphy ÁC, Weyhenmeyer B, Schmid J, et al: Activation of executioner caspases is a predictor of progression-free survival in glioblastoma patients: A systems medicine approach. Cell Death Dis 4:e629, 2013
30. R Core Team: R: A language and environment for statistical computing, 2017. http://www.R-project.org
31. Leek JT, Johnson WE, Parker HS, et al: sva: Surrogate variable analysis, 2018. R package version 3.28.0
32. Therneau TM: A package for survival analysis in S, 2017
33. Therneau TM, Grambsch PM: Modeling Survival Data: Extending the {C}ox Model. New York, NY, Springer, 2000
34. Gansner ER, North SC: An open graph visualization system and its applications to software engineering. Softw Pract Exper 30:1203-1233, 2000
35. Pierobon M, Wulfkuhle J, Liotta L, et al: Application of molecular technologies for phosphoproteomic analysis of clinical samples. Oncogene 34:805-814, 2015
36. Wachter A, Bernhardt S, Beissbarth T, et al: Analysis of reverse phase protein array data: From experimental design towards targeted biomarker discovery. Microarrays (Basel) 4:520-539, 2015
37. Negm OH, Muftah AA, Aleskandarany MA, et al: Clinical utility of reverse phase protein array for molecular classification of breast cancer. Breast Cancer Res Treat 155:25-35, 2016
38. Henderson D, Ogilvie LA, Hoyle N, et al: Personalized medicine approaches for colon cancer driven by genomics and systems biology: OncoTrack. Biotechnol J 9:1104-1114, 2014
39. Kolch W, Fey D: Personalized computational models as biomarkers. J Pers Med 7:9, 2017
40. Crawford N, Salvucci M, Hellwig CT, et al: Simulating and predicting cellular and in vivo responses of colon cancer to combined treatment with chemotherapy and IAP antagonist Birinapant/TL32711. Cell Death Differ [Epub ahead of print on March 2, 2018]
41. Akbani R, Ng PKS, Werner HMJ, et al: A pan-cancer proteomic perspective on The Cancer Genome Atlas. Nat Commun 5:3887, 2014 [Erratum: Nat Commun 6:4852, 2015]
42. Kosti I, Jain N, Aran D, et al: Cross-tissue analysis of gene and protein expression in normal and cancer tissues. Sci Rep 6:24799, 2016

APPENDIX

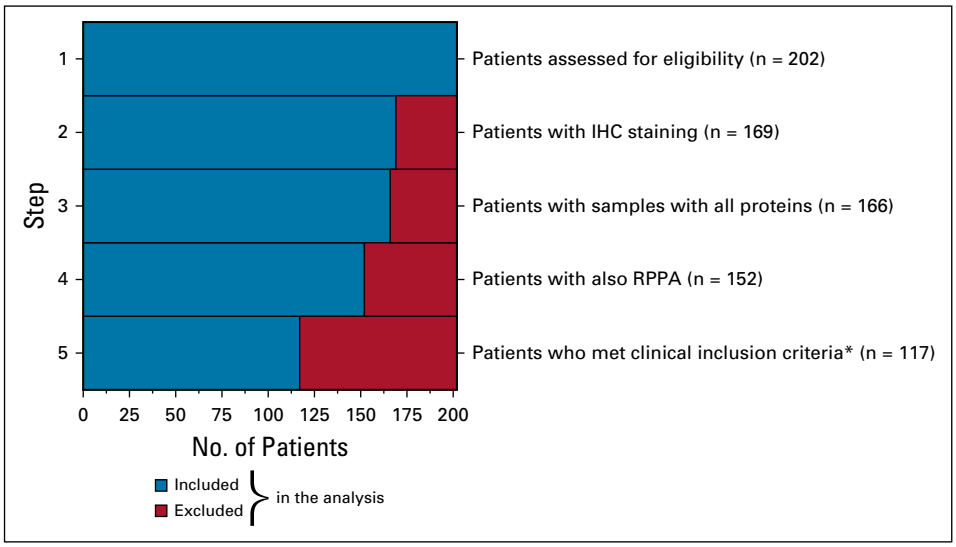


FIG A1. Flow diagram that depicts the total number of patients available at each phase of the data analysis for immunohistochemistry (IHC)-based protein quantification. (*) Clinical inclusion criteria were diagnosis of stage III colorectal cancer, treatment with fluorouracil-based chemotherapy, clean resection margins (R0), and follow-up of at least 1 month after resection surgery (Salvucci M, Würstle ML, Morgan C, et al: Clin Cancer Res 23:1200-1212, 2016). RPPA, reverse phase protein array.

