Original Article

Stomach Dose–Volume Predicts Acute Gastrointestinal Toxicity in Chemoradiotherapy for Locally Advanced Pancreatic Cancer

D.L.P. Holyoake*,†

D.R. Warren*

C. Hurt‡

M. Aznar§

M. Partridge*

S. Mukherjee*,†

M.A. Hawkins*,†,*

maria.hawkins@oncology.ox.ac.uk

*CRUK/MRC Oxford Institute for Radiation Oncology, Department of Oncology, University of Oxford, Oxford, UK

†Oxford University Hospitals NHS Foundation Trust, Oxford, UK

‡Centre for Trials Research, Cardiff University, Cardiff, Wales, UK

§Nuffield Department of Population Health, University of Oxford, Oxford, UK

*Author for correspondence: M.A. Hawkins, CRUK/MRC Oxford Institute for Radiation Oncology, Department of Oncology, University of Oxford, Old Road Campus Research Building, Roosevelt Drive, Oxford OX3 7DQ, UK. Tel: +44-1865 225842; Fax: +44-1865-857127.

Abstract

Aims

Gastrointestinal toxicity impedes dose escalation in chemoradiotherapy for hepatobiliary malignancies. Toxicity risk depends on clinical and radiotherapy metrics. We aimed to identify predictive factors using data from two prospective phase II clinical trials of locally advanced pancreatic cancer (LAPC).

Materials and methods

Ninety-one patients with available data from the ARCII (59.4 Gy in 33 fractions with gemcitabine, cisplatin and nelfinavir, n = 23) and SCALOP (50.4 Gy in 28 fractions with capecitabine or gemcitabine, n = 74) trials were studied. The independent variables analysed comprised age, sex, performance status, baseline symptoms, tumour size, weight loss, chemotherapy regimen and dose–volume histogram of stomach and duodenum in 5 Gy bins. The outcome measures used were Common Terminology Criteria of Adverse Events (CTCAE) grade and risk of CTCAE grade ≥2 acute upper gastrointestinal toxicity (anorexia, pain, nausea and/or vomiting). The risk of CTCAE grade ≥2 events was modelled using multivariable logistic regression and prediction of severity grade using ordinal regression.

Results

CTCAE grade ≥2 symptoms occurred in 38 patients (42%). On univariate analysis, stomach $V_{35–45Gy}$ was predictive of risk (odds ratio 1.035, 95% confidence interval 1.007–1.063) and grade (1.023, 1.003–1.044) of toxicity. The area under the curve was 0.632 (0.516–0.747) with toxicity risk 33/66 (50%) above and 5/25 (20%) below the optimal discriminatory threshold (7.1 cm^3). Using a threshold of 30 cm^3, risk was 13/20 (65%) versus 25/71 (35%). The optimal multivariable logistic regression model incorporated patient sex, chemotherapy regimen and stomach $V_{35–45Gy}$. Receiving gemcitabine rather than capecitabine (odds ratio 3.965, 95% confidence interval 1.274–12.342) and weight loss during induction chemotherapy (1.216, 1.043–1.419) were significant predictors for the SCALOP cohort, whereas~~wheras~~ age predicted toxicity risk in ARCII only (1.344, 1.015–1.780). Duodenum dose–volume did not predict toxicity risk or severity in any cohort.

Conclusions

In chemoradiotherapy for LAPC the volume of stomach irradiated to a moderately high dose (35–45 Gy) predicts the incidence and severity of acute toxicity. Other predictive factors can include age, sex, recent weight loss and concomitant chemotherapy agents.

Key words: NTCP; pancreatic cancer; stomach; toxicity

Introduction

Patients with locally advanced pancreatic cancer (LAPC) who do not progress on initial treatment with chemotherapy may achieve improved local control with chemoradiotherapy [1]. There is evidence for a dose–response relationship in pancreatic cancer [2], hence an increased dose could achieve better tumour control. However, the radiotherapy dose that can be delivered is limited by gastrointestinal toxicity [3–5], the risk of which also increases with dose [6].

Clinical data on the radiotherapy tolerances for the stomach and duodenum remain sparse, but some studies have confirmed the association between organ at risk radiotherapy parameters and subsequent risk of toxicity [5,7–12].

We analysed toxicity outcomes in patients with LAPC treated in two prospective phase II clinical trials. Patients in the SCALOP study (NCT 01032057, $n = 74$) with stable or responding disease after 12 weeks of induction gemcitabine and capecitabine chemotherapy were randomised to receive either gemcitabine or capecitabine alongside 50.4 Gy in 28 fractions [13]. Patients in the single-arm open-label ARCII study (EudraCT 2008-006302-42, $n = 23$) received 59.4 Gy in 33 fractions during concomitant chemoradiotherapy with gemcitabine, cisplatin and nelfinavir (a hypoxia modifier) [14].

This analysis aimed to: (i) identify normal tissue dose–volume histogram (DVH) parameters associated with increased risk of toxicity; (ii) develop and validate predictive multivariable models for personalised estimation of risk that might be utilised in the clinic; (iii) investigate possible associations of toxicity and survival outcomes.

Materials and Methods

Patient Data

The trial eligibility criteria, treatment details and outcomes have been reported previously [13,14]. Toxicity events were prospectively recorded according to Common Terminology Criteria of Adverse Events (CTCAE), version 3.0 [15] in SCALOP and version 4.0 [16] in ARCII and both studies recorded baseline symptoms. In ARCII, clinical assessments were weekly during radiotherapy, 6–8 weeks after radiotherapy and 3 monthly until 12 months. In SCALOP, assessments were monthly during induction chemotherapy, weekly during radiotherapy, 2 weeks after radiotherapy, and 3 and 6 months later. In both studies, patients were prescribed prophylactic anti-emetics and acid-suppressant medication.

