

## Scientific Article

# A simulation study to assess the potential impact of developing normal tissue complication probability models with accumulated dose

Molly M. McCulloch MS <sup>a,b,c,\*</sup>, Daniel G. Muenz PhD <sup>d</sup>,  
Matthew J. Schipper PhD <sup>b,d</sup>, Michael Velec PhD <sup>e,f,g</sup>,  
Laura A. Dawson MD <sup>e,f,g</sup>, Kristy K. Brock PhD <sup>c</sup>

<sup>a</sup> Department of Nuclear Engineering and Radiological Sciences, University of Michigan, Ann Arbor, Michigan

<sup>b</sup> Department of Radiation Oncology, University of Michigan, Ann Arbor, Michigan

<sup>c</sup> Department of Imaging Physics, The University of Texas, MD Anderson Cancer Center, Houston, Texas

<sup>d</sup> Department of Biostatistics, University of Michigan, Ann Arbor, Michigan

<sup>e</sup> Radiation Medicine Program, Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada

<sup>f</sup> Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada

<sup>g</sup> Department of Radiation Oncology, University of Toronto, Toronto, Ontario, Canada

Received 7 May 2018; accepted 11 May 2018

## Abstract

**Purpose:** This study aimed to analyze the potential clinical impact of the differences between planned and accumulated doses on the development and use of normal tissue complication probability (NTCP) models.

**Methods and Materials:** Thirty patients who were previously treated with stereotactic body radiation therapy for liver cancer and for whom the accumulated dose was computed were assessed retrospectively. The linear quadratic equivalent dose at 2 Gy per fraction and generalized equivalent uniform dose were calculated for planned and accumulated doses. Stomach and duodenal Lyman-Kutcher-Burman NTCP models ( $\alpha/\beta = 2.5$ ;  $n = .09$ ) were developed on the basis of planned and accumulated generalized equivalent uniform doses and the differences between the models assessed. In addition, the error in determining the probability of toxicity on the basis of the planned dose was evaluated by comparing planned doses in the NTCP model that were created from accumulated doses.

**Results:** The standard, planned-dose NTCP model overestimates toxicity risk for both the duodenal and stomach models at doses that are below approximately 20 Gy (6 fractions) and underestimates toxicity risk for doses above approximately 20 Gy (6 fractions). Building NTCP models with accumulated rather than planned doses changes the predicted risk by up to 16% (mean: 6%; standard

Conflicts of interest: Dr. Kristy K. Brock reports a Licensing Agreement with RaySearch Laboratories. Dr. Laura A. Dawson reports other from Merc and Sirtex outside of the submitted work and has a patent licensing agreement that is licensed to Raysearch Oncology.

Sources of support: This work is supported by Grant No. 5R01CA124714-02 from the U.S. National Institutes of Health. The clinical trials were supported by Grant No. 18207 from the National Cancer Institute of Canada and Elekta Oncology Systems, and Grant No. P01-CA059872.

\* Corresponding author. The University of Texas, MD Anderson Cancer Center, 3SCR2.3623, 1881 East Rd, Houston, TX 77030.

E-mail address: [mollyma@umich.edu](mailto:mollyma@umich.edu) (M.M. McCulloch).

<https://doi.org/10.1016/j.adro.2018.05.003>

2452-1094/© 2018 The Author(s). Published by Elsevier Inc. on behalf of the American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

deviation: 7%) for duodenal toxicity and 6% (mean: 2%; standard deviation: 2%) for stomach toxicity. For a protocol that plans a 10% iso-toxicity risk to the duodenum, a 15.7 Gy (6 fractions) maximum dose constraint would be necessary when using standard NTCP models on the basis of a planned dose and a 17.6 Gy (6 fractions) maximum dose would be allowed when using NTCP models on the basis of accumulated doses.

**Conclusions:** Assuming that accumulated dose is a more accurate representation of the true delivered dose than the planned dose, this simulation study indicates the need for prospective clinical trials to evaluate the impact of building NTCP models on the basis of accumulated doses.

© 2018 The Author(s). Published by Elsevier Inc. on behalf of the American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

Stereotactic body radiation therapy (SBRT) delivers high-precision external beam radiation therapy treatment in 2 to 5 fractions. The goal of SBRT is to deliver an ablative dose to the tumor while sparing normal tissue, which leads to lower toxicity.<sup>1</sup> Even though SBRT has been demonstrated to increase local control of liver cancer,<sup>2</sup> toxicity risks must still be acknowledged for liver and luminal gastrointestinal (GI) structures.<sup>3</sup> Dose escalation has the potential to improve local control.<sup>4</sup> However, the increase in dose is often limited by normal tissue toxicity risk. Previous studies<sup>5</sup> indicate that approximately 30% of patients are dose limited based on GI toxicity.

Overestimating toxicity risk can lead to conservative treatment for the patient and potentially lead to lower chances of tumor control. Underestimating toxicity risk can subject the patient to unplanned risks. Therefore, an accurate understanding of toxicity to normal tissue is critical. Significant efforts have been made in the development and validation of accurate normal tissue complication probability (NTCP) models,<sup>6-9</sup> which aim to characterize the correlation between dose and the likelihood of side effects.<sup>10</sup> Specifically, Lyman-Kutcher-Burman (LKB) NTCP models have been used to investigate the dose-volume response for liver cancer.<sup>11</sup>

LKB NTCP models have been developed previously for duodenal toxicity,<sup>7</sup> which demonstrates that the model can predict outcomes after SBRT. In addition, a separate investigation showed that the Lyman NTCP model can predict gastric bleeding,<sup>6</sup> which demonstrates that patients with cirrhosis are at increased risk. Both studies built the NTCP models on the basis of the planned dose for each patient. Recent retrospective studies have shown that the planned dose differs from the accumulated dose<sup>5,12,13</sup> especially when evaluating the maximum dose to luminal GI structures such as the duodenum and stomach for which the differences could reach 42% and 14%, respectively, of the prescribed dose.<sup>5</sup> Building NTCP models on the basis of accumulated doses has been hypothesized to lead to a substantially different model, which may have a clinical impact. The goal of this work is to evaluate how the known uncertainty between planned and delivered dose limits translates into potential uncertainties in NTCP modeling and determine

the differences in the model if accumulated dose is used in the derivation of the model parameters.

