

ARTICLE OPEN



Molecular Diagnostics

Gene-expression profiles of pretreatment biopsies predict complete response of rectal cancer patients to preoperative chemoradiotherapy

Georg Emons^{1,2,11}, Noam Auslander^{1,3,4,11}, Peter Jo^{2,11}, Julia Kitz⁵, Azadeh Azizian², Yue Hu¹, Clemens F. Hess⁶, Claus Roedel⁷, Ulrich Sax⁸, Gabriela Salinas⁹, Philipp Stroebel⁵, Frank Kramer¹⁰, Tim Beissbarth¹⁰, Marian Grade², Michael Ghadimi², Eytan Ruppin³, Thomas Ried^{1,11} and Jochen Gaedcke^{2,11}✉

© The Author(s) 2022

PURPOSE: Preoperative (neoadjuvant) chemoradiotherapy (CRT) and total mesorectal excision is the standard treatment for rectal cancer patients (UICC stage II/III). Up to one-third of patients treated with CRT achieve a pathological complete response (pCR). These patients could be spared from surgery and its associated morbidity and mortality, and assigned to a “watch and wait” strategy. However, reliably identifying pCR based on clinical or imaging parameters remains challenging.

EXPERIMENTAL DESIGN: We generated gene-expression profiles of 175 patients with locally advanced rectal cancer enrolled in the CAO/ARO/AIO-94 and -04 trials. One hundred and sixty-one samples were used for building, training and validating a predictor of pCR using a machine learning algorithm. The performance of the classifier was validated in three independent cohorts, comprising 76 patients from (i) the CAO/ARO/AIO-94 and -04 trials ($n = 14$), (ii) a publicly available dataset ($n = 38$) and (iii) in 24 prospectively collected samples from the TransValid A trial.

RESULTS: A 21-transcript signature yielded the best classification of pCR in 161 patients (Sensitivity: 0.31; AUC: 0.81), when not allowing misclassification of non-complete-responders (False-positive rate = 0). The classifier remained robust when applied to three independent datasets ($n = 76$).

CONCLUSION: The classifier can identify >1/3 of rectal cancer patients with a pCR while never classifying patients with an incomplete response as having pCR. Importantly, we could validate this finding in three independent datasets, including a prospectively collected cohort. Therefore, this classifier could help select rectal cancer patients for a “watch and wait” strategy.

TRANSLATIONAL RELEVANCE: Forgoing surgery with its associated side effects could be an option for rectal cancer patients if the prediction of a pathological complete response (pCR) after preoperative chemoradiotherapy would be possible. Based on gene-expression profiles of 161 patients a classifier was developed and validated in three independent datasets ($n = 76$), identifying over 1/3 of patients with pCR, while never misclassifying a non-complete-responder. Therefore, the classifier can identify patients suited for “watch and wait”.

British Journal of Cancer (2022) 127:766–775; <https://doi.org/10.1038/s41416-022-01842-2>

INTRODUCTION

Colorectal cancer is the third most common cancer in the United States (145,600 cases), of which approximately one-third are located in the rectum (SEER Cancer Stat Facts 2019). Rectal cancer has been treated primarily by surgery for over 200 years. Attributable to effective screening programs, the introduction of chemoradiotherapy (CRT), and improved surgical techniques [1],

the mortality rate for rectal cancers has been significantly reduced [2, 3]. The current therapy for rectal cancer (UICC II/III) consists of neoadjuvant CRT, followed by standardised radical surgery including total mesorectal excision (TME) [4–8]. While this treatment results in improved oncological control, surgical resection may have severe complications, including anastomotic leakage, which occurs in 10–20% of patients [9–11], accompanied

¹Section of Cancer Genomics, Genetics Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. ²Department of General, Visceral and Pediatric Surgery, University Medical Center, Göttingen, Germany. ³Cancer Data Science Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. ⁴Program in Molecular and Cellular Oncogenesis, The Wistar Institute, Philadelphia, PA, USA. ⁵Department of Pathology, University Medical Center, Göttingen, Germany. ⁶Department of Radiotherapy and Radio-oncology, University Medical Center, Göttingen, Germany. ⁷Department of Radiation Oncology, University Hospital Johann Wolfgang Goethe University, Frankfurt, Germany. ⁸Department of Medical Informatics, University Medical Center, Göttingen, Germany. ⁹Transcriptome and Genome Analysis Laboratory (TAL), Department of Developmental Biochemistry, University of Göttingen, Göttingen, Germany. ¹⁰Department of Medical Statistics, University Medical Center, Göttingen, Germany. ¹¹These authors contributed equally: Georg Emons, Noam Auslander, Peter Jo, Thomas Ried, Jochen Gaedcke. ✉email: jochen.gaedcke@med.uni-goettingen.de

Received: 4 November 2021 Revised: 19 March 2022 Accepted: 4 May 2022

Published online: 21 May 2022

by increased morbidity and mortality. Furthermore, many patients are suffering from the side effects of surgery: up to 27% of patients need a permanent stoma [12], 39% experience faecal incontinence [13, 14], 30% urinary [15] and 67% sexual dysfunctions [16, 17], all resulting in a severe reduction in quality of life.

The response to preoperative CRT is very heterogeneous, ranging from a pathological complete response (pCR) with no viable tumour cells left (10–25% of patients) to virtually no tumour regression at all or even tumour progression during therapy [5, 18, 19]. Patients with a pronounced tumour regression have a favourable oncological outcome, with a 5-year overall survival probability of more than 90% [20]. Importantly, residual lymph node metastases, even after complete regression of the primary tumour, indicate a poor prognosis [21, 22]. In 2004, Habr-Gamma et al. presented data of a selected group of rectal cancer patients, assessed to be clinically complete responders (cCR) and consecutively spared from surgery (“watch and wait”). These patients had a similar oncological outcome compared to standard treatment [23]. Forgoing surgery spares patients with a cCR from operation associated morbidity and mortality as later shown by Maas et al. [18, 24]. Conversely, as retrospectively collected data revealed an adverse outcome for “watch and wait” patients compared to the initial report of Habr-Gama et al. [25], this concept was rejected by a majority of surgeons. It was later shown that the data were too heterogeneous, possibly due to an incorrect assessment of cCR. This emphasises the importance of a stringent assessment of cCR as the key to the acceptance of a “watch and wait” strategy. Currently the selection of patients is imaging-based and incorporates clinical and endoscopic parameters. Nevertheless, the rate of local recurrence within the first two years is ~25%, as recently published by the international “watch and wait” consortium [26]. Molecular markers, as outlined by Yoo et al. [27], could augment the development of pretreatment response-prediction models improving the selection of patients for a “watch and wait” strategy.

This dilemma precipitated a plethora of studies trying to predict response to CRT by gene-expression profiling of pretreatment biopsies. However, efforts to establish universal molecular response predictors were thus far not successful [28, 29] because the validation of proposed predictors in independent datasets mostly failed. Consequently, none of these classifiers ever translated into routine clinical practice.

Here we present data from gene-expression profiles of pretreatment rectal cancer biopsies from 161 patients that were treated within the CAO/ARO/AIO-94 and CAO/ARO/AIO-04 trials using a machine learning approach to build a classifier predicting pCR to preoperative CRT. The classifier was validated in three independent cohorts ($n = 76$).

MATERIALS AND METHODS

Patients, clinical staging and treatment

Pretreatment biopsies were collected during the initial staging endoscopy between 2001 and 2012 at six surgical departments, all members of the “German Rectal Cancer Study Group (GRCSG)”. A total of 175 patients from a consecutive series of 236 patients with locally advanced rectal cancer (cUICC II/III) fulfilled inclusion requirements, regarding RNA quality, tumour cell content, complete clinical annotation and array quality (CONSORT Diagram, Fig. S1A). The study was approved by the local ethics committees, and all participants gave written informed consent. Detailed patient characteristics can be found in Table S1A. Precise staging and treatment procedures are described in Fig. S2 and in Supplementary Materials and Methods. Macroscopic and microscopic assessment of tumour stage was performed as previously described [4]. Residual tumour after CRT was assessed as a ratio of remaining viable tumour cells and the area covered by the tumour before treatment as described previously by Roedel et al. [5]. However, the tumour regression grading was reported in percent (TRG in %) to achieve a more accurate spectrum of response. Additionally, TRG was evaluated using the classification of Dworak et al.

