

ARTICLE



Clinical Studies

Development and external validation of a prediction model for overall survival after resection of distal cholangiocarcinoma

Ali Belkous¹, Stijn Van Roessel², Marin Strijker², Jacob L. van Dam³, Lois Daamen⁴, Lydia G. van der Geest⁵, Alberto Balduzzi⁶, Andrea Benedetti Cacciaguerra⁷, Susan van Dieren², Quintus Molenaar⁴, Bas Groot Koerkamp³, Joanne Verheij⁸, Elizabeth Van Eycken⁹, Giuseppe Malleo⁶, Mohammed Abu Hilal⁷, Martijn G. H. van Oijen¹, Ivan Borbath¹⁰, Chris Verslype¹¹, Cornelis J. A. Punt¹, Marc G. Besselink², Heinz-Josef Klumpen¹✉ and Dutch Pancreatic Cancer Group (DPCG)*

© The Author(s), under exclusive licence to Springer Nature Limited 2022

BACKGROUND: Various prognostic factors are associated with overall survival (OS) after resection of distal cholangiocarcinoma (dCCA). The objective of this study was to develop and validate a prediction model for 3-year OS after pancreatoduodenectomy for dCCA.

METHODS: The derivation cohort consisted of all patients who underwent pancreatoduodenectomy for dCCA in the Netherlands (2009–2016). Clinically relevant variables were selected based on the Akaike information criterion using a multivariate Cox proportional hazards regression model, with model performance being assessed by concordance index (C-index) and calibration plots. External validation was performed using patients from the Belgium Cancer Registry (2008–2016), and patients from two university hospitals of Southampton (U.K.) and Verona (Italy).

RESULTS: Independent prognostic factors for OS in the derivation cohort of 454 patients after pancreatoduodenectomy for dCCA were age (HR 1.02, 95% CI 1.01–1.03), pT (HR 1.43, 95% CI 1.07–1.90) and pN category (pN1: HR 1.78, 95% CI 1.37–2.32; pN2: HR 2.21, 95% CI 1.63–3.01), resection margin status (HR 1.79, 95% CI 1.39–2.29) and tumour differentiation (HR 2.02, 95% CI 1.62–2.53). The prediction model was based on these prognostic factors. The optimism-adjusted C-indices were similar in the derivation cohort (0.69), and in the Belgian (0.66) and Southampton-Verona (0.68) validation cohorts. Calibration was accurate in the Belgian validation cohort (slope = 0.93, intercept = 0.12), but slightly less optimal in the Southampton-Verona validation cohort (slope = 0.88, intercept = 0.32). Based on this model, three risk groups with different prognoses were identified (3-year OS of 65.4%, 33.2% and 11.8%).

CONCLUSIONS: The prediction model for 3-year OS after resection of dCCA had reasonable performance in both the derivation and geographically external validation cohort. Calibration slightly differed between validation cohorts. The model is readily available via www.pancreascalcator.com to inform patients from Western European countries on their prognosis, and may be used to stratify patients for clinical trials.

British Journal of Cancer (2022) 126:1280–1288; <https://doi.org/10.1038/s41416-021-01687-1>

BACKGROUND

Distal cholangiocarcinoma (dCCA) arises from the epithelium of the common bile duct distal to the confluence of the cystic duct and common hepatic duct and proximal from the ampulla of Vater [1, 2]. Surgical resection is the only treatment with curative intent, but the majority of patients with dCCA present at a late stage with locally advanced or metastatic disease [1]. Surgical treatment options, morbidity and mortality of dCCA differ from other types

of biliary tract cancer (i.e. intrahepatic cholangiocarcinoma, perihilar cholangiocarcinoma and ampullary carcinoma) [3]. Although a radical resection (R0) is achieved in more than 80% of the patients, the recurrence rate remains as high as 82% at 3 year [4, 5]. The 5-year overall survival (OS) after surgical resection is ~29–37% [3, 4]. Adjuvant treatment following a curative-intent resection has shown conflicting results [6, 7]. A meta-analysis of prospective and retrospective studies found no benefit in OS for

¹Department of Medical Oncology, Cancer Center Amsterdam, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands. ²Department of Surgery, Cancer Center Amsterdam, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands. ³Department of Surgery, Erasmus MC Cancer Institute, Rotterdam, The Netherlands. ⁴Department of Surgery, University Medical Center Utrecht, Utrecht, The Netherlands. ⁵Department of research, Netherlands Comprehensive Cancer Organization (IKNL), Utrecht, The Netherlands. ⁶Department of General and Pancreatic Surgery, Integrated University Hospital of Verona, Verona, Italy. ⁷Department of Surgery, University Hospital Southampton NHS Foundation Trust, Southampton, United Kingdom. ⁸Department of Pathology, Cancer Center Amsterdam, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands. ⁹Research Department, Belgian Cancer Registry, Brussels, Belgium. ¹⁰Department of Gastroenterology and Hepatology, University Hospitals St-Luc, Brussels, Belgium. ¹¹Department of Gastroenterology/Digestive Oncology, University Hospitals Gasthuisberg/Leuven & KU Leuven, Leuven, Belgium. *A list of authors and their affiliations appears at the end of the paper. ✉email: h.klumpen@amsterdamumc.nl

Received: 9 January 2021 Revised: 1 October 2021 Accepted: 23 December 2021

Published online: 17 January 2022

adjuvant chemotherapy versus observation after surgical resection of dCCA [6]. Capecitabine was recently adapted by the American Society of Clinical Oncology (ASCO) clinical practice guideline as the standard adjuvant therapy based on the results of the BILCAP trial [8, 9]. Even though, in the subgroup of 156 patients with dCCA of the BILCAP study, OS was not significantly superior with adjuvant capecitabine (HR 0.70, 95% CI 0.47–1.06), as compared to observation only [9].

