Cell-free DNA promoter hypermethylation as a diagnostic marker for pancreatic ductal adenocarcinoma – An external validation study

Henriksen, Stine D.; Stubbe, Benjamin E.; Madsen, Poul H.; Johansen, Julia S.; Jensen, Benny V.; Hansen, Carsten P.; Johansen, Martin N.; Pedersen, Inge S.; Krarup, Henrik; Thorlacius-Ussing, Ole

Published in:
Pancreatology

DOI (link to publication from Publisher):
10.1016/j.pan.2021.05.003

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Pancreatology 21 (2021) 1081–1091

Contents lists available at ScienceDirect

Pancreatology

journal homepage: www.elsevier.com/locate/pan

Cell-free DNA promoter hypermethylation as a diagnostic marker for pancreatic ductal adenocarcinoma – An external validation study

Stine D. Henriksen [a, b, c, *], Benjamin E. Stubbe [a], Poul H. Madsen [d], Julia S. Johansen [e, f, g], Benny V. Jensen [e], Carsten P. Hansen [h], Martin N. Johansen [i], Inge S. Pedersen [b, c, d], Henrik Krarup [c, d], Ole Thorlacius-Ussing [a, b, c]

[a] Department of Gastrointestinal Surgery, Aalborg University Hospital, Denmark
[b] Department of Clinical Medicine, Aalborg University, Denmark
[c] Clinical Cancer Research Center, Aalborg University Hospital, Denmark
[d] Department of Molecular Diagnostics, Aalborg University Hospital, Denmark
[e] Department of Oncology, Herlev and Gentofte Hospital, Copenhagen University Hospital, Denmark
[f] Department of Medicine, Herlev and Gentofte Hospital, Copenhagen University Hospital, Denmark
[g] Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark
[h] Department of Surgery, Rigshospitalet, Copenhagen University Hospital, Denmark
[i] Unit of Clinical Biostatistics, Aalborg University Hospital, Denmark

ARTICLE INFO

Article history:
Received 5 January 2021
Received in revised form
27 April 2021
Accepted 4 May 2021
Available online 8 May 2021

Keywords:
Biomarker
Cell-free DNA
Epigenetic
Methylation
Pancreatic cancer

ABSTRACT

Background: We recently identified a diagnostic prediction model based on promoter hypermethylation of eight selected genes in plasma cell-free (cf) DNA, which showed promising results as a diagnostic biomarker for pancreatic ductal adenocarcinoma (PDAC). The aim of the present study was to validate this biomarker profile in an external patient cohort and examine any additional effect of serum CA 19-9.
Methods: Patients with PDAC (n = 346, stage I-IV) and chronic pancreatitis (n = 25) were included. Methylation-specific PCR of a 28-gene panel was performed on serum cfDNA samples. The previously developed diagnostic prediction model (age>65 years, *BMP3, RASSF1A, BNC1, MESTv2, TFPI2, APC, SFRP1* and *SFRP2*) was validated alone and in combination with serum CA 19-9 in this external patient cohort.
Results: Patients with PDAC had a higher number of hypermethylated genes (mean 8.11, 95% CI 7.70–8.52) than patients with chronic pancreatitis (mean 5.60, 95% CI 4.42–6.78, p = 0.011). Validation of the diagnostic prediction model yielded an AUC of 0.77 (95% CI 0.69–0.84). The combination of serum CA 19-9 and our test had an AUC of 0.93 (95% CI 0.89–0.96) in the primary study and 0.85 (95% CI 0.79–0.91) in the validation study.
Conclusion: In this validation study, PDAC was associated with a higher number of hypermethylated genes in serum cfDNA than chronic pancreatitis. Our diagnostic test was superior to the predictive value of serum CA 19-9 alone in both the primary and the validation study. The combination of our test with CA 19-9 may serve as a clinically useful diagnostic biomarker for PDAC.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is one of the leading causes of cancer death in the world, with a current five-year survival rate of 8% [1]. Difficulties in detecting the disease at an early stage is one of the main reasons for its poor prognosis. This clearly indicates the need for additional diagnostic tools. Currently, the only clinical biomarker for PDAC is serum CA 19-9 which is approved for prognostic surveillance. Unfortunately, due to several limitations, serum CA 19-9 is not useful as a diagnostic biomarker

Abbreviations: AUC, area under the receiver-operating characteristic curve; BIOPAC, Biomarkers in patients with Pancreatic Cancer; cf, cell free; CI, confidence interval; CpG, cytosine preceding a guanosine; Ct, cycle threshold; IPMN, intraductal papillary mucinous neoplasm; OR, odds ratio; PDAC, pancreatic ductal adenocarcinoma; ROC, receiver-operating characteristic.

* Corresponding author. Department of Gastrointestinal Surgery, Clinical Cancer Research Center, Aalborg University Hospital, Department of Clinical Medicine, Aalborg University, Hobrovej 18-22, 9000, Aalborg, Denmark.
E-mail address: stdh@rn.dk (S.D. Henriksen).

https://doi.org/10.1016/j.pan.2021.05.003

and no approved test exists for early detection of PDAC [2,3].

Epigenetic modifications such as DNA hypermethylation of the promoter sequence are known to be aspects of early cancer development [4,5]. Promoter hypermethylation in tumour suppressor genes can downregulate gene function [6–8]. A progressive increase in DNA methylation has been demonstrated with increasing dysplasia [6,8,9]. Furthermore, changes in DNA methylation have been shown in pancreatic tissue containing intraductal papillary mucinous neoplasm (IPMN) precursor lesions [10], indicating that such lesions occur in early tumour development. Such changes therefore hold potential as diagnostic biomarkers for PDAC. Epigenetic alterations can be analysed in cell-free (cf)DNA, which potentially is tumour specific [11–14].

In recent years, cfDNA has gained major interest as a tool for minimally invasive diagnostics. It is an alternative approach to cancer tissue biopsy for analysing genetic and epigenetic modifications. Several studies have shown that circulating tumour DNA fragments contain biological alterations identical to those in the primary tumour [11–16]. However, tumours are usually heterogenic containing a mixture of different cancer cell clones and normal cells [11,13,14]. A major challenge is to differentiate circulating tumour DNA from circulating non-tumour DNA [17]. This challenge is accentuated by the fact that several benign conditions, such as inflammatory disease, e. g. pancreatitis and sepsis, are also associated with an increased level of cfDNA [18] and even with changes in the methylation profile [9,19,20]. To develop a diagnostic biomarker for PDAC, it is crucial to be able to differentiate between potentially cancer-specific hypermethylation and hypermethylation that occurs as part of the inflammatory response in the pancreas. Therefore, it is important to include clinically relevant control groups, such as patients with benign pancreatic disease.

Unfortunately, previous research into DNA methylation as a diagnostic marker for PDAC has primarily differentiated between cancer patients and healthy controls [20,21]. In addition, the majority of the studies lack validation [4,19,22].

Our group has previously examined promoter hypermethylation in a panel of genes in cfDNA from patients with PDAC and in patients with benign pancreatic disease [4]. We recently published a diagnostic prediction model based on hypermethylation of eight selected genes. The prediction model yielded promising results as a diagnostic biomarker for PDAC. With high performance, the test facilitated differentiation of patients with PDAC from patients with chronic pancreatitis or symptoms mimicking upper gastrointestinal cancer [4]. However, our primary study was based on training data only and lacked external validation.

The primary aim of the present study was to conduct an external validation of our previously published diagnostic prediction model for PDAC. The secondary purpose was to examine the additional effect of serum CA 19-9 on the predictive performance of the diagnostic test.