[30]. Pathological complete response (pCR) was defined as the absence of viable tumour cells in the primary tumour and lymph nodes (TRG = 100%, TRG 4, ypT0 pN0). Four to six weeks after surgery patients received an adjuvant therapy of either 5-FU alone or in combination with folinic acid and Oxaliplatin [4, 8].

For 14/175 patients treated at three German surgical centres we were unable to obtain TRG in %, so these cases were only used as the first independent validation dataset (Table S1B).

For the second independent validation dataset, clinical and gene-expression data of additional 38 patients with local advanced rectal cancer treated at the Department of Surgery, University of Padua, Italy, were acquired from the GEO database [31] (Table S1C).

For a third independent validation, a clinical trial was set up, named TransValid A (German Clinical Trial Register No 00003659). The goal of this open, non-randomised, prospective exploratory multicenter validation study was to collect biomaterial from patients with stage II/III rectal cancer treated according to established multimodal therapy. Inclusion criteria were adjusted to the previous trials of the GRCSG [7, 8]. Overall, 25 patients from 6 institutions were included. Applying the same quality criteria as for the first two sample sets 24 patients could be included in the final analysis (Fig. S1B and Table S1D).

Pretreatment tumour biopsies, RNA isolation and gene-expression microarray analysis

Biopsies were collected during the initial staging rectoscopy and immediately stored in RNAlater as previously described [32, 33]. The tumour content of the biopsies was analysed by a pathologist, and only samples containing more than 50% of tumour tissue were used in this analysis. Gene-expression microarray (Human 4 × 44 K v2 gene-expression array, Agilent Technologies (G4845A)) analysis was performed per manufacturer’s instructions as previously described [34–36]. Gene-expression data were deposited to Gene Expression Omnibus (GSE87211). Details are described in Supplementary Methods.

Statistical methods

Metric, ordinal and categorical parameters were compared between complete and partial responders by the two-sample t -, Mann–Whitney U - and Fisher’s exact test, respectively. Survival times were compared by Kaplan–Meier curves. Independent cohorts of Microarray gene-expression data were log-transformed and normalised using the quantile method. The normalisation algorithm was initially applied to the first dataset, and the quantiles of the validation datasets were then aligned to the same distribution. MATLAB R2018b Support Vector Machine functions were used to train and test classifiers. Performance was evaluated via area under the curve (AUC) resulting from receiver operating characteristic (ROC) curves. All programming work in this study was done in MATLAB 2018b (The MathWorks Inc., Natick, MA, 2000).

Evaluation of classifiers

The classification performance was evaluated via two measures. The first is the area under the curve (AUC) of the receiver operating characteristic curves (ROC) of classification. The second is the maximal sensitivity (True positive rate) reached when the false-positive rate is zero (i.e. the maximal rate of true pCR captured when no patients were misclassified as pCR). For simplicity, this measure is termed sensitivity throughout the manuscript.

Feature selection, cross-validation and support vector machine (SVM) classifier training

For feature selection, out of the 161 cases enrolled in this study, 32 positive (pCR, TRG = 100%) and 32 negative (poor response, TRG < 45%) cases were used. The samples were randomly divided 500 times into a training set (3/4 of the samples) and a test set (the remaining 1/4). Each time we performed a differential gene-expression analysis. A hill-climbing [37] procedure was applied to each of the resulting 500 gene sets to generate Support Vector Machines (SVMs). To evaluate the predictive power, each of the 500 resulting SVMs was applied to their corresponding test sets. Finally, genes that were selected a significant number of times (P -value < 0.05, Supp. Methods) as having optimal sensitivity on the test sets, i.e. a set achieving the highest percentage of correctly identified positive samples, while not allowing any false positives in the test set, were used as features for the classifier. With the resulting 21 transcripts, we performed a four-fold cross-validation. The full set of 64 samples and the classifier signature were used to train a final SVM which was applied to all 161 cases, with a tumour

Table 1. Comparison of the performance (in terms of AUC) of published gene signatures (GS) identifying pCR when applied to our primary patient cohort ($n = 161$) and the one from Millino et al. [31] ($n = 38$).

Test set	our signature	GS1	GS2	GS3	GS4	GS5
161 samples	0.81	0.51	0.55	0.61	0.27	0.59
38	0.76	0.48	0.68	0.63	0.74	0.4

GS1 was published by Lopes-Ramos et al. [29]; GS2 by Ghadimi et al. [39]; GS3 by Empuko et al. [40]; GS4 by Watanabe et al. [38] and GS5 by Kim et al. [41].

regression grade ranging from 10 to 100%. Subsequently the classifier was applied to independent datasets. A detailed description, as well as a graphic visualisation of the feature selection process, is shown in Fig. S4 and Supplementary Methods.

Evaluating previously published signatures for predicting response to CRT

Gene-expression signatures from five studies, published by us and others [29, 38–41], that aimed to predict response to preoperative CRT, were analysed. The gene lists were adapted to unify annotation. The predictive performance of each signature was independently evaluated in our dataset of 161 patients and in a second dataset [31]. Therefore, a SVM classifier with a linear kernel was used via a leave-one-out cross-validation, as performed in these studies. The resulting performance (in terms of AUC) for each of these classifiers was established (Table 1). Additionally, a heatmap and PCA were plotted for each signature (Fig. S5).

Validation of our SVM classifier in three independent rectal cancer datasets

Independent dataset I: 14 patients, all treated within the CAO/ARO/AIO-04 or -04 study. These patients were not included in the initial analysis because we were unable to obtain the percentage of TRG. However, as TRG according to the classification of Dworak [30] was available, these patients were used as an independent validation set. Samples with TRG = 4 were considered pCR ($n = 4$).

Independent dataset II: Gene-expression and clinical data of 38 rectal cancer patients published by Millino et al. [31]. Samples with TRG = 1 per the classification of Mandard [42] were considered pCR ($n = 8$), patients with TRG > 1 were considered incomplete responders ($n = 30$). Additional patients published as a validation dataset by Millino et al. [31] were not analysed, as there was no annotation of pCR.

Independent dataset III (TransValid A trial): Additional pretreatment biopsies and clinical data were collected after 2012 and analysed prospectively on a day-to-day basis ($n = 25$). Four patients had a pCR (TRG = 100%), 20 were incomplete responders (TRG < 100%).

All gene-expression data were quantile-normalised using the mean expression values of the initial gene-expression data ($n = 161$) as reference distribution and the predictor was applied to the normalised gene-expression patterns.

Building a second, not normalisation-dependent, score-classifier based on the 21 genes

SVM classifiers require a data distribution similar to their training set and are hence normalisation dependent. Normalisation introduces additional bias and requires trained personnel. To demonstrate the usefulness of the 21 genes, a second classifier, that does not require any transformation of the data, was trained. The second classifier is a score, based on the expression levels of the 21 genes selected for the SVM classifier. First, 64 cases, comprising 32 positive (pCR, TRG = 100%) and 32 negative (poor response, TRG < 45%) cases, were randomly divided into a training and test set with two-fold cross-validation. In 100 iterations, a simulated annealing procedure was used to train the score. The procedure starts with a random assignment of genes to the numerator and/or denominator of the score. In each step, all possible modifications of the score (additions or removals to or from the numerator or denominator) are evaluated. The score is adjusted in a way that maximises the sensitivity of the training (until reaching a maxima). The 100 trained scores are then evaluated in the test set. The score-classifier with maximal test sensitivity is selected:

$$\frac{\sum_{i=IAA1598, ASPM, TMPO, HOMER1, CXCL10, CENPL, BRCA1, FZD10, C19orf51, C20orf26, CASC5, CCN81, FANCM, BLM} \text{Exp}_i}{\sum_{i=CGREF1, TNPO3, XPO1, TSNAX, CENPL, FANCM, FZD10, C19orf51, C20orf26, CCN81, C5orf1, BLM} \text{Exp}_i}$$

Analog to the SVM-classifier, the final score-classifier is applied to all 161 cases (tumour regression grade ranging from 10 to 100%) and to three independent datasets comprising additional 78 patients. Additionally the score classifier was applied to the expression data of 28 patients from Canto et al. [43] (GSE123390), the resulting ROC curve can be found in the Supplementary Fig. SF7.

