Protein Arginine Deiminase 2 and 4 (PAD2 & PAD4) are independent predictors of good prognosis in colorectal cancer

Mohamed Gijon[1], Rachael L Metheringham[1], Samantha J Paston[1], Lindy G Durrant[1, 2*]

[1] Scancell Limited, Biodiscovery Institute, University Park, Science Road, Nottingham NG7 2RD.
[2]University of Nottingham, Biodiscovery Institute, Division of Cancer and Stem Cells, University Park, Science Road, Nottingham NG7 2RD.

Short title: PAD2, PAD4 and Colorectal Cancer
Abstract word count = 209
Total word count =3567

*Corresponding author: Prof. L.G. Durrant, University of Nottingham, Biodiscovery Institute, Division of Cancer and Stem Cells, University Park, Science Road, Nottingham NG7 2RD; Lindy.Durrant@nottingham.ac.uk.

RLM, MG, LGD designed research
MG performed research,
RLM, MG, & LGD analysed data,
MG, RLM, SJP and LGD wrote the paper with contribution from all authors.

ABSTRACT

Objective: Protein arginine deiminase (PADs) are a family of enzymes that catalyse the post translational modification (PTM) of proteins. In this study the prognostic significance of PAD2 and PAD4 in colorectal cancer (CRC) was assessed.

Design: PAD2 and PAD4 expression was assessed immunohistochemically in a cohort of CRC patients. Association between PAD expression with clinicopathology, protein expression and outcome was determined.

Results: CRC tissues expressed variable levels of PAD2 which was mainly localised in the cytoplasm and correlated with patient survival (p=0.005); high expression increased survival time from 43.5 to 67.6 months. Expression of cytoplasmic PAD2 correlated with expression of nuclear β catenin, PAD4 and alpha-enolase. In contrast expression of nuclear PAD2 correlated with a decrease in survival (p=0.010), with high expression decreasing survival from 76.4 to 42.9 months.

CRC tissues expressed variable levels of PAD4 in both the nucleus and cytoplasm. Expression of cytoplasmic PAD4 correlated with survival (p=0.001) with high expression increasing survival time from 48.1 to 71.8 months. Expression of cytoplasmic PAD4 correlated with expression of, nuclear β catenin, alpha-enolase ($p \leq 0.0001$, p=0.002) and the apoptotic related protein, Bcl-2. Expression of nuclear PAD4 also correlated with survival (p=0.011) with high expression of nuclear PAD4 increasing survival time from 55.4 to 74 months. Expression of nuclear PAD4 correlated with p53, alpha-enolase and Bcl-2. Multivariate analysis showed that TNM stage and cytoplasmic PAD2 and PAD4 remained independent prognostic factors in CRC.

Conclusion: Both PAD2 and PAD4 are good prognostic factors in CRC.

INTRODUCTION

Worldwide, CRC is the third commonly diagnosed cancer after lung and female breast cancer.[1] Moreover, CRC is the second most frequent cancer in Europe and the fourth most diagnosed cancer in the United Stated.[2], [3] Despite substantial improvements in the diagnosis and treatment of CRC it still has a relatively low survival rate with 65% surviving 5 years in high-income countries and less than 50% in low income countries.[4] The early detection and diagnosis of CRC could improve the outcome for the majority of patients; however, early stage CRC patients often do not exhibit any symptoms or have non-specific symptoms such as lower abdominal pain or weight loss. The diagnosis of CRC is often based on results from a colonoscopy, an invasive procedure with limitations since specific symptoms as indicators of CRC are not well established or are compatible with anorectal benign lesions.[5]

The majority of CRCs (nearly 85%) arise from adenomas,[6] the identification of the primary tumour, BRAF gene mutation status and tumour microsatellite instability (MSI) which needs to be determined to help guide clinician on the best treatment. For non-metastatic cases, colectomy and adjuvant chemotherapy is generally used.[7] However, most colon cancer patients will be diagnosed with regional or distant metastasis, thus adjuvant chemotherapy after surgery or palliative chemotherapy will often be needed.[8] The treatment of dispersed metastatic colon cancer includes several active drugs, used alone or in combination including, cetuximab, ramucirumab, bevacizumab, pembrolizumab among others, the treatment often being determined according to the progression of the disease, timing of the prior therapy, mutation profile and the toxicity profile of the corresponding drugs.[9] A small subset of CRC patients have tumours (4-5%) that are microsatellite instability-high (MSI-H) or mismatch repair deficient (dMMR), these patients often have a poor response to chemotherapy[10] but have been shown to have a clinical benefit following treatment with immune checkpoint inhibitors of PD-1/PD-L1 or in combination ipilimumab.[11], [12] A new approach for categorising colorectal cancer is to generate the Immune Score, which is defined by the number and localisation of tumour infiltrating lymphocytes. This score has a dual advantage: firstly, it appears to be the strongest prognostic factor for disease free and overall survival particularly in early stage cancers, and secondly, it provides a tool or a target for novel therapeutic approaches, including immunotherapy.[13] In line with these studies, T cell infiltration was a biomarker of good prognosis in this CRC cohort.[14]

PAD enzymes catalyse the PTM of proteins, in a process called citrullination. The exact biological function of citrullination remains obscure: arginine deamination is recognised to be central for translational regulation of gene expression; to initiate tumour degradation, for formation of neutrophil extracellular traps (NET), and presentation of antigens within stressed cells to the immune system. Five highly conserved PAD enzymes, (1-4 and 6) have been reported in mammals, each exhibiting tissue specific distribution, subcellular localisation and substrate specificity. PAD2 and PAD4 have widespread distribution and are the predominant types in immune cells.[15] The substrate for PAD2 includes the intermediate filament protein

vimentin that is expressed by all mesenchymal cells, myelin basic protein (MBP), actin and keratins.[16] The expression of PAD2 is mainly found in the cytoplasm, but has also been detected in the nucleus despite its lack of a classic nuclear translocation site.[17] β-catenin is critical for tumour development and is commonly activated in CRC. Abnormally high expression of β-catenin can lead to the development of various diseases including cancer where it acts as a transcriptional co-regulator and an adaptor protein for intracellular adhesion.[19] It has been shown that PAD2 negatively regulates β-catenin signalling.[18] PAD4 can also be translocated to the nucleus to citrullinate histones and some other transcription factors, it has also been detected within cytoplasmic granules.[20], [21] PAD4 acts mainly as a transcriptional co-regulator; its substrates include multiple proteins involved in gene regulation of the p53 pathway such as ING4,[22] p300,[23] P21[24] and histones (H3, H2A and H4).[25]–[27] Other substrates for PAD4 includes nucleophosmin[28] and nuclear lamin C,[29] both involved in apoptosis. The expression of PAD2 and PAD4 has been associated with cancer, increased expression has been detected in malignant tumours compared to benign tumours and non-tumour inflamed tissues, suggesting citrullination has induced tumourigenesis.[30] In colon cancer PAD2 is downregulated which is indicative of poor prognosis.[31], [32]