Previous work supports the hypothesis that accumulated dose can improve the understanding of clinical outcomes of SBRT, which demonstrates that the accumulated dose to the gross tumor volume more strongly predicts for total time to local progression.<sup>13</sup> The current study builds on this work and assesses the impact of accumulated dose on NTCP models by comparing LKB NTCP models on the basis of planned doses with those based on accumulated doses. Preceding clinical trials, the aim of this study is to assess the potential differences in the development of the NTCP models using accumulated versus planned doses. The second aim is to apply these models for patient-specific assessments.

## Methods and materials

### Patient data

A previous study retrospectively evaluated the deviations between planned and accumulated doses to tumors and normal tissues in the liver SBRT of 30 patients.<sup>5</sup> Patients in a trial that was approved by an institutional review board were treated under free-breathing conditions with 6 fractions to an individualized, risk-based dose of 27 Gy to 30 Gy. Daily cone beam computed tomography (CT) guidance was performed. For each patient, the dose was accumulated by registering the daily inhale/exhale cone beam CT to the planning CT scan with Morfeus, which is an in-house, biomechanical, model-based, deformable registration algorithm (DIR).<sup>14</sup> For each of the 30 patients, the planning and accumulated doses for the stomach and duodenum were used in the Lyman NTCP model.

### Dose accumulation

The dose accumulation was previously reported<sup>5</sup> and will be briefly described here for completeness. Morfeus was used to accumulate the dose for each patient. The dose distribution from the static radiation treatment plan was

calculated in the treatment planning system on the exhale and inhale planning CT scans. Both dose distribution files were imported into Morfeus. In Morfeus, the organs are described using finite element models, which represent substructure within the organs through tetrahedral elements. Through the DIR, the locations of the tetrahedral elements can be tracked between the exhale and inhale images of 4-dimensional scans. The interpolation of dose matrices onto the position of each element at exhale, inhale, and 4 intermediate phases is performed to accumulate the dose. The weighting of each phase was determined by the time spent at that phase and the elemental position in the breathing cycle. Finally, a summation of the elemental dose over the breathing cycle was calculated to determine the accumulated dose.

The planned dose is defined as the static clinical planned dose, which was found by interpolating the dose matrix from the exhale CT scan onto the initial tetrahedral mesh that was constructed from the anatomy on the exhale CT scan. The calculation of this dose does not include any changes due to breathing motion or setup error.

Accumulated dose refers to the dose accumulated over SBRT and accounts for residual setup errors (eg, errors still present after daily image guidance), respiratory motion, shifting of the liver, and deformation. To account for setup errors and organ deformations that are present at each fraction, the exhale CT scan is deformed to the exhale cone beam CT scan of each fraction. Next, to account for daily breathing motion, the exhale CT scan is deformed from the exhale to the inhale cone beam CT scan of each fraction. To accumulate the dose, dose matrices from the exhale and inhale CT scans are interpolated onto the deformation map from the exhale-to-inhale cone beam CT scan of each fraction. Finally, the doses from the 6 fractions of treatment are summed. To accumulate the dose, DIR tracked anatomical motion and deformation in the dose matrices of the initial planning 4-dimensional CT.

### Differences between planned and accumulated doses

The percent change (PC) from planned to accumulated mean dose was analyzed for the duodenum ( $n = 30$ ), stomach for patients with primary (non-cirrhotic) liver cancer ( $n = 15$ ), and stomach for cirrhotic patients ( $n = 15$ ) and calculated by the equation:

$$PC_{dose} = \frac{Accumulated\ Dose - Planned\ Dose}{Planned\ Dose} * 100 \quad (1)$$

### Change in probability of toxicity

The delivered and accumulated doses (in the form of tabular dose-volume histograms) were biologically corrected

to the linear quadratic equivalent dose at 2 Gy per fraction using an  $\alpha/\beta$  ratio of 2.5.<sup>6</sup> The generalized equivalent uniform dose (gEUD) was calculated with  $n = .09$ . The Lyman NTCP model as shown in Equation 2 was used for toxicity modeling in this study:

$$NTCP = \Phi\left(\frac{gEUD - TD_{50}}{m \cdot TD_{50}}\right) \quad (2)$$

The function  $\Phi$  represents the NTCP model<sup>15</sup> where gEUD is evaluated using  $TD_{50} = 24.6$  and  $m = .23$  for duodenum,<sup>7</sup>  $TD_{50} = 22$  and  $m = .21$  for stomach and cirrhosis,<sup>6</sup> and  $TD_{50} = 56$  and  $m = .21$  for stomach and no cirrhosis.<sup>6</sup>

The PC in the NTCP model from planned to accumulated dose was calculated using the equation:

$$PC_{NTCP} = \frac{NTCP\ Accumulated - NTCP\ Planned}{NTCP\ Planned} * 100 \quad (3)$$

### Simulated toxicities for the duodenum and stomach

Toxicity models were simulated for the duodenum and stomach with cirrhosis. Stomach without cirrhosis was not modeled because there were no patients with an absolute change in NTCP >5% between the accumulated and planned doses. To simulate toxicity models for a larger cohort of patients than was used in the study, a resampling process was developed.