Methods

In order to validate the findings from the primary study [4], we conducted a validation study on an external patient cohort obtained from the Danish Biomarkers in patients with Pancreatic Cancer (BIOPAC) study (NCT03311776) [23]. BIOPAC is a comprehensive Danish pancreatic cancer biobank established in 2008 through nationwide collaboration, and inclusion is still ongoing. The validation study was designed as a retrospective cross-sectional cohort study.

Participants

The validation study included patients with PDAC diagnosed from July 2008 to August 2016 in the BIOPAC study. Patients were included from two (out of seven) participating hospitals in Denmark (Herlev Hospital and Copenhagen University Hospital (Rigshospitalet)). Patients had histologically verified PDAC (resected specimen from patients with stage I-II and biopsies from patients with stage III-IV). Patients were staged according to the American Joint Committee on Cancer (AJCC), 7th edition. Patients were included before any treatment and within 2–4 weeks from the time of diagnosis. All patients were above 18 years of age and had signed an informed consent form.

A subgroup of patients in the BIOPAC study were included on suspicion of pancreatic cancer. Subsequently, pancreatic cancer was ruled out and these patients were diagnosed with chronic pancreatitis. This subgroup of patients was included as a benign control group of patients diagnosed with chronic pancreatitis.

All clinical data were registered in the BIOPAC database.

The BIOPAC study was approved by the Regional Ethics Committee (VEK ref. KA-20060113) and the Danish Data Protection Agency (r.no. 2012-58-0004; HGH-2015-027; I-Suite j.no. 03960).

The present validation study was approved by the Research Ethics Committee for the North Denmark Region (N-20130037).

Blood sampling

Standard operating procedures were used for handling of blood samples in the BIOPAC study. Within 30–120 min after sampling, the blood was centrifuged at 2330 g for 10 min at 4 °C and serum was then aliquoted and stored at −80 °C until analysis.

Methylation analysis

All serum samples were analysed at the Department of Molecular Diagnostics, Aalborg University Hospital, Denmark.

The methylation analyses were performed blinded without knowledge of clinical data by a senior laboratory scientist. Extraction and deamination of serum cfDNA was performed as described previously [4,24].

In brief, a first-round PCR amplification was conducted to expand the amount of relevant deaminated DNA. A mix of outer methylation-specific primers was used for all the tested promoter regions. Subsequently, in individual reactions, a second round of PCR was performed using each of the inner methylation-specific primers and methylation-specific probes. In both the first and second round of PCR, the hemi-methylated *MEST transcript variant 1* was used as the reference gene [4].

A panel of 28 selected genes (the promoter sequences) (Supplementary Table 1) was tested. We have previously described in detail the selection of the 28 genes [4].

CA 19-9

CA 19-9 analysis was performed on serum samples from all patients in the validation study at the time of inclusion in the BIOPAC study. Serum CA 19-9 was analysed at Herlev Hospital using the Immulite 2000 GI-MA assay (Siemens, Catalogue Number L2KG12), a solid-phase, 2-site sequential chemiluminescent immunometric assay. Elevated serum CA 19-9 is defined as > 37 U/ml.

In the context of the primary study [4], CA 19-9 was not available as this test was not implemented at our department at the time of patient inclusion. Subsequently, it has become possible to analyse serum CA-19-9 for all patients included in the primary study [4]. Serum CA 19-9 analysis was performed using Cobas 6000 (Roche, module 601E, IUPAC code NPU01450).

Outcome

The primary outcome of this study was PDAC and the performance of the diagnostic prediction model in the external patient cohort. The secondary outcome was the potential effect of serum CA 19-9.

Statistical methods

Each gene in the panel was analysed following dichotomization. A detectable cycle threshold (Ct) was interpreted as a gene with promoter hypermethylation and no Ct as a non-methylated gene promoter. Validation of dichotomous data has previously been described [4].

The methylation frequency of each gene and the (exact) 95% confidence interval (CI) were calculated for each patient group. The mean number of hypermethylated genes and the 95% CI were calculated. The means were compared as numerical data using the nonparametric Wilcoxon rank sum test. P-values below 0.05 were considered statistically significant.

Logistic regression was performed separately for all gene variables. The p value and the area under the receiver-operating characteristic curve (AUC) were calculated to measure the discriminative ability of the variables. For variables with complete separation between the patient groups, the chi squared test served as significance test.

Furthermore, simple logistic regression was performed for dichotomized serum CA 19-9 (cut-off 37 kU/l), age > 65 years, gender and smoking status.

In our primary study, a prediction model for PDAC was developed (*BMP3, RASSF1A, BNC1, MESTv2, TFPI2, APC, SFRP1, SFRP2* and age > 65 years) [4]. The diagnostic prediction model from the primary study (both the selected variables and their coefficients) was validated on the total cohort, on the subgroup of patients with stage I-II disease and on the subgroup of patients with stage IV disease. A probability score was calculated for each patient.

Subsequently, the additional effect of serum CA 19-9 on the diagnostic prediction model (*BMP3, RASSF1A, BNC1, MESTv2, TFPI2, APC, SFRP1, SFRP2* and age > 65 years) was tested in the total patient cohort from the primary study (See supplementary text with description of the patient cohort in the primary study) [4]. A new prediction model was developed based on the same gene variables; however, the new model contained new coefficients due to the added effect of serum CA 19-9. A probability score was calculated for each patient. The model performance was evaluated in the total patient cohort, in the subgroup of patients with stage I-II disease and in the subgroup of patients with stage IV disease.

The new diagnostic prediction model combining the predictive effect of serum CA 19-9 and the hypermethylation status of serum cfDNA *BMP3, RASSF1A, BNC1, MESTv2, TFPI2, APC, SFRP1, SFRP2* and age >65 years was validated on the total validation cohort, on the subgroup of patients with stage I -II disease and on the subgroup of patients with stage IV disease from the validation cohort. For each patient, a probability score was calculated.

For all the prediction models mentioned above, the AUC was calculated to measure the discriminative ability of the model; a ROC curve was used to illustrate model performance and a calibration plot was conducted to illustrate model calibration.

Stata 16.0 software (Stata-Corp LP, TX) was used for data analysis. The Stata module pmcalplot was used to produce the calibration plot.

All authors had access to the study data and reviewed and approved the final manuscript.

Results

Descriptive data

In total, 346 patients with PDAC (stage I; n = 11, stage II; n = 165, stage III; n = 33 and stage IV; n = 137) and 25 patients with chronic pancreatitis were included in the validation study. Descriptive data are shown in Tables 1 and 2.

The mean number of hypermethylated genes

The mean number of hypermethylated genes of the 28-gene panel in serum cfDNA was calculated. Patients with PDAC had a mean 8.11 (95% CI 7.70–8.52) hypermethylated genes compared with 5.60 (95% CI 4.42–6.78) in the control group of patients with chronic pancreatitis (p = 0.011).

The hypermethylation frequency and simple logistic regression

The frequency of hypermethylation for each of the 28 genes in serum cfDNA for each patient group is shown in Table 3 and the odds ratio (OR) for each variable in the gene panel is shown in Table 4. A significant difference between the cancer group and the control group was demonstrated for three gene variables; *APC, ESR1* and *TAC1* with an OR of 3.58 (95% CI 1.50–8.56), 4.22 (95% CI 1.72–10.38) and 2.63 (95% CI 1.16–5.96), respectively (Table 4). In addition, a significant difference was demonstrated for *EYA2* hypermethylation (OR of 0.39 (95% CI 0.17–0.87)); however, with the control group being more frequently hypermethylated than the cancer group. The promoter sequence of *SFRP2, TFPI2, CHFR, GSTP1, HIC1, SEPT9v2, VIM and CDKN2A* was hypermethylated only in the cancer group. Due to complete separation, these variables could not be analysed by logistic regression. Instead, the chi-squared test was used and a significant difference in hypermethylation frequency of *SFRP2* was found, whereas the other completely separated variables did not reach statistical significance (Table 3).

