Int. J. Radiation Oncology Biol. Phys., Vol. 53, No. 4, pp. 810–821, 2002

0360-3016/02/$–see front matter

ELSEVIER

PII S0360-3016(02)02846-8

CLINICAL INVESTIGATION **Normal Tissue**

ANALYSIS OF RADIATION-INDUCED LIVER DISEASE USING THE LYMAN NTCP MODEL

LAURA A. DAWSON, M.D., DANIEL NORMOLLE, PH.D., JAMES M. BALTER, PH.D., CORNELIUS J. MCGINN, M.D., THEODORE S. LAWRENCE, M.D., PH.D., AND RANDALL K. TEN HAKEN, PH.D.

Department of Radiation Oncology, University of Michigan, Ann Arbor, MI

Purpose: To describe the dose–volume tolerance for radiation-induced liver disease (RILD) using the Lyman–Kutcher–Burman (LKB) normal tissue complication probability (NTCP) model.

Methods and Materials: A total of 203 patients treated with conformal liver radiotherapy and concurrent hepatic arterial chemotherapy were prospectively followed for RILD. Normal liver dose–volume histograms and RILD status for these patients were used as input data for determination of LKB model parameters. A complication was defined as Radiation Therapy Oncology Group Grade 3 or higher RILD ≤4 months after completion of radiotherapy. A maximal likelihood analysis yielded best estimates for the LKB NTCP model parameters for the liver for the entire patient population. A multivariate analysis of the potential factors associated with RILD was also completed, and refined LKB model parameters were obtained for patient subgroups with different risks of RILD.

Results: Of 203 patients treated with focal liver irradiation, 19 developed RILD. The LKB NTCP model fit the complication data for the entire group. The "n" parameter was larger than previously described, suggesting a strong volume effect for RILD and a correlation of NTCP with the mean liver dose. No cases of RILD were observed when the mean liver dose was <31 Gy. Multivariate analysis demonstrated that in addition to NTCP and the mean liver dose, a primary hepatobiliary cancer diagnosis (vs. liver metastases), bromodeoxyuridine hepatic artery chemotherapy (vs. fluorodeoxyuridine chemotherapy), and male gender were associated with RILD. For 169 patients treated with fluorodeoxyuridine, the refined LKB model parameters were $n = 0.97$, $m = 0.12$, tolerance dose for 50% complication risk for whole organ irradiated uniformly $[TD_{50}(1)] = 45.8$ Gy for patients with liver metastases, and $TD_{50}(1) = 39.8$ Gy for patients with primary hepatobiliary cancer.

Conclusion: These data demonstrate that the liver exhibits a large volume effect for RILD, suggesting that the mean liver dose may be useful in ranking radiation plans. The inclusion of clinical factors, especially the diagnosis of primary hepatobiliary cancer vs. liver metastases, improves the estimation of NTCP over that obtained solely by the use of dose–volume data. These findings should facilitate the application of focal liver irradiation in future clinical trials. © 2002 Elsevier Science Inc.

Liver cancer, Conformal radiotherapy, Radiation toxicity, Radiation-induced liver disease, NTCP, Lyman model.

INTRODUCTION

Radiation-induced liver disease (RILD) is a dose-limiting complication of liver irradiation. RILD is a clinical syndrome of anicteric hepatomegaly, ascites, and elevated liver enzymes (particularly serum alkaline phosphatase) occurring typically 2 weeks to 4 months after completion of hepatic irradiation. RILD resembles the suprahepatic vein obstruction and hepatic toxicity seen after high-dose chemotherapy (with or without total body irradiation) for bone marrow transplantation. The pathologic lesion in RILD is venoocclusive disease, characterized by areas of marked venous congestion in the central portion of each lobule, with sparing of the larger veins. Unfortunately, the treatment options for RILD are limited, and, in severe cases, liver failure and death can occur (1).

The tolerance of the whole liver to radiation is low, and RILD is seen in 5–10% of patients treated with 30–35 Gy to the whole liver. For this reason, radiation has traditionally had a limited role in the treatment of intrahepatic cancers. However, treatment of parts of the liver with higher radiation doses is possible without adverse consequences as long as an adequate volume of normal liver is not irradiated to high doses (2–5). Patients with focal unresectable intrahepatic malignancies treated with higher radiation doses have

Reprint requests to: Laura A. Dawson, M.D., Department of Radiation Oncology, University of Michigan, UH-B2C447, 1500 E. Medical Center Drive, Ann Arbor, MI 48109-0010. Tel: (734) 926-7810; Fax: (734) 763-7370; E-mail: dawson@umich.edu

Supported in part by NIH Grants P01 CA59827, R01 CA85684, and M01-RR00042.