Citrullination can also occur as a result of a degradation and recycling process called autophagy, and is induced in stressed cells,[33] including cancer cells. Citrullination is catalysed exclusively by the family of Ca^{2+} dependent PAD enzymes that convert positively charged arginine residues within substrate proteins to neutrally charged citrulline, this modification can impact the protein structure, induce changes that result in protein denaturation and potentially alter the function of the modified proteins.[34], [35] PAD enzymes are activated by a high concentration of calcium, typically in the millimolar range[33]. In viable cells Ca^{2+} is maintained at low concentrations and under these conditions the PAD enzymes are inactive,[34] however, high concentrations of calcium can occur in double membrane vesicles within viable cells. Citrullination within the nucleus results in gene regulation and can have a role in tumour progression.[16] Citrullination can also occur within autophagosomes as a result of autophagy, here PAD enzymes become activated and citrullinate engulfed proteins from the cytoplasm,[36], [37] this process is induced in stressed cells[17] including cancer cells. In the presence of inflammation, epitopes are citrullinated and can be subsequently presented by MHC class II (MHC-II) molecules for recognition by CD4 T cells[17]. CD4 T cell responses can be detected to citrullinated proteins including vimentin and enolase in autoimmune disease patients such as rheumatoid arthritis.[38]–[41] We have recently shown that citrullinated α-enolase and vimentin are excellent targets for anti-tumour immunity inducing potent immune responses[38], [42], [43] and *in vivo* tumour rejection.[44]

To date the prognostic value of PAD2 and PAD4 in colorectal cancer remains unclear. The authors have investigated the prognostic value of PAD2 and PAD4 expression (total of 247 and 231 cases respectively), the correlation of PAD2 and PAD4 expression relative to relevant clinical and other relevant markers, and

the association of PAD2 and PAD4 expression on survival in patients undergoing surgical tumour resection for CRC.

PATIENTS AND METHODS

Patient samples

This study included 247 valid CRC primary tumours stained for PAD2 of which 231 were also stained for PAD4, these were obtained from patients undergoing elective surgical resection of histologically proven primary colorectal carcinoma at Nottingham University Hospitals, Nottingham between January 1994 and December 2000, and constructed as a tissue microarray (TMA). Clinical details of the study cohort have been previously described [14], [45] and are summarised in Table 1.

Patients with lymph node positive disease routinely received adjuvant chemotherapy with 5- fluorouracil and folinic acid. The mean follow up period was 42 months (range 1-116). Disease specific survival (DSS) was determined from the date of resection of the primary tumour until December 31st 2003 when any survivors were censored.

Immunohistochemistry (IHC)

To validate the use of TMAs for immuno-phenotyping, full-face sections of 40 randomly selected cases were stained and the protein expression levels were compared. The concordance between TMAs and full-face sections was excellent using Cohen's kappa test (kappa=0.8). Immunohistochemical staining was performed using a routine streptavidin–biotin peroxidase method in conjunction with a Leica autostainer XL for tissue dewaxing and rehydration, and the Novolink polymer detection system kit (Leica Biosystems, Newcastle-upon-Tyne, UK). Full face and TMA sections were first deparaffinised with xylene, rehydrated through graded alcohol and immersed in Peroxidase block (Novolink) for 5 mins. Antigen retrieval was achieved by immersing sections in pH 6.0 citrate buffer and heating in a 800W microwave for 10 mins at high power and 10 mins at low power. Endogenous avidin/biotin binding was blocked (Protein blocking kit, Novolink), followed by addition of 100 µL Protein Block (Novolink) for 5 minutes. Sections were incubated for one hour with polyclonal rabbit anti-PAD2 antibody (Abcam, cat no 16478, 1:150 dilution) or monoclonal rabbit anti-PAD4 antibody (Abcam cat no ab128086 1:700 dilution) for a 1 hr incubation at room temperature. After washing with tris buffered saline (TBS), sections were incubated with 100 µL of post primary antibody (Novolink kit) for 30 mins. After washing in TBS sections were incubated with 100 µL Novolink polymer for 30 mins. Sections were washed in TBS. Visualisation was achieved using 3, 3′-diaminobenzidine tetra hydrochloride (DAB, Novolink), and then counterstained with haematoxylin (Novolink), dehydrated in alcohol, cleared in xylene (Genta Medica, York, UK), and mounted with distyrene, plasticiser and xylene (DPX–BDH, Poole, UK).

Slide scanning, viewing and scoring

TMA stained slides were digitalised using NanoZoomer (Hamamatsu, Bridgewater, NJ, USA) high throughput scanner. Images were visualised using Aperio ImageScope v11.2.0.780 software. Cytoplasmic PAD2 and cytoplasmic and nuclear PAD4 proteins expression was assessed using the percentage of positivity (0-100%) and staining intensity (Negative: Score 0, Weak: Score 1, Moderate: Score 2, and Strong: Score 3) following the Histo-score system (H-score 0 to 300. Negative=0; weak=1-100; moderate=101-200 and strong=201-300). Observers were blind to clinical and pathological parameters.

Statistical analysis

Statistical analysis was performed using SPSS21 statistical software (SPSS Inc., Chicago, IL, USA). Pearson's $\chi 2$-tests and Fisher's exact tests were used to determine the significance of associations between categorical variables. For sorting of the data in two groups, X-Tile (X-Tile Bioinformatics Software, Yale University, version 3.6.1) was used. Survival rates were calculated using the Kaplan–Meier method; differences between survival curves were examined using the log-rank test. The Cox regression model was used for multivariate analysis in order to calculate the Hazard ratios and independent significance of individual factors. In all cases two-sided P-values of <0.05 were considered as statistically significant.

RESULTS

Frequency and distribution of PAD2 and PAD4 expression in the study cohort

The patient cohort consisted of 247 and 231 valid tumour sections from patients with CRC stained for PAD2 and PAD4 expression respectively after exclusion of uninformative cores (i.e., loss or folding of cores), details of the TMA samples are shown in Table 1. The results from the IHC staining showed that 99% of the cases expressed PAD2 in the cytoplasm, 28% of these also stained positive for nuclear expression of PAD2. All tumours in the cohort showed strong PAD4 expression, the localisation of the staining was both cytoplasmic and nuclear. The IHC staining also revealed that some surrounding inflammatory cells and fibroblasts also expressed PAD2 and PAD4, however, this was not evaluated further in this study.

Using the H-score assessment the expression of cytoplasmic and nuclear PAD2 and PAD4 was evaluated. According to the H-score assessment 62% tumours showed strong expression of cytoplasmic PAD2, 28% moderate expression, 9% weak expression and only 1% were negative for PAD2, representative staining is shown in Figure 1A-C. For nuclear PAD2 30% tumours showed strong expression, 32% moderate expression, 34% weak expression and only 4% were negative for nuclear PAD2. The H-score assessment for cytoplasmic PAD4 showed all tumours expressed PAD4, 11% strong expression, 61% moderate expression and 28% weak expression, representative staining is shown in Figure 1D-F. All tumours expressed nuclear PAD4, 77% strong expression, 19% moderate expression and 4% showed weak expression.

The cohort was categorised into two groups by defining the optimal cut-off points for cytoplasmic PAD2 (H-score 150), nuclear PAD2 (H-score 160), cytoplasmic PAD4 (H-score 100) and nuclear PAD4 (H-score 260) expression by X-tile software analysis (Table 2).