Symptoms of acute toxicity (nausea, vomiting, abdominal pain, gastrointestinal bleeding/perforation, bowel obstruction, anorexia ± weight loss) were pooled to generate a single end point of ‘upper gastrointestinal toxicity’ (UGIT) [17]. The maximal grade of any of these symptoms suffered by each patient during 3 months (90 days) from the onset of radiotherapy was collated. The toxicity outcome was dichotomised according to a threshold of grade ≥2, chosen because this indicates requirement for medical intervention.

Two patients from ARCII were excluded as they received only one radiotherapy fraction. For one patient this was due to disease progression and in the other was due to unrelated medical comorbidity. One patient from SCALOP was excluded due to gastric outlet obstruction on planning computed tomography causing abnormal stomach dilatation (measured stomach volume was 2954 cm^3). The median stomach volume was 361 cm^3 (interquartile range 255–541 cm^3). Radiotherapy dose data were not available for three patients from the SCALOP study. In total, 91 patients (70 from SCALOP, 21 from ARCII) were included in the final analysis.

Disease and patient characteristics collected included age, sex, performance status, body mass index, tumour volume and tumour location as indicated by centre of mass (head, neck or other). In ARCII the Karnofsky performance status had been recorded and values were converted to equivalent Eastern Cooperative Oncology Group (ECOG) grade [18]. For SCALOP patients, weight loss during induction chemotherapy was calculated in kilograms.

Radiotherapy Data

The details of radiotherapy delivery in the two studies, including radiotherapy trials quality assurance, have been described elsewhere [13,14,19,20]. In SCALOP, the gross tumour volume (GTV) was defined as tumour visualised on computed tomography with lymph nodes > 1 cm diameter; the planning target volume was defined as the GTV plus a 20 mm margin in the craniocaudal direction and a 15 mm margin otherwise. All patients were prescribed 50.4 Gy in 28 daily fractions in a single phase with three-dimensional conformal radiotherapy. In ARCII, radiotherapy was delivered in two phases: 50.4 Gy in 28 daily fractions was prescribed to the primary tumour and draining lymph node

regions followed by a sequential boost of 9 Gy in five fractions to the primary tumour planning target volume (also defined as GTV plus 20 mm in the craniocaudal direction and 15 mm in other directions). Phase 1 was delivered using intensity-modulated radiotherapy and phase 2 using conformal planning. No dose–volume constraints for the stomach or duodenum were set in either study. The use of intravenous contrast and oral water contrast (100–200 ml) for treatment planning imaging was specified in both studies. For ARCII, patients were fasted for 2 h before planning and treatment. Two patients in ARCII underwent re-planning due to weight loss during radiotherapy. The two partial courses were summed using deformable registration in Mirada RTx (Mirada Medical, Oxford, UK). Doses were recalculated to reflect the delivered dose if patients did not complete their prescribed treatment (four patients in SCALOP and two in ARCII discontinued radiotherapy early due to toxicity, whereas overall 95% of planned fractions were delivered in SCALOP and 99% in ARCII). For both trials the prophylactic use of a proton pump inhibitor or histamine receptor blocker and appropriate anti-emetics during radiotherapy were mandatory, unless contraindicated.

The stomach and duodenum were contoured retrospectively according the Radiation Therapy Oncology Group atlas and guidance [21], with specialist radiologist support. Where these structures had been previously contoured by treating clinicians (in all ARCII patients and in nine SCALOP patients) they were modified as necessary.

Radiotherapy planning computed tomography and dose data were anonymised and imported into the Computational Environment for Radiotherapy Research software package [22] and cumulative absolute dose–volume data were exported in 5 Gy bins (i.e. $V_{5Gy,}$ V_{10Gy} etc.).

Statistics

Radiotherapy dose–volume data were not normally distributed, hence non-parametric tests were used for the assessment of associations and correlation. Differences in continuous variables between groups were assessed using the Mann–Whitney–Wilcoxon test and differences in categorical frequencies using chi-squared. Tests for correlations used Kendall tau-b for dichotomous variables and Spearman's rank for continuous variables. Binary logistic regression was used to test relationships of predictive variables with risk of dichotomised outcomes and ordinal regression for grade of toxicity. Optimal risk thresholds for division of patients into groups according to continuous variables were derived using receiver operator characteristic (ROC) analysis and Youden's index [23]. Factors showing significant associations on univariate analysis and those with sound clinical rationale for relationships with toxicity were incorporated into multivariate logistic regression, which utilised backwards stepwise selection with criterion for retention in the final model $P < 0.1$. Optimal models were selected by maximal classification accuracy and area under the curve (AUC), with minimum Akaike Information Criterion and five-fold internal cross-validation was used to estimate generalisability. Overall and progression-free survival were compared between groups using Log-rank (Mantel-Cox).

Results

CTCAE grade ≥2 UGIT symptoms occurred in 38 patients (42%): grade 1 in 36 (39.6%), grade 2 in 26 (28.6%) and grade 3 in 12 (13.2%). No grade 4 or 5 toxicity occurred. The risk of grade ≥2 UGIT was higher among the ARCII cohort than among the SCALOP patients, but not significantly ($\chi^2 = 1.267$, $P = 0.260$), as were the proportion of patients with performance status ≥ 1 ($\chi^2 = 2.837$, $P = 0.092$) and with the tumour centre of mass in the neck or body of the pancreas, rather than the head ($\chi^2 = 6.773$, $P = 0.335$). Other clinical and treatment parameters were equivalent across the two cohorts (Table 1).