A number of prognostic factors for OS after surgical resection of dCCA have been identified, including the number of positive regional lymph nodes, the resection margin status and the tumour differentiation [4, 6]. Currently, prediction models for survival after resection are available for intrahepatic, perihilar cholangiocarcinoma and ampullary carcinoma, but to our knowledge, not for dCCA [10, 11]. Prediction of individual patient survival could be used to identify individual patients at high risk for recurrence. Hypothetically, these patients may be more likely to benefit from adjuvant treatment. Moreover, a prediction model could be used to inform patients and to stratify patients in clinical trials, especially in a disease such as dCCA where the effect of adjuvant therapy still remains largely unknown. The objectives of this study were to derive and externally validate a prediction model for 3-year OS after resection of dCCA and to assess the effect of adjuvant treatment in different subgroups based on the final model to potentially identify patients who may benefit from adjuvant therapy.

METHODS

This study was reported in line with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis (TRIPOD) statement and performed according to the Prediction model Risk Of Bias ASsessment Tool (PROBAST) criteria [12, 13].

Patients

The derivation cohort consisted of patients who underwent pancreatoduodenectomy with curative intent for dCCA in the Netherlands from January 1, 2009 up to December 31, 2016. Patients were extracted from the prospectively maintained population-based Netherlands Cancer Registry (NCR). These data are collected by trained registry administrators across the Netherlands and updated on a regular basis, starting from nine months after diagnosis until the patient is deceased or lost to follow-up [14]. Diagnoses of patients from three university medical centers (Amsterdam UMC [location Academic Medical Center], Erasmus MC and University Medical Center Utrecht) were cross-checked with data collected from electronic patients records. Patients were excluded if they had another periampullary tumour than dCCA. Subsequently, the NCR data were linked to the nationwide network and registry of histopathology in the Netherlands (PALGA) to identify missing clinicopathologic characteristics, including pT, pN category, resection margin, tumour differentiation and tumour diameter.

We used two independent validation cohorts to assess the performance of the final model. The first validation cohort consisted of patients from the prospectively maintained population-based Belgium Cancer Registry (BCR) diagnosed with dCCA from January 1, 2008 up to December 31, 2016. A trained registry administrator from the BCR has updated histopathological characteristics from pathology reports and collected additional data on pN category, resection margin, tumour diameter and type of surgery (pancreatoduodenectomy or only bile duct resection). Subsequently, the BCR database was linked to the reimbursement data collected by the Intermutualistic Agency (IMA-AIM) to identify whether patients have received chemotherapy and/or radiotherapy. The second validation cohort consisted of patients from the prospectively maintained, retrospective database of the University Hospital Southampton (United Kingdom, 2006–2019) and Verona University Hospital (Italy, 2000–2019). These two databases are updated on a regular basis. The histopathological characteristics and survival outcomes of patients from these two university hospitals were additionally updated for our study by two researchers (A.B. from Verona and A.B.C.).

All patients included in the derivation and validation cohorts had a histologically proven diagnosis of dCCA and underwent pancreatoduodenectomy (either a classical Whipple or a pylorus-preserving

pancreatoduodenectomy). Patients were excluded if they had a T0-tumour, only a bile duct resection with or without lymphadenectomy, distant metastases or unknown metastatic status at the time of surgical resection, macroscopic positive resection margin (R2), and rare tumour histology with non-biliary origin (morphological codes 8144, 8255, 8260, 8263 and 8523). Patients were excluded if they survived less than 90 days after surgical resection to avoid including patients that deceased because of postoperative complications. Patients with missing values for clinically relevant variables, including pT category, number of evaluated or positive lymph nodes and the status of the resection margin, were excluded. To distinguish pN1 (1–3 positive lymph nodes) from pN2 category (4 or more positive lymph nodes) according to 8th edition of American Joint Committee on Cancer (AJCC) staging system, only patients with at least four evaluated lymph nodes were included. Finally, we excluded patients who received neoadjuvant or induction therapy because these treatments may influence the pathological AJCC staging by tumour and nodal downstaging as shown for perihilar cholangiocarcinoma [15].

This study was approved by the Dutch Pancreatic Cancer Group, a scientific committee for pancreatic cancer and periampullary tumours, including dCCA, in the Netherlands. We received a waiver for ethical approval of this study by the Medical Ethics Review Committee at Amsterdam UMC, location Academic Medical Center (W19_191 # 19.230).

Variables

All three databases contained data on demographic, tumour and treatment characteristics. Main clinicopathologic characteristics include tumour diameter, pT and pN category, number of harvested and positive lymph nodes, status of resection margin and tumour differentiation. Microscopic positive resection margin (R1) was defined as the presence of a residual tumour within 1 mm from the resection margin as defined by a pathologist. Tumour differentiation was missing in 55 patients (12.1%) from the derivation cohort and in 17 patients (9.1%) from the Belgian validation cohort. We imputed these missing values using multiple imputations from the ‘mice’ package in R. Ten datasets were developed after multiple imputations and all steps for the prediction model derivation and validation were analysed in these datasets. Since in previous AJCC staging systems, invasive depth was not a requested item for staging, this parameter was not recorded in the majority of the pathology reports. Therefore, we were unable to retrospectively classify pT category according to the 8th edition of the AJCC staging system. In the final model, pT category according to the 7th edition of the AJCC staging system is used. For pN category, we retrospectively categorised patients according to the 8th edition of the AJCC staging system based on the number of positive lymph nodes.

Treatment characteristics of interest were type, year and location of surgical resection and adjuvant chemotherapy and/or radiotherapy. In the NCR derivation cohort and the Southampton-Verona validation cohort, adjuvant chemotherapy or radiotherapy was defined as the administration of at least one cycle of chemotherapy after surgical resection and before disease recurrence. In the Belgian validation cohort, one cycle of chemotherapy or radiotherapy administered within 3 months after surgical resection was considered as adjuvant chemotherapy or radiotherapy.

The primary outcome was the 3-year OS rate. The OS was assessed from the date of diagnosis to the date of death or last follow-up.

Statistical analysis

Statistical analyses were performed in R version 3.4.3 (cran.rproject.org). Means were compared by *t*-test and groups by Fischer’s exact or χ^2 test. Time-to-event data were analysed with Kaplan–Meier curves and compared with log-rank test.