Ingenuity pathway analysis (IPA)

The 21 transcripts comprising the classifier were uploaded into the IPA® software (Ingenuity, QIAGEN, Hilden, Germany), and a pathway, as well as a network analysis, was performed. Only experimentally validated edges were considered.

Testing the predictive value of the classifier in rectal cancer patients

As a response to preoperative CRT is associated with a favourable outcome, we speculated that the classifier should also identify patients with a good clinical prognosis. Therefore, patient data from the “The Cancer Genome Atlas” (TCGA) were analysed. We used the classification scores obtained from applying the trained classifier to the gene-expression data of the 21-transcript used to train the classifiers (applied to the mRNA values, as is, using *RNA_Seq_v2_expression_median*). The classifier was applied to the 161 patients in our rectal cancer dataset, and the resulting classification scores were correlated with the TRG in percent of the patient are plotted.

Then, cancer survival and gene-expression data from 468 colon and rectal cancer patients were retrieved from the READ-TCGA [44], and the classifier was applied to this data as well. The top and bottom 50% of the resulting classification scores were used to plot the Kaplan–Meier survival curve assessing overall survival (OS), and disease-free survival (DFS).

RESULTS

Patient characteristics

Patient characteristics for the test and training as well as the independent validation sets are summarised in Table S1A–C and Fig. S3. For neither dataset was there a significant correlation of response with age, sex, tumour stage and addition of oxaliplatin to CRT. There was also no significant ($p = 0.346$) distribution difference between responders and nonresponders according to their Consensus molecular subgroups (CMS) [45], see Table S1F.

Identification of a gene-expression signature that predicts pathological complete response (pCR)

As the purpose of this study is to reliably identify pathological complete response (pCR) without misclassifying incomplete responders (false-positive rate = 0), we performed feature selection aiming to maximise the performance of the classifier. Thirty-two positive (pCR, TRG = 100%) and 32 negative (poor response, TRG < 45%) cases were chosen for feature selection and generation of the classifier. These 64 cases were randomly divided into training and test sets. Differentially expressed genes between pCR patients and incomplete responders were identified in the training set. A hill-climbing feature selection [37] was applied, gradually adding genes/features, that improve the sensitivity of the resulting classifier when applied to the test set. The procedure of splitting 64 samples into two groups, performing differential gene-expression analysis and generating a classifier with maximal sensitivity when applied to the test set

was repeated 500 times. A signature of 21 transcripts, that were consistently selected (P -value < 0.05) to have maximal sensitivity during the 500 repetitions, was chosen to generate a final classifier (Table 2). A four-fold cross-validation procedure on 32 positive and 32 negative cases resulted in a sensitivity of 0.4 and an AUC of 0.75 (Fig. 1a). Based on these 64 cases, an SVM approach was applied to all 161 cases. Encouragingly, applying our classifier to the entire dataset, it achieved a sensitivity of 0.31 and an AUC of 0.81 (Fig. 1b), indicating that even when considering the full range of TRG (10–100%), our SVM-classifier can correctly identify more than a third of complete responders without error. The tradeoff between precision and recall is visualised in Fig. 1c.

Previously published signatures are poor predictors of response in independent datasets

A major issue impeding the emergence of robust prognostic or diagnostic tools from molecular signatures is the poor reproducibility in independent datasets [46]. We therefore compared our SVM-classifier with classifiers published earlier by us and others [29, 38–41] on our dataset of 161 patients as well as on the 38 patients published by Millino et al. [31]. As shown in Table 1 and visualised in Fig. S5, none of the published signatures was useful for response prediction.

The 21-transcript classifier robustly identifies over 1/3 of patients with pCR in independent datasets

To test if our SVM-classifier can identify patients with pCR in independent datasets we applied it without any further training, to three additional patient cohorts. The first one comprised 14 patients treated within the CAO/ARO/AIO-94 or 04 studies. The classifier identified two out of four patients with pCR (Sensitivity = 0.5, AUC = 0.7), while none of the patients was wrongly classified as pCR (Fig. 2a, b). In a dataset of 38 patients (8 pCR), recently published by Millino et al. [31], the classifier correctly identified four patients to have pCR (Sensitivity = 0.5, AUC = 0.76), again not misclassifying any patient incorrectly as pCR (Fig. 2c, d). Finally, the classifier was applied to 25 prospectively collected patients and could correctly identify two out of five patients with pCR (Sensitivity = 0.4 AUC = 0.81), again not misclassifying any patient as pCR when they were not (Fig. 2e, f). A graphic illustration of the support vector values used for the SVM-classifier can be found in Fig. S6 and the full support vector is provided in Table S2. The classification was independent of demographic parameters (Fig. S3).

To further demonstrate the clinical relevance of the 21 genes, we developed a second, simple, score-based classifier using the 21 genes (Methods, Fig. 2g). This simple classifier does not require normalisation yet allowed good identification of a considerable percentage of the pCR patients, while not misclassifying any patient incorrectly as pCR (Fig. 2h). It identified more than 30% of all pCR samples in the independent datasets, without misclassifying any patient incorrectly as pCR (Fig. 2i, j).

Expression of the 21-transcript score correlates with treatment response and survival of colorectal cancer patients

pCR is associated with improved survival and a more favourable oncological outcome [18, 47]. We therefore hypothesised that our classifier would also be useful for disease prognostication. To this end, we first used the classifier score based on the expression of the 21 transcripts on the 161 rectal cancer patients. As expected, we observed a significant (Spearman $\rho = 0.36$, $p = 5e-06$) correlation with tumour regression (Fig. 3a). We then applied the classifier to 468 colon and rectal cancer patients from the TCGA database to obtain classification scores and found that high classification scores were significantly ($p = 0.0007$) associated with longer overall survival (Fig. 3b) and a significantly ($p = 0.0002$) longer disease-free survival (Fig. 3c).

Ingenuity pathway analysis (IPA) identifies enrichment of radiation response pathway genes

To gain insight into the biological mechanisms of the 21 transcripts comprising the classifier an IPA®-pathway analysis was performed. There was significant enrichment for genes involved in DNA-damage control pathways; the “Role of *BRCA1* in DNA Damage Response” ($p = 0.002239$) pathway, “DNA G2/M DNA Damage Checkpoint regulation” ($p = 0.012589$) and “ATM signaling” ($p = 0.002239$) were, after correction for multiple testing, significantly over-represented (Table S3).

DISCUSSION

Sparing rectal cancer patients from severe side effects and complications associated with radical surgery or even amputation (removal of the anus, the rectum, partly the sigmoid colon and permanent colostomy) is an attractive option for patients with a cCR after preoperative CRT. Unfortunately, the number of local relapses is still high, indicating shortcomings in the initial assessment of cCR which requires patients to undergo a strictly standardised follow-up, including frequent clinical, endoscopic and image-based surveillance, in highly specialised centres [26]. Even though the majority of patients with tumour regrowth can be salvaged with definite surgery after “watch and wait” [48], there is an urgent clinical need for additional parameters to correctly identify complete responders.

Recent clinical trials in other tumour entities demonstrated that gene-expression profiling can be useful. Examples are the MammaPrint [49, 50] signature for breast cancer, the OncoTypeDX® [51] and ColoPrint® for colorectal cancer [52, 53].

Exploring the utility of gene-expression profiling for the prediction of response to preoperative CRT in rectal cancers has been actively pursued. However, despite several studies reporting promising results, none has translated into a clinical test to guide individualised therapy [54, 55]. Two studies tried to identify markers that predict pCR in rectal cancer. The first, published by Kim et al. [41], using gene-expression profiling of pretreatment ($n = 31/11$ pCR) biopsies to build a classifier. When applied to an independent dataset ($n = 15/4$ pCR), the signature identified all four pCR patients correctly, but also identified two incomplete responders as pCR. The second study was performed by Brettingham-Moore, analyzing 51 cases (7 pCR). However, they could not identify reproducible predictive signatures [28].