Table 1 Collated patient characteristics and acute toxicity incidence

N		Combined cohort	ARCII	SCALOP
		91	21	70
Age, mean (range)		64.0 (43.4–78.8)	64.5 (43.4–78.8)	63.8 (45.9–77.9)
Sex	Male	50 (54.9%)	11 (52.4%)	39 (55.7%)
	Female	41 (45.1%)	10 (47.6%)	31 (44.3%)
BMI, mean (range)		23.9 (16.4–44.4)	23.6 (17.4–34.2)	24.0 (16.4–44.4)
ECOG performance status	0	45 (49.5%)	7 (33%)	38 (54.3%)
	1	39 (42.9%)	10 (47.6%)	29 (41.4%)
	2	7 (7.7%)	4 (19.0%)	3 (4.3%)
Tumour volume [cm^3], mean (range)		35.9 (1.4–114.2)	29.8 (6.6–74.0)	37.7 (1.4–114.2)

Tumour location	Body	8 (8.8%)	4 (19.0%)	4 (5.7%)
	Head	64 (70.3%)	13 (61.9%)	51 (72.9%)
	Neck	19 (20.9%)	4 (19.0%)	45 (21.4%)
Concomitant chemotherapy			Gemcitabine, cisplatin and	Gemcitabine 36 (51.4%)
			nelfinavir (100%)	Capecitabine 34 (48.6%)
Grade 0 upper gastrointestinal toxicity		17 (18.7%)	3 (14.3%)	14 (20.0%)
Grade 1 upper gastrointestinal toxicity		36 (39.6%)	7 (33.3%)	29 (41.4%)
Grade 2 upper gastrointestinal toxicity		26 (28.6%)	8 (38.1%)	18 (25.7%)
Grade 3 upper gastrointestinal toxicity		12 (13.2%)	3 (14.3%)	9 (12.9%)
Grade ≥2 upper gastrointestinal toxicity		38 (41.8%)	11 (52.4%)	27 (38.6%)
Grade ≥2 nausea or vomiting		27 (29.7%)	12 (57.1%)	15 (21.4%)
Grade ≥2 anorexia		13 (14.3%)	3 (14.3%)	10 (14.3%)
Grade ≥2 abdominal pain		13 (14.3%)	3 (14.3%)	10 (14.3%)
Grade ≥2 weight loss		8 (8.8%)	1 (4.8%)	7 (10%)
Grade ≥2 gastrointestinal bleeding		9 (9.9%)	0 (0%)	9 (12.9%)
Baseline grade ≥2 upper gastrointestinal toxicity		10 (11.0%)	1 (4.8%)	9 (12.9%)
Baseline grade ≥2 nausea or vomiting		2 (2.2%)	1 (4.8%)	1 (1.4%)
Baseline grade ≥2 abdominal pain		5 (5.5%)	0	5 (7.1%)
Baseline grade ≥2 weight loss		3 (3.3%)	0	3 (4.3%)

BMI, body mass index; ECOG, Eastern Cooperative Oncology Group.

On univariate analysis, the stomach cumulative V_{35Gy} (volume receiving 35 Gy or more) was significantly higher for patients with toxicity: median for patients with grade 0–1 UGIT = 30.9 cm^3 (interquartile range 12.0–62.0) compared with median for patients with grade ≥2 = 39.4 cm^3 (24.8–101.7), $P = 0.036$. For the differential DVH parameter V_{35Gy} (volume receiving between 35 and 40 Gy) the statistical significance of the difference between these groups was greater (median 6.36 cm^3 [interquartile range 1.89–18.4] versus 10.2 cm^3 [5.24–18.4], $P = 0.035$) and the strongest association with toxicity risk was seen for differential DVH region $V_{35–45Gy}$ (median 12.9 cm^3 [interquartile range 3.56–23.4] versus 16.4 cm^3 [9.75–32.0], $P = 0.033$). On univariate logistic regression, the stomach $V_{35–45Gy}$ was predictive of risk of UGIT grade ≥2 (odds ratio 1.035, 95% confidence interval 1.007–1.063, $P = 0.014$). The median stomach $V_{35–45Gy}$ for patients increased sequentially with toxicity grade (see Table 2 and Figure 1) and on univariate ordinal regression was predictive of UGIT grade (odds ratio 1.023, 1.003–1.044, $P = 0.022$). As a predictor of the risk of UGIT grade ≥2, the AUC was 0.632 (0.516–0.747). Toxicity incidence was 33/66 (50%) for patients with $V_{35–45Gy}$ above the optimal discriminatory threshold of 7.1 cm^3, and 5/25 (20%) below, with sensitivity of 0.868 and specificity 0.377. Using an alternative threshold of 30 cm^3 (the volume threshold with the second-highest Youden index), the risk was 13/20 (65%) above and 25/71 (35%) below. Duodenum dose–volume parameters did not predict toxicity risk or severity in any cohort.

Table 2 Stomach $V_{35–45Gy}$ according to grade of acute upper gastrointestinal toxicity, for the pooled cohort ($n = 91$)

Toxicity grade		0	1	2	3
Stomach $V_{35–45Gy}$ [cm^3]	Median	9.70	12.93	15.31	28.02

	Interquartile range	1.72–22.00	5.54–23.37	9.95–31.51	9.74–32.26

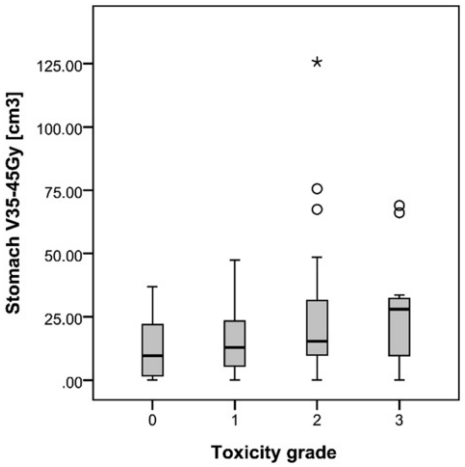


Fig 1 Box and whisker plot of stomach $V_{35–45Gy}$ against maximal acute upper gastrointestinal toxicity grade.

Table 3 shows the results of univariate logistic regression for risk of UGIT grade ≥2 for clinical factors. Risk was higher in patients with ECOG performance status 1 or higher, but the difference was not significant (odds ratio 1.661, 95% confidence interval 0.717–3.852, $P = 0.235$). For patients in the SCALOP study the amount of weight loss during induction chemotherapy and concomitant chemotherapy (risk was highest for patients receiving gemcitabine) was a significant predictor of risk and severity of UGIT: odds ratio with 95% confidence interval 1.199 (1.040–1.382) and 3.632 (1.300–10.151), respectively. Increasing age was a significant predictor of toxicity risk and severity only for the ARCII cohort; odds ratio 1.344 (95% confidence interval 1.015–1.780).