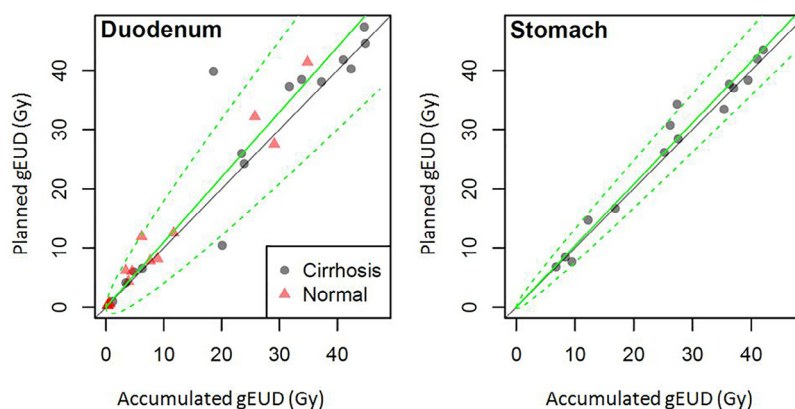
Figure 1 depicts the planned dose over the accumulated dose for the original 30 patients for the duodenum and the 15 patients with cirrhosis for the stomach. On the basis of the trends in Figure 1, a model was fit to predict the planned dose on the basis of the accumulated dose to resample data points:

$$d_{plan} \sim N(d_{acc} + \beta d_{acc}, \sigma^2 d_{acc}) \quad (4)$$

where  $d_{plan}$  is the planned dose,  $d_{acc}$  is the accumulated dose, and  $\beta$  (bias function of dose) and  $\sigma^2$  are parameters estimated by the linear regression model in Figure 1. These parameters were estimated via maximum likelihood.

NTCP models are typically estimated using planned dose values. To quantify the effect of the differences between the planned and accumulated doses on model parameter estimation, the following steps were performed. First, the distribution of the planned dose was estimated as a function of the accumulated dose, which involved both estimation of systematic bias (ie,  $\beta \neq 1$  in Eq. 4) and random variation of the planned dose about the accumulated dose. If the 2 dose terms agree perfectly, there is neither bias (ie,  $\beta = 1$ ) nor variation (ie,  $\sigma = 0$ ).

Second, both toxicity values were simulated (using the accumulated dose NTCP model) and planned dose values (using Eq. 4) from the observed delivered dose values in



**Figure 1** Planned versus accumulated generalized equivalent uniform dose for the duodenum and stomach with cirrhosis. The solid green line depicts the mean, the dashed line shows the 95% confidence limits for regression, and the solid gray line represents that the planned and accumulated doses are equal.

the study ( $n = 30$ ). Larger sample sizes can overcome variations between the planned and accumulated doses. However, larger sample sizes do not overcome systematic bias. Thus, studies of 30, 150, or 600 patients were simulated by using each observed patient's delivered dose value 1, 5, or 20 times. Because a random toxicity outcome and random planned dose value were simulated using the relationship established in Equation 4, the multiple outcomes simulated from a single patient in the 30-patient dataset were distinct.

Third, for each of the simulated studies, the NTCP model parameters were estimated on the basis of the planned dose values. Finally, steps 2 and 3 were repeated 2000 times and the mean NTCP curve was plotted as well as the 10<sup>th</sup> and 90<sup>th</sup> percentiles of the curves.

### Deviations in probability of toxicity

The deviations in the probabilities of toxicity between the standard and accumulated NTCP models were quantified. The true probability of toxicity risk is unknown; therefore, this was estimated with the probability of toxicity risk of the  $d_{acc}$  derived from the accumulated dose model (NTCP<sub>1</sub>). The  $d_{plan}$  and  $d_{acc}$  (in maximum dose to 0.5 cc) of each patient in the study were correlated to the probability of toxicity risk using each NTCP model.

NTCP<sub>2</sub> was defined as the probability of toxicity risk of  $d_{plan}$  derived from the standard model and NTCP<sub>3</sub> was defined as the probability of toxicity risk of  $d_{plan}$  derived from the accumulated model. The error in the standard model was calculated as the difference between NTCP<sub>2</sub> and NTCP<sub>1</sub> and the error in the accumulated model between NTCP<sub>3</sub> and NTCP<sub>1</sub>.

### Results

The previously reported deviations of  $\geq 5\%$  in 70% of patients<sup>5</sup> translates into a deviation of  $>5\%$  NTCP for 57%

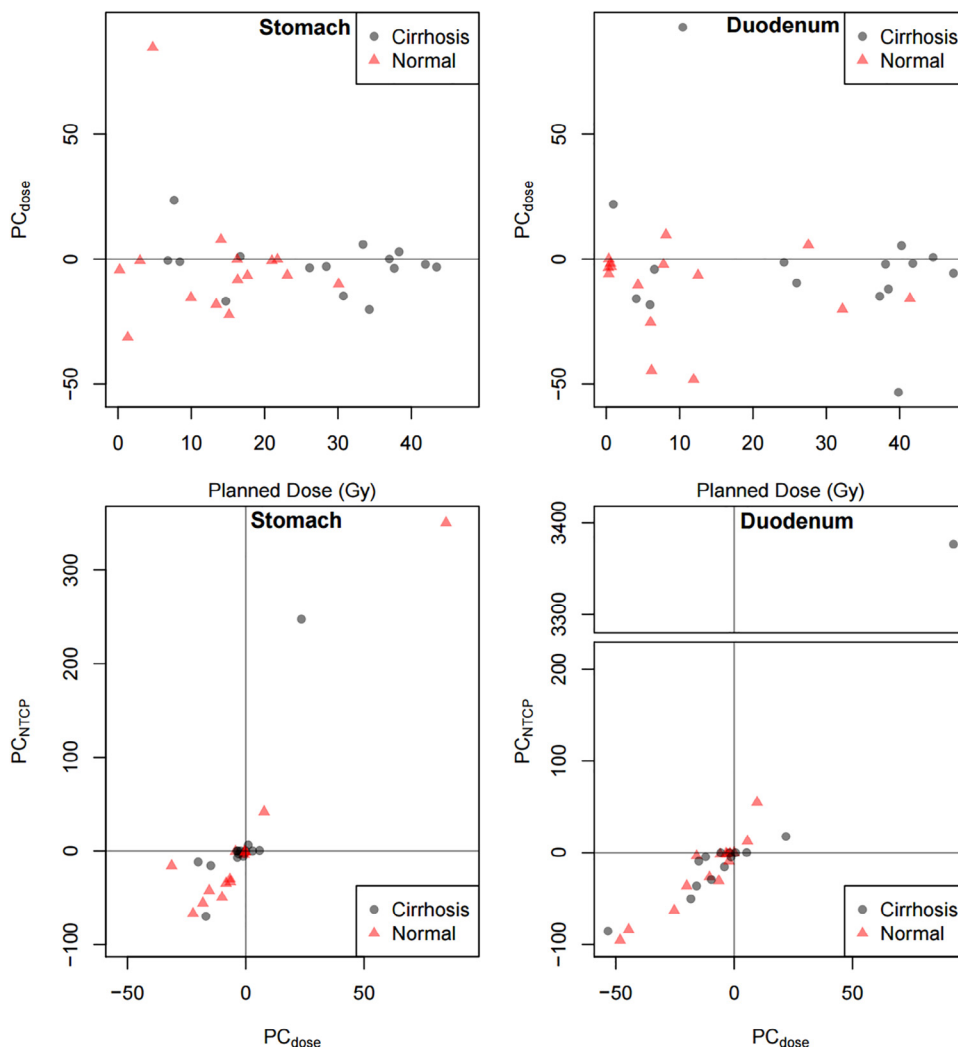
of patients for the duodenum and 60% of patients for the stomach with cirrhosis in the current study. Figures 2a and B depict the PC from planned to accumulated dose for the stomach and duodenum. The PC in dose versus planned dose (Gy) is plotted for both the stomach and duodenum in Figure 2. For both organs, the PC between the planned and accumulated doses is within 50% for all but 1 patient.