All clinically relevant variables were entered into the Cox proportional hazard regression model and prognostic variables were selected using backward stepwise selection with the Akaike information criterion (AIC). Coefficients, hazard ratios and 95% confidence intervals of each predictor of the final model are presented. The performance of the final model was assessed by the Harrell concordance index (C-index) and calibrations plots. Calibration of the prediction model was assessed to evaluate how well predicted survival for each patient corresponded with the actual observed survival. The cohort was divided into 10 risk groups of similar size according to the risk score. The 3-year OS rate of each risk group was then compared with the predicted 3-year OS rate and visualised in a calibration plot. Perfect calibration (i.e. exact agreement between predicted and observed) is characterised by a calibration slope of 1 and an intercept of 0.

Table 1. Baseline characteristics of patients with distal cholangiocarcinoma.

Baseline characteristics	Netherlands Cancer Registry Derivation Cohort (N = 454)	Belgian Cancer Registry Validation Cohort (N = 187)		Southampton-Verona Validation Cohort (N = 111)	
	N (%)	N (%)	P-value	N (%)	P-value
Age in years (median, IQR)	67 (59–73)	68 (61–73)	0.515	72 (65–78)	<0.001
<65 years	186 (41.0%)	70 (37.4%)	0.458	27 (24.3%)	0.002
≥65 years	268 (59.0%)	117 (62.6%)		84 (75.7%)	
Sex			0.275		0.175
Male	281 (61.9%)	125 (66.8%)		77 (69.4%)	
Female	173 (38.1%)	62 (33.2%)		34 (30.6%)	
Year of diagnosis (median, IQR)	2013 (2011–2015)	2013 (2010–2015)	0.025	2012 (2009–2016)	0.012
≤2012	174 (38.3%)	91 (48.7%)	0.020	62 (55.9%)	0.001
≥2013	280 (61.7%)	96 (51.3%)		49 (44.1%)	
Year of surgery (median, IQR)	2014 (2011–2015)	2013 (2011–2015)	0.014	2012 (2009–2016)	0.013
≤2012	171 (37.7%)	88 (47.1%)	0.034	62 (55.9%)	<0.001
≥2013	283 (62.3%)	99 (52.9%)		49 (44.1%)	
Resection at an academic center	253 (55.7%)	88 (47.1%)	0.056	111 (100.0%)	<0.001
Type of pancreatoduodenectomy			<0.001		0.188
Classical Whipple	271 (59.7%)	169 (90.4%)		58 (52.3%)	
Pylorus preserving	183 (40.3%)	18 (9.6%)		53 (47.7%)	
pT category (7th AJCC)			0.937		0.588
pT1/T2	105 (23.1%)	42 (22.5%)		29 (26.1%)	
pT3/T4	349 (76.9%)	145 (77.5%)		82 (73.9%)	
Tumour size in millimetres ^a , median (IQR)	20 (15–30)	20 (15–28)	0.937	20 (15–24)	0.005
pN category (8th AJCC)			0.089		0.040
pN0	182 (40.1%)	87 (46.5%)		55 (49.5%)	
pN1	168 (37.0%)	71 (38.0%)		42 (37.8%)	
pN2	104 (22.9%)	29 (15.5%)		14 (12.6%)	
Lymph nodes evaluated, median (IQR)	12 (8–16)	14 (10–20)	<0.001	21 (14–29)	<0.001
Positive lymph nodes, median (IQR)	1 (0–3)	1 (0–3)	0.059	1 (0–3)	0.082
Resection margin			0.215		0.303
R0	375 (78.5%)	152 (81.3%)		79 (71.2%)	
R1	103 (21.5%)	35 (18.7%)		32 (28.8%)	
Tumour differentiation			0.510		0.147
Well/ moderately differentiated	252 (55.5%)	113 (60.4%)		79 (71.2%)	
Poorly differentiated	147 (32.4%)	57 (30.5%)		32 (28.8%)	
Unknown	55 (12.1%)	17 (9.1%)		0	
Adjuvant chemotherapy	35 (7.7%)	66 (35.3%)	<0.001	73 (65.8%)	<0.001
Adjuvant radiotherapy	0	6 (3.2%)	<0.001	5 (4.5%)	<0.001

Statistically significant *P*-values are in bold.

CI confidence interval, IQR interquartile range, N number of patients.

^aThe tumour size was missing in 35 and 66 patients from the derivation and BCR validation cohort, respectively.

Patients were stratified into three risk groups of equal size based on the risk score as measured by the final model. The final model was internally validated using bootstrap cross-validation (number of repetitions 200) and externally validated in two independent international cohorts. The model was made available online via www.pancreascalculator.com.

RESULTS

Derivation cohort

A total of 548 patients were derived from the NCR, of whom 454 met the inclusion criteria of this study (Supplementary Fig. 1). The median age in the derivation cohort was 67 years (IQR 59–73), and 61.9% of the patients were male. Most patients had T3 or T4

tumours (76.9%), positive lymph nodes (59.9%) and negative microscopic resection margins (78.5%). Patients underwent a classical Whipple (59.7%) or pylorus-preserving pancreatoduodenectomy (40.3%). A total of 35 patients (7.7%) received adjuvant chemotherapy in the derivation cohort (Table 1).

OS and predictors of OS

After a median follow-up of 65.8 months (IQR 54.3–74.3) in the derivation cohort, 339 (74.7%) patients had deceased and median OS was 24.9 months (IQR 22.3–27.2). The number of events at 3 year was 313 (68.9%). The 3-year and 5-year OS were 36.8% (95% CI, 32.5–41.6%) and 24.7% (95% CI, 20.7–29.4%), respectively.

Table 2. Cox proportional hazards regression model in the derivation cohort.