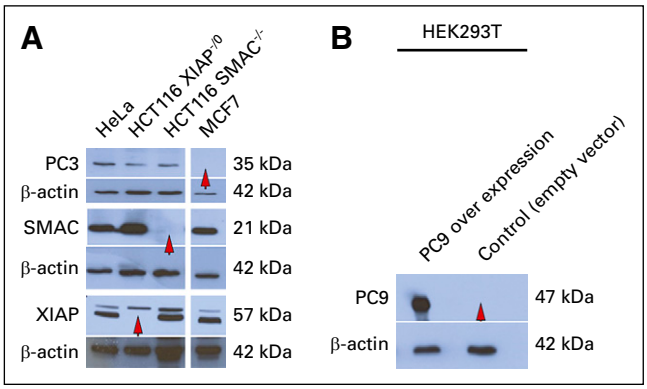


FIG A2. Antibody specificity tested by Western blotting. (A) and (B) Western blots demonstrate antibody quality for proteins of interest. β-actin served as the loading control.

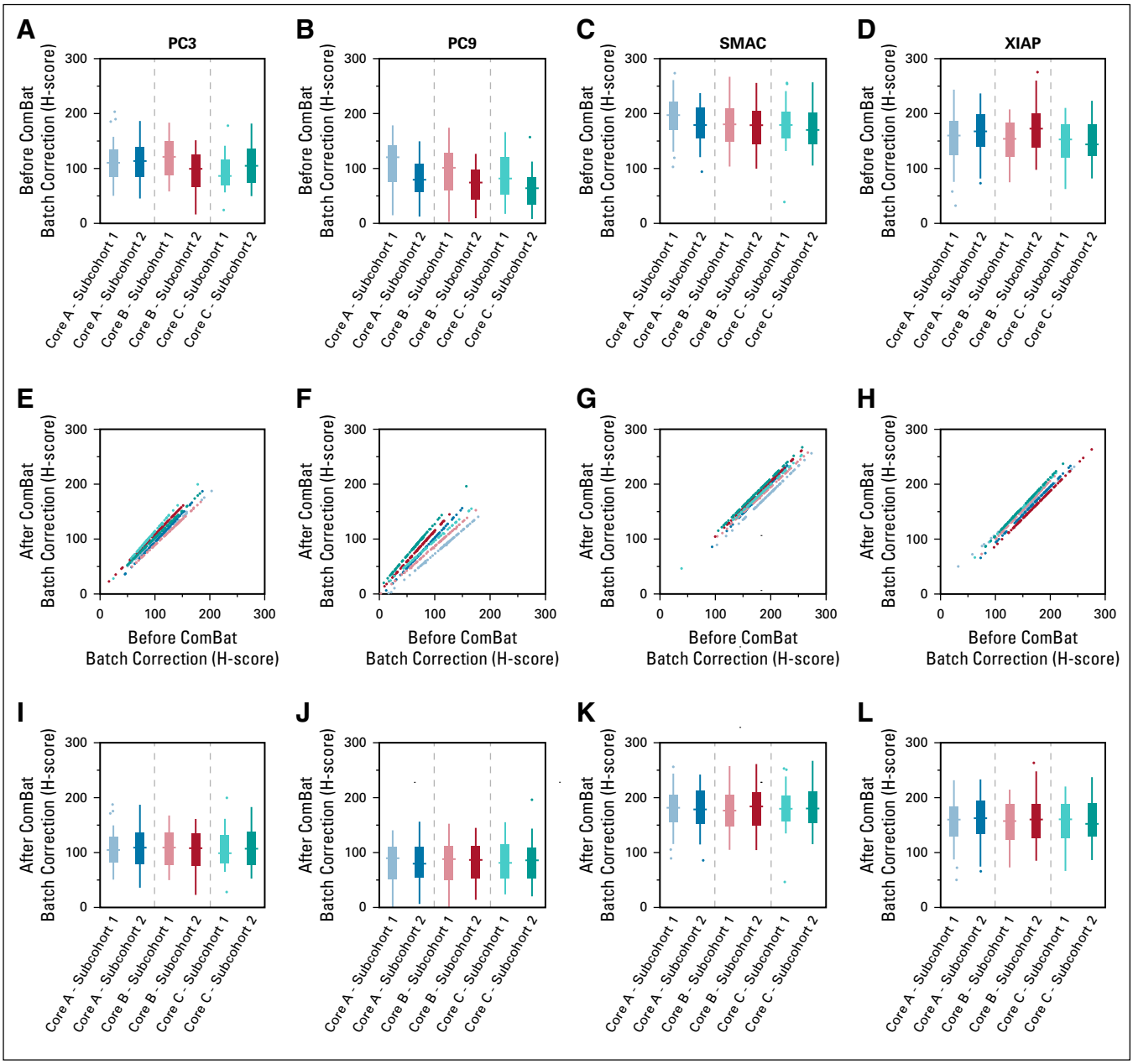


FIG A3. Batch correction of protein expression quantifications in tumor cores from 166 patients. Owing to the large number of patients, two slides were required for each of the three cores analyzed (core A, core B, and core C). Protein expression (histoscore [H-score]) for (A) procaspase-3 (PC3), (B) procaspase-9 (PC9), (C) SMAC, and (D) XIAP in tumor tissue samples showed evidence of batch effects across slides (subcohort 1 and subcohort 2). (E-H) Batch effects were corrected by applying ComBat (Leek JT, Johnson WE, Parker HS, et al: sva: Surrogate Variable Analysis, 2017. R package version 3.26.0. http://bioconductor.org/packages/release/bioc/html/sva.html) with stage and chemotherapy administration as covariates. Corresponding batch-corrected H-scores are shown for (I) PC3, (J) PC9, (K) SMAC, and (L) XIAP.