Figures 2c and D illustrate the PC in NTCP by PC in dose for both the stomach and duodenum. The PC in NTCP versus PC in dose is shown for organs of both patients with cirrhosis and non-cirrhosis. The PC in NTCP is potentially clinically meaningful for the duodenum (mean: 6%; standard deviation [SD]: 7%) but not for the stomach (mean: 2%; SD: 2%). The range of absolute magnitude of NTCP change is 0% to 16% for the duodenum and 0% to 6% for the stomach.

### Normal tissue complication probability models for the duodenum and stomach

For duodenal toxicity, the difference between NTCP on the basis of accumulated and planned doses is substantial (Fig 3). Table 1 shows the results from deriving error from the toxicity models for the duodenum for the 30 patients with dose accumulation. The toxicity risk using the accumulated dose value and the probability of toxicity risk that was derived from the accumulated model were assumed to be the most accurate assessment of toxicity risk. The error in using the planned dose with the standard model was 6% (SD: 7%) and with the accumulated model 4% (SD: 7%).

For stomach toxicity, the difference between NTCP on the basis of accumulated and planned doses is less substantial (Fig 3). Table 2 shows the results from deriving error from the toxicity models for stomach for the 30 patients with dose accumulation. The error in using the planned dose with the standard model was 3% (SD: 3%) and with the accumulated model 3% (SD: 4%).



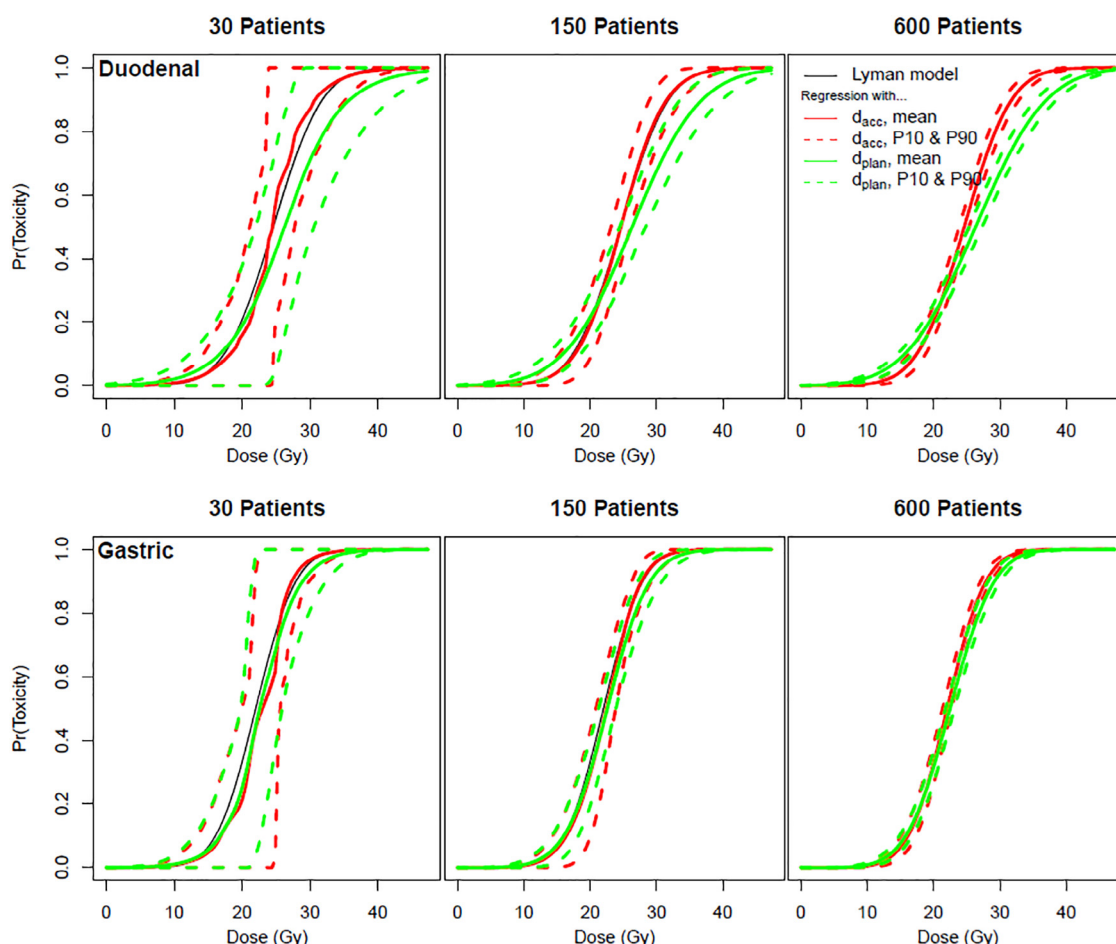
**Figure 2** Percent change from planned to accumulated dose (PC<sub>dose</sub>) by planned dose (Gy) for the stomach and duodenum and percent change in normal tissue complication probability by percent change in dose (PC<sub>NTCP</sub>) for the stomach and duodenum. The solid gray line represents 0 percent change in dose.