Table 3 Collated results of univariate logistic regression of clinical factors against risk of grade ≥2 upper gastrointestinal toxicity, showing odds ratio (with 95% confidence interval)

	Combined cohort	ARCII	SCALOP
Age	1.007 (0.955–1.060)	1.344 (1.015–1.780)*	0.969 (0.913–1.027)
BMI	0.970 (0.881–1.068)	1.082 (0.861–1.359)	0.945 (0.843–1.060)
Tumour volume	1.014 (0.995–1.033)	1.007 (0.962–1.055)	1.007 (0.989–1.026)
ECOG performance status (0–1 versus 2)	1.961 (0.412–9.322)	3.375 (0.290–39.322)	0.788 (0.068–9.139)
Patient sex†	0.591 (0.255–1.370	0.556 (0.098–3.148)	0.607 (0.230–1.604)
Tumour location (head versus other)	1.795 (0.723–4.454)	4.800 (0.682–33.798)	1.225 (0.419–3.582)
Baseline upper gastrointestinal toxicity (any grade)	1.540 (0.649–3.656)	1.500 (0.195–11.536)	0.800 (0.155–4.123)
Weight loss during induction chemotherapy			1.199 (1.040–1.382)*
Concomitant chemotherapy‡			3.632 (1.300–10.151)*

BMI, body mass index; ECOG, Eastern Cooperative Oncology Group.

* Significance at $P < 0.05$ level.

† Risk is higher for females, in both studies.

‡ Gemcitabine versus capecitabine, risk is highest for gemcitabine arm.

The optimal multivariate model incorporated patient sex, chemotherapy regimen (gemcitabine versus capecitabine versus ARCII regimen) and stomach $V_{35–45Gy}$ as a continuous variable. This achieved overall predictive accuracy of 71.4% and AUC of 0.745; the ROC curve is shown in Figure 2. Model coefficients are included in the supplementary material (see Appendix). Age, performance status, tumour volume and tumour location were not retained in the model. Figure 3 shows the observed against predicted incidence of toxicity for three evenly sized groups divided by predicted risk. On five-fold internal cross-validation the mean AUC on the training set was 0.736 (standard deviation 0.035) but on an unseen test cohort was 0.598 (standard deviation 0.188). The predictive factor most frequently retained in the model was chemotherapy group (retained in four models), followed by stomach $V_{35–45Gy}$ (three models), sex and GTV volume (two models each). Tumour volume and stomach $V_{35–45Gy}$ were significantly correlated with each other (Spearman's correlation coefficient 0.211, $P = 0.044$). Stomach $V_{35–45Gy}$ was highest in ARCII (median 21.0, interquartile range 10.6–30.1), next highest in the gemcitabine arm (14.5, 7.9–28.7) and lowest in the capecitabine arm (9.8, 2.9–22.3), and significantly different between these groups (Kruskal–Wallis test, $\chi^2 = 7.197$, $P = 0.027$).

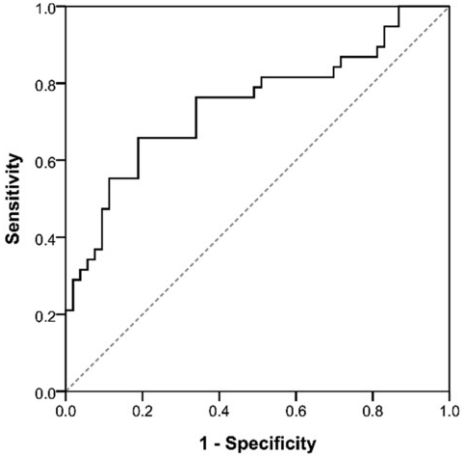


Fig 2 Receiver operator characteristic (ROC) curve for multivariable logistic regression model incorporating patient sex, chemotherapy treatment arm and stomach $V_{35–45Gy}$.

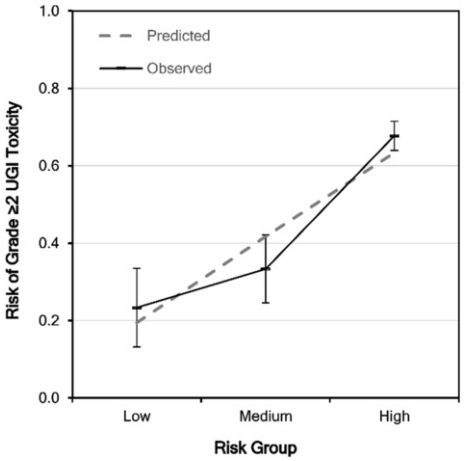


Fig 3 Plot of observed proportional incidence of acute upper gastrointestinal toxicity grade ≥2 among three evenly sized cohorts divided according to risk (low, medium, high) as predicted by multivariable logistic regression model incorporating patient sex, chemotherapy treatment arm and stomach $V_{35–45Gy}$.

Overall survival was significantly worse in patients who experienced grade ≥2 UGIT: median 10.8 months (95% confidence interval 11.9–17.7) versus 14.8 months (8.5–13.1), Log-rank 5.637, $P = 0.018$, and this association persisted when patients who went on to suffer progressive disease within 90 days of starting radiotherapy were excluded from the analysis (this was the case for three patients from ARCII and three from SCALOP). Patients with acute grade ≥2 UGIT were not less likely to complete radiotherapy: 41/53 patients (71.1%) received all fractions compared with 27/38 (77.4%), $\chi^2 = 0.466$, $P = 0.495$. Progression-free survival was also worse for patients with acute UGIT,

but the difference was not significant: 6.1 months (95% confidence interval 4.1–8.2) compared with 8.2 (6.6–9), $P = 0.052$.

Discussion

We have analysed data from two prospective trials of chemoradiotherapy for LAPC and shown dose–volume and patient-related factors influencing risk and severity of acute gastrointestinal toxicity, which itself was associated with worse overall survival.