Since previously published gene-expression signatures for response prediction lacked reproducibility [29], we re-tested several of these classifiers and could also not observe a reliable prediction of response in our patient cohort (Table 1, Fig. S5). A possible explanation for this is that feature selection was performed in parallel to the training procedure, which increases the risk of overfitting [56, 57]. Our model was initially designed to discriminate between the extremes in response to CRT (pCR vs poor response), but it remained robust when applied to the full range (0–100%) of response (Fig. 1). It identifies over 1/3 of the patients with pCR in all datasets examined and, most importantly, when considering a possible clinical application, never classifies any patient with an incomplete response as pCR. Unlike previous studies we aimed to predict pCR with the highest accuracy, at the cost of sensitivity [58]. Allowing mistakes when training a classifier would increase the rate of identifying patients with pCR; however, the clinical benefit of such a predictor would be limited. Finally, unlike previous studies, we show that using the same 21-transcript signature in independent, comparable datasets previously published by others (Fig. 2c, d), robustly classifies pCR, which is essential for clinical applicability. Additionally, we show that our classifier predicts response in routinely collected samples: for the TransValid trial we did not collect a set of pretreatment biopsies and analysed them in one batch. For every four patients, RNA was isolated and gene-expression levels analysed. This method introduced additional bias,

Table 2. List of 21 transcripts in the classifier.

Gene symbol	Probe	Systematic name	Chromosome	Probe sequence 3'-5'
CGREF1	A_33_P3281850	CR623121	2	GGTGGTCACTTCTAACTCTGTCATTACCAACAGCAGCCTAGCGTTGCCCCATTACAACA
FZD10	A_23_P203972	NM_007197	12	CTTCACAGTCCAGGAAAGAGTGGTTCCTGCGTGTGTATATTTGTAATATATGATATTTT
ASPM	A_23_P52017	NM_018136	1	ATCACAAATCCCTGCAAGCTATTCAAATGGTGGATGATACGCTTGGCATTCTCTATTAG
SH1N1	A_23_P202587	NM_018330	10	GTAATAATTGCAGTAGTGTATTGATTGATTTTTCACGCTGTGGTAAGCATAGGCTTG
CCNB1	A_23_P122197	NM_031966	5	GGACACCAACTCTACAACATTACCTGTCATATACTGAAAGAATCTCTCTCCAGTTATGC
CXCL10	A_24_P303091	NM_001565	4	GTCAGCCATAATTGTTCTTAGTTTGCAGTTACACTAAAGGTGACCAATGATGGTCACC
CFAP61	A_32_P4262	NM_015585	12	TTTACAATCCAGGCATGTTCTGTGCGAATAATCCATTCAGCAATAAATGAGATC
XPO1	A_23_P40078	NM_003400	10	GGGTATTTGTCGACCAAAATGATGCCAAATTTGTAAATTAATAATGTCACCTAGTGGCCCTT
BRCA1	A_23_P207400	NM_007300	17	AAAGTGTTCATATAACCCATTATCCAGGACTGTTATAGCTGTTGGAAGGACTAGGTCT
TMPO	A_23_P325040	NM_003276	12	AAAGTAATGCCTGTGTAACACTACTGTCTTTCTAAAGATTTGCGTAGATAGGAAGCC
CASC5	A_23_P100127	NM_170589	15	CGGTCTCTAGCAAAGATTCAGGCATTGGATCTGTTGCAGGTAAACTGAACCTAAAGTCCTT
FANCM	A_32_P106732	NM_020937	14	AATCAAGCTGCTCAAGATGGGTTTTCAAAGACCTCTACAATAATAAATGACACTTCAAT
BLM	A_23_P88630	NM_000057	15	TATGCAITCTCATAACAACCCGAATCTCAATGTACATAGACCCCTCTTCTTGTGTTGTCAGC
HOMER1	A_33_P3372257	NM_004272	5	GAGCACCAAGTTTTAATTTAAATAGGAGATTAAACACTAGGGATCAGGGAGTTTAGTATG
TMPO3	A_33_P3370132	NM_012470	7	TAAAAGGTTGCCCCAAGGAAACAAACCGTGGGAGCCGTCACAGTGACACACAAACAACCTTA
CENPL	A_23_P126120	NM_033319	1	TTTCAAAATTCATTTATCAGCCACAAGATTAGTTCGTGTTTCAACATCTGTAGCTTCAGC
CSPP1	A_23_P71537	NM_001077204	8	TGGTTGACCCCTGATGACATCATGAAACACATAGGGGATGACGGATCAAACCTCTGTAGCAA
STAR3	A_33_P3246804	NM_001165937	17	CCTGTCACCCGTGTGAAGATGAAGGGCTCTTCATCTGCTGCCCTCGCTCGGTTTTTTT
DNAAF3	A_33_P3286349	NM_178837	19	ATTGTTCTCAGAAATCCAAATTCACCTCTGAGGCTTAAGCCAGCCAGGATCTGA
MCM5	A_23_P132277	NM_006739	22	CAGCATCATCAAGGACTTCACCAAGCAGAAATACCCGGAGCAGCCATCCACAAGGTGCT
TSNAX	A_24_P148151	NM_005999	1	AAAGTTGAGTTATATACTTGTACATACAATGGAATGCTTTTAGTAGTATTATTAGCA