Presented in part at the American Society for Therapeutic Radiology and Oncology (ASTRO) Annual Meeting, San Francisco, California, November 2001.

Acknowledgments—The authors thank Steve Kronenberg for assistance with figure preparation.

Received Dec 6, 2001, and in revised form Mar 13, 2002. Accepted for publication Mar 19, 2002.

Table 1. Treatment regimen

Patient group	Patients (*n*)	Whole liver dose (Gy)	Total tumor dose (Gy)	Chemotherapy	Patients with RILD (*n*)
1	19	33	33	HA FUdR	0
2	13	36	36	HA FUdR	3
3	11	30	45 or 48	HA FUdR	4
4	12	0	48 or 52.8	HA FUdR	0
5	9	30	60 or 66	HA FUdR	2
6	15	0	66 or 72.6	HA FUdR	0
7	14	0	49.8–72.6	HA FUdR	0
8	14	24–33	24–33	HA BUdR	3
9	20	0	48–66	HA BUdR	2
10	76	0	50–90	HA FUdR	5
Total	203	0–36	24–90		19

Abbreviations: RILD = radiation-induced liver disease; HA FUdR = hepatic artery fluorodeoxyuridine; HA BUdR = hepatic artery bromodeoxyuridine.

better response rates, symptom improvement, and survival rates than do patients treated with lower doses (6–8). Additional knowledge of the partial liver tolerance to radiation may permit safer dose escalation and lead to improvements in clinical outcomes for patients with intrahepatic malignancies. In addition, knowledge of partial organ tolerances to radiation is required for successful implementation of novel radiotherapy (RT) strategies such as automated optimization for intensity-modulated RT.

A number of models estimating the volume dependence of normal tissue toxicity have been used to compare the relative merits of competing three-dimensional RT plans (9–12). The Lyman model assumes a sigmoid relationship between a dose of uniform radiation given to a volume of an organ and the chance of a complication occurring (9) (see "Methods and Materials" below). We have used the Lyman model (13) clinically, because it is relatively simple (containing 3 parameters) and, when implemented using the Kutcher–Burman (KB) effective volume (V_{eff}) dose–volume histogram (DVH) reduction scheme (14), it permits comparisons between plans based on DVHs before assigning the dose (15). Our previous effort (16) to reestimate the parameters for the Lyman model was limited by a relatively small number of patients, few cases of RILD, and a simplified statistical analysis. The purpose of this study was to describe more quantitatively the dose–volume relationship of the liver to RILD, based on dose–volume data from more than twice as many patients as previously analyzed, using the LKB normal tissue complication probability (NTCP) model. When a multivariate analysis demonstrated that RILD was associated with nondosimetric factors such as diagnosis, the LKB model was used to describe the dose–volume effects for patient subgroups with different risks of RILD.

METHODS AND MATERIALS

Patients

All patients included in this analysis had unresectable intrahepatic cancer (hepatocellular carcinoma, cholangiocarcinoma, or colorectal carcinoma metastatic to the liver) and were treated in prospective clinical trials (Table 1). Ninety-three patients included in this analysis have been previously described (16, 17).

To be eligible for treatment with liver RT, patients had to have an estimated life expectancy of ≥12 weeks, be ≥18 years old, and have had normal liver function (prothrombin and partial thromboplastin time normal or correctable with vitamin K), renal function, and bone marrow function. Informed consent was obtained in accordance with the procedures of the Institutional Review Board of the University of Michigan. Ineligible patients included those with prior upper abdominal irradiation, a history of bleeding from esophageal varices, or other serious intercurrent illnesses.

To be eligible for the present RILD analysis, patients had to be assessable for RILD (minimal follow-up of 4 months, with no evidence of hepatic progression). Approximately three-quarters of all treated patients were eligible for RILD evaluation.

Radiation-induced liver disease evaluation

All 203 patients were prospectively followed for RILD. CT scans were routinely ordered at 2 and 4 months after RT, and additional scans were ordered if clinically indicated. Patients were followed for a minimum of 4 months. A complication (yes vs. no) was defined as Radiation Therapy Oncology Group Grade 3 or higher RILD (clinical RILD requiring treatment). Alkaline phosphatase had to be elevated by at least a factor of 2, together with nonmalignant ascites, detected either clinically or by CT, in the absence of documented disease progression. Patients who developed hepatic tumor progression associated with ascites or impaired liver function within 4 months after RT and those who were not seen in follow-up 4 months after RT were not eligible for this analysis.