In an iso-toxicity protocol with a 10% risk to the duodenum, 17.6 Gy (6 fractions) would be allowed if the accumulated model was applied while 15.7 Gy (6 fractions) would be allowed if the standard model was applied. The stomach model shows only slight differences between the standard and accumulated NTCP models at this probability of toxicity risk but the standard model still overestimates toxicity risk. With a 10% limit on toxicity risk to the stomach, the accumulated model would allow 16.1 Gy (6 fractions) but the standard model would allow 15.8 Gy (6 fractions).

The bias in the planned doses in Figure 3 is reflected in the differences between the solid red and green lines. Increasing the sample size does not diminish this difference. The dashed green lines indicate variation in the fitted NTCP curves between trials. As the sample size increases

from 30 to 600 patients, the variation between the hypothetical trials becomes very small so that almost any 600-patient study will result in nearly the same biased NTCP curve (solid green line).

The greatest deviation between the probability of toxicity risk on the basis of the accumulated and planned models occurs at a planned dose of approximately 30 Gy (6 fractions) for the duodenum (Fig 4a) and 25 Gy (6 fractions) for the stomach (Fig 4b). This deviation is greater for the duodenum (16% greater probability of toxicity risk on the basis of the accumulated model) than for the stomach (6% greater probability of toxicity risk on the basis of the accumulated model). For both the duodenum and stomach models, the planned model overestimates the probability of toxicity risk for planned doses until the crossover point of 21 Gy (6 fractions) for the duodenum and 18 Gy



**Figure 3** Probability of the duodenal and gastric (with cirrhosis) toxicity. Probability for simulations of 30, 150, and 600 patients. The black line represents the Lyman normal tissue complication probability (NTCP) by dose. The solid red line represents the mean NTCP on the basis of simulations with accumulated doses ( $d_{acc}$ ). The green lines represent the mean NTCP on the basis of simulations with planned doses ( $d_{plan}$ ). The dashed lines represent the 10<sup>th</sup> and 90<sup>th</sup> percentiles.

(6 fractions) for the stomach and overestimates the probability of toxicity after the crossover point.

### Discussion

The potential impact of the deviation between the planned and delivered doses on the development of NTCP models was investigated for patients who received liver SBRT. Deviations were observed between the NTCP curves on the basis of accumulated doses and standard NTCP curves for duodenal toxicity, which indicate the potential impact that updated NTCP curves could have on patient treatment. For gastric toxicity, only modest deviations were observed with a maximum deviation of 14%. However, more substantial and potentially clinically significant deviations were seen for the duodenum with an average deviation of 6% and a maximum deviation of 21%.

For both duodenal and gastric toxicity, standard NTCP curves overestimated toxicity risk at lower doses (doses

up to 21.5 Gy in 6 fractions for the duodenum and up to 19 Gy in 6 fractions for the stomach). Deriving probability of toxicity risk for the planned dose using simulations of NTCP curves on the basis of accumulated doses yields a smaller error than deriving probability of toxicity risk for planned doses using the standard, planned dose NTCP model.

This study was based on 30 original patients who were resampled to simulate a cohort of 150 and 600 patients to build statistical Lyman NTCP models. Due to the limited number of patients in the original analysis,<sup>5</sup> true models could vary from the simulations. In addition, these 30 patients were treated at 1 institution, which indicates the possibility of dependence of the results on the treatment setup and treatment planning tendencies. As with all models, there is some uncertainty and this NTCP model has been shown to potentially overestimate the toxicity risk.<sup>16</sup> However, the focus of this paper is to investigate the potential difference between NTCP models given the known differences between accumulated and planned doses to