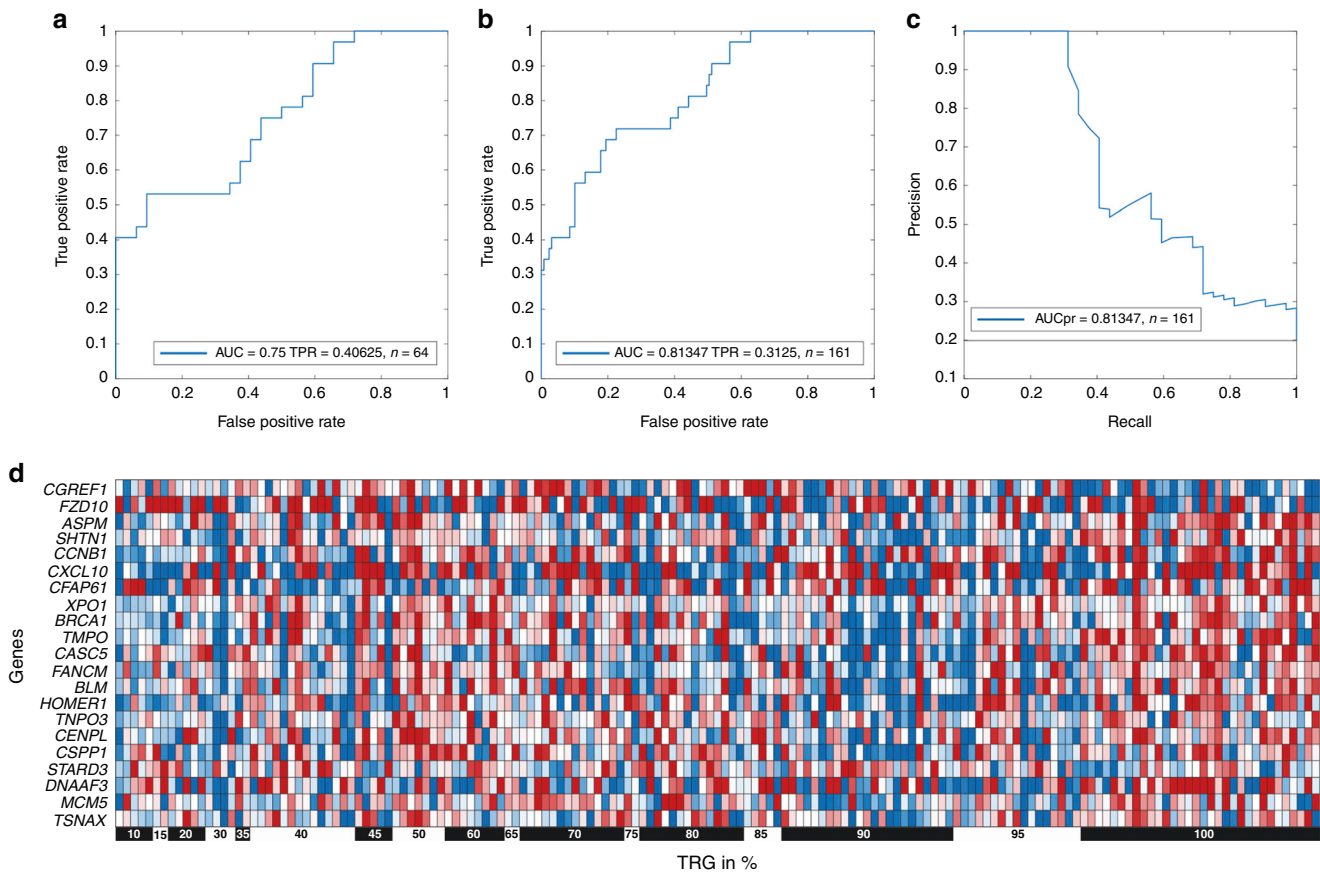


Fig. 1 A 21-transcript signature robustly identifies complete responders. **a** Receiver Operating Characteristic (ROC) curve resulting from a four-fold cross-validation using 32 positive and 32 negative cases. The area under the curve (AUC) and the True positive rate (TPR) are reported. **b** ROC plot of the classifier when tested on 161 cases, with a full range of tumour regression grading in percent (TRG 10–100%) and **c** precision recall curve. **d** Heatmap visualisation of the expression of the 21-transcript signature in 161 pretreatment biopsies of rectal cancer patients.

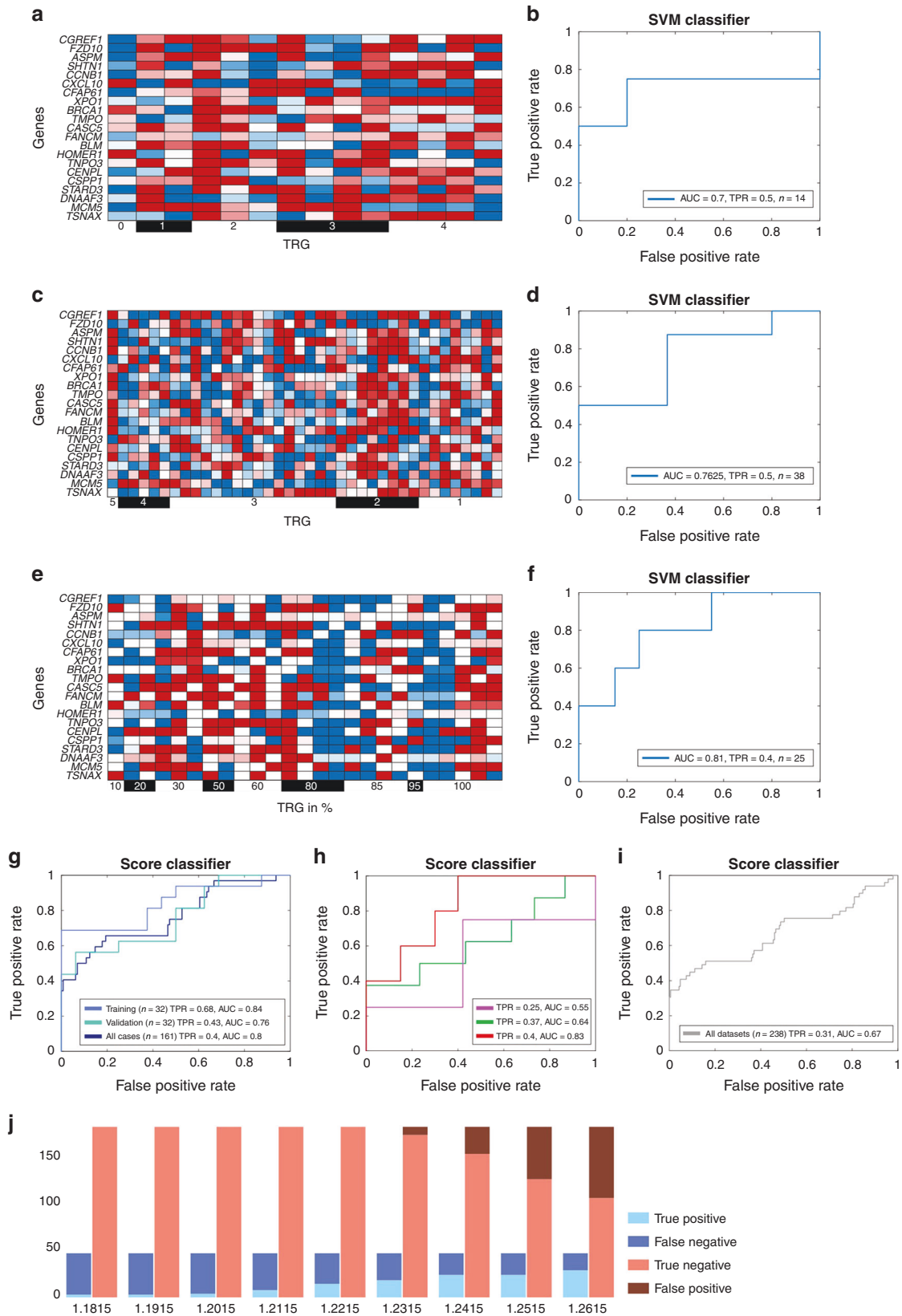
but resembles the infrastructure of routine day-to-day clinical testing (Fig. 2e, f). To overcome these issues, the test microarray datasets were quantile-normalised before applying the SVM classifier, which introduces a limitation to the approach which is, necessarily, dependent upon normalisation.

To demonstrate that the 21 gene signature is also useful without any data transformation, we trained an additional classifier. The second classification score was developed using the same 21 genes, it does not require normalisation and can be applied to any data as is (Fig. 2g–j). Both classifiers robustly predicted pCR in all test datasets. Even when applied to additional independent data, published by Canto et al. [43], missing values for some of the 21 genes, the score classifier stayed robust (Supplementary Fig. SF7) underscoring a potential clinical utility.

Investigating the cases that were not identified as pCR in our data by the two classifiers, we did not identify genes that were up or downregulated in these samples. We additionally did not identify differences in stage or age between the correctly classified and misclassified samples, but we did find that the misclassified samples included more female patients (46%, compared to 32% in the total set of patients). It is possible that the performance for female patients is impaired because the classifier was trained using predominantly male patient samples (as a result of the increased frequency of rectal cancer in male compare to female patients). Future studies are therefore warranted to thoroughly investigate the need for sex-balanced patient cohorts when training pCR classifiers.

To further strengthen the potency of the classifier we used it as a predictor of prognosis. Knowing that pCR is associated with a good prognosis determined by disease-free survival and overall survival [18, 59], we speculated that patients with a high score based on the classifier genes should also have a favourable outcome. As the follow-up data for our 161 patients is relatively short, and to test our classifier in another independent patient cohort, we applied it to 468 colon and rectal cancer patients from the TCGA database. Tumour regression for these patients is not reported; we therefore used overall survival as a surrogate for pCR. We found that the 50% of patients with the highest “pCR score”, as determined by the expression of the 21 genes, had an excellent outcome (Fig. 3). This shows the clinical utility of the classifier in general, even though it was trained for the specific task of pCR prediction.

From a mechanistic point of view, the classifier contains, intuitively, genes associated with DNA-Damage control and repair, underscoring the relevance of the signature. Most prominent are *BRCA1*, *BLM* and *FANCM*, all associated with the same DNA-damage response signaling pathway. *CCNB1* (Cyclin B1) and other genes are associated with cell cycle control, ATM signaling and DNA-damage checkpoint regulation. Furthermore, the classifier includes *FZD10*, a gene coding for one of the Frizzled receptors, activating the canonical Wnt/beta catenin pathway. Recently, we and others provided experimental proof that the Wnt-pathway mediates treatment resistance in rectal cancer [60–62]. Other genes are involved in response to cellular stress and other pathways associated with cellular survival. We assume that genes



in the classifier might be suitable targets for overcoming treatment resistance.