By simple logistic regression, CA 19-9 (cut-off 37 kU/l) had an OR of 5.30 (95% CI 2.23–12.60) and reached an AUC of 0.68. Eight patients in the validation study lacked the serum CA 19-9 result; six cancer patients (three stage II, one stage III and two stage IV) and two patients with chronic pancreatitis. Patients with missing CA 19-9 values were excluded in the validation analysis concerning the combination of the prediction model and CA 19-9, however included in the validation analysis solely concerning the prediction model.

In the univariate analysis, male gender was a significant risk factor (OR 2.69 (95% CI 1.05–6.91)), age > 65 years reached an OR of 4.39 (95% CI 1.71–11.26) and smoking reached an OR of 0.56 (95% CI 0.33–0.95).

Table 1
Descriptive data on patients with pancreatic ductal adenocarcinoma (PDAC) and chronic pancreatitis.

		PDAC		Chronic pancreatitis	
N		346		25	
Median age (range)		67	(37–86)	61	(33–85)
Sex (% men)		187	(54)	19	(76)
Smoking status	current (%)	105	(30)	15	(60)
	previous (%)	118	(34)	3	(12)
	never (%)	123	(36)	7	(28)
AJCC/UICC staging	I (IA and IB) (%)	11	(3)		
	II (IIA and IIB) (%)	165	(48)		
	III (%)	33	(9)		
	IV (%)	137	(40)		

Table 2
Descriptive data on the patients with pancreatic ductal adenocarcinoma according to cancer stage.

	Stage	I (Ia+Ib)		II (IIa+IIb)		III		IV		*Chronic pancreatitis*	
	N	11		165		33		137		25	
	Age (median) (range)	68	(53–84)	67	(37–81)	67	(42–86)	68	(40–84)	61	(33–85)
	Sex (men:women)	4:7		93:72		16:17		74:63		19:6	
PS (number, %)	0	4	36%	61	37%	14	42%	56	41%	0	–
	1	3	27%	41	25%	9	27%	58	42%	1	4%
	2	0	–	9	5%	3	9%	9	7%	0	–
	3	0	–	2	1%	0	–	0	–	0	–
	unknown	4	36%	52	32%	7	21%	14	10%	24	96%
Surgical treatment number (%)	None	0	–	5	3%	14	42%	71	52%	11	44%
	Whipple	7	64%	88	53%	0	–	2	1%	7	28%
	Distal resection	2	18%	16	10%	0	–	0	–	4	16%
	Total resection	1	9%	48	29%	0	–	1	1%	1	4%
	Explorative lap	1	9%	4	3%	19	58%	60	44%	1	4%
	Unknown	0	–	3	2%	0	–	3	2%	1	4%

PS, WHO performance status
Note: Stage according to the American Joint Committee on Cancer (AJCC) stage classification

Table 3
Hypermethylation frequencies for each gene in each group.

	Pancreatic ductal adenocarcinoma (N = 346)			Chronic pancreatitis (N = 25)		
Gene	%	n	95% CI	%	n	95% CI
ALX4	14.2	49	(10.9–18.3)	12.0	3	(3.7–32.7)
APC	86.4	299	(82.4–89.6)	64.0	16	(42.9–80.8)
BMP3	45.7	158	(40.5–51.0)	32.0	8	(16.3–53.3)
BNC1	28.6	99	(24.1–33.6)	24.0	6	(10.7–45.4)
BRCA1	30.4	105	(25.7–35.4)	40.0	10	(22.3–60.8)
CDKN2A	10.1	35	(7.30–13.8)	0	0	–
CDKN2B	8.1	28	(5.6–11.5)	4.0	1	(0.5–25.5)
CHFR	5.8	20	(3.8–8.8)	0	0	–
ESR1	62.1	215	(56.9–67.1)	28.0	7	(13.4–49.4)
EYA2	33.0	114	(28.2–38.1)	56.0	14	(35.7–74.5)
GSTP1	1.2	4	(0.4–3.0)	0	0	–
HIC1	4.3	15	(2.6–7.1)	0	0	–
MESTv2	12.4	43	(9.3–16.4)	12.0	3	(3.7–32.7)
MGMT	4.9	17	(3.1–7.8)	8.0	2	(1.9–28.5)
MLH1	42.5	147	(37.4–47.8)	32.0	8	(16.3–53.3)
NPTX2	75.7	262	(70.9–80.0)	60.0	15	(39.2–77.7)
NEUROG1	44.5	154	(39.3–49.8)	32.0	8	(16.3–53.3)
RARB	60.7	210	(55.4–65.7)	60.0	15	(39.2–77.7)
RASSF1A	35.0	121	(30.1–40.2)	4.0	1	(0.5–25.5)
SFRP1	26.3	91	(21.9–31.2)	12.0	3	(3.7–32.7)
SFRP2	28.9	100	(24.4–33.9)	0	0	–
SEPT9v2	8.7	30	(6.1–12.1)	0	0	–
SST	39.3	136	(34.3–44.6)	20.0	5	(8.2–41.2)
TFPI2	13.0	45	(9.8–17.0)	0	0	–
TAC1	70.8	245	(65.8–75.4)	48.0	12	(28.8–67.8)
VIM	3.8	13	(2.2–6.4)	0	0	–
WNT5A	6.9	24	(4.7–10.2)	4.0	1	(0.5–25.5)
PENK	7.8	27	(5.4–11.2)	8.0	2	(1.9–28.5)

Table 4
Simple logistic regression.

Gene	OR	95% CI	P-value	AUC
ALX4	1.21	(0.35–4.20)	0.764	0.511
APC	**3.58**	**(1.50-8.56)**	**0.004**	**0.612**
BMP3	1.79	(0.75–4.25)	0.190	0.568
BNC1	1.27	(0.49–3.27)	0.622	0.523
BRCA1	0.65	(0.28–1.50)	0.317	0.548
*CDKN2A**	–	–	–	–
CDKN2B	2.11	(0.28–16.21)	0.472	0.520
*CHFR**	–	–	–	–
ESR1	**4.22**	**(1.72-10.38)**	**0.002**	**0.671**
*EYA2***	0.39	(0.17-.88)	0.023	0.615
*GSTP1**	–	–	–	–
*HIC1**	–	–	–	–
MESTv2	1.04	(0.30–3.63)	0.950	0.502
MGMT	0.59	(0.13–2.73)	0.504	0.515
MLH1	1.57	(0.66–3.74)	0.308	0.552
NPTX2	2.08	(0.90–4.80)	0.087	0.579
NEUROG1	1.70	(0.72–4.06)	0.228	0.563
RARB	1.03	(0.45–2.36)	0.945	0.503
RASSF1A	12.91	(1.73–96.57)	**0.013**	0.655
SFRP1	2.62	(0.77–8.95)	0.125	0.572
SFRP2*	**-**	**-**	**-**	**-**
*SEPT9v2**	–	–	0.269	0.543
SST	2.59	(0.95–7.07)	0.063	0.597
*TFPI2**	–	–	–	–
TAC1	**2.63**	**(1.16-5.96)**	**0.021**	**0.614**
*VIM**	–	–	–	–
WNT5A	1.79	(0.23–13.80)	0.577	0.515
PENK	0.97	(0.22–4.35)	0.972	0.501
sex	**2.69**	**(1.05-6.91)**	**0.039**	**0.610**
age 65	**4.39**	**(1.71-11.26)**	**0.002**	**0.670**

All variables were analysed by simple logistic regression comparing the pancreatic ductal adenocarcinoma group and the chronic pancreatitis group. Bold marks the genes where there was significant difference ($p < 0.05$) between the cancer group and the control group.
**CDKN2A, SFRP2, TFPI2, CHFR, GSTP1, HIC1, SEPT9v2* and *VIM* were only hypermethylated in the cancer group. Due to complete separation of the variables, logistic regression could not be performed. Instead, the chi-squared test was used and a significant difference in hypermethylation of *SFRP2* was established between the groups. The other completely separated variables did not reach statistical significance.
***A significant difference was also demonstrated for *EYA2;* however, with the control group being more frequently hypermethylated than the cancer group
OR, odds ratio; CI, confidential interval; AUC, area under the receiver-operating characteristic curve.