Correlation of PAD2 and PAD4 expression with standard clinicopathological variables and gene expression

A univariate analysis was performed to determine whether the expression of PAD2 or PAD4 correlated with other markers involved in cancer, cell survival and progression (Table 2). This analysis showed that the expression nuclear PAD2 significantly correlated with the TNM stage (p=0.017), no significant association were observed between nuclear PAD4, cytoplasmic PAD2 or PAD4 expression with any of the other clinicopathological parameters.

The results from the univariate analysis for cytoplasmic PAD2 with protein expression showed that there is a significant correlation with nuclear β-catenin (p=0.020), vascular endothelial growth factor A (VEGFA) (p=0.001) and CEA (p=0.042), cytoplasmic and nuclear PAD4 ($p \leq 0.0001$, p=0.048) and cytoplasmic alpha-enolase ($p \leq 0.001$). The results from the analysis for nuclear PAD2 showed only a correlation with the tumour suppressor protein, p53 (p=0.024). The results from the univariate analysis for cytoplasmic PAD4 showed there is a significant correlation with nuclear β-catenin (p=0.004), nuclear PAD4 ($p \leq 0.0001$), cytoplasmic and nuclear alpha-enolase ($p \leq 0.0001$, $p \leq 0.002$) and the apoptotic related protein, Bcl-2 (p=0.032). The results from the analysis for nuclear PAD4 showed there is a significant correlation with nuclear alpha-enolase (p=0.011), p53 (p=0.016) and Bcl-2 (p=0.022) (Table 3).

Prognosis outcome according to PAD2 and PAD4 expression

We next examined the relationship between the expression of PAD2 or PAD4 and overall survival. The Kaplan–Meier survival curve for cytoplasmic PAD2 (Figure 2A) showed there was a significant correlation with survival (log rank test, p=0.005) where the expression of cytoplasmic PAD2 increased survival time from 43.5 months (95% CI 29.2 to 57.8) to 67.6 months (95% CI 60.5 to 74.8). In contrast the Kaplan–Meier survival curve for nuclear PAD2, (Figure 2B) showed high expression of nuclear PAD2 correlated with a decrease in survival (log rank test, p=0.010), decreasing from 76.4 months (95% CI 59.5 to 93.45) to 42.9 months (95% CI 28.3 to 57.6, supplementary data Table 1). The Kaplan-Meier survival curve for cytoplasmic and nuclear PAD4 (Figure 2C and 2D) showed there was a significant correlation with survival with tumours expressing cytoplasmic PAD4 and nuclear PAD4 (log rank test, p=0.001 and p=0.011 respectively). The expression of cytoplasmic PAD4 increased survival time from 48.1 months (95% CI 35.2 to 60.1) to 71.8 months (95% CI 64.0 to 79.5) and the expression of nuclear PAD4 increased survival time from 55.4 months (95% CI 45.9 to 65.0) to 74.0 months (95% CI 64.7 to 83.4, supplementary data Table 1).

Cytoplasmic PAD2 and PAD4 expression by colorectal tumours are independent prognostic markers

We next investigated if cytoplasmic PAD2 and cytoplasmic PAD4 were independent prognostic markers. A multivariate analysis was performed using Cox regression including the clinicopathological variables TNM stage and vascular invasion. This analysis showed that the expression of cytoplasmic PAD2 (Table 4) was a prognostic protective marker that was independent of clinicopathology, in addition, the analysis also revealed that TNM stage (95% CI 1689 to 3.105, $p \leq 0.001$) was identified as having independent prognostic significance. The expression of cytoplasmic PAD4 (Table 4) also conferred tumour protection that was independent of clinicopathology. The TNM stage (95% CI 1.613 to 3.104, $p \leq 0.0001$) was identified as being an independent marker of good prognosis in CRC cancer. These results indicate that both cytoplasmic PAD2 and cytoplasmic PAD4 would make good prognostic markers in CRC cancer.

The multivariate analysis for cytoplamsic PAD2 and cytoplasmic PAD4 showed that both were independent prognostic markers (Table 4). This analysis showed that the expression of nuclear PAD2 and nuclear PAD4 were not independent of clinicopathology, with TNM stage (p=0.001) being identified as having independent prognostic significance.

DISCUSSION

The early diagnosis and treatment of CRC is critical to improve the outcome. Many prognostic makers have been investigated to try and identify patients at high risk of relapse. The identification of a reliable prognostic marker has been challenging. This study investigated the prognostic value of measuring the expression of PAD2 and PAD4 in a cohort of colorectal cancer patients. Post translational modifications of the human proteome have a major impact on health and disease, expanding the functional diversity of proteins by adding new functional groups. There is accumulating evidence that citrullination is implicated in cancer development and can be an excellent target to alert the immune response to recognise and remove stressed cells including cancer cells.

Our data shows that 99% of CRC tumours express PAD2 and all tumours in our cohort expressed PAD4. Increasing evidence suggests that the PAD enzymes play a role normal cell transformation and tumour progression. We also show that in the majority of cases PAD4 was expressed in both the cytoplasm and nucleus, whereas PAD2 was mainly located in the cytoplasm with only 28% of tumours also showing nuclear expression. Analysis of the clinicopathological parameters for the cohort stained with PAD2 and PAD4 revealed that only nuclear PAD2 showed a correlation with a clinical parameter, notably TNM stage. In healthy tissue strong nuclear and cytoplasmic PAD2 staining has been observed, but in tumours and as they progress there is a general reduction in the nuclear localisation of PAD2.[46] We also confirm this finding in our cohort, where only 28% of tumours expressed nuclear PAD2.

We investigated if the expression of PAD2 or PAD4 correlated with other markers involved in cancer, cell survival and progression. Nuclear PAD2 correlated with p53 expression and both conferred a worse prognosis in this cohort.[47] In contrast nuclear PAD4 conferred a good prognosis and inversely correlated with p53 expression. In line with this observation Anwar *et al*, described 35 studies (24 using immunohistochemical methods) in which p53 was found to be predictive of a worse outcome. Multiple interactions between PAD4 and p53 have been reported, suggesting the importance of PAD-induced citrullination in apoptosis. Expression of a major p53 target gene OKL38 was repressed by a p53-mediated recruitment of PAD4 to the promoter of OKL38 and subsequent removal of histone arginine methylation mark, thus directly modulating apoptosis.[48] PAD4 was found to bind and subsequently citrullinate the inhibitor of growth 4 (ING4), another tumour suppressor protein known to bind p53. PAD4-driven citrullination of ING4 at the nuclear localisation sequence region prevented p53-to-ING4 binding, repressed p53 acetylation, and subsequently inhibited downstream p21 expression.[22] Nuclear and cytoplasmic PAD4 conferred a good prognosis and correlated with expression of Bcl-2 which is also a predictor of good prognosis in this array.[47] Despite the defined role of Bcl-2 protein in suppression of apoptosis, in colorectal cancer the evidence suggests that Bcl-2 expression is correlated with favourable parameters[49] and a better prognosis.[50] This may be explained by the finding that although Bcl-2 inhibits apoptosis it may also slow cell growth.[51] It has been reported that Bcl-2 contains an anti-proliferative domain, distinct from the domains required for its anti-apoptotic activity,[52] and also Bcl-2 can be converted to a BAX-like death effector by cleavage of a regulatory loop domain by caspases.[53] Cytoplasmic PAD2 and PAD4 expression correlated with nuclear β-catenin expression and both correlated with good prognosis in this cohort. Wnt/β-catenin signalling is important for carcinogenesis of certain tumours, and PAD2-mediated citrullination was recently linked to this pathway.[18] The anti-parasitic drug temozolomide (NTZ) inhibited Wnt/β-catenin signalling through directly targeting PAD2 as demonstrated by co-immunoprecipitation studies. The addition of NTZ resulted in a significant increase of β-catenin citrullination by PAD2 and also substantially extended the half-life of PAD2 protein, suggesting that NTZ enhances the stability of PAD2. Furthermore, colony growth assays suggested that PAD2- mediated citrullination of β-catenin could limit the proliferation of colon cancer cells by inhibiting the Wnt pathway. As a result, this elegant study uncovered a hitherto unrecognised mechanism of Wnt signalling inhibition via PAD2.