Currently, watch and wait strategies are validated in clinical trials such as the ACO/ARO/AIO-18.1 randomised trial

(NCT04246684). According to the International Watch & Wait Database (IWWD), local regrowth is expected in up to 25% of patients with clinical complete response (cCR) [63], indicating that a non-surgical approach for all cCR patients diagnosed on

Fig. 2 The SVM classifier and the score classifier robustly predict pathological complete response in three independent patient cohorts, while never misclassifying a partial responder as complete. **a** Heatmap visualisation and **b** ROC curve of the SVM-classifier performance 14 patients from the CAO/ARO/AIO-94 or 04 trial, not examined in the first analysis. Two out of 4 patients with pCR were identified. **c** Heatmap visualisation and **d** ROC curve of SVM-classifier performance for 38 patients published by Millino et al. [31]. Four out of 8 patients with pCR were identified. **e** Visualisation and **f** ROC curve of SVM-classifier performance when applied to 25 prospectively collected specimens from the TransValidA trial. Two out of 5 patients with pCR were identified. **g** The score-classifier performance for the training, validation and all 161 cases, with a full range of tumour regression grading in percent. **h** ROC curve of the score-classifier performance for 14 patients from the CAO/ARO/AIO-94 or 04 trial (purple, one out of 4 patients with pCR were identified), for 38 patients published by Millino et al. [31] (three out of 8 patients with pCR were identified) and for the 25 prospectively collected specimens from the TransValidA trial (red, two out of 5 patients with pCR were identified). **i** ROC curve of the score-classifier performance for the integration of all 238 samples in this study. **j** Bar plots showing the prediction accuracy and error types for different score thresholds on the aggregate compendium of the 238 patients included in all datasets studied (161 + 14 patients from the CAO/ARO/AIO-94 or -04 trial, 38 cases published by Millino et al. [31], 25 patients included in the TransValidA trial).

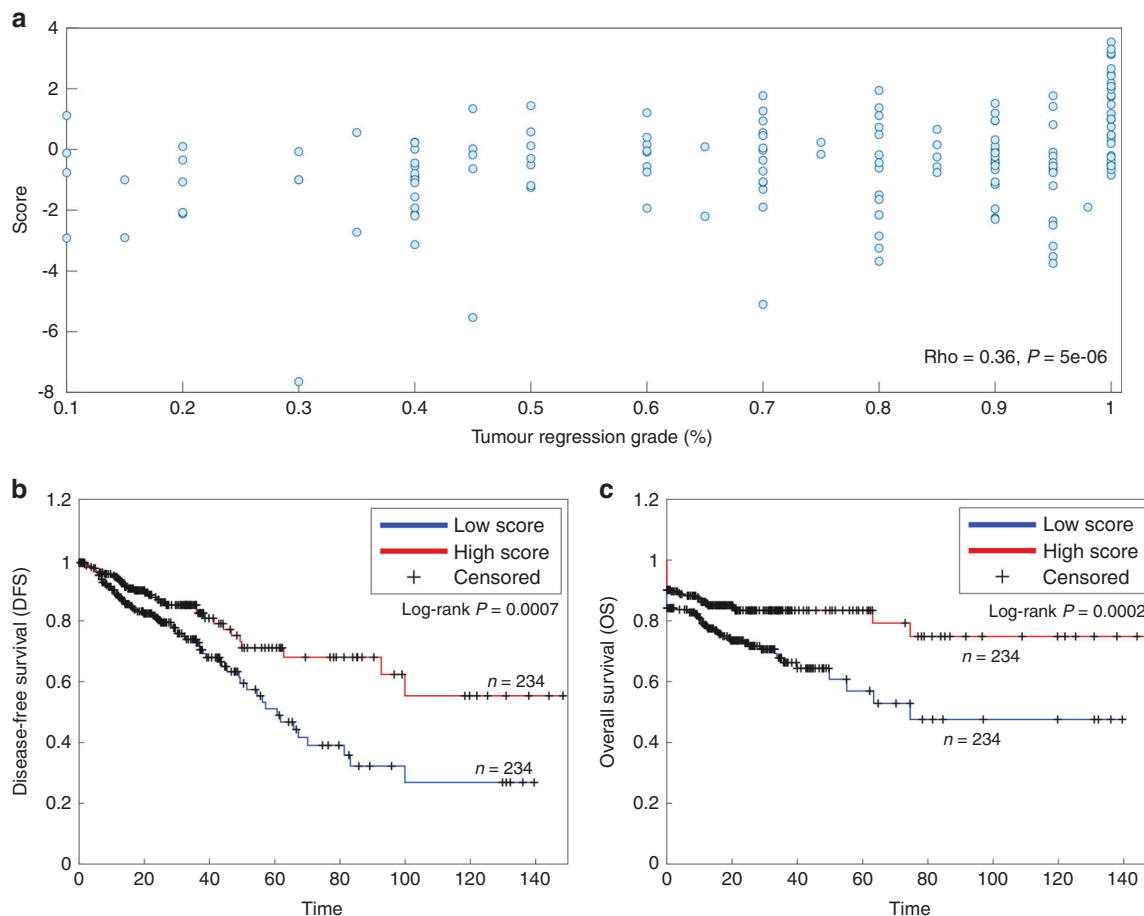


Fig. 3 The classifier score based on the expression of the 21 transcripts correlates with TRG and is associated with outcome in an independent dataset. **a** Correlation of the classifier score and TRG in 161 rectal cancer patients. **b** Kaplan–Meier curves of colon and rectal cancer patients from the TCGA ($n = 468$) comparing overall survival (OS) of 50% of patients with the highest vs. 50% with the lowest classifier score. **c** Kaplan–Meier curves of colon and rectal cancer patients from the TCGA ($n = 468$) comparing disease-free survival (DFS) of 50% of patients with the highest vs. 50% with the lowest classifier score.

the basis of radiological and clinical data is unsupportable. Therefore, the identified signature and other molecular markers associated with treatment response, e.g. *TP53* mutations [64], or analysis of the tumour secretome [43], should be added to the decision making and validated in clinical trials [22]. Our data and the complex situation of neoadjuvant therapy do not support the reverse conclusion that non-pCR patients should be spared from preoperative CRT. We are currently exploring the conversion of the classifier to a NanoString® platform, which has been shown to robustly quantify gene-expression levels in routinely collected, formalin-fixed paraffin-embedded tissue.

In summary, we describe a 21-transcript signature that robustly predicts a pathological complete response in rectal cancer patients based on gene-expression profiles of pretreatment biopsies.

DATA AVAILABILITY

Gene-expression data were deposited to Gene Expression Omnibus (GSE87211).

REFERENCES

1. Heald RJ, Husband EM, Ryall RD. The mesorectum in rectal cancer surgery—the clue to pelvic recurrence? *Br J Surg.* 1982;69:613–6.