Radiotherapy

All patients were treated on consecutive liver radiation dose-escalation protocols at the University of Michigan

from 1987 to 1999 (Table 1). Forty-one patients were treated with whole liver irradiation; 20 patients were treated with whole liver irradiation followed by a boost of radiation to a partial liver volume, and 142 patients were treated with partial liver irradiation alone. Radiation was delivered twice daily in 11 fractions/wk, 1.5–1.65 Gy/fraction, with a minimal interfraction interval of 4–6 h. The median dose of radiation delivered was 52.5 Gy (range 24–90). Concurrent continuous infusion hepatic artery fluorodeoxyuridine (FUdR) ($n = 169$) or bromodeoxyuridine (BUdR) ($n = 34$) was used as a radiation sensitizer for the first 4 weeks of RT. There was a 2-week break in RT after 2 weeks to minimize the risk of complications related to the hepatic artery catheter used for chemotherapy delivery.

All patients were treated using three-dimensional conformal RT techniques (18). Target and normal liver volumes were contoured on axial CT cuts, and the targets were expanded to account for occult disease, setup uncertainty, and breathing motion (13, 19). Treatment planning was performed using the University of Michigan planning system (20). The LKB NTCP model with revised parameters (16) was used most recently in a dose-escalation trial, in which each patient was subjected to a fixed predicted risk of RILD and the V_{eff} of liver treated ranged from 20% to 90% (13, 15).

Normal liver (liver minus gross tumor) DVHs were obtained for all 203 patients. The physical dose values in the three-dimensional dose distributions for each treatment course of all patients were converted to normalized isobiologic effective doses at 1.5 Gy/fraction using the linear quadratic model ($\alpha/\beta = 2$ Gy) before computation of the composite dose distributions from which the DVHs were computed. Because uniform dose–volume distributions are required in the Lyman NTCP model, the Kutcher and Burman V_{eff} DVH reduction scheme was used to convert the nonuniform complex dose distributions into "equivalent" uniform dose distributions. V_{eff} is defined as the normal liver volume, which, if irradiated uniformly to the reference dose, would be associated with the same NTCP as the nonuniform dose distribution actually delivered (14).

The mean dose to the normal liver was calculated by summing the dose values for all voxels within the liver (liver volume minus the gross tumor volume) and dividing the sum by the number of voxels.

Lyman NTCP model

Data were fit to the empiric Lyman NTCP model, which describes the probability of a complication after uniform radiation of a fractional volume of normal tissue (v) to a dose (D), assuming a sigmoid dose–response relationship, with no threshold, as follows (9):

$$NTCP = \phi(t) = 1 \Big/ \sqrt{2\pi} \int_{-\infty}^{t} e^{-x^2/2} dx$$

where

$$t = (D - TD_{50}(v))/(m \cdot TD_{50}(v))$$

and $TD_{50}(v)$ represents the tolerances doses associated with a 50% chance of complications for uniform partial liver irradiation, where $TD_{50}(v)$ is related to the whole liver ($v = 1$) tolerance through the power law relationship:

$$TD_{50}(v) = TD_{50}(1) \cdot v^{-n}$$

and $TD_{50}(1)$ represents the tolerance of the whole organ to irradiation, m characterizes the steepness of the dose–response at $TD_{50}(1)$, and n represents the volume effect, which relates the tolerance doses of uniform whole organ irradiation to uniform partial organ irradiation. When n is near 1, the volume effect is large and when it is near 0, the volume effect is small.

The partial volume-dose–complication risk relationship for the liver was graphically displayed in dose–NTCP plots (as a function of V_{eff}), V_{eff}–NTCP plots (as a function of dose), and dose–V_{eff} plots (as a function of NTCP).

Maximal likelihood estimation

Normal liver DVHs and the occurrence or lack of occurrence of RILD from 203 patients comprised the input data for the determination of the LKB model parameters using a maximal likelihood analysis. The NTCP model parameters $TD_{50}(1)$, m, and n were adjusted to maximize the probability of predicting complications for those patients who experienced complications and maximize the probability of predicting no complications for those patients who were complication free, as detailed in Appendix 1.