The PAD enzymes have been described as playing a role in early tumour cell transformation and tumour progression. In addition, the PAD enzymes also catalyse the PTM of proteins under conditions of cellular stress, in the process of citrullination. Citrullination occurs as a result of the degradation and recycling process of autophagy. In the presence of inflammation, epitopes are citrullinated and can be subsequently presented by MHC-II molecules for recognition by CD4 T cells.[17] We have recently shown that citrullinated α-enolase and vimentin are excellent targets for anti-tumour immunity inducing potent immune responses[38], [42], [43] and *in vivo* tumour rejection.[44] It was therefore of interest that cytoplasmic PAD2 correlated with cytoplasmic and nuclear alpha-enolase, and cytoplasmic PAD4 correlated with

vimentin and cytoplasmic and nuclear alpha-enolase suggesting that these may be good targets for citrullination in colorectal cancer, these tumour cells can then be recognised and removed by the immune system.

In conclusion our results show that high expression of cytoplasmic PAD2, cytoplasmic PAD4 and nuclear PAD4 were all associated with an increase in overall survival, this ranged from an 18 to 24 months improvement over patients whose tumours expressed low levels of these markers. Nuclear expression of PAD4 may confer a good prognosis by regulating expression of genes such as the p53/Bcl-2. Cytoplasmic PAD2 and or PAD4 may confer a good prognosis by altering cell signalling pathways such as via β-catenin or may be involved in protein degradation which can lead to presentation of citrullinated epitopes for detection by cytotoxic CD4 T cells. This finding is consistent with a previous study which showed that CRC tumours tend to lose expression of the PAD enzymes thus allowing them to evade immune recognition. However, cancer vaccines targeting citrullinated vimentin and/or enolase maybe a good therapy for CRC patients expressing high levels of PAD enzymes.[15]

Acknowledgements

The authors would also like to thank Dr Tina Parsons for critical reading of the manuscript and Maria Diez and Chris Nolan for their technical support and guidance.

Funding: This work was funded by Scancell Ltd.

Conflict of interest

LG Durrant is a director and shareholder in Scancell Ltd. All authors are employees of Scancell Ltd.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA. Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, 2018, doi: 10.3322/caac.21492.

[2] M. Zavoral *et al.*, "Colorectal cancer screening in Europe," *World J. Gastroenterol.*, vol. 15, no. 47, pp. 5907–5915, 2009, doi: 10.3748/wjg.15.5907.

[3] J. K. Lee, E. G. Liles, S. Bent, T. R. Levin, and D. A. Corley, "Accuracy of fecal immunochemical tests for colorectal cancer: Systematic review and meta-analysis," *Ann. Intern. Med.*, vol. 160, no. 3, pp. 171–181, 2014, doi: 10.7326/M13-1484.

[4] H. Brenner, M. Kloor, and C. P. Pox, "Colorectal cancer," *Lancet*, vol. 383, no. 9927, pp. 1490–1502, 2014, doi: 10.1016/S0140-6736(13)61649-9.

[5] P. Vega, F. Valentín, and J. Cubiella, "Colorectal cancer diagnosis: Pitfalls and opportunities," *World J. Gastrointest. Oncol.*, vol. 7, no. 12, p. 422, 2015, doi: 10.4251/wjgo.v7.i12.422.

[6] M. L. Dwyer, A. Colombo, and G. Bozzi, "A Genetic Model for Colorectal Tumorigenesis," *G. Ital. Cardiol.*, vol. 19, no. 2, pp. 170–172, 1989.

[7] A. B. Benson *et al.*, "NCCN Guidelines ® Insights Colon Cancer, Version 2.2018 Featured Updates to the NCCN Guidelines," *JNCCN J. Natl. Compr. Cancer Netw.*, vol. 16, no. 4, pp. 359–369, 2018, doi: 10.6004/jnccn.2018.0021.

[8] C. Wu, "Systemic Therapy for Colon Cancer," *Surg. Oncol. Clin. N. Am.*, vol. 27, no. 2, pp. 235–242, 2018, doi: 10.1016/j.soc.2017.11.001.

[9] A. B. Benson *et al.*, "Colon cancer, version 1.2017: Clinical practice guidelines in oncology," *JNCCN J. Natl. Compr. Cancer Netw.*, vol. 15, no. 3, pp. 370–398, 2017, doi: 10.6004/jnccn.2017.0036.

[10] J. Taieb *et al.*, "473OThree versus six months' adjuvant oxaliplatin-based chemotherapy for patients with stage III colon cancer: Per-protocol, subgroups and long-lasting neuropathy results," *Ann. Oncol.*, vol. 28, no. suppl_5, p. 2017, 2017, doi: 10.1093/annonc/mdx393.

[11] L. Marcus, S. J. Lemery, P. Keegan, and R. Pazdur, "FDA approval summary: Pembrolizumab for the treatment of microsatellite instability-high solid tumors," *Clin. Cancer Res.*, vol. 25, no. 13, pp. 3753–3758, 2019, doi: 10.1158/1078-0432.CCR-18-4070.

[12] S. Kamatham, F. Shahjehan, and P. M. Kasi, "Immune Checkpoint Inhibitors in Metastatic Colorectal Cancer: Current Status, Recent Advances, and Future Directions," *Curr. Colorectal Cancer Rep.*, vol. 15, no. 4, pp. 112–121, 2019, doi: 10.1007/s11888-019-00437-z.

[13] H. K. Angell, D. Bruni, J. Carl Barrett, R. Herbst, and J. Galon, "The immunoscore: Colon cancer and beyond a C," *Clin. Cancer Res.*, vol. 26, no. 2, pp. 332–339, 2020, doi: 10.1158/1078-0432.CCR-18-1851.

[14] J. A. D. Simpson, A. Al-Attar, N. F. S. Watson, J. H. Scholefield, M. Ilyas, and L. G. Durrant, "Intratumoral T cell infiltration, MHC class I and STAT1 as biomarkers of good prognosis in colorectal cancer," *Gut*, vol. 59, no. 7, pp. 926–933, 2010, doi: 10.1136/gut.2009.194472.