	Univariate model		Multivariate model		
	Hazard ratio (95% CI)	P-value	Hazard ratio (95% CI)	P-value	Coefficient (β)
<i>Selected factors</i>					
Age in years	1.01 (0.99–1.02)	0.28	1.02 (1.01–1.03)	0.001	0.019
pT category (7th AJCC)					
pT1/T2	Ref		Ref		
pT3/T4	1.79 (1.36–2.36)	<0.001	1.43 (1.07–1.90)	0.015	0.355
pN category (8th AJCC)					
pN0	Ref		Ref		
pN1	2.05 (1.59–2.65)	<0.001	1.78 (1.37–2.32)	<0.001	0.578
pN2	3.17 (2.39–4.22)	<0.001	2.21 (1.63–3.01)	<0.001	0.794
Resection margin					
R0	Ref		Ref		
R1	2.09 (1.65–2.65)	<0.001	1.79 (1.39–2.29)	<0.001	0.580
Tumour differentiation					
Well/ moderately differentiated	Ref		Ref		
Poorly differentiated	2.15 (1.73–2.67)	<0.001	2.02 (1.62–2.53)	<0.001	0.704
<i>Other factors</i>					
Age in years					
<65 years	Ref				
≥65 years	1.07 (0.86–1.33)	0.527			
Sex					
Male	Ref				
Female	1.18 (0.95–1.47)	0.130			
Year of diagnosis (median, IQR)					
2009–2012	Ref				
2013–2016	1.08 (0.86–1.35)	0.492			
Year of surgery (median, IQR)					
2009–2012	Ref				
2013–2017	1.03 (0.83–1.29)	0.780			
Resection at an academic center (ref. no)					
	0.97 (0.78–1.20)	0.765			
Type of pancreatoduodenectomy					
Classical Whipple	Ref				
Pylorus preserving	0.98 (0.79–1.23)	0.888			
Tumour size in millimetres, median (IQR)					
<20 mm	Ref				
≥20 mm	1.65 (1.31–2.09)	<0.001			
Lymph nodes evaluated, median (IQR)					
	1.02 (1.01–1.04)	0.006			
Positive lymph nodes, median (IQR)					
	1.13 (1.10–1.16)	<0.001			
Adjuvant chemotherapy (ref. no)					
	1.16 (0.78–1.73)	0.451			

Statistically significant *P*-values are in bold.

CI confidence interval, IQR interquartile range, *N* number of patients, *Ref* reference.

In the multivariable Cox proportional hazards regression model of the derivation cohort, age at diagnosis, pT category (T1/T2 versus T3/T4) according to 7th edition of AJCC staging system and pN category (N0 versus N1 versus N2) according to 8th edition of AJCC staging system, resection margin status (R0 versus R1) and tumour differentiation (well/moderately versus poorly differentiated) were identified as independent prognostic factors for OS (Table 2 and Supplementary Fig. 2). These five variables were included in the final model and a nomogram was constructed (Fig. 1). The estimated baseline hazard function from the Cox

proportional hazards regression model was 0.901 (the formula for 3-years OS prediction is presented in Fig. 1).

The calibration plot comparing predicted with observed OS in the derivation cohort is presented in Table 3. The overall C-index of this prediction model was 0.69. Internal validation using bootstrapping demonstrated an optimism-corrected C-index of 0.69. This final model had a better performance compared to the 7th edition of AJCC staging system, which had a C-index of 0.64. Based on this final model, three risk groups were identified with a 3-year OS of 65.4%, 33.2% and 11.8%, respectively (Fig. 2). The nomogram to predict

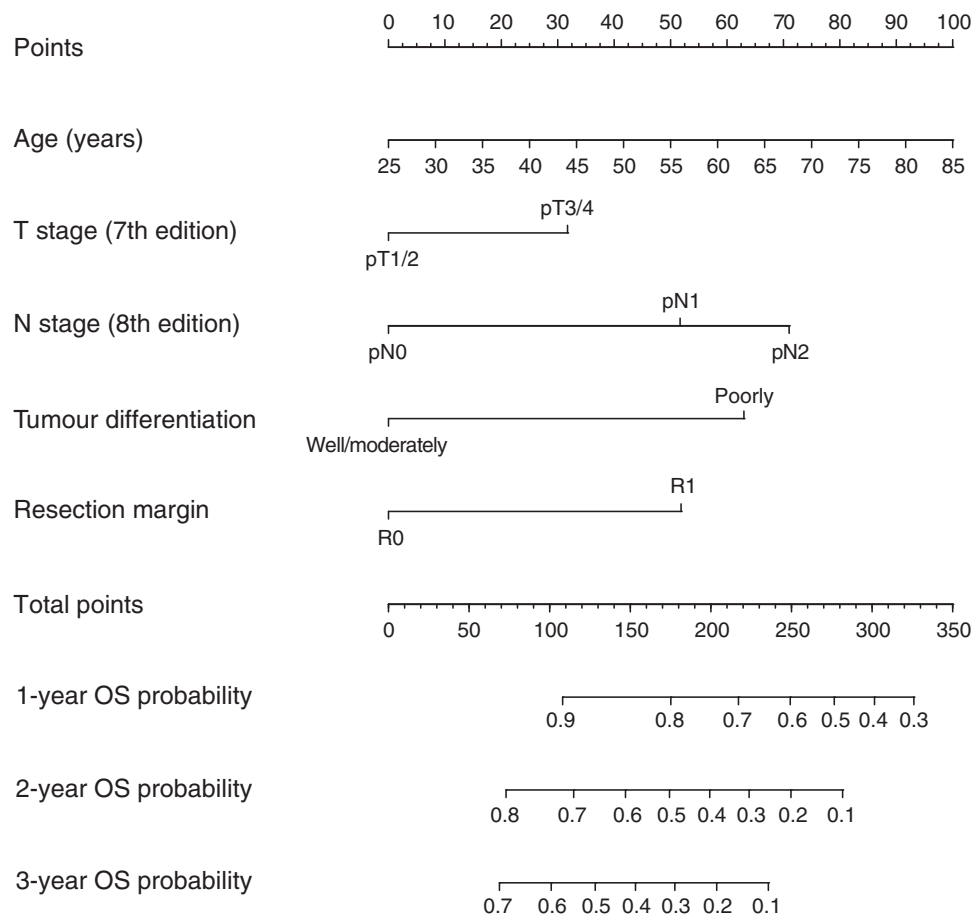


Fig. 1 Nomogram for prediction of 1, 2 and 3-year overall survival. Description of Fig. 1: The formula for 3-years OS prediction: $S(t) = 0.901^{(0.019 * \text{age} + 0.355 * \text{pT} + 0.578 * \text{pN1} + 0.794 * \text{pN2} + 0.580 * \text{R1} + 0.704 * \text{tumour differentiation})}$.