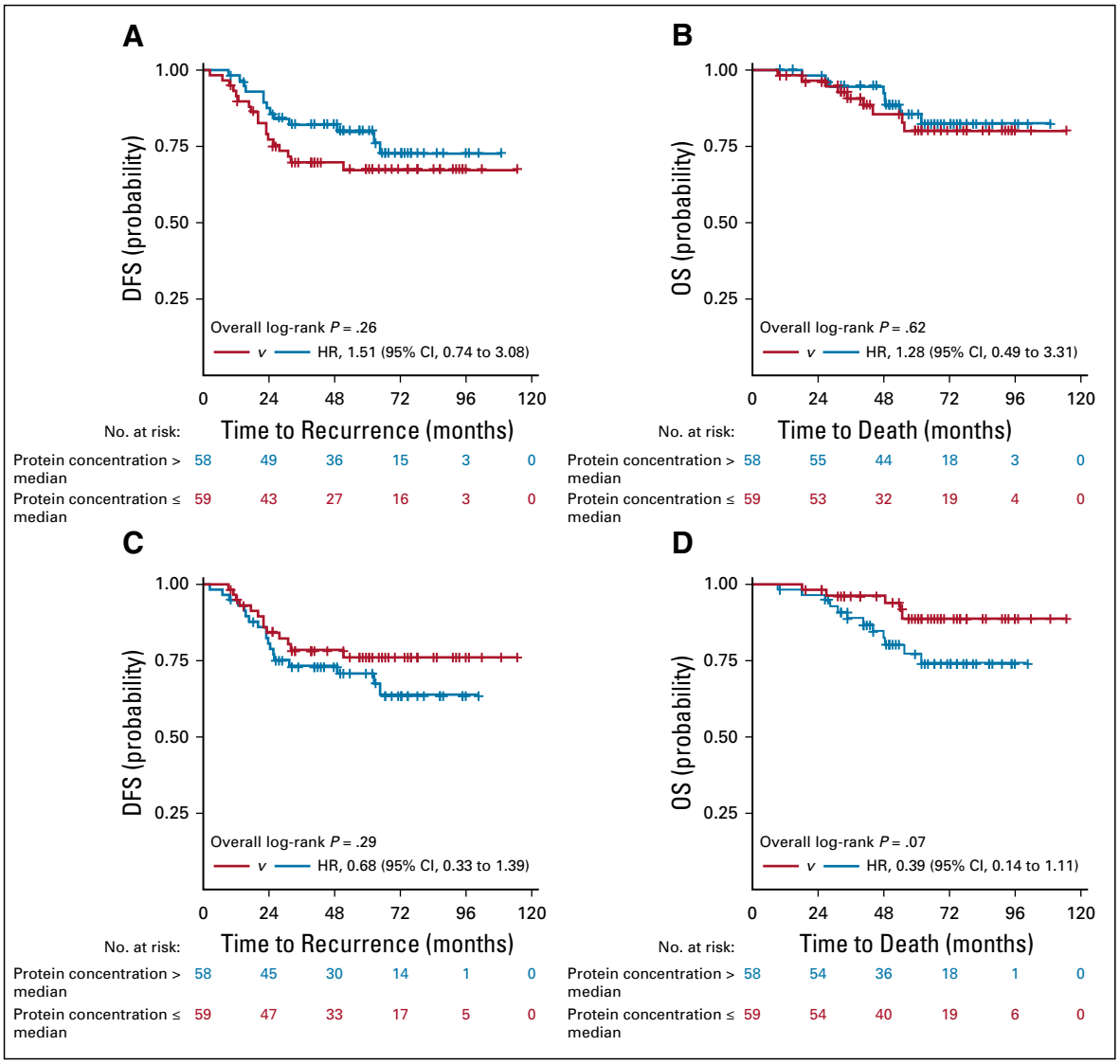


FIG A4. Assessment of the prognostic utility of procaspase-3 (PC3) and XIAP as single-protein biomarkers (n = 117). Kaplan-Meier estimates for disease-free survival (DFS) and overall survival (OS) among patients with stage III colorectal cancer of the in-house cohort by PC3 (A-B) and by XIAP (C-D). Patients were dichotomized using the population median expression as cutoff. HR, hazard ratio.