Validation of the diagnostic prediction model

The previously developed diagnostic prediction model contained the covariate age >65 years and the hypermethylation status of the promoter sequence of eight gene variables; *BMP3, RASSF1A, BNC1, MESTv2, TFPI2, APC, SFRP1,* and *SFRP2.* [4].

Validation of the diagnostic prediction model in the external patient cohort yielded an AUC of 0.77 (95% CI 0.69–0.84) (Fig. 1a). The mean probability score in the cancer group was 0.54 (95% CI 0.50–0.57) compared with 0.24 (95% CI 0.16–0.32) in the control group (Fig. 1b).

Testing the model in the subgroup of patients with early-stage tumours (stage I and II; n = 201) yielded an AUC of 0.73 (95% CI 0.64–0.82) (Fig. 1c), with a mean probability score of 0.48 (95% CI 0.43–0.52) (Fig. 1d).

Testing the model in the subgroup of patients with stage IV disease (n = 132) yielded an AUC of 0.82 (95% CI 0.75–0.89)

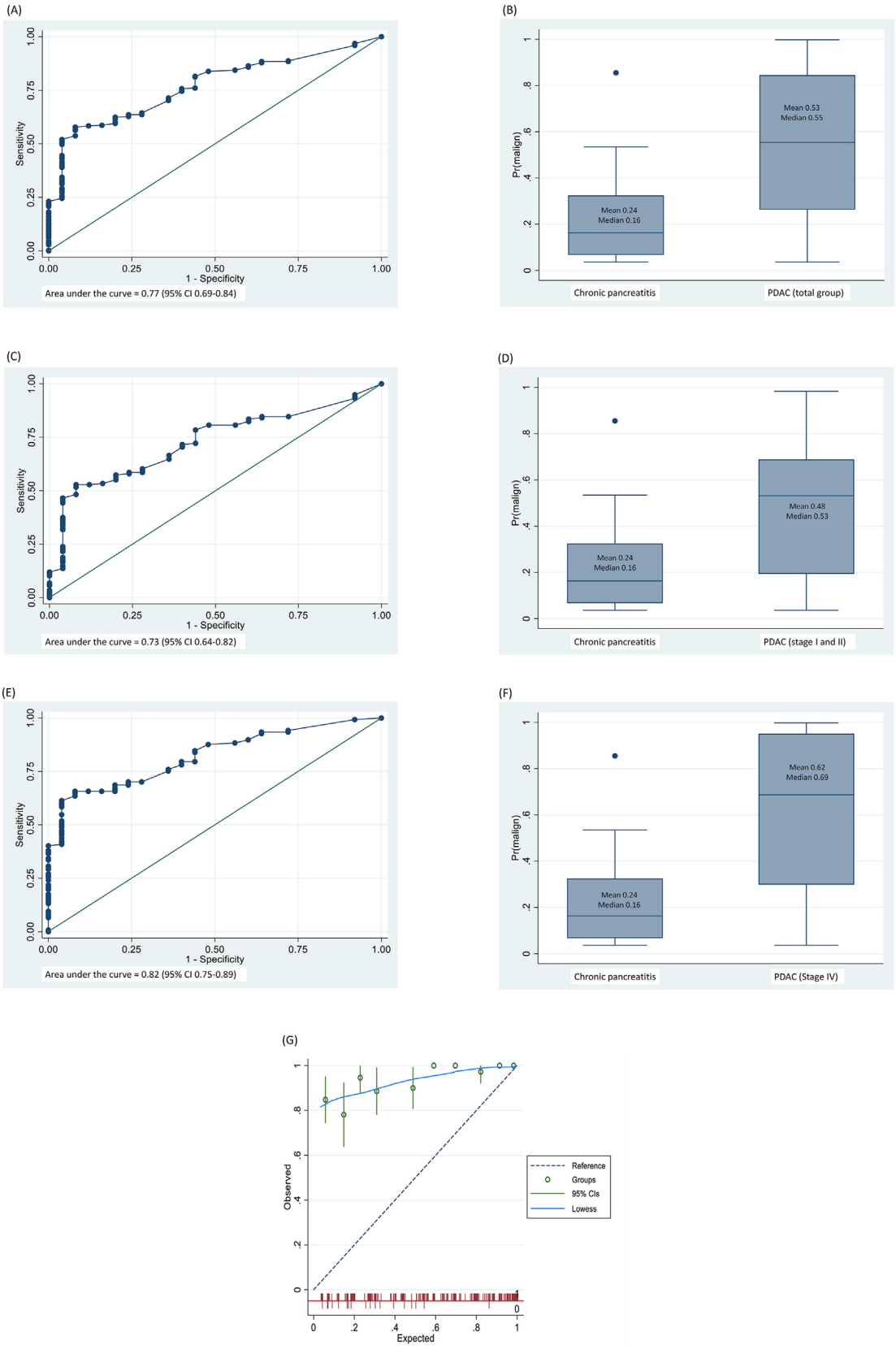

Fig. 1. Validation of the diagnostic prediction model (*BMP3, RASSF1A, BNC1, MESTv2, TFPI2, APC, SFRP1, SFRP2* and age > 65 years) in the validation cohort. 1a) Performance of the diagnostic prediction model in the total validation cohort; 1b) Box plot of the distribution of probability sore in each patient group in the total validation cohort; 1c) Performance of the diagnostic prediction model in patients with stage I and II PDAC; 1d) Box plot of the distribution of probability sore in patients with stage I and II PDAC and chronic pancreatitis; 1e) Performance of the diagnostic prediction model in patients with stage IV PDAC; 1f) Box plot of the distribution of probability sore in patients with stage IV PDAC and chronic pancreatitis; and 1 g) Calibration plot.

(Fig. 1e), with a mean probability score of 0.62 (95% CI 0.57-0-68) (Fig. 1f).

Fig. 1g presents the calibration plot, which illustrates insufficient model calibration with the observed values being higher than the expected values.

Prediction model performance and the combined effect of serum CA 19-9

In our primary study, dichotomized serum CA 19-9 had an OR of 14.79 (95% CI 7.60–28.78) by simple logistic regression and reached an AUC of 0.80. The new prediction model (including the same variables with new coefficients): age > 65 years, *BMP3, RASSF1A, BNC1, MESTv2, TFPI2, APC, SFRP1, SFRP2* and serum CA 19-9 yielded an AUC of 0.93 (95% CI 0.89–0.96) in the primary study (Fig. 2a). The mean probability score was 0.76 (95% CI 0.71–0.81) in the cancer group versus 0.18 (95% CI 0.14–0.22) in the control group (Fig. 2b).

In the subgroup of PDAC patients with stage I and II disease, serum CA 19-9 had an OR of 6.85 (95% CI 3.12–15.02) with a corresponding AUC of 0.72. The combination of serum CA 19-9 and the prediction model reached an AUC of 0.89 (95% CI 0.83–0.95) for early-stage PDAC (Fig. 2c), with a mean probability score of 0.68 (95% CI 0.59–0.78) in the cancer group (Fig. 2d).