3-year OS after resection of dCCA was also made available through a web-based calculator, www.pancreascalculator.com.

External validation

Relevant clinicopathological and treatment characteristics of patients from the validation cohorts are presented in Table 1. Patients in the Belgian validation cohort and the derivation cohort had comparable baseline characteristics, except that more adjuvant chemotherapy and radiotherapy was administered in the Belgian validation cohort (Table 1).

Patients in the Southampton-Verona validation cohort were older, had less frequently N2 disease and received more frequently adjuvant chemotherapy and/or radiotherapy than patients in the derivation cohort (Table 1). The univariate analyses of the validation cohorts are shown in supplementary Table 1.

After a median follow-up of 64.7 months (95% CI, 54.2–71.3 months) in the Belgian validation cohort, 115 (61.5%) patients had deceased and median OS was 33.8 months (95% CI, 30.0–51.0). The number of events at 3 years was 120 (64.2%). The 3-year and 5-year OS in the Belgian validation cohort were 48.4% (95% CI, 41.4–56.5%) and 36.6% (95% CI, 29.5–45.4%), respectively. In the Southampton-Verona validation cohort, 49 patients (44.1%) deceased during the median follow-up of 53.4 months (95% CI, 39.8–69.0 months). The number of events at 3 year was 64 (57.7%). Median, 3-year and 5-year OS in the Southampton-Verona validation cohort were 52.6 months (95% CI, 40.0–65.0 months), 66.5% (95% CI, 57.1–77.4%) and 43.9% (95% CI, 33.5–57.4%), respectively. The OS was significantly longer in the validation cohorts compared to the derivation cohort.

Validation of the final model in the Belgian validation cohort showed a C-index of 0.66 compared to a C-index of 0.62 of the 7th edition of the AJCC staging system, with a good calibration (slope 0.93, intercept 0.12, Table 3 and Supplementary Fig. 3). In the Southampton-Verona validation cohort, the final model had reasonable discrimination (C-index 0.68), but suboptimal calibration. The slope of 0.88 indicates that the order of predictions is fairly accurate, but the intercept of 0.32 indicates that the observed OS was better than predicted and that the survival predictions were systematically too low in this cohort (Table 3 and Supplementary Fig. 3).

Adjuvant therapy

Adjuvant therapy was used more frequently in the validation cohorts compared to the derivation cohort (Table 1). Adjuvant therapy was not associated with better OS in the derivation cohort (HR 1.26, 95% CI 0.92–1.71) or validation cohorts (HR 1.16, 95% CI 0.78–1.73) compared to only surgical resection (Fig. 3a, e). In the low- and intermediate-risk groups, adjuvant therapy was not significantly associated with OS (Fig. 3b, f, c, g, respectively). In the high-risk group, a non-statistical survival difference in OS was observed in patients treated with adjuvant therapy in the derivation cohort (HR 0.69, 95% CI 0.39–1.24, Fig. 3d) and validation cohorts (HR 0.85, 95% CI 0.52–1.40, Fig. 3h) compared to only surgical resection.

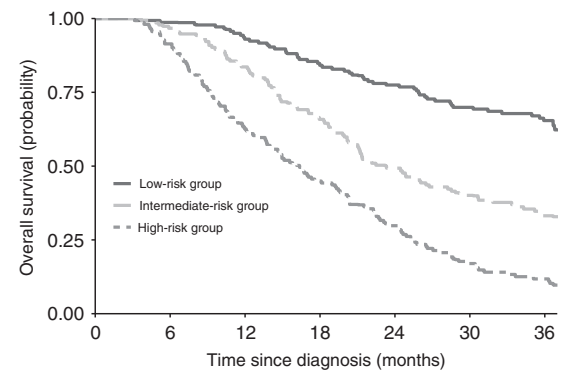
DISCUSSION

This first prediction model for 3-year OS after pancreatoduodenectomy for dCCA was derived using five commonly available and

Table 3. Performance and calibration of the prediction model in the derivation and validation cohorts.

	Netherlands Cancer Registry Derivation Cohort (N = 454)	Belgian Cancer Registry Validation Cohort (N = 187)	Southampton-Verona Validation Cohort (N = 111)
C-index (95% CI)	0.69 (0.68 to 0.70)	0.66 (0.64 to 0.68)	0.68 (0.58 to 0.77)
Calibration slope (95% CI)	1.08 (0.95 to 1.22)	0.93 (0.43 to 1.43)	0.88 (0.44 to 1.32)
Calibration intercept (95% CI)	-0.03 (-0.09 to 0.03)	0.12 (-0.10 to 0.34)	0.32 (0.13 to 0.50)
Calibration deviance (95% CI)	0.04 (0.01 to 0.08)	0.14 (0.01 to 0.25)	0.28 (0.11 to 0.41)
Quartiles based on the nomogram score			
3-year OS (%)			
	Predicted	Observed	Observed
1st quartile	22.0	19.6	30.0
2nd quartile	39.5	35.3	52.3
3rd quartile	57.0	58.9	64.0
4th quartile	74.4	83.5	100.0
	Predicted	Predicted	Predicted
	22.8	40.3	57.7
	21.9	75.1	56.0
	21.9	75.1	0.73
	21.9	75.1	100.0

CI confidence interval, N number of patients, OS overall survival.



Number at risk

Low-risk group	151	149	141	127	117	95	82
Intermediate-risk group	151	146	126	100	75	54	44
High-risk group	151	138	95	68	45	24	15

Fig. 2 Three risk groups based on the prediction model score. The line in dark colour represents the low-risk group; the dashed line in light colour represents the intermediate-risk group; the dash-dotted line in dark colour represents the high-risk group.

relevant clinicopathological characteristics. The final model included age at diagnosis, pT category (AJCC 7th), pN category (AJCC 8th), the status of the resection margin and tumour differentiation and was externally validated in two independent cohorts. The final model had reasonable discrimination and outperformed the AJCC staging system (7th edition) in all three cohorts. The calibration of this model was good in the derivation and the Belgian Cancer Registry validation cohort, but was suboptimal in the Southampton-Verona validation cohort because of better OS compared to the derivation cohort. This prediction model may be used to inform patients from Western European countries on their prognosis, and to stratify patients in clinical trials.