We found stomach dose–volume was predictive of toxicity risk and have derived thresholds that could be used to inform radiotherapy planning: for our patients, if $V_{35–45Gy}$ was kept below 30 cm^3, the risk of grade ≥2 toxicity was 35% or less; if kept below 7.1 cm^3, the risk was 20% or less. Similar findings have previously been reported (Table 4), however we believe our study is the first using pancreatic cancer chemoradiotherapy data to show with ordinal regression that increasing volume of stomach irradiated predicts increasing severity of toxicity. This result emphasises the fact that the stomach should always be accurately contoured for patients receiving this treatment and suggests that the more effort is made to minimise stomach dose–volume, the greater the benefit to the patient in terms of both reduced risk and severity of toxicity. Although tumour volume (as indicated by GTV volume) and stomach $V_{35–45Gy}$ were significantly correlated with each other, as would be expected, the evidence indicates that the stomach dose–volume is associated with increased risk of toxicity independently of tumour size. As shown in Table 3, GTV was not itself associated with increased toxicity on univariate analysis, and although it did get retained in two of the five multivariable models generated using the internal cross-validation, it was not mutually exclusive for retention with the stomach $V_{35–45Gy}$, which would be expected if the two parameters were simply surrogates for each other.

Table 4 Findings of published analyses of the dose–volume relationships of gastrointestinal organs at risk in pancreatic cancer radiotherapy, or of treatment of other conditions in which relationships with stomach dose–volume histograms are identified. Risk comparison, where reported, indicates proportional incidence of specified toxicity for patients whose radiotherapy plans achieved or did not achieve the specified threshold value

Reference	Cancer site	*n*	Concurrent chemotherapy	Radiotherapy dose schedule	Radiotherapy technique	Toxicity outcome	Predictive dose–volume parameter and threshold values (with risk comparison) where reported
[7]	LAPC	74	Gemcitabine	36 Gy in 15 fractions	3D	'Duodenal'	Duodenum gEUD
[9]	LAPC	46	Gemcitabine ± erlotinib	36 Gy in 15 fractions	3D (40)/IMRT (6)	Grade ≥3 gastrointestinal	All patients: Duodenum V_{25Gy} 45% (8% versus 48%) Non-erlotinib: Duodenum V_{35Gy} 20% (0% versus 41%)
[10]	LAPC	40	Gemcitabine	54 Gy in 30 fractions	3D + IMRT boost	Acute gastrointestinal and upper gastrointestinal bleed	Stomach V_{50Gy} 16 cm^3 (9% versus 61%) Stomach D_{2cm3} 53.6 Gy~~ Gy~~ (0% versus 57%) StoDuo V_{50Gy} 33 cm^3 (0% versus 44%)
[5]	LAPC	106	Gemcitabine ± 5-FU/capecitabine ± EGFRi	50.4 Gy in 28 fractions	3D (75)/IMRT (31)	Grade ≥ 2 'Duodenal'	Duodenum V_{55Gy} 1 cm^3
[11]	LAPC	61	Capecitabine or 5-FU	44 ± 15 Gy boost, in 15 fractions	IMRT	Grade ≥2 gastrointestinal	Stomach V_{20Gy} 31% Duodenum V_{40Gy} 16% Duodenum V_{45Gy} 2.6%
[8]	HCC	73	None	36 Gy in 12 fractions	3D	Grade ≥ 3 'gastroduodenal'	Gastroduodenal~~Gastroduodenum~~ V_{35Gy} 5% (4% vs 48%)
[12]	HCC	90	None	33 Gy in 11 fractions	3D	Grade ≥ 2 'gastroduodenal'	Stomach V_{25Gy} 6.3% (2.9% vs 57.1%) Duodenum V_{35Gy} 5.4% (9.4% vs 45.9%)

[30]	Hepatic tumours	92	Hepatic arterial chemotherapy	1.5 Gy per fraction BD with chemotherapy or 1.8–3 Gy per fraction QDS without	3D	Upper gastrointestinal bleed	Stomach TD_{50} 62 Gy Duodenum TD_{50} 180 Gy
[31]	Hepatic tumours	116	IV 5-FU or hepatic FUDR	54 Gy in 28 fractions or 1.5 Gy per fraction BD regimen with FUDR	3D	Gastric bleeding	Stomach D_{max}

LAPC, locally advanced pancreatic cancer; 3D, three-dimensional conformal radiotherapy; gEUD, generalised equivalent uniform dose; IMRT, intensity-modulated radiotherapy; 5-FU, 5-fluorouracil; EGFRi, epithelial growth factor receptor inhibitor; HCC, hepatocellular carcinoma; StoDuo, combined stomach and duodenum volume; IV, intravenous; hepatic FUDR, intrahepatic infusion of 5-fluorodeoxyuridine; NS, not specified; D_{max}, maximum dose; D_{mean}, mean dose.

We note the slightly higher predictive performance of a differential DVH parameter rather than a conventional cumulative DVH and suggest that a more specific measurement has shown a clearer association with outcomes. Relative (proportional) volume parameters were also investigated, but due to very small differences in values between patients, these parameters led to unstable results in regression modelling.

The incidence of toxicity was somewhat different between these two clinical trials, although the causal relationship for this is difficult to elucidate, due to possible confounding differences in the treatments delivered. Not only was the concomitant chemotherapy different, but in ARCII the radiotherapy field was larger in order to irradiate draining lymph nodes, and the dose to the tumour was higher, while the SCALOP patients had received induction chemotherapy. Despite this heterogeneity, we were able to show that patient sex and the stomach dose–volume remained significant contributors to overall risk in the multivariable model for the pooled cohort. It is not clear why, in both studies, women were at greater risk of toxicity than men, and there were no significant associations between sex and other factors, such as age, performance status, tumour size/location or presence of baseline symptoms (included in the supplementary data; see Appendix). In the SCALOP cohort, weight loss during induction chemotherapy was associated with increasing risk of acute toxicity after concomitant radiotherapy and the risk of both acute and late gastrointestinal toxicity after pelvic radiotherapy has previously been shown to be higher in women with low body weight [24,25].