In the subgroup of PDAC patients with stage IV disease, serum CA 19-9 had an OR of 36.87 (95% CI 11.77–115.50) with a corresponding AUC of 0.85. The combined predictive performance of serum CA 19-9 and the prediction model in stage IV patients reached an AUC of 0.95 (95% CI 0.92–0.98) (Fig. 2e), with a mean probability score of 0.82 (95% CI 0.74–0.89) (Fig. 2f).

Validation of the combined predictive performance in the total validation cohort reached an AUC of 0.85 (95% CI 0.79–0.91) (Fig. 3a). The mean probability score was 0.65 (95% CI 0.62–0.68) in the cancer group versus 0.22 (95% CI 0.14–0.31) in the control group (Fig. 3b).

Testing the combined model in patients with stage I and II tumours from the validation cohort showed an AUC of 0.82 (0.75–0.89) (Fig. 3c) with a mean probability score of 0.59 (95% CI 0.55–0.64) in early-stage cancer patients (Fig. 3d).

In the subgroup of PDAC patients with stage IV disease from the validation cohort, the combined predictive performance yielded an AUC of 0.90 (95% CI 0.84–0.95) (Fig. 3e) for stage IV PDAC with a mean probability of 0.73 (85% CI 0.68–0.78) (Fig. 3f).

A calibration plot was produced to assess the performance of the prediction model combined with serum CA 19-9 (Fig. 3g). Similar to the previous calibration plot, the observed values were higher than the expected values.

Discussion

PDAC is a highly aggressive disease resistant to most oncological therapies and usually diagnosed at an advanced stage, resulting in a poor prognosis. Diagnosing PDAC is challenged by lack of symptoms in the early disease stages. Even if symptoms are present, they are likely to be unspecific such as abdominal pain or discomfort, weight loss, fatigue and jaundice [25–27]. Such symptoms are also seen in chronic pancreatitis, an essential differential diagnosis and risk factor for PDAC [28]. The diagnostic challenge resulting in a dismal prognosis and a high mortality rate stresses the need for reliable, robust, non-invasive, early detection methods for PDAC.

Recently, our group examined promoter hypermethylation in plasma cfDNA from patients with pancreatic disease, both malignant and benign [4]. We demonstrated that plasma cfDNA hypermethylation was detectable in both malignant and benign pancreatic disease. Additionally, we showed that patients with PDAC had a significantly higher number of hypermethylated genes than a benign control group. We developed a diagnostic prediction model containing the hypermethylation status of eight promoter sequences in plasma cfDNA which showed promising results. Demonstrating high performance, the model differentiated patients with PDAC from a relevant control group [4]. However, showing that a diagnostic prediction model successfully predicts outcome in the initial data is not tantamount to demonstrating its diagnostic value. Evidence is needed that the model performs well for a similar patient group in a different cancer centre [29].

In the present study, we present the results from an external validation of our diagnostic test for PDAC. We confirmed that patients with PDAC have a significantly higher number of hypermethylated genes in serum cfDNA than patients with chronic pancreatitis. In the validation study, cancer patients had a mean of 8.11 (95% CI 7.70–8.52) hypermethylated genes in plasma cfDNA, which was similar to the findings in cancer patients in the primary study with a mean of 8.41 (95% CI 7.62–9.20) hypermethylated genes [4]. Likewise, patients in the control group of the validation study had a mean of 5.60 (95% CI 4.42–6.78) hypermethylated genes compared with 4.46 (95% CI 4.04–4.88) in the primary study [4].

Several genes in serum cfDNA were more frequently hypermethylated in the validation cancer group than in the control group. *APC, ESR1, TAC1 and SFRP2* reached statistical significance, which was also the case in our primary study [4]. However, in the primary study, a significant difference in hypermethylation frequency was also demonstrated for several other genes; a finding not reproduced in the validation study.

In the primary study, age >65 years had an OR of 4.14 (95% CI 2.33–7.33) by simple logistic regression. As epigenetic change is a natural part of ageing [30], patient age > 65 years was included as a covariate in the multivariable logistic regression analysis [4]. A similar OR (4.39, 95% CI 1.71–11.26) for age > 65 years was found in the validation study.

Similar to the primary study, smoking was a preventive factor for PDAC in the validation study when comparing PDAC patients and patients with chronic pancreatitis. Smoking has the potential to influence DNA methylation [31]. To address this problem in the primary study, smoking was excluded from the diagnostic prediction model [4], as smoking is a known risk factor for cancer development.

The validation of the diagnostic prediction model reached an AUC of 0.77 (95% CI 0.69–0.84) compared with an AUC of 0.86 (95% CI 0.81–0.91) in the primary study [4]. The difference in performance between the two studies is most likely partly due to the fact that our first study was based on training data only, which is known to produce an overestimation of test performance due to overfitting [29,32]. To account for optimism in the model development, internal validation using a bootstrapping procedure was performed [32], which estimated an optimism in the AUC of 0.03 [4]. This indicates that the amount of optimism cannot entirely explain the total difference in performance between the primary study and the validation study. The composition of the patient groups may also play a role. The group of PDAC patients in the two studies was very similar; however, the distribution of cancer stages within the studies was slightly different. In the validation study, 50% of the patients had stage I and II disease versus 42% of the patients in the primary study [4]. Unfortunately, early-stage cancer is likely to be more challenging to diagnose as epigenetic changes accumulate with progression of carcinogenesis [9]. In addition, the difference in the control groups might affect the predictive power of the diagnostic test. In the primary study, the control group encompassed a combination of patients with chronic pancreatitis included from the outpatient clinic (n = 97) and patients referred to the hospital