[15] Y. Zhou *et al.*, "Spontaneous secretion of the citrullination enzyme PAD2 and cell surface exposure of PAD4 by neutrophils," *Front. Immunol.*, vol. 8, no. SEP, pp. 1–16, 2017, doi: 10.3389/fimmu.2017.01200.

[16] A. E. Yuzhalin, "Citrullination in cancer," *Cancer Res.*, vol. 79, no. 7, pp. 1274–1284, 2019, doi: 10.1158/0008-5472.CAN-18-2797.

[17] V. A. Brentville, M. Vankemmelbeke, R. L. Metheringham, and L. G. Durrant, "Post-translational modifications such as citrullination are excellent targets for cancer therapy," *Semin. Immunol.*, vol. 47, no. October 2019, p. 101393, 2020, doi: 10.1016/j.smim.2020.101393.

[18] Y. Qu *et al.*, "Small molecule promotes b-catenin citrullination and inhibits Wnt signaling in cancer," *Nat. Chem. Biol.*, vol. 14, no. 94, 2018, doi: 10.1038/nchembio.2510.

[19] S. Shang, F. Hua, and Z. W. Hu, "The regulation of β-catenin activity and function in cancer:

Therapeutic opportunities," *Oncotarget*, vol. 8, no. 20, pp. 33972–33989, 2017, doi: 10.18632/oncotarget.15687.

[20] H. Asaga, K. Nakashima, T. Senshu, and A. Ishigami, "Immunocytochemical localization of peptidylarginine deiminase in human eosinophils and neutrophils Abstract : Peptidylarginine deiminase , registered banks , is expressed in HL-60 cells differentiated blood cell fractions . PAD activity with similar sub- ," *J. Leukoc. Biol.*, 2001.

[21] K. Nakashima, T. Hagiwara, and M. Yamada, "Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes," *J. Biol. Chem.*, vol. 277, no. 51, pp. 49562–49568, 2002, doi: 10.1074/jbc.M208795200.

[22] Q. Guo and W. Fast, "Citrullination of Inhibitor of Growth 4 (ING4) by Peptidylarginine Deminase 4 (PAD4) disrupts the interaction between ING4 and p53," *J. Biol. Chem.*, vol. 286, no. 19, pp. 17069–17078, 2011, doi: 10.1074/jbc.M111.230961.

[23] Y. H. Lee, S. A. Coonrod, W. L. Kraus, M. A. Jelinek, and M. R. Stallcup, "Regulation of coactivator complex assembly and function by protein arginine methylation and demethylimination," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 10, pp. 3611–3616, 2005, doi: 10.1073/pnas.0407159102.

[24] P. Li *et al.*, "Regulation of p53 Target Gene Expression by Peptidylarginine Deiminase 4," *Mol. Cell. Biol.*, vol. 28, no. 15, pp. 4745–4758, 2008, doi: 10.1128/mcb.01747-07.

[25] T. Hagiwara, Y. Hidaka, and M. Yamada, "Deimination of histone H2A and H4 at arginine 3 in HL-60 granulocytes," *Biochemistry*, vol. 44, no. 15, pp. 5827–5834, 2005, doi: 10.1021/bi047505c.

[26] G. L. Cuthbert *et al.*, "Histone deimination antagonizes arginine methylation," *Cell*, vol. 118, no. 5, pp. 545–553, 2004, doi: 10.1016/j.cell.2004.08.020.

[27] Y. Wang *et al.*, "Human PAD4 regulates histone arginine methylation levels via demethylimination," *Science (80-.).*, vol. 306, no. 5694, pp. 279–283, 2004, doi: 10.1126/science.1101400.

[28] T. Hagiwara, K. Nakashima, H. Hirano, T. Senshu, and M. Yamada, "Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes," *Biochem. Biophys. Res. Commun.*, vol. 290, no. 3, pp. 979–983, 2002, doi: 10.1006/bbrc.2001.6303.

[29] C. Tanikawa *et al.*, "Regulation of histone modification and chromatin structure by the p53-PADI4 pathway," *Nat. Commun.*, vol. 3, 2012, doi: 10.1038/ncomms1676.

[30] X. Chang, J. Han, L. Pang, Y. Zhao, Y. Yang, and Z. Shen, "Increased PADI4 expression in blood and tissues of patients with malignant tumors," *BMC Cancer*, vol. 9, pp. 1–11, 2009, doi: 10.1186/1471-2407-9-40.

[31] R. Funayama *et al.*, "Protein-arginine deiminase 2 suppresses proliferation of colon cancer cells through protein citrullination," *Cancer Sci.*, vol. 108, no. 4, pp. 713–718, 2017, doi: 10.1111/cas.13179.

[32] N. Cantarino *et al.*, "Downregulation of the deiminase PADI2 is an early event in colorectal carcinogenesis and indicates poor prognosis," *Mol. Cancer Res.*, vol. 14, no. 9, pp. 841–848, 2016, doi: 10.1158/1541-7786.MCR-16-0034.

[33] V. A. Brentville *et al.*, "Combination vaccine based on citrullinated vimentin and enolase mediated peptides induces potent CD4-mediated- anti--tumor responses," pp. 6–9, 2020, doi: 10.1136/jitc-2020-000560.

[34] M. Heilbron Ferrer, "Protein Arginine Deiminases and Associated Citrullination: Physiological Functions and Diseases Associated with Dysregulation," *Curr Drug Targets. 2015 ;*, vol. 16, no. 57, pp. 199–227, 2002, doi: 10.1016/j.bbamem.2015.02.010.Cationic.

[35] M. Alghamdi *et al.*, "An overview of the intrinsic role of citrullination in autoimmune disorders," *J. Immunol. Res.*, vol. 2019, 2019, doi: 10.1155/2019/7592851.

[36] J. M. Ireland and E. R. Unanue, "Autophagy in antigen-presenting cells results in presentation of citrullinated peptides to CD4 T cells," *J. Exp. Med.*, vol. 208, no. 13, pp. 2625–2632, 2011, doi: 10.1084/jem.20110640.

[37] M. Sorice *et al.*, "Autophagy generates citrullinated peptides in human synoviocytes: A possible

trigger for anti-citrullinated peptide antibodies," *Rheumatol. (United Kingdom)*, vol. 55, no. 8, pp. 1374–1385, 2016, doi: 10.1093/rheumatology/kew178.

[38] A. L. Feitsma *et al.*, "Identification of citrullinated vimentin peptides as T cell epitopes in HLA-DR4-positive patients with rheumatoid arthritis," *Arthritis Rheum.*, vol. 62, no. 1, pp. 117–125, 2010, doi: 10.1002/art.25059.

[39] C. Gerstner *et al.*, "Functional and structural characterization of a novel HLA-DRB1*04: 01-restricted α-enolase T cell epitope in rheumatoid arthritis," *Front. Immunol.*, vol. 7, no. NOV, pp. 1–12, 2016, doi: 10.3389/fimmu.2016.00494.

[40] E. A. James *et al.*, "Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy," *Arthritis Rheumatol.*, vol. 66, no. 7, pp. 1712–1722, 2014, doi: 10.1002/art.38637.

[41] O. Snir *et al.*, "Identification and functional characterization of T cells reactive to citrullinated vimentin in HLA-DRB1*0401-positive humanized mice and rheumatoid arthritis patients," *Arthritis Rheum.*, vol. 63, no. 10, pp. 2873–2883, 2011, doi: 10.1002/art.30445.