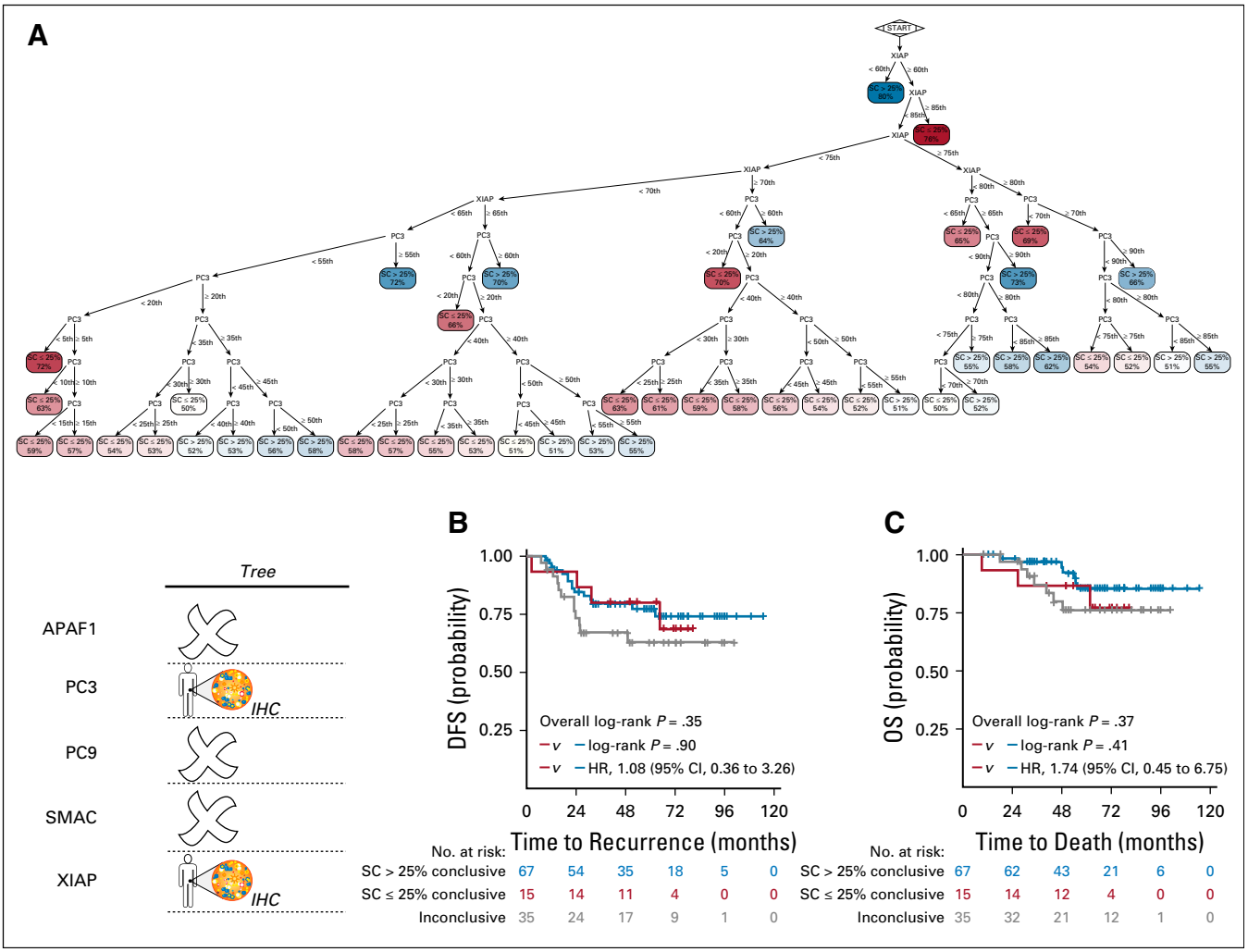


FIG A5. Application of Tree with patient-specific protein profiling performed by immunohistochemistry (IHC). (A) Best trade-off classification decision tree for APOPTO-CELL built with a penalized Gini index cost function and customized to use procaspase-3 (PC3) + XIAP protein expression (Tree-2P) assessed by immunohistochemistry (IHC). Kaplan-Meier survival estimates for (B) disease-free survival (DFS) and (C) overall survival (OS) for patients with stage III colorectal cancer grouped by the Tree-2P signature. SC, substrate cleavage, HR, hazard ratio.