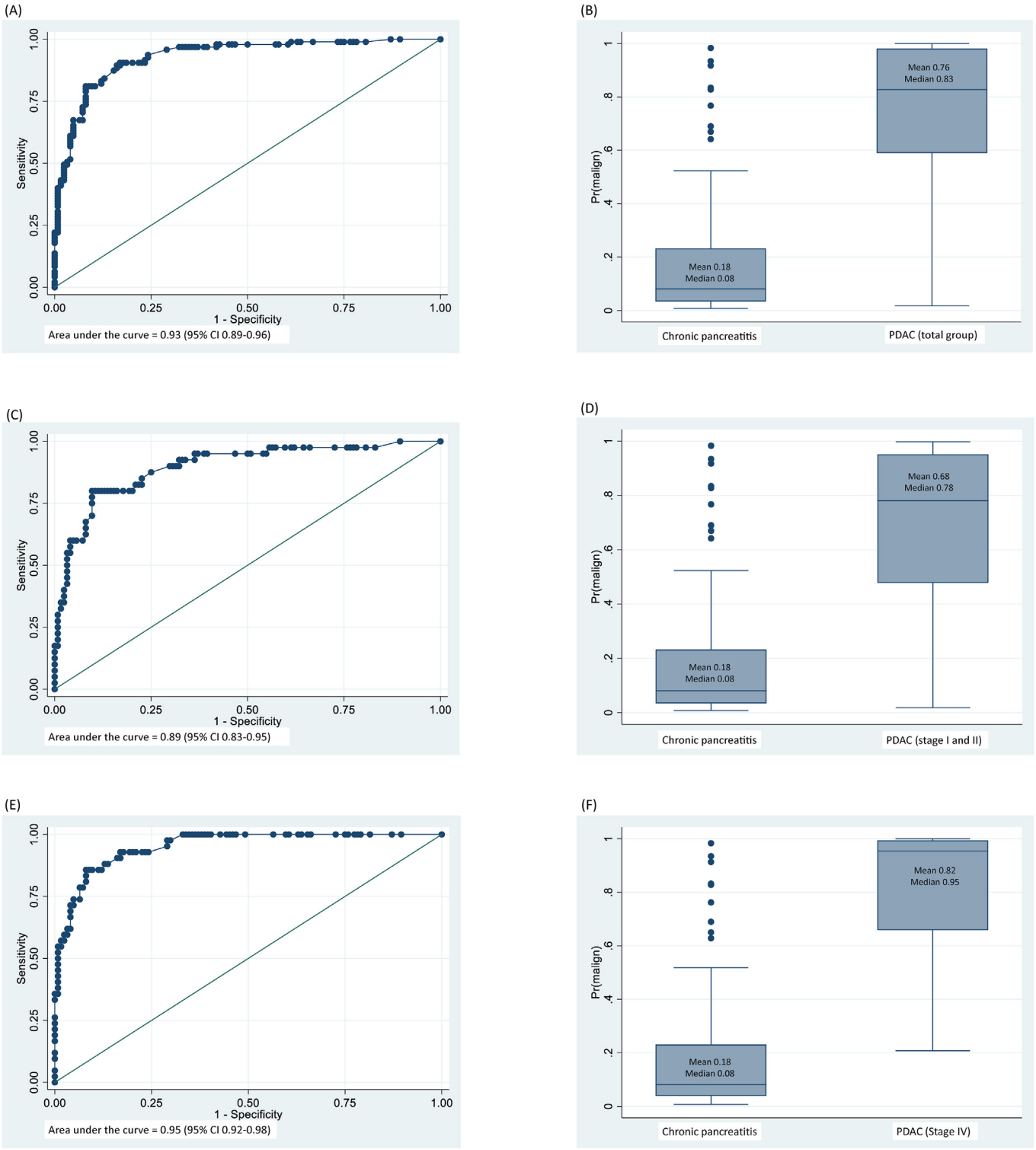

Fig. 2. Combined predictive effect of serum CA 19-9 and the diagnostic prediction model (*BMP3, RASSF1A, BNC1, MESTv2, TFPI2, APC, SFRP1, SFRP2* and age > 65 years) in the primary study.
2a) Performance of the diagnostic prediction model combined with CA 19-9 in the total primary patient cohort; 2b) Box plot of the distribution of probability score in each patient group in the total primary patient cohort; 2c) Performance of the diagnostic prediction model combined with serum CA 19-9 in patients with stage I and II PDAC from the primary study; 2d) Box plot of the distribution of probability score in patients with stage I and II PDAC and chronic pancreatitis from the primary study; 2e) Performance of the diagnostic prediction model combined with serum CA 19-9 in patients with stage IV PDAC from the primary study; and 2f) Box plot of the distribution of probability score in patients with stage IV PDAC and chronic pancreatitis from the primary study.

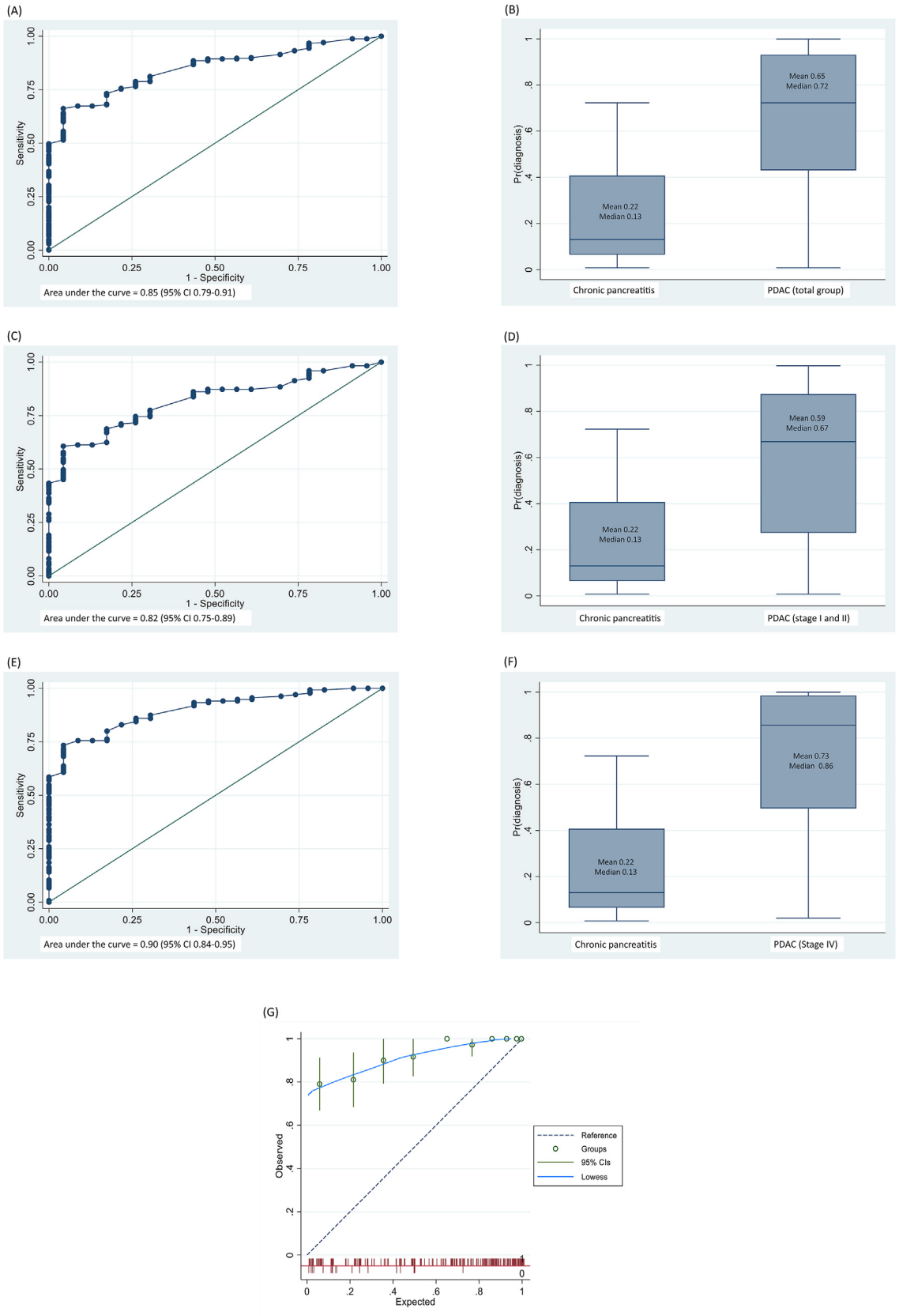

Fig. 3. Combined predictive effect of serum CA 19-9 and the diagnostic prediction model (*BMP3, RASSF1A, BNC1, MESTv2, TFPI2, APC, SFRP1, SFRP2* and age > 65 years) in the validation study. 3a) Performance of the diagnostic prediction model combined with CA 19-9 in the total validation patient cohort; 3b) Box plot of the distribution of probability score in each patient group in the total validation patient cohort; 3c) Performance of the diagnostic prediction model combined with serum CA 19-9 in patients with stage I and II PDAC and chronic pancreatitis in the validation study; 3d) Box plot of the distribution of probability score in patients with stage I and II PDAC and chronic pancreatitis from the

on suspicion of upper gastrointestinal cancer, but without any evidence of cancer (n = 27). In the latter group, 15 out of 27 patients were diagnosed with chronic pancreatitis. Only a very small proportion of the control patients included in the primary study had undergone pancreatic resection on suspicion of pancreatic cancer [4]. In the validation study, all patients in the control group were included due to suspicion of pancreatic cancer and they were subsequently diagnosed with chronic pancreatitis. Approximately half of the patients were diagnosed after surgical resection of the pancreas. This indicates that half of the patients in the control group of the validation study formed a particular subgroup of patients with chronic pancreatitis of high clinical relevance, as it was impossible preoperatively to rule out pancreatic cancer with extant diagnostic tools.