**Table 1** Duodenal normal tissue complication probability results

Patient	Dose <sub>planned</sub> maximum dose to 0.5 cc (Gy)	NTCP <sub>2</sub> Dose <sub>planned</sub> , Model <sub>std</sub> (%)	NTCP <sub>3</sub> Dose <sub>planned</sub> , Model <sub>acc</sub> (%)	Dose <sub>acc</sub> max dose to 0.5 cc (Gy)	NTCP <sub>1</sub> <sup>a</sup> Dose <sub>acc</sub> , Model <sub>acc</sub> (%)	Difference NTCP <sub>2</sub> - NTCP <sub>1</sub>	Difference NTCP <sub>3</sub> - NTCP <sub>1</sub>
1	29.91	66.61	82.93	29.41	80.55	13.94	2.38
2	25.44	45.72	55.86	20.91	25.18	20.54	30.68
3	8.63	1.74	0.28	6.67	0.10	1.63	0.18
4	6.67	0.96	0.10	5.86	0.07	0.90	0.04
5	15.32	9.34	4.91	15.12	4.58	4.76	0.33
7	29.70	65.66	81.94	29.74	82.14	16.48	0.19
8	0.79	0.13	0.00	0.77	0.00	0.12	0.00
9	10.42	2.87	0.66	9.04	0.34	2.53	0.32
11	0.69	0.12	0.00	0.64	0.00	0.12	0.00
12	9.77	2.40	0.49	9.69	0.47	1.93	0.02
13	8.70	1.77	0.29	5.65	0.06	1.71	0.23
14	0.64	0.12	0.00	0.63	0.00	0.12	0.00
15	1.94	0.19	0.01	1.59	0.01	0.19	0.00
16	9.61	2.29	0.45	9.29	0.39	1.90	0.06
17	19.07	19.16	15.94	12.36	1.55	17.61	14.39
18	32.28	76.18	91.47	27.76	71.44	4.74	20.03
19	31.22	72.08	88.16	30.92	87.08	15.00	1.08
20	32.35	76.42	91.64	30.06	83.59	7.17	8.06
21	29.59	65.20	81.45	29.52	81.07	15.88	0.37
22	30.08	67.35	83.70	28.75	77.15	9.81	6.55
23	1.32	0.16	0.00	1.26	0.00	0.15	0.00
24	6.56	0.93	0.10	5.79	0.06	0.87	0.03
25	1.47	0.16	0.00	1.30	0.00	0.16	0.00
26	22.39	31.78	34.34	20.75	24.26	7.51	10.07
27	23.26	35.59	40.27	21.12	26.39	9.20	13.88
28	24.19	39.80	46.83	22.88	37.62	2.18	9.21
29	33.22	79.52	93.76	33.82	94.94	15.43	1.18
30	31.25	72.20	88.26	30.20	84.19	12.00	4.07
31	10.15	2.66	0.58	7.00	0.12	2.54	0.46
32	9.57	2.27	0.44	8.49	0.26	2.01	0.18
					Average	6.30	4.13
					SD	6.53	7.31

NTCP, normal tissue complication probability; SD, standard deviation.

<sup>a</sup> Assumption that this model is the closest to the true risk on the basis of the most accurate measurement of the delivered dose.

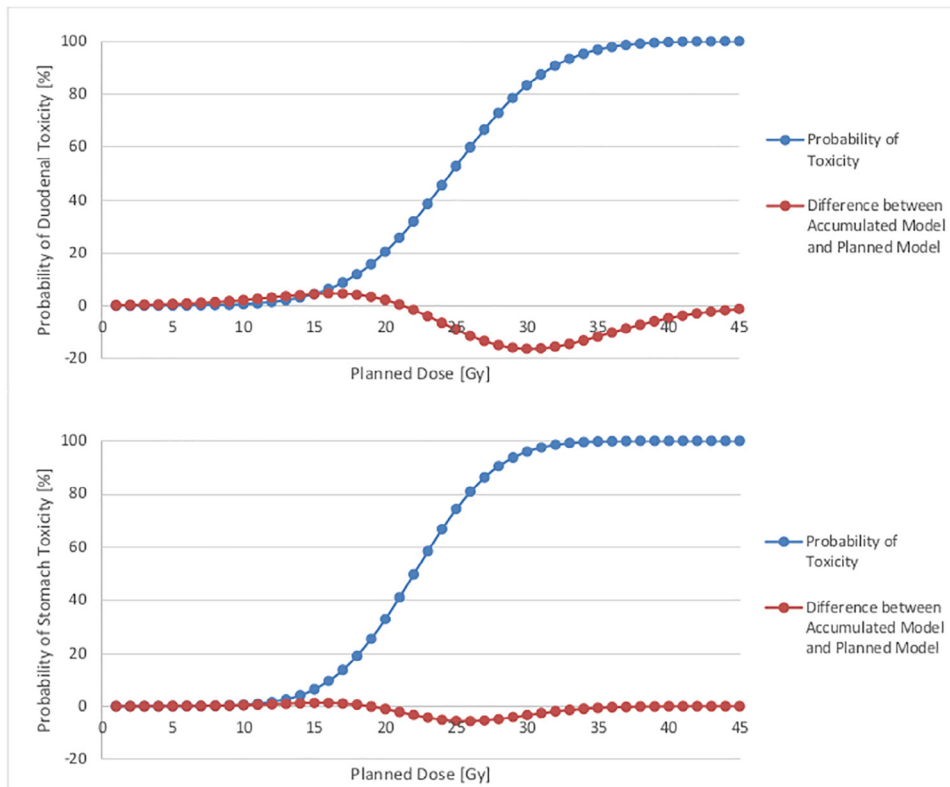
**Table 2** Stomach normal tissue complication probability results

Patient	Dose <sub>planned</sub> maximum dose to 0.5 cc (Gy)	NTCP <sub>2</sub> Dose <sub>planned</sub> , Model <sub>std</sub> (%)	NTCP <sub>3</sub> Dose <sub>planned</sub> , Model <sub>acc</sub> (%)	Dose <sub>acc</sub> max dose to 0.5 cc (Gy)	NTCP <sub>1</sub> <sup>a</sup> Dose <sub>acc</sub> , Model <sub>acc</sub> (%)	Difference NTCP <sub>2</sub> - NTCP <sub>1</sub>	Difference NTCP <sub>3</sub> - NTCP <sub>1</sub>
1	29.21	90.23	94.19	29.24	94.26	4.03	0.08
2	7.10	0.19	0.09	9.08	0.31	0.11	0.22
3	10.31	1.06	0.64	10.09	0.56	0.51	0.08
4	17.63	17.65	16.85	17.88	18.28	0.63	1.43
5	2.86	0.01	0.00	3.22	0.01	0.01	0.00
7	20.57	35.83	37.46	19.51	29.06	6.77	8.40
8	13.94	5.17	4.05	12.36	1.91	3.26	2.15
9	19.89	31.05	31.97	19.67	30.24	0.81	1.73
11	0.61	0.00	0.00	0.61	0.00	0.00	0.00
12	12.31	2.66	1.86	11.96	1.55	1.11	0.31
13	23.55	58.30	63.03	22.20	51.50	6.80	11.53
14	23.05	54.58	58.89	22.91	57.63	3.05	1.27
15	26.44	77.80	83.33	23.63	63.77	14.03	19.56
16	19.46	28.22	28.72	19.96	32.47	4.25	3.75
17	28.18	86.43	91.14	27.78	89.65	3.22	1.48
18	28.59	88.03	92.45	26.32	82.73	5.30	9.72
19	5.62	0.08	0.03	5.56	0.03	0.05	0.00
20	31.28	95.40	97.79	29.89	95.68	0.28	2.10
21	31.07	95.02	97.56	31.65	98.20	3.17	0.64
22	32.56	97.30	98.88	31.98	98.47	1.17	0.42
23	29.23	90.30	94.24	30.09	96.09	5.79	1.85
24	31.74	96.18	98.25	31.15	97.65	1.48	0.60
25	18.63	23.06	22.86	17.96	18.72	4.34	4.14
26	27.12	81.49	86.79	26.54	83.89	2.40	2.90
27	16.79	13.84	12.74	15.43	7.66	6.17	5.07
28	22.86	53.06	57.19	22.80	56.70	3.64	0.49
29	11.14	1.58	1.01	12.15	1.72	0.14	0.71
30	30.28	93.27	96.40	28.59	92.45	0.83	3.95
31	17.66	17.82	17.04	16.96	13.49	4.33	3.54
32	17.12	15.28	14.28	15.68	8.45	6.83	5.83
					Average	3.15	3.13
					SD	3.08	4.30