Predictive modelling of radiotherapy toxicity can be prone to overfitting, with disappointing results on external validation [26]. We believe our work is the first in this clinical treatment paradigm to undertake internal cross-validation to assess generalisability and restrain potentially over-optimistic estimates of prospective model performance. Our results indicate a degree of overfitting may have occurred, with model performance on unseen data weaker than on training cohorts. This has been recognised to be a possible drawback of using stepwise predictor selection methods [27].

The association between toxicity and survival does not seem to be explained by any single factor, such as failure to complete radiotherapy, and instead may represent a combination of effects that are not individually significant. For multiple clinical factors (tumour size, age, body mass index, performance status, recent weight loss, symptoms at baseline) the direction of effect associated with risk of toxicity would also be expected to be associated with worse outcomes overall (Table 2).

Several publications have analysed gastrointestinal organ at risk dose–volumes in chemoradiotherapy for pancreatic cancer (Table 3) and a number have identified stomach dose–volume as predictive of toxicity. These include Nakamura *et al.* [10] who found the best predictors of grade ≥ 2 acute gastrointestinal toxicity were stomach V_{50Gy} and D_{2cm3} (highest dose to ≥2 cm^3). Cattaneo *et al.* [11] analysed outcomes for patients treated with relatively hypofractionated chemoradiotherapy (prescribed dose 44.25 Gy in 15 fractions) and found stomach V_{20Gy} was the best predictor of acute toxicity, whereas duodenum V40Gy and V45Gy were correlated specifically with ‘anatomical’ damage. Recently Shinoto *et al.* [28] have published analysis of carbon ion chemoradiotherapy dose–volume parameters, although the values themselves may not be directly comparable with those from photon radiotherapy. It is worth noting that in our analysis we failed to show associations of gastrointestinal toxicity with duodenum dose–volume, despite other studies having shown evidence of a relationship [29]. Murphy *et al.* [7] showed a trend for association of increased duodenal dose–volume with grade ≥ 3 gastrointestinal toxicity in patients treated with concomitant gemcitabine, but Huang *et al.* [9] found that the best predictor of toxicity in similar patients was the duodenum V_{35Gy} (V_{25Gy} when including patients who received concomitant erlotinib). Kelly *et al.* [5] examined the largest cohort in this group and found duodenum V_{55Gy} to be the best predictor of risk. Stomach dose–volume parameters have also been shown to predict toxicity risk in radiotherapy for hepatic tumours [30]: the combined ‘gastroduodenal~~gastroduodenum~~’ V_{35Gy} [8] and the stomach V_{25Gy} [12] have been shown to predict ‘gastroduodenal’ toxicity, and stomach D_{max} has shown an association with gastric bleeding [31].

The stomach is where ingested food bathes in acidic secretions before passing into the duodenum. Irradiation of the stomach can cause loss of appetite, nausea and vomiting within hours. The mechanism is thought to involve $5\text{-}HT_3$ secretion by enterochromaffin cells and vagal neural impulses to the brainstem vomiting centre, which stimulates expulsion of the stomach contents [32]. Gastric mucosal inflammation develops over subsequent days, with denudation of the gastric mucosa leading to painful superficial ulceration with possible gastrointestinal bleeding [33]. We have added to the growing evidence base for the importance of stomach dose–volume in predicting toxicity in chemoradiotherapy for LAPC, although with limitations. Due to the number of parameters that are tested in DVH analysis, there is a risk of type I error (a ‘false positive’, in which the null hypothesis is rejected despite it being true). Conversely, when analysing observational data, negative results may be due to insufficient cases or insufficient variation in the independent factors. However, neither trial protocol enforced radiotherapy planning constraints for the

stomach or duodenum, potentially allowing the dose–volume to vary between patients. Although this dataset comprises the largest prospective pancreas radiotherapy trial cohort analysed to date, the sample size remains relatively small and the method of pooling data from two trials increases risk heterogeneity and of confounds. True external validation of these findings is required and is intended to be conducted using data from an ongoing randomised study investigating chemoradiotherapy in LAPC.

An analysis of toxicity is hindered by the complex relationship of symptoms with the effects of treatment and disease. Conflating multiple symptoms to derive a single end point will obscure causes of specific symptoms, but although end points such as ulceration on endoscopy are highly objective, these severe events represent only a proportion of all relevant morbidity. Patient-reported outcomes may represent an improvement over conventional methods but are not established in upper abdominal radiotherapy toxicity assessment. Spatial parameters describing dose to the structure wall could be more appropriate than dose–volume parameters [34], but due to interfractional motion and deformation the delivered dose to the stomach may significantly differ from the planned dose, even if breath-hold methods are implemented [35], and ideally outcomes would be analysed against delivered dose.

Conclusions

In chemoradiotherapy for LAPC, the volume of stomach irradiated to a moderately high dose (35–45 Gy) predicts the incidence and severity of acute toxicity, whereas other factors predictive of risk include sex and concomitant chemotherapy agents. Reducing the volume of stomach irradiated to doses in this range may reduce the risk and severity of toxicity for these patients.

Acknowledgements

DW is funded by CRUK, grant C5255/A15935. MA is funded by CRUK, grant C8225/A21133. MP is funded by CRUK, grant C5255/A15935. SM is part-funded by Oxford Biomedical Research Centre. MAH is funded by MRC, grant MC_UU_0001 ~~MC_PC_12001/2~~. SCALOP data were supplied by Cardiff University South East Wales Trials Unit. ARCII data were supplied by Oxford University Clinical Trials Research Unit.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.clon.2018.02.067.

References

[1] P. Hammel, F. Huguet, J.L. van Laethem, D. Goldstein, B. Glimelius, P. Artru, et al., Effect of chemoradiotherapy vs chemotherapy on survival in patients with locally advanced pancreatic cancer controlled after 4 months of gemcitabine with or without erlotinib: the LAP07 randomized clinical trial, *JAMA* **315**, 2016, 1844–1853.