2. Galler AS, Petrelli NJ, Shakamuri SP. Rectal cancer surgery: a brief history. *Surg Oncol.* 2011;20:223–30.
3. Czito BG, Willett CG. Thirty years of rectal cancer research: a brief history. *Oncology (Williston Park).* 2008;22:1441–2:1444.
4. Rodel C, Liersch T, Becker H, Fietkau R, Hohenberger W, Hothorn T, et al. Preoperative chemoradiotherapy and postoperative chemotherapy with fluorouracil and oxaliplatin versus fluorouracil alone in locally advanced rectal cancer: initial results of the German CAO/ARO/AIO-04 randomised phase 3 trial. *Lancet Oncol.* 2012;13:679–87.
5. Rodel C, Martus P, Papadopoulos T, Fuzesi L, Klimpfnger M, Fietkau R, et al. Prognostic significance of tumor regression after preoperative chemoradiotherapy for rectal cancer. *J Clin Oncol.* 2005;23:8688–96.
6. Roedel C, Becker H, Fietkau R, Graeven U, Hohenberger W, Hothorn T, et al. Preoperative chemoradiotherapy and postoperative chemotherapy with 5-fluorouracil and oxaliplatin versus 5-fluorouracil alone in locally advanced rectal cancer: First results of the German CAO/ARO/AIO-04 randomized phase III trial. *J Clin Oncol.* 2011;29:LBA3505.
7. Sauer R, Becker H, Hohenberger W, Rodel C, Wittekind C, Fietkau R, et al. Preoperative versus postoperative chemoradiotherapy for rectal cancer. *N Engl J Med.* 2004;351:1731–40.
8. Sauer R, Liersch T, Merkel S, Fietkau R, Hohenberger W, Hess C, et al. Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. *J Clin Oncol.* 2012;30:1926–33.
9. Bertelsen CA, Andreasen AH, Jorgensen T, Harling H, Danish Colorectal Cancer, G. Anastomotic leakage after anterior resection for rectal cancer: risk factors. *Colorectal Dis.* 2010;12:37–43.
10. Tortorelli AP, Alfieri S, Sanchez AM, Rosa F, Papa V, Di Miceli D, et al. Anastomotic leakage after anterior resection for rectal cancer with mesorectal excision: incidence, risk factors, and management. *Am Surg.* 2015;81:41–47.
11. Denost Q, Rouanet P, Faucheron JL, Panis Y, Meunier B, Cotte E, et al. To drain or not to drain infraperitoneal anastomosis after rectal excision for cancer: the GRECCAR 5 randomized trial. *Ann Surg.* 2017;265:474–80.
12. Rutegard M, Bostrom P, Haapamaki M, Matthiessen P, Rutegard J. Current use of diverting stoma in anterior resection for cancer: population-based cohort study of total and partial mesorectal excision. *Int J Colorectal Dis.* 2016;31:579–85.
13. Lange MM, den Dulk M, Bossema ER, Maas CP, Peeters KC, Rutten HJ, et al. Risk factors for faecal incontinence after rectal cancer treatment. *Br J Surg.* 2007;94:1278–84.
14. Bryant CL, Lunniss PJ, Knowles CH, Thaha MA, Chan CL. Anterior resection syndrome. *Lancet Oncol.* 2012;13:e403–408.
15. Lange MM, Maas CP, Marijnen CA, Wiggers T, Rutten HJ, Kranenbarg EK, et al. Urinary dysfunction after rectal cancer treatment is mainly caused by surgery. *Br J Surg.* 2008;95:1020–8.
16. Nishizawa Y, Ito M, Saito N, Suzuki T, Sugito M, Tanaka T. Male sexual dysfunction after rectal cancer surgery. *Int J Colorectal Dis.* 2011;26:1541–8.
17. Saito S, Fujita S, Mizusawa J, Kanemitsu Y, Saito N, Kinugasa Y, et al. Male sexual dysfunction after rectal cancer surgery: Results of a randomized trial comparing mesorectal excision with and without lateral lymph node dissection for patients with lower rectal cancer: Japan Clinical Oncology Group Study JCOG0212. *Eur J Surg Oncol.* 2016;42:1851–8.
18. Maas M, Nelemans PJ, Valentini V, Das P, Rodel C, Kuo LJ, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *Lancet Oncol.* 2010;11:835–44.
19. Sanghera P, Wong DW, McConkey CC, Geh JI, Hartley A. Chemoradiotherapy for rectal cancer: an updated analysis of factors affecting pathological response. *Clin Oncol (R Coll Radio).* 2008;20:176–83.
20. Fokas E, Strobil P, Fietkau R, Ghadimi M, Liersch T, Grabenbauer GG, et al. Tumor regression grading after preoperative chemoradiotherapy as a prognostic factor and individual-level surrogate for disease-free survival in rectal cancer. *J Natl Cancer Inst.* 2017;109:1–10.
21. Fokas E, Liersch T, Fietkau R, Hohenberger W, Beissbarth T, Hess C, et al. Tumor regression grading after preoperative chemoradiotherapy for locally advanced rectal carcinoma revisited: updated results of the CAO/ARO/AIO-94 trial. *J Clin Oncol.* 2014;32:1554–62.
22. Yeo SG, Kim DY, Kim TH, Chang HJ, Oh JH, Park W, et al. Pathologic complete response of primary tumor following preoperative chemoradiotherapy for locally advanced rectal cancer: long-term outcomes and prognostic significance of pathologic nodal status (KROG 09-01). *Ann Surg.* 2010;252:998–1004.
23. Habr-Gama A, Perez RO, Nadalin W, Sabbaga J, Ribeiro U Jr, Silva e Sousa AH Jr, et al. Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Ann Surg.* 2004;240:711–7.
24. Maas M, Lambregts DM, Nelemans PJ, Heijnen LA, Martens MH, Leijtens JW, et al. Assessment of clinical complete response after chemoradiation for rectal cancer with digital rectal examination, endoscopy, and MRI: selection for organ-saving treatment. *Ann Surg Oncol.* 2015;22:3873–80.
25. Lambregts DM, Maas M, Bakers FC, Cappendijk VC, Lammering G, Beets GL, et al. Long-term follow-up features on rectal MRI during a wait-and-see approach after a clinical complete response in patients with rectal cancer treated with chemoradiotherapy. *Dis Colon Rectum.* 2011;54:1521–8.
26. van der Valk MJM, Hilling DE, Bastiaannet E, Meershoek-Klein Kranenbarg E, Beets GL, Figueiredo NL, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IIWWD): an international multicentre registry study. *Lancet.* 2018;391:2537–45.
27. Yoo BC, Yeo SG. Clinical utility of pretreatment prediction of chemoradiotherapy response in rectal cancer: a review. *EPMA J.* 2017;8:61–67.
28. Brettingham-Moore KH, Duong CP, Greenawald DM, Heriot AG, Ellul J, Dow CA, et al. Pretreatment transcriptional profiling for predicting response to neoadjuvant chemoradiotherapy in rectal adenocarcinoma. *Clin Cancer Res.* 2011;17:3039–47.
29. Lopes-Ramos C, Koyama FC, Habr-Gama A, Salim AC, Bettoni F, Asprino PF, et al. Comprehensive evaluation of the effectiveness of gene expression signatures to predict complete response to neoadjuvant chemoradiotherapy and guide surgical intervention in rectal cancer. *Cancer Genet.* 2015;208:319–26.
30. Dworak O, Keilholz L, Hoffmann A. Pathological features of rectal cancer after preoperative radiochemotherapy. *Int J Colorectal Dis.* 1997;12:19–23.
31. Millino C, Maretto I, Pacchioni B, Digito M, De Paoli A, Canzonieri V, et al. Gene and microRNA expression are predictive of tumor response in rectal adenocarcinoma patients treated with preoperative chemoradiotherapy. *J Cell Physiol.* 2017;232:426–35.
32. Gaedcke J, Grade M, Camps J, Sokilde R, Kaczkowski B, Schetter AJ, et al. The rectal cancer microRNAome-microRNA expression in rectal cancer and matched normal mucosa. *Clin Cancer Res.* 2012;18:4919–30.
33. Jo P, Nietert M, Gusky L, Kitz J, Conradi LC, Muller-Dornieden A, et al. Neoadjuvant therapy in rectal cancer - biobanking of preoperative tumor biopsies. *Sci Rep.* 2016;6:35589.
34. Grade M, Hummon AB, Camps J, Emons G, Spitzner M, Gaedcke J, et al. A genomic strategy for the functional validation of colorectal cancer genes identifies potential therapeutic targets. *Int J Cancer.* 2011;128:1069–79.
35. Camps J, Pitt JJ, Emons G, Hummon AB, Case CM, Grade M, et al. Genetic amplification of the NOTCH modulator LNX2 upregulates the WNT/beta-catenin pathway in colorectal cancer. *Cancer Res.* 2013;73:2003–13.
36. Ghadimi BM, Jo P. Cancer gene profiling for response prediction. *Methods Mol Biol.* 2016;1381:163–79.
37. Kohavi R, John GH. Wrappers for feature subset selection. *Artif Intell.* 1997;97:273–324.
38. Watanabe T, Komuro Y, Kiyomatsu T, Kanazawa T, Kazama Y, Tanaka J, et al. Prediction of sensitivity of rectal cancer cells in response to preoperative radiotherapy by DNA microarray analysis of gene expression profiles. *Cancer Res.* 2006;66:3370–4.
39. Ghadimi BM, Grade M, Difilippantonio MJ, Varma S, Simon R, Montagna C, et al. Effectiveness of gene expression profiling for response prediction of rectal adenocarcinomas to preoperative chemoradiotherapy. *J Clin Oncol.* 2005;23:1826–38.
40. Empuku S, Nakajima K, Akagi T, Kaneko K, Hijiya N, Etoh T, et al. An 80-gene set to predict response to preoperative chemoradiotherapy for rectal cancer by principle component analysis. *Mol Clin Oncol.* 2016;4:733–9.
41. Kim IJ, Lim SB, Kang HC, Chang HJ, Ahn SA, Park HW, et al. Microarray gene expression profiling for predicting complete response to preoperative chemoradiotherapy in patients with advanced rectal cancer. *Dis Colon Rectum.* 2007;50:1342–53.
42. Mandard AM, Dalibard F, Mandard JC, Marnay J, Henry-Amar M, Petiot JF, et al. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer.* 1994;73:2680–6.
43. Canto LMD, Cury SS, Barros-Filho MC, Kupper BEC, Begnami M, Scapulatempo-Neto C, et al. Locally advanced rectal cancer transcriptomic-based secretome analysis reveals novel biomarkers useful to identify patients according to neoadjuvant chemoradiotherapy response. *Sci Rep.* 2019;9:8702.
44. Cancer Genome Atlas, N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012;487:330–7.
45. Piskol R, Huw L, Sergin I, Kljin C, Modrusan Z, Kim D, et al. A clinically applicable gene-expression classifier reveals intrinsic and extrinsic contributions to consensus molecular subtypes in primary and metastatic colon cancer. *Clin Cancer Res.* 2019;25:4431–42.
46. Conde-Muino R, Cuadros M, Zambudio N, Segura-Jimenez I, Cano C, Palma P. Predictive biomarkers to chemoradiation in locally advanced rectal cancer. *Biomed Res Int.* 2015;2015:921435.