A higher proportion of patients in the validation cohorts received adjuvant therapy compared to the derivation cohort, but this treatment did not result in a significantly longer OS in these cohorts separately or combined. In fact, adjuvant therapy was associated with significantly worse OS in the Belgian validation cohort compared to only surgical resection, indicating that patients who received adjuvant therapy had probably a more advanced disease (data not shown). The addition of adjuvant therapy as a variable into the prediction model did not improve its performance. Patients in the high-risk group apparently carry unfavourable tumour characteristics (pT3/4 and pN2 category, R1-resection and poor tumour differentiation) and may benefit from adjuvant therapy as shown by a recent Phase II trial [16]. We feel that advanced dCCA is not only a matter of stage (time) but also a matter of biology (behaviour). In that sense, the high-risk group might show more similarity with pancreatic cancer and therefore be more responsive to adjuvant chemotherapy as proven in several randomised controlled trials for pancreatic cancer [16]. The BCAT (gemcitabine) and PRODIGE 12 (gemcitabine plus oxaliplatin) trials did not show a survival benefit from adjuvant chemotherapy compared to observation, but the BILCAP (capecitabine versus observation) trial found longer OS in the per-protocol-analysis, but not in the intention-to-treat analysis [9, 17, 18]. In the post-hoc analysis of the BILCAP study, dCCA patients treated with capecitabine did not have significantly longer OS compared to those in the observation arm. The ongoing ACTICCA-1 and ASCOT trials will provide more data on the value of adjuvant chemotherapy [19, 20].

Previous studies have also shown that older age at diagnosis, pT3/4 category, positive regional lymph nodes, positive resection margin and moderate/poor tumour differentiation are independently associated with poor OS in dCCA [6, 14]. The pT category according to the 8th edition of the AJCC staging system, which

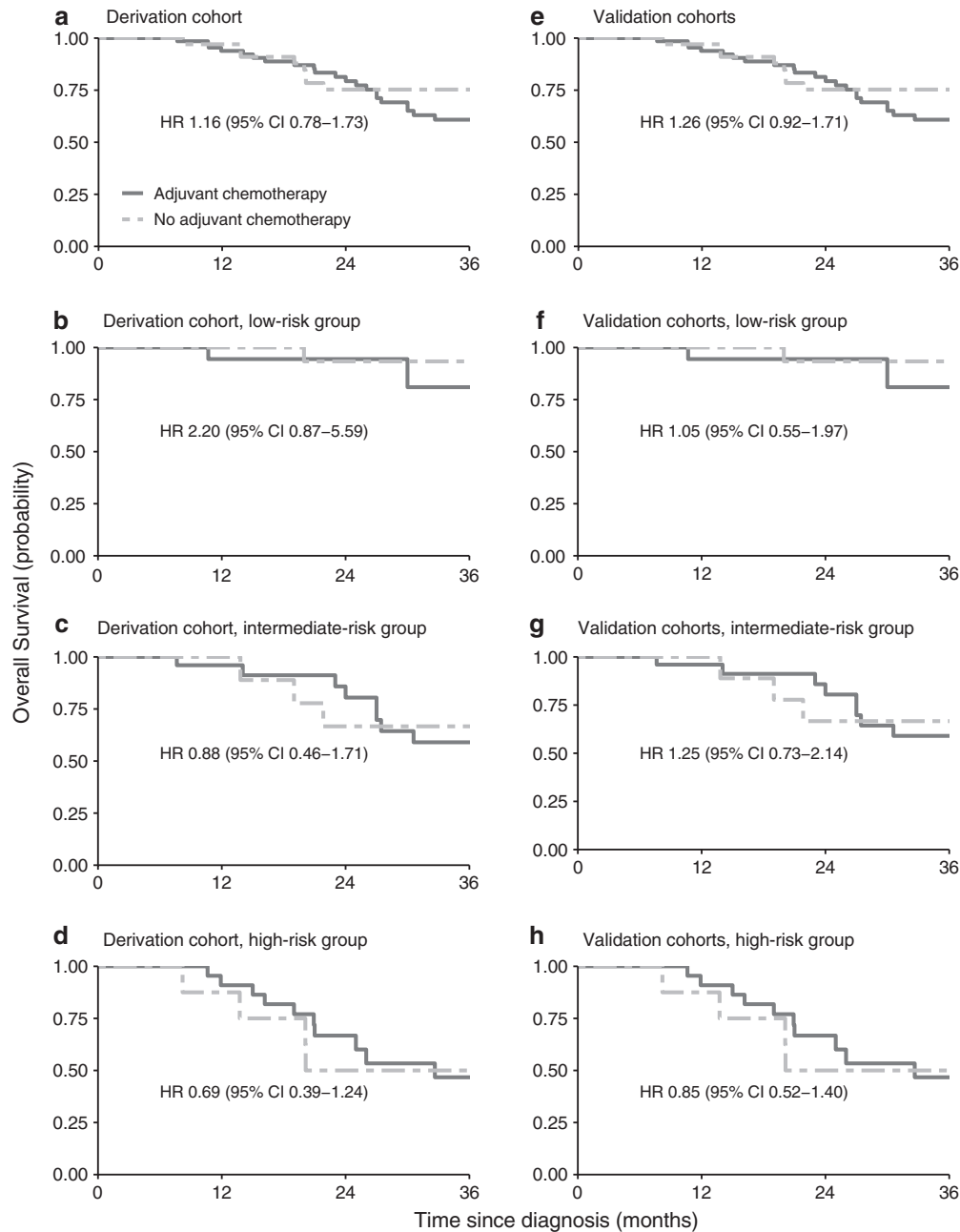


Fig. 3 Kaplan–Meier curves by adjuvant chemotherapy for the derivation and validation cohorts, stratified by risk group. The line in dark colour represents patients treated with adjuvant chemotherapy; the dashed line in light colour represents patients not treated with adjuvant chemotherapy.

was adapted in 2016, was not available for our derivation cohort (2009–2016), because tumour invasion depth was not a requested parameter for AJCC staging during this period. Recent studies showed that assessment of tumour invasion depth was not feasible in approximately half of the number of patients, because of the absence of basal lamina, which is frequently affected by tumour invasiveness or the surgical procedure [21]. A number of studies have shown that tumour diameter is also an independent prognostic factor for OS [22, 23]. Therefore, we analysed the prognostic value of the tumour diameter as measured by the pathologist instead of tumour invasion depth. The tumour diameter was not an independent prognostic factor in the multivariate Cox proportional hazards regression model and did not improve the performance of the final model.