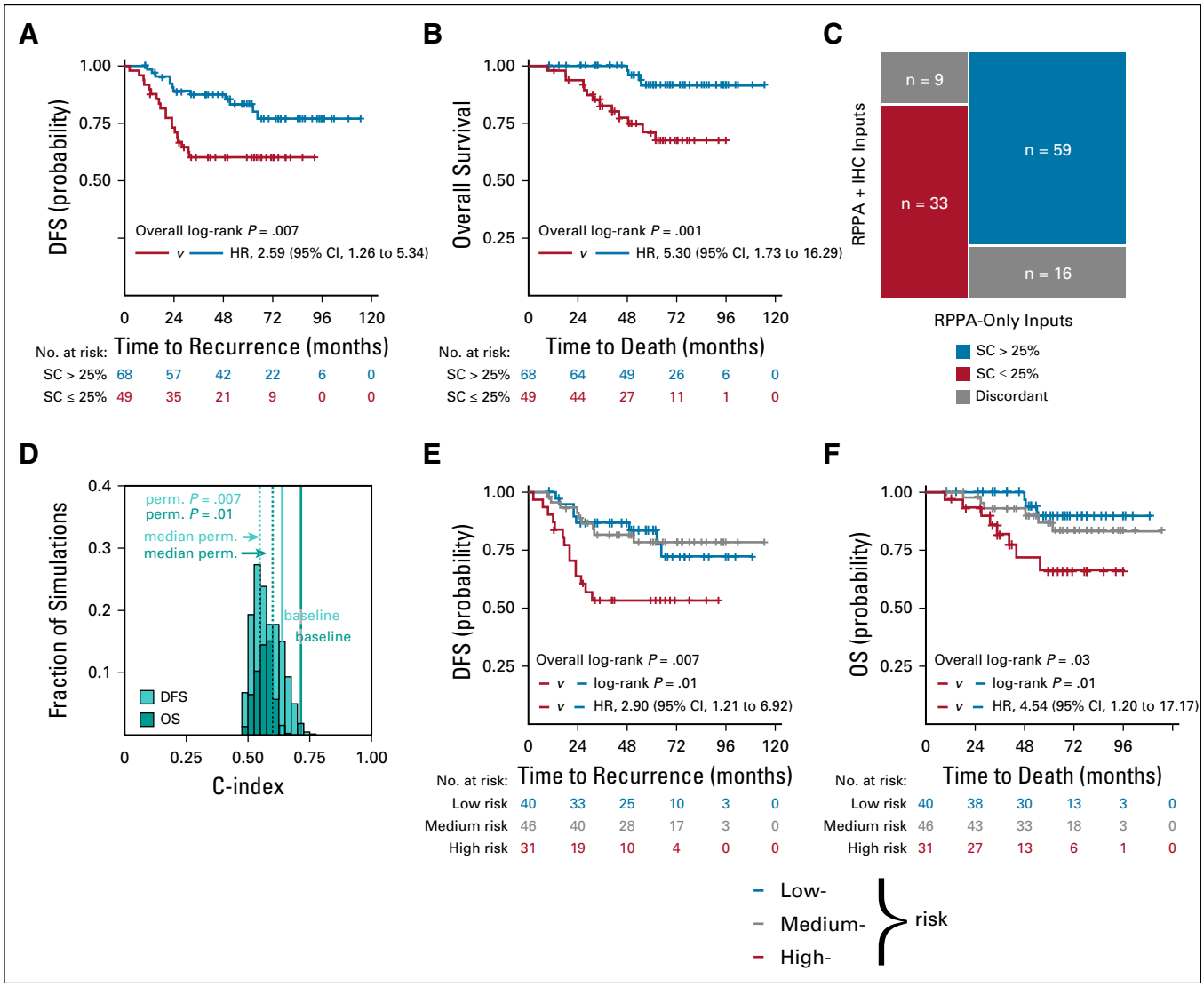


FIG A6. Procaspase-3 (PC3) and XIAP measured by immunohistochemistry (IHC) critically contribute to the APOPTO-CELL signatures in patients with stage III colorectal cancer (CRC) of the in-house cohort (n = 117). Kaplan-Meier estimates for (A) disease-free survival (DFS) and (B) overall survival (OS) that compare patients categorized as apoptosis sensitive or apoptosis resistant by the APOPTO-CELL signature. As simulation inputs, we used the population median (Salvucci M,Würstle ML, Morgan C, et al: Clin Cancer Res 23:1200-1212, 2016; Hector S, Rehm M, Schmid J, et al: Gut 61:725-733, 2012) for APAF1 for all patients and personalized concentrations for PC3 + XIAP (by IHC; Figs 6A and 6B) and for procaspase-9 (PC9) + SMAC (by reverse phase protein assay [RPPA; Salvucci M,Würstle ML, Morgan C, et al: Clin Cancer Res 23:1200-1212, 2016); (C) Mosaic plot comparing the APOPTO-CELL signature obtained using patient-specific inputs assayed by RPPA-only versus a combination of RPPA- and IHC-based measurements for PC9 + SMAC and PC3 + XIAP, respectively. (D) Distribution of the concordance index (c-index) computed by univariate Cox proportional hazards regression models for DFS and OS obtained by permutation (Perm; n = 10,000) to test the contribution of IHC-based inputs to the prognostic value of the APOPTO-CELL signature. Baseline (ie, unperturbed model) and median permuted c-indices are indicated by solid and dotted lines in light and dark blue for DFS and OS, respectively. (E) and (F) Kaplan-Meier estimates for DFS and OS that compare patients categorized as low risk, medium risk, and high risk by the APOPTO-CELL PC3 signature (Salvucci M,Würstle ML, Morgan C, et al: Clin Cancer Res 23:1200-1212, 2016) computed using the APOPTO-CELL predictions determined in (A) and (B) and PC3 expression by IHC (Fig 6A).