47. Zorcolo L, Rosman AS, Restivo A, Pisano M, Nigri GR, Fancellu A, et al. Complete pathologic response after combined modality treatment for rectal cancer and long-term survival: a meta-analysis. *Ann Surg Oncol*. 2012;19:2822–32.
48. Dossa F, Chesney TR, Acuna SA, Baxter NN. A watch-and-wait approach for locally advanced rectal cancer after a clinical complete response following neoadjuvant chemoradiation: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol*. 2017;2:501–13.
49. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delalogue S, et al. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med*. 2016;375:717–29.
50. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*. 2002;347:1999–2009.
51. Clark-Langone KM, Sangli C, Krishnakumar J, Watson D. Translating tumor biology into personalized treatment planning: analytical performance characteristics of the Oncotype DX Colon Cancer Assay. *BMC Cancer*. 2010;10:691.
52. Gray RG, Quirke P, Handley K, Lopatin M, Magill L, Baehner FL, et al. Validation study of a quantitative multigene reverse transcriptase-polymerase chain reaction assay for assessment of recurrence risk in patients with stage II colon cancer. *J Clin Oncol*. 2011;29:4611–9.
53. Kopetz S, Tabernero J, Rosenberg R, Jiang ZQ, Moreno V, Bachleitner-Hofmann T, et al. Genomic classifier ColoPrint predicts recurrence in stage II colorectal cancer patients more accurately than clinical factors. *Oncologist*. 2015;20:127–33.
54. Ryan JE, Warrier SK, Lynch AC, Ramsay RG, Phillips WA, Heriot AG. Predicting pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a systematic review. *Colorectal Dis*. 2016;18:234–46.
55. Kern SE. Why your new cancer biomarker may never work: recurrent patterns and remarkable diversity in biomarker failures. *Cancer Res*. 2012;72:6097–101.
56. Syed AR. A Review of Cross Validation and Adaptive Model Selection. Thesis, Georgia State University; 2011.
57. Tang J, Alelyani S, Liu H. Feature Selection for Classification: A Review. In: Aggarwal CC, Data Classification: Algorithms and Applications. Data Mining and Knowledge Discovery Series (Chapman and Hall/CRC: New York; 2014) pp 37–64.
58. Jung K, Grade M, Gaedcke J, Jo P, Opitz L, Becker H, et al. A new sensitivity-preferred strategy to build prediction rules for therapy response of cancer patients using gene expression data. *Comput Methods Prog Biomed*. 2010;100:132–9.
59. Runau F, Collins A, Fenech GA, Ford E, Dimitriou N, Chaudhri S, et al. A single institution's long-term follow-up of patients with pathological complete response in locally advanced rectal adenocarcinoma following neoadjuvant chemoradiotherapy. *Int J Colorectal Dis*. 2017;32:341–8.
60. Jun S, Jung YS, Suh HN, Wang W, Kim MJ, Oh YS, et al. LIG4 mediates Wnt signalling-induced radioresistance. *Nat Commun*. 2016;7:10994.
61. Kendziorra E, Ahlborn K, Spitzner M, Rave-Frank M, Emons G, Gaedcke J, et al. Silencing of the Wnt transcription factor TCF4 sensitizes colorectal cancer cells to (chemo-) radiotherapy. *Carcinogenesis*. 2011;32:1824–31.
62. Emons G, Spitzner M, Reineke S, Moller J, Auslander N, Kramer F, et al. Chemoradiotherapy resistance in colorectal cancer cells is mediated by Wnt/beta-catenin signaling. *Mol Cancer Res*. 2017;15:1481–90.
63. Fernandez LM, Sao Juliao GP, Figueiredo NL, Beets GL, van der Valk MJM, Bahadoer RR, et al. Conditional recurrence-free survival of clinical complete responders managed by watch and wait after neoadjuvant chemoradiotherapy for rectal cancer in the International Watch & Wait Database: a retrospective, international, multicentre registry study. *Lancet Oncol*. 2021;22:43–50.
64. Chen MB, Wu XY, Yu R, Li C, Wang LQ, Shen W, et al. P53 status as a predictive biomarker for patients receiving neoadjuvant radiation-based treatment: a meta-analysis in rectal cancer. *PLoS ONE*. 2012;7:e45388.

ACKNOWLEDGEMENTS

We are indebted to all the patients who participated in this analysis. For technical and logistical support and acquisition of biomaterial and follow-up data we thank Chan Rong Lai, Jessica Eggert, Sarah Wachenhausen, Uli Hampacher, Stefanie Zischkau, Birgit Jünemann, Elke Stauffer, Johanna Kreutzer and Philipp Schüler. We acknowledge the efforts of our clinical colleagues Germer, Isbert, Gassmann, Maschuw,

Kessler, Franck and Ludwig that contributed to this work. We thank Reinhard Ebner for critical comments on the manuscript and Buddy Chen for editorial assistance. The work is part of the Clinical Research Unit (KFO 179) and was supported by the Deutsche Forschungsgemeinschaft and by the Intramural Research Program, NIH/National Cancer Institute. NA is supported by the NCI-University of Maryland partnership program for Integrative Cancer Research. GE received a stipend from the Deutsche Forschungsgemeinschaft.

AUTHOR CONTRIBUTIONS

GE, NA, PJ, TR and JG designed the study. PJ, AA, CFH, CR, PS, MG, MG and JG gathered clinical data, collected samples and contributed to the clinical trial. PJ, GS and JG performed the gene-expression arrays. JK and PS analysed pretreatment patient biopsies and generated tumour regression scores. NA and ER implemented the machine learning algorithm. NA, YH, FK and TB executed statistical analysis. US provided IT-infrastructure. GE, NA, YH, FK, TB and ER analysed the data. GE, NA, PJ, TR and JG prepared the manuscript. All authors discussed the results and implications and commented on the manuscript at all stages.

FUNDING

This work was supported by the Deutsche Forschungsgemeinschaft (Klinische Forschergruppe 179) and the Intramural Research Program of the National Institutes of Health, National Cancer Institute. Open Access funding enabled and organized by Projekt DEAL.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the Ethics Committee of the University of Goettingen. All patients provided written informed consent.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41416-022-01842-2>.

Correspondence and requests for materials should be addressed to Jochen Gaedcke.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022