To our knowledge, this is the first prediction model for OS in dCCA following surgical resection. Our prediction model is derived based on a relatively large population-based cohort and is validated in two independent international cohorts. This study has a number of limitations. The calibration of the final model was suboptimal in the Southampton-Verona validation cohort because of a statistically significant difference in OS compared to the derivation cohort. The suboptimal calibration is regularly observed in external validation cohorts, as shown recently by a prediction model for perihilar cholangiocarcinoma, because of differences in clinicopathological characteristics and treatment options between the derivation and external validation cohort [10]. However, in our study most baseline characteristics were almost comparable between the derivation and validation cohorts, except that adjuvant chemotherapy was

frequently used in the validation cohorts and significantly less patients had an N2 category in the Southampton-Verona validation cohort compared to the derivation cohort. We did not perform a propensity score matching using adjuvant chemotherapy as a confounder because the addition of adjuvant chemotherapy as a variable to the prediction model did not improve its performance. The large variation in OS (5-year OS 13–54%) of patients treated with surgical resection for dCCA was described previously in a systemic review of 39 studies with 3258 patients [4]. The difference in OS between the derivation and validation cohort may be clarified by the fact that it is challenging to distinguish dCCA from other periampullary tumours. Besides, the pathological examination was not fully standardised among the different countries, which might have its impact on the inclusion of patients. It is, therefore, possible that patients with ampullary cancer might be included in the validation cohorts misclassified as having dCCA, accounting for the differences in outcome between cohorts. Therefore, we have cross-checked the diagnosis of patients from three large university medical centres in the Netherlands and found that only one patient with an ampullary carcinoma (0.6%) was misclassified as dCCA. We used two geographically distinct cohorts for the validation of the final model, which may have introduced some bias related to geographical differences. However, external validation in different geographical cohorts is a more robust way to assess model performance than using the entire cohort for both model development and internal validation. The derivation cohort was limited to patients diagnosed in the period 2009–2016, because dCCA and perihilar cholangiocarcinoma were not registered as separate diseases in NCR before 2009. In the Belgian validation cohort, we included patients diagnosed in the period 2008–2016 and in the Southampton-Verona validation cohort (2000–2019) we did not limit the inclusion period to the same period as the derivation cohort. However, the median and 3-year OS of the Belgian and Southampton-Verona validation cohorts were comparable before and after excluding patients diagnosed before 2009 or after 2016 in the sensitivity analysis (data not shown). Although it is possible that patients in the Southampton-Verona validation cohort may have benefited from treatment in high-volume hospitals which may have resulted in longer OS. This association was not observed in the derivation cohort given that the surgical treatment of dCCA in the Netherlands from 2011 is centralised in 17 pancreatic centres. A recent study with 363 patients with biliary tract cancer did not find an association between treatment in a high-volume hospital and OS [24]. Finally, other prognostic variables such as perineural and pancreatic invasion that may clarify the difference in OS between the derivation and validation cohorts were not available for the majority of patients [25]. This prediction model may be improved in the future by incorporating clinicopathological characteristics with molecular data and treatment characteristics [26].

In conclusion, our final model predicts 3-year OS after resection of dCCA based on age at diagnosis, pT and pN category, resection margin status and tumour differentiation. The final model had reasonable discrimination and calibration in patients from the Netherlands and Belgium, but had suboptimal calibration in the Southampton-Verona validation cohort. This prediction model may be used to inform patients from Western European countries about their prognosis and to stratify patients in clinical trials.

DATA AVAILABILITY

Anonymous individual data could be requested from the corresponding author.