[42] V. A. Brentville *et al.*, "Citrullinated vimentin presented on MHC-II in tumor cells is a target for CD4+ T-Cell-mediated antitumor immunity," *Cancer Res.*, vol. 76, no. 3, pp. 548–560, 2016, doi: 10.1158/0008-5472.CAN-15-1085.

[43] L. G. Durrant, R. L. Metheringham, and V. A. Brentville, "Autophagy, citrullination and cancer," *Autophagy*, vol. 12, no. 6, pp. 1055–1056, 2016, doi: 10.1080/15548627.2016.1166326.

[44] V. A. Brentville *et al.*, "Citrullinated vimentin presented on MHC-II in tumor cells is a target for CD4+ T-Cell-mediated antitumor immunity," *Cancer Res.*, vol. 76, no. 3, pp. 548–560, 2016, doi: 10.1158/0008-5472.CAN-15-1085.

[45] N. F. S. Watson *et al.*, "Immunosurveillance is active in colorectal cancer as downregulation but not complete loss of MHC class I expression correlates with a poor prognosis," *Int. J. Cancer*, vol. 118, no. 1, pp. 6–10, 2006, doi: 10.1002/ijc.21303.

[46] B. D. Cherrington *et al.*, "Comparative Analysis of Peptidylarginine Deiminase-2 Expression in Canine, Feline and Human Mammary Tumours," *J. Comp. Pathol.*, vol. 147, no. 2–3, pp. 139–146, 2012, doi: 10.1016/j.jcpa.2012.01.021.

[47] J. H. S. and L. G. D. Nicholas FS Watson*1,2, Zahra Madjd1, Duncan Scrimegour1, Ian Spendlove1, Ian O Ellis3, "Evidence that the p53 negative / Bcl-2 positive phenotype is an independent indicator of good prognosis in colorectal cancer: A tissue microarray study of 460 patients," *Tandlægebladet*, vol. 9, no. 10, pp. 790–796, 2005, doi: 10.1186/1477-7819-3-Received.

[48] H. Yao *et al.*, "Histone Arg modifications and p53 regulate the expression of OKL38, a mediator of apoptosis," *J. Biol. Chem.*, vol. 283, no. 29, pp. 20060–20068, 2008, doi: 10.1074/jbc.M802940200.

[49] F. A. Sinicrope, J. Hart, F. Michelassi, and J. Jack Lee, "Prognostic Value of bcl-2 Oncoprotein Expression in Stage II Colon Carcinoma," *Clin. Cancer Res.*, vol. 1, no. 10, pp. 1103–1110, 1995.

[50] D. Öfner *et al.*, "Immunohistochemically detectable bcl-2 expression in colorectal carcinoma: Correlation with tumour stage and patient survival," *Br. J. Cancer*, vol. 72, no. 4, pp. 981–985, 1995, doi: 10.1038/bjc.1995.446.

[51] J. A. Pietenpol, N. Papadopoulos, S. Markowitz, J. K. V Willson, K. W. Kinzler, and B. Vogelstein, "Paradoxical Inhibition of Solid Tumor Cell Growth by bcl2 '," *Cancer Res.*, vol. 54, no. 14, pp. 3714–3717, 1994.

[52] P. Theodorakis, C. D'Sa-Eipper, T. Subramanian, and G. Chinnadurai, "Unmasking of a proliferation-restraining activity of the anti-apoptosis protein EBV BHRF1," *Oncogene*, vol. 12, no. 8, pp. 1707–1713, 1996.

[53] E. H. Cheng *et al.*, "Conversion of Bcl-2 to a Bax-like Death Effector by Caspases," *Science (80-.).*, vol. 278, no. December, pp. 2–5, 1997.

Table 1: Details of the colorectal tumour microarray

	PAD2	***(%)***	***PAD4***	***(%)***
Total cases:	247		231	
Age at operation (years):				
Median	72		71	
Range	57-88		57-87	
Gender:				
Male	149	60	136	59
Female	98	40	95	41
Histologic tumour type				
Adenocarcinoma	213	86	198	86
Mucinous	31	13	29	13
Unknown	3	1	4	1
Tumour differentiation				
Well	19	8	15	6
Moderate	175	71	169	73
Poor	49	20	43	19
Unknown	4	1	4	2
Tumour site				
Right colon	76	31	75	32
Left colon	55	22	44	19
Rectal	94	38	87	38
Unknown	22	9	25	11
TNM stage				
I	35	14	30	13
II	95	39	85	37
III	85	34	84	36
IV	28	11	27	12
*Unknown**	4	2	5	2

*Unknown cases were excluded from analysis

Table 2. Clinicopathological variables for the patient cohort stained for cytoplasmic and nuclear PAD2 and PAD4

PAD2	**Cytoplasmic PAD2 (%)**			**Nuclear PAD2 (%)**		
Variable	**Low PAD2**	**High PAD2**	**χ2 test (*p* Value)**	**Low PAD4**	**High PAD4**	**χ2 test (*p* Value)**
Age						
≤ 72	22 (16.2)	114 (83.8)	0.702	18 (43.9)	23 (56.1)	0.800
> 72	20 (18.0)	91 (82.0)		15 (46.9)	17 (53.1)	
Gender						
Male	24 (16.1)	125 (83.9)	0.644	20 (51.3)	19 (48.7)	0.264
Female	18 (18.4)	80 (81.6)		13 (38.2)	21 (61.8)	
Histologic tumour type						
Adenocarcinoma	34 (16.0)	179 (84.0)	0.634	20 (50.0)	19 (50.0)	0.264
Mucinous	6 (19.4)	25 (80.6)		13 (20.0)	21 (80.0)	
Tumour differentiation						
Well	2 (10.5)	17 (89.5)		3 (42.9)	4 (57.1)	
Moderate	28 (16.0)	147 (84.0)	0.586	24 (50.0)	24 (50.0)	0.571
Poor	10 (20.4)	39 (79.6)		6 (35.3)	11 (64.7)	
Tumour site						
Right colon	14 (18.4)	62 (81.6)		10 (45.5)	12 (54.5)	
Left colon	10 (18.2)	45 (81.8)	0.898	9 (60.0)	10 (40.0)	0.544
Rectal	15 (16.0)	79 (84.0)		12 (42.9)	16 (57.1)	
TNM stage						
I - II	20 (15.4)	110 (84.6)	0.475	24 (57.1)	18 (42.9)	**0.017***
III - IV	22 (18.8)	95 (81.2)		9 (29.0)	22 (71.0)	
PAD4	**Cytoplasmic PAD4 (%)**			**Nuclear PAD4 (%)**		
Variable	**Low PAD2**	**High PAD2**	**χ2 test (*p* Value)**	**Low PAD4**	**High PAD4**	**χ2 test (*p* Value)**
Age						
≤ 72	35 (27.1)	94 (72.9)	0.827	60 (46.5)	69 (53.5)	0.598
> 72	29 (28.4)	73 (71.6)		51 (50.0)	51 (50.0)	
Gender						
Male	34 (25.0)	102 (75.0)	0.272	62 (25.0)	74 (54.4)	0.370
Female	30 (31.6)	65 (68.4)		49 (31.6)	46 (48.4)	
Histologic tumour type						
Adenocarcinoma	56 (28.3)	142 (71.7)	0.391	93 (47.0)	105 (53.0)	0.370
Mucinous	6 (20.7)	23 (79.3)		16 (55.2)	13 (44.8)	
Tumour differentiation						
Well	4 (26.7)	11 (73.3)	0.691	7 (46.7)	8 (53.3)	0.336
Moderate	44 (26.0)	125 (74.0)		77 (45.6)	92 (54.4)	
Poor	14 (32.6)	29 (67.4)		25 (58.1)	18 (41.9)	
Tumour site						
Right colon	24 (32.0)	51 (68.0)	0.498	41 (54.7)	34 (45.3)	0.240
Left colon	11 (25.0)	33 (75.0)		21 (47.7)	23 (52.3)	
Rectal	21 (24.1)	66 (75.9)		36 (41.4)	51 (58.6)	
TNM stage						
I - II	29 (25.2)	86 (74.8)	0.400	50 (43.5)	65 (56.5)	0.166
III - IV	35 (30.2)	81 (69.8)		61 (52.6)	55 (47.4)	