NTCP, normal tissue complication probability; SD, standard deviation.

<sup>a</sup> Assumption that this model is the closest to the true risk on the basis of the most accurate measurement of the delivered dose.





**Figure 4** Probability of duodenal toxicity (percentage; on the basis of accumulated model) for the duodenum and stomach. The blue line represents the probability of toxicity derived from the normal tissue complication probability model on the basis of accumulated versus planned doses. The orange line represents the difference between probability of toxicity using the accumulated and planned models. The largest difference for the duodenum occurs at a planned dose of approximately 30 Gy (toxicity risk of approximately 80%). The largest difference for the stomach occurs at a planned dose of approximately 26 Gy (toxicity risk of approximately 80%).

determine if there is a strong need to gather multi-institutional data to ensure the accuracy of these models.

The results of this simulation study requires confirmation through prospective studies. Further affirmations are necessary for the impact assessment on normal tissue complications of the stomach and duodenum. In an image guided radiation therapy protocol, the goals normally target tumor alignment and avoid normal tissue. A complicated relationship exists between the planned and accumulated doses due to dose distribution and high-dose regions and a study with >30 patients is necessary to understand this relationship.

The 30 patients in the current study were originally selected for a secondary analysis of SBRT dose accumulation<sup>5</sup> from patients who were treated in phase 1/2 clinical trials<sup>4,17,18</sup> on the basis of available cone beam CT imaging and a breathing motion of >5 mm. They were not specifically selected on the basis of any clinical outcome including toxicity. There were no grade  $\geq 3$  GI toxicities in this limited sample of 30 patients. There were 2 patients with Grade 3 platelet counts, which are likely not related to duodenum or stomach toxicity; however, these events are related to liver function and possibly non-classic, radiation-induced liver disease.

Of the 30 patients who were included in this study, 9 patients had dose-limiting normal tissue (ie, planning target

volume coverage or prescription dose was limited due to normal tissue toxicity risks).<sup>19</sup> The toxicity models that were developed for this study showed that the planned-dose model overestimates toxicity risk in the average dose range, which implies that the dose-limiting organs could potentially receive a higher dose with the same risk of toxicity. The 30 patients in this study did not have GI toxicity; however, this may have been the result of a conservative tumor dose.

Understanding the dose-toxicity relationship is important especially in escalation studies because underdosing due to misinterpreted normal tissue complication risk can be detrimental to local control because the tumor will not receive as high a dose as possible. The QUANTEC report<sup>16</sup> on stomach toxicity acknowledged that there is limited data on GI toxicity, which indicates a need for a better understanding of GI toxicity models and especially with the introduction of molecular agents in radiation therapy. An understanding of the delivered dose is crucial so that the models used for radiation therapy are as accurate as possible and enable the highest therapeutic ratio possible.

This study demonstrates the potential clinical importance of including accumulated doses in the development of NTCP models in future clinical trials. Since existing NTCP models are based on patient data that preceded volumetric daily imaging, they are based on only planned doses.

Currently, equipment and software for volumetric daily imaging are available at the clinic and new DIR tools can be applied to perform dose accumulation. Future clinical trials should prospectively record delivered (accumulated) doses and toxicity outcomes in an effort to improve understanding of the dose-response relationship. For example, in a phase 3 study<sup>20</sup> of 368 patients, more accurate toxicity data could be developed if accumulated and planned dose information was collected. This would require the collection of daily images of all patients to correlate the true delivered doses with toxicity. This study supports the need for these trials to include the collection of imaging obtained at treatment delivery so that accumulated doses can be calculated. This simulation study suggests that deriving new NTCP models on the basis of accumulated doses will yield clinically significant results compared with the current models on the basis of planned doses.

The simulation data support the need for accumulated dose calculation in clinical trials. One such trial design is an observational study in which patients are treated per standard of care and both planned and accumulated dose values are recorded along with clinical toxicity. The number of patients that is required will depend on many factors including the overall toxicity rate. In patients with primary liver cancer who are treated with radiation, approximately 20% will experience toxicity that is defined as a change in Child-Pugh score of  $\geq 2$  points. Demonstrating improved predictive performance of accumulated dose models would be a primary aim of such a trial. The size of the trial would be determined to achieve the desired level of precision to estimate the proportion of patients with deviations. As an example, if the true proportion of patients with deviations was 0.10, a trial of 100 patients would result in a standard error of an estimated proportion equal to 0.03.