The predictive power of our test decreased in the validation study. However, the result remains promising when keeping in mind that this prediction model was developed to differentiate patients with PDAC from patients with chronic pancreatitis, which is known to be a very challenging task.

Previous studies evaluating cfDNA hypermethylation in patients with PDAC have shown a significant difference between cancer patients and healthy controls [19–21]. Unfortunately, the clinically relevant aspect in differentiating between malignant and benign disease has been associated with difficulties. Eissa et al. published a validation study of promoter methylation of *ADAMTS1* and *BNC1* as potential blood-based biomarkers for early detection of PDAC [20]. In the primary study, 42 PDAC patients and 26 healthy volunteers were included. *ADAMTS1* reached 48% sensitivity and 92% specificity, and *BNC1* reached 79% sensitivity and 89% specificity [21]. In the validation study, 39 patients with PDAC and 95 age-matched controls were included, and the combination of *ADAMTS1* and *BNC1* showed a convincing AUC of 0.95 (0.91–0.98) [20]. The panel was also analysed in eight patients with chronic pancreatitis. Unfortunately, seven of the eight patients were hypermethylated in either of the two genes [20]. Once again, this demonstrates the challenges in differentiating between malignant and benign pancreatic disease. Consequently, the authors concluded that their biomarker would be more useful in cases where chronic pancreatitis can be excluded [20]. This conclusion makes one doubt the clinical usefulness of their biomarker as chronic pancreatitis is an important risk factor for PDAC and differentiation of patients with PDAC from patients with chronic pancreatitis is a known clinical challenge.

Besides validating our diagnostic test, we evaluated serum CA 19-9 in both the primary study and the validation study. In both studies, our diagnostic prediction model was superior to the predictive value of CA 19-9. In addition, we tested the combined predictive effect of serum CA 19-9 and our diagnostic prediction model. We found an additive effect of using CA 19-9 on the predictive performance in both studies, with an AUC of 0.93 (95% CI 0.89–0.96) in the primary study and an AUC of 0.85 (95% CI 0.79–0.91) in the validation study. The combined predictive effect of CA 19-9 and plasma cfDNA hypermethylation is a promising result in the search for a clinically useful diagnostic biomarker for PDAC.

The most clinically relevant aspect of a diagnostic biomarker for pancreatic cancer is the ability to diagnose potentially resectable disease (stage I and II) as these patients can be offered curative treatment. We assessed the performance of our diagnostic test alone and in combination with serum CA 19-9 in the subgroup of PDAC patients with early-stage disease. In general, the performance declined for patients with early-stage PDAC compared with the total patient cohort. However, the performance of our test remained high even in early-stage PDAC and it was superior to that of serum CA 19-9.

Furthermore, we assessed the performance of our diagnostic test alone and in combination with serum CA 19-9 in the subgroup of PDAC patients with stage IV disease. As expected, the diagnostic performance increased significantly in patients with stage IV disease compared to patients with early stage disease, which is in concordance with current literature illustrating accumulation of epigenetic changes with aggravating cancer stage [9].

Our study has some limitations. The control group of patients with chronic pancreatitis only included 25 patients. A larger control group would have been favourable as it would have improved the certainty of the results. The fact that the ratios of cases and controls in the primary study and the validation study are not equal complicates model calibration. This might result in the observed values being much higher than the expected values as the proportion of cases was much larger in the validation study than in the primary study.

Eight patients (six PDAC patients and two controls) lacked CA 19-9 analysis and was excluded from the analysis of the prediction model performance and the combined effect of serum CA 19-9. This is a limitation and reduces the power of the results.

The cases and controls were not matched according to age, which can be a potential disadvanges because epigenetic changes are a part of ageing. This problem was addressed in the primary study [4], as age was incorporated as a covariate in the diagnostic prediction model to avoid variable selection to be driven by possible differences in general methylation status between patients of different ages.

In addition, the cases and controls were not matched according to smoking status, which also could be a disadvanges, as smoking has the potential to affect methylation status [31]. However, a previous study on pancreatic cancer did not find an association between DNA methylation of ppENK and p16 and smoking [33]. In this validation study a larger proportion of controls were current smokers compared to cases. Theoretically, this should result in the methylation status of the controls being more pronounced due to smoking compared to the methylation status of the cases, which potentially would lead to decreased discriminative ability of the diagnostic test.

Regarding the methylation analysis, cfDNA was extracted from plasma in the primary study, whereas serum samples were used in the validation study. It is important to be aware of this as there may be a slight difference in the cfDNA concentration of plasma and serum [34,35].

Furthermore, our study has several strengths. We developed a minimally invasive blood-based diagnostic test for PDAC. Unlike tissue-based markers, blood-based markers involve less risk of iatrogenic harm as they are minimally invasive. They are particularly useful in pancreatic disease, as tissue biopsies may be difficult to obtain. We performed validation of our previous findings on a large external cohort of patients with PDAC that was almost identical to the primary patient group and representative of patients with PDAC in general in Denmark. Similar to the primary study, we used a clinical relevant control group of patients who were suspected of cancer, but who were eventually diagnosed with chronic pancreatitis.

We were able to evaluate our biomarker in combination with serum CA 19-9 in both the primary and the validation study. All

validation study; 3e) Performance of the diagnostic prediction model combined with serum CA 19-9 in patients with stage IV PDAC and chronic pancreatitis in the validation study; 3f) Box plot of the distribution of probability score in patients with stage IV PDAC and chronic pancreatitis from the validation study; and 3 g) Calibration plot.

samples in the validation study were analysed at Aalborg University Hospital, Denmark, by the same laboratory scientist who performed the methylation analysis in the primary study. The analyses were performed blinded and conducted in the exact same manner in both studies.

The performance of our test in the validation study illustrates that we have not reached the goal of developing a stand-alone diagnostic test for PDAC based on plasma/serum cfDNA hypermethylation. However, the high performance demonstrated by combining serum CA 19-9 and our diagnostic test was encouraging and indicates that a multi-target approach is needed to achieve a high predictive power.

Conclusion

We previously published a diagnostic test for PDAC [4]. The PDAC test has now been validated in a large external patient cohort. The results are promising as our test of promoter hypermethylation of eight selected genes in cfDNA in combination with serum CA-19-9 facilitates differentiation between patients with PDAC and patients with chronic pancreatitis with a high performance demonstrated by an AUC of 0.93 and 0.85 in the primary study and validation study, respectively. This result is convincing and brings us one step closer to a clinically useful diagnostic biomarker for PDAC, enabling the important differentiation of malignant from benign pancreatic disease. The promising result of this study must lead to further research into the development a clinical useful diagnostic test based on the combination of cfDNA hypermethylation and CA 19-9. Future research should include longitudinal retrospective studies and well-designed prospective studies [36,37].

Declaration of competing interest

Aalborg University Hospital is currently applying for patents relating to the gene panel.

All authors declare that they have no competing interests.

Acknowledgements

We would like to express our gratitude to all the patients who participated in the study.