Table 3. Univariate analysis of cytoplasmic and nuclear PAD2 and PAD4 expression in correlation with other tumour markers using the χ2 or Fisher's exact test.

Variable	PAD2 cytoplasmic	PAD2 nuclear	PAD4 cytoplasmic	PAD4 nuclear
	χ2 test (*p* Value)	**χ2 test (*p* Value)**	**χ2 test (*p* Value)**	**χ2 test (*p* Value)**
PAD2 cytoplasmic		0.394	**<0.0001****	**0.048***
PAD2 nuclear	0.394		0.145	0.336
PAD4 cytoplasmic	**<0.0001****	0.145		**<0.0001****
PAD4 Nuclear	**0.048***	0.336	**<0.0001****	
MHC class II in tumour	0.407	0.476	0.398	0.370
Vimentin	1.000	0.094	0.529	0.554
α-Enolase cytoplasmic	**0.001****	0.729	**<0.0001****	0.469
α-Enolase nuclear	0.088	0.670	**0.002****	**0.011***
Distant metastasis	0.687	0.658	0.059	0.143
Nuclear β-catenin	**0.020***	0.988	**0.004****	0.549
P53	0.645	**0.024***	0.922	**0.016***
Ki-67	0.368	0.474	0.224	0.765
Bcl2	0.975	0.764	**0.032***	**0.022***
VEGFA	**0.001***	0.933	0.363	0.267
CEA	**0.042**	0.243	0.522	0.091
Chemotherapy	0.840	0.500	0.846	0.599
Radiotherapy	0.938	0.351	0.424	0.123

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Table 4. Multivariate analysis results (Cox regression model) for TNM stage, Vascular Invasion, Nuclear β-catenin and cytoplasmic and nuclear PAD2 and PAD4 expression.

PAD2		Cytoplasmic PAD2			Nuclear PAD2		
VARIABLE	**CATEGORY**	**HAZARD RATIO (HR)**	**95% CI**	**P VALUE**	**HAZARD RATIO (HR)**	**95% CI**	**P VALUE**
TNM Stage	I and II III and IV	2.290	1.689 to 3.105	**<0.001***	2.651	1.590 to 4.421	**<0.001***
Vascular Invasion	Negative Positive	1.437	0.937 to 2.204	0.097	2.152	0.934 to 4.957	0.072
PAD2 (Cytoplasmic)	Low High	0.458	0.288 to 0.728	**0.001**	N/A		
PAD2 (nuclear)	Low High	N/A			1.677	0.755 to 3.722	0.204
PAD4		**Cytoplasmic PAD4**			**Nuclear PAD4**		
VARIABLE	**CATEGORY**	**HAZARD RATIO (HR)**	**95% CI**	**P VALUE**	**HAZARD RATIO (HR)**	**95% CI**	**P VALUE**
TNM Stage	I and II III and IV	2.238	1.613 to 3.104	**<0.001***	2.303	1.658 to 3.200	**<0.001***
Vascular Invasion	Negative Positive	1.493	0.953 to 2.339	0.080	1.435	0.917 to 2.247	0.114
PAD4 (Cytoplasmic)	Low High	0.629	0.398 to 0.994	**0.047***	N/A		
PAD4 (nuclear)	Low High	N/A			0.727	0.469 to 1.126	0.153

Significant p values are in bold

Figure 1. Representative immunohistochemistry staining of PAD2 (A-C) and PAD4 (D-F) in colorectal tissue microarray images. The level of expression ranged from weak (A, D), moderate (B, E) and strong expression (C, F) according to the H-score.