REFERENCES

- Rizvi S, Khan SA, Hallemeier CL, Kelley RK, Gores GJ. Cholangiocarcinoma—evolving concepts and therapeutic strategies. *Nat Rev Clin Oncol*. 2018;15:95–111.
- Razumilava N, Gores GJ. Classification, diagnosis, and management of cholangiocarcinoma. *Clin Gastroenterol Hepatol*. 2013;11:13–21 e11.
- Roos E, Strijker M, Franken LC, Busch OR, van Hooft JE, Klumpen HJ, et al. Comparison of short- and long-term outcomes between anatomical subtypes of resected biliary tract cancer in a Western high-volume center. *HPB (Oxf)*. 2020;22:405–14.
- Zhou Y, Liu S, Wu L, Wan T. Survival after surgical resection of distal cholangiocarcinoma: a systematic review and meta-analysis of prognostic factors. *Asian J Surg*. 2017;40:129–38.
- Byrling J, Andersson R, Sasor A, Lindell G, Ansari D, Nilsson J, et al. Outcome and evaluation of prognostic factors after pancreaticoduodenectomy for distal cholangiocarcinoma. *Ann Gastroenterol*. 2017;30:571–7.
- Wellner UF, Shen Y, Keck T, Jin W, Xu Z. The survival outcome and prognostic factors for distal cholangiocarcinoma following surgical resection: a meta-analysis for the 5-year survival. *Surg Today*. 2017;47:271–9.
- Belkous A, Wilmink JW, Haj Mohammad N, Hagendoorn J, de Vos-Geelen J, Dejong CHC, et al. Advances in adjuvant therapy of biliary tract cancer: an overview of current clinical evidence based on phase II and III trials. *Crit Rev Oncol Hematol*. 2020;151:102975.
- Shroff RT, Kennedy EB, Bachini M, Bekaii-Saab T, Crane C, Edeline J, et al. Adjuvant therapy for resected biliary tract cancer: ASCO Clinical Practice Guideline. *J Clin Oncol*. 2019;37:1015–27.
- Primrose JN, Fox RP, Palmer DH, Malik HZ, Prasad R, Mirza D, et al. Capecitabine compared with observation in resected biliary tract cancer (BILCAP): a randomised, controlled, multicentre, phase 3 study. *Lancet Oncol*. 2019;20:663–73.
- Groot Koerkamp B, Wiggers JK, Gonen M, Doussot A, Allen PJ, Besselink MGH, et al. Survival after resection of perihilar cholangiocarcinoma—development and external validation of a prognostic nomogram. *Ann Oncol*. 2015;26:1930–5.
- Hyder O, Marques H, Pulitano C, Marsh JW, Alexandrescu S, Bauer TW, et al. A nomogram to predict long-term survival after resection for intrahepatic cholangiocarcinoma: an Eastern and Western experience. *JAMA Surg*. 2014;149:432–8.
- Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med*. 2019;170:W1–W33.
- Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162:55–63.
- Strijker M, Belkous A, van der Geest LG, van Gulik TM, van Hooft JE, de Meijer VE, et al. Treatment and survival of resected and unresected distal cholangiocarcinoma: a nationwide study. *Acta Oncol*. 2019;58:1048–55.
- Frosio F, Mocchegiani F, Conte G, Bona ED, Vecchi A, Nicolini D, et al. Neoadjuvant therapy in the treatment of hilar cholangiocarcinoma: review of the literature. *World J Gastrointest Surg*. 2019;11:279–86.
- Seita K, Ebata T, Mizuno T, Maeda A, Yamaguchi R, Kurumiya Y, et al. Phase 2 trial of adjuvant chemotherapy with S-1 for node-positive biliary tract cancer (N-SOG 09). *Ann Surg Oncol*. 2020;27:2348–56.
- Ebata T, Hirano S, Konishi M, Uesaka K, Tsuchiya Y, Ohtsuka M, et al. Randomized clinical trial of adjuvant gemcitabine chemotherapy versus observation in resected bile duct cancer. *Br J Surg*. 2018;105:192–202.
- Edeline J, Benabdelghani M, Bertaut A, Watelet J, Hammel P, Joly JP, et al. Gemcitabine and oxaliplatin chemotherapy or surveillance in resected biliary tract cancer (PRODIGE 12-ACCORD 18-UNICANCER GI): a randomized phase III study. *J Clin Oncol*. 2019;37:658–67.
- Nakachi K, Konishi M, Ikeda M, Mizusawa J, Eba J, Okusaka T, et al. A randomized Phase III trial of adjuvant S-1 therapy vs. observation alone in resected biliary tract cancer: Japan Clinical Oncology Group Study (JCOG1202, ASCOT). *Jpn J Clin Oncol*. 2018;48:392–5.
- Stein A, Arnold D, Bridgewater J, Goldstein D, Jensen LH, Klumpen HJ, et al. Adjuvant chemotherapy with gemcitabine and cisplatin compared to observation after curative intent resection of cholangiocarcinoma and muscle invasive gallbladder carcinoma (ACTICA-1 trial)—a randomized, multidisciplinary, multinational phase III trial. *BMC Cancer*. 2015;15:564.
- Moon A, Choi DW, Choi SH, Heo JS, Jang KT. Validation of T stage according to depth of invasion and N stage subclassification based on number of metastatic lymph nodes for distal extrahepatic bile duct (EBD) carcinoma. *Medicine (Baltim)*. 2015;94:e2064.
- Postlewait LM, Ethun CG, Le N, Pawlik TM, Buettner S, Poultides G, et al. Proposal for a new T-stage classification system for distal cholangiocarcinoma: a 10-institution study from the U.S. Extrahepatic Biliary Malignancy Consortium. *HPB (Oxf)*. 2016;18:793–9.
- Li X, Lin H, Sun Y, Gong J, Feng H, Tu J. Prognostic significance of the lymph node ratio in surgical patients with distal cholangiocarcinoma. *J Surg Res*. 2019;236:2–11.
- Gottlieb-Vedi E, Mattsson F, Lagergren P, Lagergren J. Annual hospital volume of surgery for gastrointestinal cancer in relation to prognosis. *Eur J Surg Oncol*. 2019;45:1839–46.
- Komaya K, Ebata T, Shirai K, Ohira S, Morofuji N, Akutagawa A, et al. Recurrence after resection with curative intent for distal cholangiocarcinoma. *Br J Surg*. 2017;104:426–33.

26. Strijker M, Chen JW, Mungroop TH, Jamieson NB, van Eijck CH, Steyerberg EW, et al. Systematic review of clinical prediction models for survival after surgery for resectable pancreatic cancer. *Br J Surg.* 2019;106:342–54.

ACKNOWLEDGEMENTS

This study was conducted on behalf of the Dutch Pancreatic Cancer Group (DPCG).

AUTHOR CONTRIBUTIONS

AB, SR, IB, CV, CP, MB and HK designed the study. AB, SR, CD, LD, A. Balduzzi, A.B.C. and a trained registry administrator from the BCR collected and updated the necessary data. A.B. performed the statistical analyses with support from S.R. and M. O. which were interpreted by all authors. A.B. drafted the manuscript and all authors revised the manuscript and approved the final manuscript for publication.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

COMPETING INTERESTS

The authors declare no competing interests.

DUTCH PANCREATIC CANCER GROUP (DPCG)

Quintus Molenaar⁴, Bas Groot Koerkamp³, Joanne Verheij⁸ and Marc G. Besselink²

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the Dutch Pancreatic Cancer Group. We received a waiver for ethical approval of this study by the Medical Ethics Review Committee at Amsterdam UMC, location Academic Medical Center (W19_191 # 19.230).

CONSENT TO PUBLISH

This manuscript does not contain any individual person's data in any form.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41416-021-01687-1>.

Correspondence and requests for materials should be addressed to Heinz-Josef Klumpen.

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

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