Also, we thank the private foundations that supported the study: the Medical Advancement of Science Foundation, Speciallæge Heinrick Koops Foundation, Aase and Ejnar Danielsens Foundation, Marie Pedersen and Jensine Heibergs Foundation, Beckett Foundation, Blegdalens Foundation and the Resident Foundation at Aalborg University Hospital. The foundations had no influence on the study design, data analysis, data interpretation or manuscript writing.

Likewise we express our gratitude to: Department of Gastrointestinal Surgery, Aalborg University Hospital, Department of Clinical Medicine, Aalborg University, Department of Molecular Diagnostics, Aalborg University Hospital and Department of Oncology, Herlev and Gentofte Hospital, Copenhagen University Hospital.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pan.2021.05.003.

References

[1] Siegel RL, Miller K, Jamal A. Cancer statistics, 2020. CA A Cancer J Clin 2020;70:7–30.
[2] Haab BB, Huang Y, Balasenthil S, Partyka K, Tang H, Anderson M, et al. Definitive characterization of CA 19-9 in resectable pancreatic cancer using a reference set of serum and plasma specimens. PLoS One 2015;10:1–18.
[3] Hartwig W, Strobel O, Hinz U, Fritz S, Hackert T, Roth C, et al. CA19-9 in potentially resectable pancreatic cancer: perspective to adjust surgical and perioperative therapy. Ann Surg Oncol 2013;20:2188–96.
[4] Henriksen SD, Madsen PH, Larsen AC, Johansen MB, Drewes AM, Pedersen IS, et al. Cell-free DNA promoter hypermethylation in plasma as a diagnostic marker for pancreatic adenocarcinoma. Clin Epigenet 2016;8:117.
[5] Henriksen SD, Madsen PH, Krarup H, Thorlacius-Ussing O. DNA hypermethylation as a blood-based marker for pancreatic cancer: a literature review. Pancreas 2015;44:1036–45.
[6] Delpu Y, Hanoun N, Lulka H, Sicard F, Selves J, Buscail L, et al. Genetic and epigenetic alterations in pancreatic carcinogenesis. Curr Genom 2011;12:15–24.
[7] Costa FF. Epigenomics in cancer management. Canc Manag Res 2010;2:255–65.
[8] Lomberk GA. Epigenetic silencing of tumor suppressor genes in pancreatic cancer. J Gastrointest Canc 2011;42:93–9.
[9] Henriksen SD, Madsen PH, Larsen AC, Johansen MB, Pedersen IS, Krarup H, et al. Promoter hypermethylation in plasma-derived cell-free DNA as a prognostic marker for pancreatic adenocarcinoma staging. Int J Canc 2017:141.
[10] Sato N, Fukushima N, Hruban RH, Goggins M. CpG island methylation profile of pancreatic intraepithelial neoplasia. Mod Pathol 2008;21:238–44.
[11] Schwarzenbach H, Hoon DSB, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. Nat Rev Canc 2011;11:426–37.
[12] Bettagowda C, Sausen M, Leary R, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014;6:69–122.
[13] Siravegna G, Mussolin B, Venesio T, Marsoni S, Seoane J, Dive C, et al. How liquid biopsies can change clinical practice in oncology. Ann Oncol 2019;30:1580–90.
[14] Gall TMH, Belete S, Khanderia E, Frampton AE, Jiao LR. Circulating tumor cells and cell-free DNA in pancreatic ductal adenocarcinoma. Am J Pathol 2019;189:71–81.
[15] Francis G, Stein S. Circulating cell-free tumour DNA in the management of cancer. Int J Mol Sci 2015;16:14122–42.
[16] Hadano N, Murakami Y, Uemura K, Hashimoto Y, Kondo N, Nakagawa N, et al. Prognostic value of circulating tumour DNA in patients undergoing curative resection for pancreatic cancer. Br J Canc 2016;115:59–65.
[17] Mouliere F, Rosenfeld N. Circulating tumor-derived DNA is shorter than somatic DNA in plasma. Proc Natl Acad Sci U S A 2015;112:3178–9.
[18] Swaminathan R, Butt AN. Circulating nucleic acids in plasma and serum: recent developments. Ann N Y Acad Sci 2006;1075:1–9.
[19] Park JW, Baek IH, Kim YT. Preliminary study analyzing the methylated genes in the plasma of patients with pancreatic cancer. Scand J Surg 2012;101:38–44.
[20] Eissa MAL, Lerner L, Abdelfatah E, Shankar N, Canner JK, Hasan NM, et al. Promoter methylation of ADAMTS1 and BNC1 as potential biomarkers for early detection of pancreatic cancer in blood. Clin Epigenet 2019;11:1–10.
[21] Yi JM, Guzzetta A a, Bailey VJ, Downing SR, Van Neste L, Chiappinelli KB, et al. Novel methylation biomarker panel for the early detection of pancreatic cancer. Clin Canc Res 2013;19:6544–55.
[22] Park JK, Ryu JK, Yoon WJ, Lee SH, Lee GY, Jeong KS-S, et al. The role of quantitative NPTX2 hypermethylation as a novel serum diagnostic marker in pancreatic cancer. Pancreas 2012;41:95–101.
[23] www.herlevhospital.dk/BIOPAC/Sider/default.aspx.
[24] Pedersen IS, Krarup HB, Thorlacius-Ussing O, Madsen PH. High recovery of cell-free methylated DNA based on a rapid bisulfite-treatment protocol. BMC Mol Biol 2012;13:12.
[25] Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. N Engl J Med 2015;371:1039–49.
[26] Kamisawa T, Wood LD, Itoi T, Takaori K. Pancreatic cancer seminar. Lancet 2016;388:73–85.
[27] Kleeff J, Korc M, Apte M, Vecchia C La, Johnson CD. Pancreatic cancer. Nat Publ Gr 2016;2:1–23.
[28] Raimondi S, Lowenfels AB, Morselli-Labate AM, Maisonneuve P, Pezzilli R. Pancreatic cancer in chronic pancreatitis; aetiology, incidence, and early detection. Best Pract Res Clin Gastroenterol 2010;24:349–58.
[29] Altman Douglas G, Vergouwe Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:1432–5.
[30] Sinsheimer JS, Bocklandt S, Lin W, Sehl ME, Sa FJ, Vilain E. Epigenetic predictor of age. PLoS One 2011;6:1–6.
[31] Christiansen C, Castillo-Fernandez JE, Domingo-Relloso A, Zhao W, El-Sayed Moustafa JS, Tsai PC, et al. Novel DNA methylation signatures of tobacco smoking with trans-ethnic effects. Clin Epigenet 2021;13:1–13.
[32] Smith GCS, Seaman SR, Wood AM, Royston P, White IR. Correcting for optimistic prediction in small data sets. Am J Epidemiol 2014;180:318–24.
[33] Jiao L, Zhu J, Hassan MM, Evans DB, Abbruzzese, James L, et al. K-ras mutation and p16 and preproenkephalin promoter hypermethylation in plasma DNA of pancreatic cancer patients: in relation to cigarette smoking. Pancreas 2007;34:55–62.

[34] Jung M, Klotzek S, Lewandowski M, Fleischhacker M, Jung K. Changes in concentration of DNA in serum and plasma during storage of blood samples. Clin Chem 2003;49:1028–9.
[35] Wong FCK, Sun K, Jiang P, Cheng YKY, Chan KCA, Leung TY, et al. Cell-free DNA in maternal plasma and serum: a comparison of quantity, quality and tissue origin using genomic and epigenomic approaches. Clin Biochem 2016;49:1379–86.
[36] Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson M Lou, Thornquist M, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst 2001;93:1054–61.
[37] Henriksen SD, Thorlacius-Ussing O. Cell-free DNA methylation as blood-based biomarkers for pancreatic adenocarcinoma—a literature update. Epigenomes 2021;5:8.