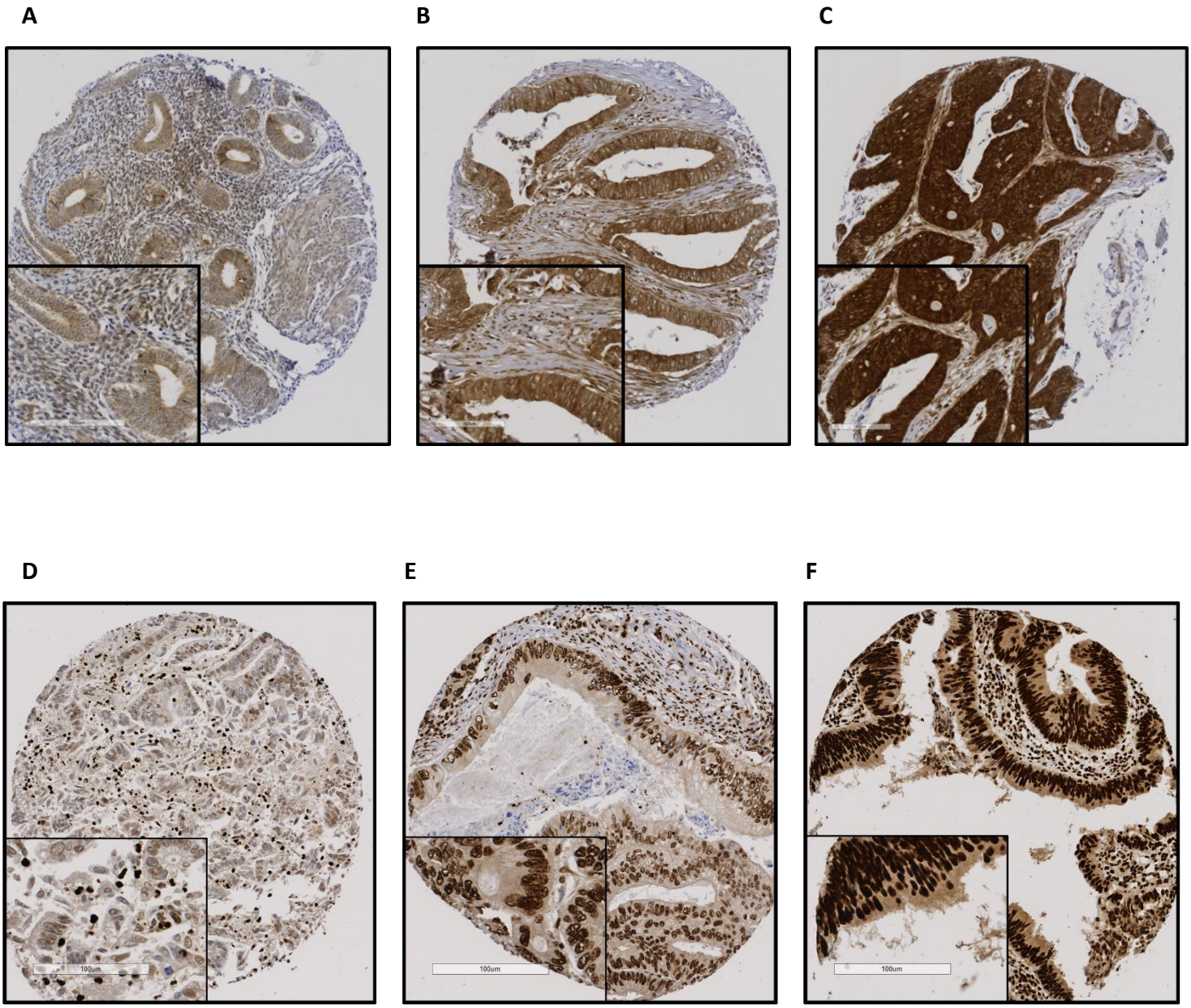
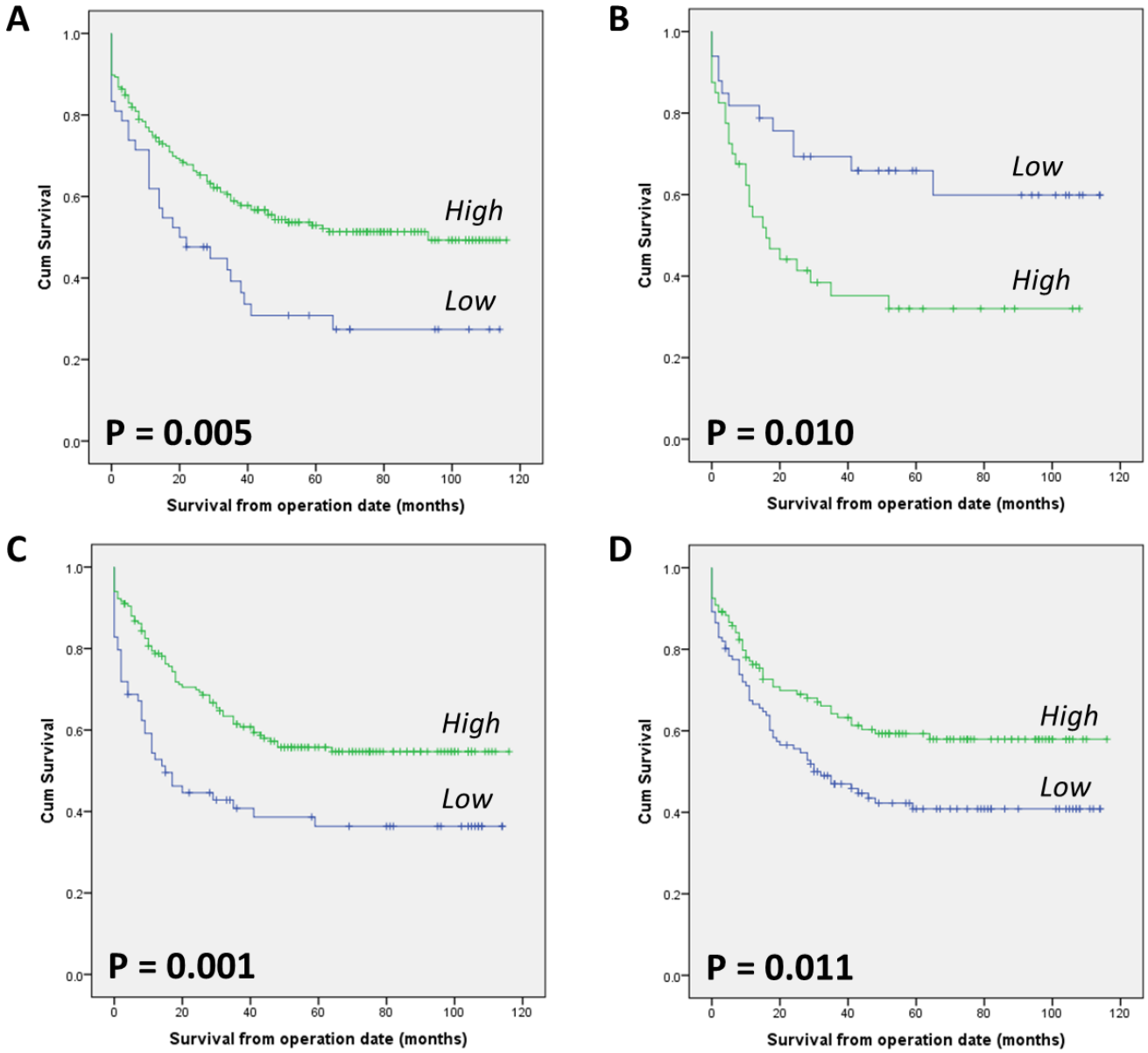


Figure 2. Kaplan-Meier curves for disease survival for CRC patients from the time of tumour resection showing high (green line) or low expression (blue line) of PAD2 or PAD4, (A) cytoplasmic PAD2 (B) nuclear PAD2 (C) cytoplasmic PAD4 (D) nuclear PAD4. Significance determined using log-rank test.



Supplementary data

Table 1 Log-rank test results for disease survival for CRC patients from the time of tumour resection.

Expression	Mean survival time (Months) in relation to cytoplasmic PAD2 expression				*p* Value
	Estimate (Months)	Std. Error	95% Confidence Interval		
			Lower Bound	Upper Bound	
Neg/Low	43.499	7.276	29.238	57.760	**0.005***
Moderate/High	67.633	3.638	60.503	74.763	
Overall	63.640	3.320	57.133	70.148	

Expression	Mean survival time (Months) in relation to nuclear PAD2 expression				*p* Value
	Estimate (Months)	Std. Error	95% Confidence Interval		
			Lower Bound	Upper Bound	
Neg/Low	76.443	8.663	59.463	93.423	**0.010***
Moderate/High	42.928	7.485	28.257	57.600	
Overall	59.220	6.140	47.186	71.255	

Expression	Mean survival time (Months) in relation to cytoplasmic PAD4 expression				*p* Value
	Estimate (Months)	Std. Error	95% Confidence Interval		
			Lower Bound	Upper Bound	
Neg/Low	48.079	6.564	35.213	60.945	**0.001***
Moderate/High	71.749	3.973	63.961	79.537	
Overall	65.322	3.486	58.489	72.155	

Expression	Mean survival time (Months) in relation to nuclear PAD2 expression				*p* Value
	Estimate (Months)	Std. Error	95% Confidence Interval		
			Lower Bound	Upper Bound	
Neg/Low	55.439	4.864	45.906	64.973	**0.011***
Moderate/High	74.042	4.758	64.717	83.368	
Overall	65.322	3.486	58.489	72.155	