[2] I.C. Moraru, A. Tai, B. Erickson and X.A. Li, Radiation dose responses for chemoradiation therapy of pancreatic cancer: an analysis of compiled clinical data using biophysical models, *Pract Radiat Oncol* **4**, 2014, 13–19.

[3] C.J. McGinn, M.M. Zalupski, I. Shureiqi, J.M. Robertson, F.E. Eckhauser, D.C. Smith, et al., Phase I trial of radiation dose escalation with concurrent weekly full-dose gemcitabine in patients with advanced pancreatic cancer, *J Clin Oncol* **19**, 2001, 4202–4208.

[4] C.H. Crane, J.A. Antolak, I.I. Rosen, K.M. Forster, D.B. Evans, N.A. Janjan, et al., Phase I study of concomitant gemcitabine and IMRT for patients with unresectable adenocarcinoma of the pancreatic head, *Int J Gastrointest Cancer* **30**, 2001, 123–132.

[5] P. Kelly, P. Das, C.C. Pinnix, S. Beddar, T. Briere, M. Pham, et al., Duodenal toxicity after fractionated chemoradiation for unresectable pancreatic cancer, *Int J Radiat Oncol Biol Phys* **85**, 2013, e143–e149.

[6] A. Elhammali, M. Patel, B. Weinberg, V. Verma, J. Liu, J.R. Olsen, et al., Late gastrointestinal tissue effects after hypofractionated radiation therapy of the pancreas, *Radiat Oncol* **10**, 2015, 186.

[7] J.D. Murphy, S. Adusumilli, K.A. Griffith, M.E. Ray, M.M. Zalupski, T.S. Lawrence, et al., Full-dose gemcitabine and concurrent radiotherapy for unresectable pancreatic cancer, *Int J Radiat Oncol Biol Phys* **68**, 2007, 801–808.

[8] H. Kim, D.H. Lim, S.W. Paik, B.C. Yoo, K.G. Koh, J.H. Lee, et al., Predictive factors of gastroduodenal toxicity in cirrhotic patients after three-dimensional conformal radiotherapy for hepatocellular carcinoma, *Radiother Oncol* **93**, 2009, 302–306.

[9] J. Huang, J.M. Robertson, H. Ye, J. Margolis, L. Nadeau and D. Yan, Dose-volume analysis of predictors for gastrointestinal toxicity after concurrent full-dose gemcitabine and radiotherapy for locally advanced pancreatic adenocarcinoma, *Int J Radiat Oncol Biol Phys* **83**, 2012, 1120–1125.

[10] A. Nakamura, K. Shibuya, Y. Matsuo, M. Nakamura, T. Shiinoki, T. Mizowaki, et al., Analysis of dosimetric parameters associated with acute gastrointestinal toxicity and upper gastrointestinal bleeding in locally

advanced pancreatic cancer patients treated with gemcitabine-based concurrent chemoradiotherapy, *Int J Radiat Oncol Biol Phys* **84**, 2012, 369–375.

[11] G.M. Cattaneo, P. Passoni, B. Longobardi, N. Slim, M. Reni, S. Cereda, et al., Dosimetric and clinical predictors of toxicity following combined chemotherapy and moderately hypofractionated rotational radiotherapy of locally advanced pancreatic adenocarcinoma, *Radiother Oncol* **108**, 2013, 66–71.

[12] H. Yoon, D. Oh, H.C. Park, S.W. Kang, Y. Han, D.H. Lim, et al., Predictive factors for gastroduodenal toxicity based on endoscopy following radiotherapy in patients with hepatocellular carcinoma, *Strahlenther Onkol* **189**, 2013, 541–546.

[13] S. Mukherjee, C.N. Hurt, J. Bridgewater, S. Falk, S. Cummins, H. Wasan, et al., Gemcitabine-based or capecitabine-based chemoradiotherapy for locally advanced pancreatic cancer (SCALOP): a multicentre, randomised, phase 2 trial, *Lancet Oncol* **14**, 2013, 317–326.

[14] J.M. Wilson, E. Fokas, S.J. Dutton, N. Patel, M.A. Hawkins, C. Eccles, et al., ARCII: a phase II trial of the HIV protease inhibitor Nelfinavir in combination with chemoradiation for locally advanced inoperable pancreatic cancer, *Radiother Oncol* **119**, 2016, 306–311.

[15] National Cancer Institute, Common terminology criteria for adverse events v3.0 (CTCAE), 2006.

[16] National Cancer Institute, Common terminology criteria for adverse events v4.0 (CTCAE), 2009.

[17] J.D. Cox, J. Stetz and T.F. Pajak, Toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European Organization for Research and Treatment of Cancer (EORTC), *Int J Radiat Oncol Biol Phys* **31**, 1995 1341–1346.

[18] M.M. Oken, R.H. Creech, D.C. Tormey, J. Horton, T.E. Davis, E.T. McFadden, et al., Toxicity and response criteria of the eastern cooperative oncology group, *Am J Clin Oncol* **5**, 1982, 649–655.

[19] E. Fokas, C. Clifford, E. Spezi, G. Joseph, J. Branagan, C. Hurt, et al., Comparison of investigator-delineated gross tumor volumes and quality assurance in pancreatic cancer: analysis of the pretrial benchmark case for the SCALOP trial, *Radiother Oncol* **117**, 2015, 432–437.

[20] E. Fokas, E. Spezi, N. Patel, C. Hurt, L. Nixon, K.Y. Chu, et al., Comparison of investigator-delineated gross tumour volumes and quality assurance in pancreatic cancer: analysis of the on-trial cases for the SCALOP trial, *Radiother Oncol* **120**, 2016, 212–216.

[21] S.K. Jabbour, S.A. Hashem, W. Bosch, T.K. Kim, S.E. Finkelstein, B.M. Anderson, et al., Upper abdominal normal organ contouring guidelines and atlas: a Radiation Therapy Oncology Group consensus, *Pract Radiat Onco* **4**, 2014, 82–89.