Possible reasons for the numerical differences between the NTCP curves of the duodenum and stomach include variations in patient's breathing magnitude and trajectory. Regions of high-dose gradient can differ on the basis of the area of the organ and how it is affected by the breathing motion. NTCP parameters, radiation treatment planning, and sensitivity of motion all can influence the final NTCP results for each organ and potentially cause the duodenum and stomach to have different numerical results. In addition, due to the limited number of patients, potential deviations for the stomach may not have been observed but may be observed in a larger population.

This study focused on liver cancer and the organs at risk for toxicity during radiation treatment. However, the results may reflect the impact of incorporating accumulated doses in the development of luminal GI toxicity models in other abdominal treatment sites such as the stomach, pancreas, or pelvic area where toxicity to GI structures are of clinical concern. Because an accumulated dose is a more accurate predictor of the delivered dose than a planned dose, NTCP models for all sites (eg, head and neck) could be improved if based on an accumulated dose instead of a planned dose.

## Conclusions

The need for prospective clinical trials to evaluate the clinical impact of developing NTCP models on the basis of accumulated rather than planned doses has been demonstrated in this study. This study used simulated data for 600 patients to develop toxicity models and indicated that historical NTCP models on the basis of planned doses may overestimate the toxicity risk for lower doses but underestimate the risk of toxicity for higher doses with errors up to 21%. The differences between NTCP models on the basis of accumulated doses compared with planned doses are greater for the duodenum than for the stomach. However, with the availability of deformable, registration-based dose accumulation using volumetric daily imaging, improved NTCP models are possible and should be included in the development of future clinical trials.

## References

- Scorsetti M, Clerici E, Comito T. Stereotactic body radiation therapy for liver metastases. *J Gastrointest Oncol*. 2014;5:190-197.
- Nouhaud E, Créange G, Cuffe A, et al. Stereotactic body radiation therapy for liver tumors with or without rotational intensity modulated radiation therapy. *BMC Res Notes*. 2013;6:492.
- Thomas TO, Hasan S, Small W Jr, et al. The tolerance of gastrointestinal organs to stereotactic body radiation therapy: What do we know so far? *J Gastrointest Oncol*. 2014;5:236-246.
- Bujold A, Massey CA, Kim JJ, et al. Sequential phase I and II trials of stereotactic body radiotherapy for locally advanced hepatocellular carcinoma. *J Clin Oncol*. 2013;31:1631-1639.
- Velec M, Moseley JL, Craig T, et al. Accumulated dose in liver stereotactic-body radiotherapy: positioning, breathing and deformation effects. *Int J Radiat Oncol Biol Phys*. 2012;83:1132-1140.
- Feng M, Normolle D, Pan CC, et al. Dosimetric analysis of radiation-induced gastric bleeding. *Int J Radiat Oncol Biol Phys*. 2012;84:e1-e6.
- Murphy JD, Christman-Skieller C, Kim J, Dieterich S, Chang DT, Koong AC. A dosimetric model of duodenal toxicity after stereotactic body radiotherapy for pancreatic cancer. *Int J Radiat Oncol Biol Phys*. 2010;78:1420-1426.
- Dean JA, Wong KH, Welsh LC, et al. Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy. *Radiother Oncol*. 2016;120:21-27.
- Holyoake DLP, Aznar M, Mukherjee S, Partridge M, Hawkins MA. Modelling duodenum radiotherapy toxicity using cohort dose-volume-histogram data. *Radiother Oncol*. 2017;123:431-437.
- Xu CJ, van der Schaaf A, van't Veld AA, Langendijk JA, Schilstra C. Statistical validation of normal tissue complication probability models. *Int J Radiat Oncol Biol Phys*. 2012;84:e123-e129.
- Dawson LA, Balter JM, McGinn CJ, Lawrence TS, Ten Haken RK. Description of radiation induced liver disease using the lyman and local damage-organ injury NTCP models. *Int J Radiat Oncol Biol Phys*. 2001;51:33-34.
- Scaife JE, Thomas SJ, Harrison K, et al. Accumulated dose to the rectum, measured using dose-volume histograms and dose-surface maps, is different from planned dose in all patients treated with radiotherapy for prostate cancer. *Br J Radiol*. 2015;88:20150243.
- Swaminath A, Massey C, Brierley JD, et al. Accumulated delivered dose response of stereotactic body radiation therapy

- for liver metastases. *Int J Radiat Oncol Biol Phys.* 2015;93:639-648.
14. Brock KK, Sharpe MB, Dawson LA, et al. Accuracy of finite element model-based multi-organ deformable image registration. *Med Phys.* 2005;32:1647-1659.
  15. Yorke ED. *Parameters for and use of NTCP models in the clinic.* Minneapolis, MN: 2007 49th Annual Meeting of the American Association of Physicists in Medicine. July 22-26.
  16. Michalski JM, Gay H, Jackson A, Tucker SL, Deasy JO. Radiation dose-volume effects in radiation-induced rectal injury. *Int J Radiat Oncol Biol Phys.* 2010;76:S123-S129.
  17. Lee MT, Kim JJ, Dinniwel R, et al. Phase 1 study of individualized stereotactic body radiotherapy of liver metastases. *J Clin Oncol.* 2009;27:1585-1591.
  18. Tse RV, Hawkins M, Lockwood G, et al. Phase 1 study of individualized stereotactic body radiotherapy for hepatocellular carcinoma and intrahepatic cholangiocarcinoma. *J Clin Oncol.* 2008;26:3911-3912.
  19. Velec M. *Deformable dose reconstruction to optimize the planning and delivery of liver cancer radiotherapy.* Toronto, Canada: University of Toronto; 2014.
  20. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29-Identifier NCT01730937, Sorafenib tosylate with or without stereotactic body radiation therapy in treating patients with liver cancer; 2012 Nov 15 [cited 2018 Mar 6]; [about 8 screens]. Available at: <https://www.clinicaltrials.gov/ct2/show/NCT01730937?term=dawson&recrs=a&phase=2&draw=2&rank=12>. Accessed June 6, 2018.