[22] J.O. Deasy, A.I. Blanco and V.H. Clark, CERR: a computational environment for radiotherapy research, *Med Phys* **30**, 2003, 979–985.

[23] W.J. Youden, Index for rating diagnostic tests, *Cancer* **3**, 1950, 32–35.

[24] P.J. Eifel, A. Jhingran, D.C. Bodurka, C. Levenback and H. Thames, Correlation of smoking history and other patient characteristics with major complications of pelvic radiation therapy for cervical cancer, *J Clin Oncol* **20**, 2002, 3651–3657.

[25] E.Y. Huang, C.C. Sung, S.F. Ko, C.J. Wang and K.D. Yang, The different volume effects of small-bowel toxicity during pelvic irradiation between gynecologic patients with and without abdominal surgery: a prospective study with computed tomography-based dosimetry, *Int J Radiat Oncol Biol Phys* **69**, 2007, 732–739.

[26] C. Mbah, H. Thierens, O. Thas, J. De Neve, J. Chang-Claude, P. Seibold, et al., Pitfalls in prediction modeling for normal tissue toxicity in radiation therapy: an illustration with the individual radiation sensitivity and mammary carcinoma risk factor investigation cohorts, *Int J Radiat Oncol Biol Phys* **95**, 2016, 1466–1476.

[27] C.J. Xu, A. van der Schaaf, C. Schilstra, J.A. Langendijk and A.A. van't Veld, Impact of statistical learning methods on the predictive power of multivariate normal tissue complication probability models, *Int J Radiat Onco Biol Phys* **82**, 2012, e677–e684.

[28] M. Shinoto, Y. Shioyama, A. Matsunobu, K. Okamoto, H. Suefuji, S. Toyama, et al., Dosimetric analysis of upper gastrointestinal ulcer after carbon-ion radiotherapy for pancreatic cancer, *Radiother Oncol* **120**, 2016, 140–144.

[29] D.L.P. Holyoake, M. Aznar, S. Mukherjee, M. Partridge and M.A. Hawkins, Modelling duodenum radiotherapy toxicity using cohort dose-volume-histogram data, *Radiother Oncol* **123**, 2017, 431–437.

[30] C.C. Pan, L.A. Dawson, C.J. McGinn, T.S. Lawrence and R.K. Ten Haken, Analysis of radiation-induced gastric and duodenal bleeds using the Lyman-Kutcher-Burman model, *Int J Radiat Oncol Biol Phys* **57**, 2003, S217–S218.

[31] M. Feng, D. Normolle, C.C. Pan, L.A. Dawson, S. Amarnath, W.D. Ensminger, et al., Dosimetric analysis of radiation-induced gastric bleeding, *Int J Radiat Oncol Biol Phys* **84**, 2012, e1–6.

[32] M. Feldman, L.S. Friedman and L.J. Brandt, Sleisenger and Fordtran's gastrointestinal and liver disease, 10th ed., 2015, Elsevier; Philadelphia.

[33] E.J. Hall and A.J. Giaccia, Radiobiology for the radiologist, 6th ed., 2006, Lippincott Williams & Wilkins; London.

[34] A. Witztum, B. George, S. Warren, M. Partridge and M.A. Hawkins, Unwrapping 3D complex hollow organs for spatial dose surface analysis, *Med Phys* **43**, 2016, 6009.

[35] A. Nakamura, K. Shibuya, M. Nakamura, Y. Matsuo, T. Shiinoki, M. Nakata, et al., Interfractional dose variations in the stomach and the bowels during breathhold intensity-modulated radiotherapy for pancreatic cancer: implications for a dose-escalation strategy, *Med Phys* **40**, 2013, 021701.

Appendix A. Supplementary data

The following are the supplementary data related to this article:

Multimedia Component 1

mmc1

Multimedia Component 2

mmc2

Highlights

- Acute gastrointestinal toxicity was analysed for chemoradiotherapy of locally advanced pancreatic cancer in two prospective phase II ~~clinica~~l trials.
- Stomach dose–volume (in the ‘moderate’ dose region of 35–45 Gy) and patient clinical factors ~~(sex and regimen of concomitant chemotherapy)~~ were identified as predictive for risk of toxicity on multivariate analysis.
- Stomach dose–volume was ~~also~~ predictive of severity of toxicity. ~~We understand~~ ~~t~~This is the first time this has been shown in pancreatic cancer chemoradiotherapy.
- ~~Internal cross-validation was carried out, and although performance of the model on unseen test partitions was poor, we found no prior publications of internal or external validation of toxicity models in pancreatic cancer chemoradiotherapy.~~
- ~~Toxicity was shown to be associated with reduced overall survival but this was not attributable to failure to complete treatment.~~ (deleted in error and now can not change)
- Radiotherapy planning optimisation objectives have been derived with the aim of ~~helping to~~ reduce~~e~~ing risk for prospective patients.

Queries and Answers

Query: The country name has been inserted for the third author affiliation. Please check, and correct if necessary.

Answer: correct

Query: Please check that the affiliation “3” link with their correct departments, institutions, and locations, and correct if necessary.

Answer: correct

Query: Please check the address for the corresponding author that has been added here, and correct if necessary.

Answer: correct

Query: Please provide the grant number for “Oxford Biomedical Research Centre”, “Cardiff University” and “Oxford University” if any.

Answer: not applicable

Query: Please check the layout of Table 1.

Answer: all good

Query: Highlights should consist of 3-5 bullet points. There are 6 bullet points provided. Please edit the highlights to meet the requirement.

Answer: bullet point removed

Query: Highlights should only consist of 125 characters per bullet point, including spaces. The highlights provided are too long; please edit them to meet the requirement.

Answer: changes made

Query: Please confirm that given names and surnames have been identified correctly and are presented in the desired order and please carefully verify the spelling of all authors’ names.

Answer: Yes