Subset specific estimates of $TD_{50}(1)$ were obtained by estimating the n and m parameters from the entire data set and allowing $TD_{50}(1)$ to vary depending on the patient subset (e.g., diagnosis).

Confidence intervals for the parameters were determined by exploring the space around the maximal likely parameter set using profile-likelihood methods (21, 22). Confidence intervals were also obtained and displayed on dose–V_{eff} iso-NTCP curves using the methods outlined in Appendix 2.

Goodness of fit

The deviance of a given set of parameters is related to the log-likelihood and can be used to assess the goodness of fit. The deviance has an approximate χ^2 distribution, where values closer to 0 indicate a significant lack of fit. Although this approximation is poor when modeling binary data with continuous variables (23), the statistic was used as a general guide. The details are outlined in Appendix 3.

p values were used to compare across sample sizes. A large goodness-of-fit p value implies a better fit than a smaller p value.

Table 2. Patient, tumor, and treatment characteristics for all patients, and for patients with and without radiation induced liver disease

Characteristic	All patients (n = 203)	Patients without RILD (n = 184, 91%)	Patients with RILD (n = 19, 9%)
Age (y)			
Median	60	60	62.5
Range	28–85	29–85	28–72
Gender			
Female	85 (29)	82 (45)	3 (16)
Male	118 (58)	102 (55)	16 (84)
Diagnosis			
Hepatocellular carcinoma	58 (29%)	46 (25%)	12 (63)
Cholangiocarcinoma	47 (23%)	45 (24%)	2 (11)
Liver metastases	98 (48%)	93 (51%)	5 (26)
Prescribed dose (Gy)			
Median	52.8	52.8	48.0
Range	24–90	24–90	38.0–60.0
Whole liver radiation			
Yes	65 (32)	59 (32)	6 (32)
No	138 (68)	125 (68)	13 (68)
LKB NTCP*			
Median	0.05	0.04	0.17
Range	0.00–0.46	0.00–0.46	0.1–0.46
Mean liver dose (Gy)			
Median	32.0	31.3	37.0
Range	14.9–44.0	14.9–44.0	31.6–43.7

*NTCP for RILD, based on LKB model fit for all 203 patients (n = 1.1, m = 0.18 and $TD_{50}(1)$ = 43.3 Gy).
Abbreviations: RILD = radiation-induced liver disease; LKB = Lyman-Kutcher-Burman; NTCP = normal tissue complication probability.

Exploration of factors associated with RILD

Logistic regression analysis was used to measure the influence of potential clinical, demographic, and treatment factors on the occurrence of RILD. In this logistic regression analysis model, age, gender, diagnosis, prescribed dose, use of chemotherapy (FUdR vs. BUdR), use of whole liver irradiation, treatment regimen, and liver volume (whole liver minus gross tumor volume) were combined with a single predictor related to NTCP (NTCP based on parameters obtained for all 203 patients [logit-transformed] or mean liver dose). If a factor was significantly associated with RILD, then, when the distribution of events permitted, the NTCP model parameters, or a subset of parameters, were reestimated within the patient subgroups as previously described.

RESULTS

A total of 203 patients were assessable for RILD analysis. Two patients, who were lost to follow-up after 2 months, were considered to be without RILD; all other patients were followed for a minimum of 4 months after RT completion. Of the 203 patients, 19 developed Grade 3 or higher RILD (Table 2). Of the 19 patients, 6 were treated with whole liver RT, 6 were treated with whole liver RT plus higher dose partial liver RT, and 7 were treated with partial liver RT alone.

Using the maximal likelihood method, an estimation of the LKB model parameters based on the DVHs and the presence or absence of RILD was completed. For the entire group of 203 patients, the calculated values of $TD_{50}(1)$ and m were 43.3 Gy (95% confidence interval [CI] 41.9–52.8) and 0.18 (95% CI 0.14–0.24), respectively. The volume effect parameter n was estimated to be 1.1 (95% CI 0.88–1.6).

A comparison of this fit of 203 patients to the original LKB model parameters estimated from the literature (24) and our initial fit of 79 patients (16) is shown in Fig. 1. The deviance (D) for the newest parameter set was 100.6 ($p > 0.99$), implying a good fit. The fit would have been much worse using the original estimates of the LKB model parameters [$n = 0.32$, $m = 0.15$, $TD_{50}(1) = 45$ Gy [25], D = 210, $p = 0.30$), but only somewhat diminished using our previously obtained parameter set based on the first 79 patients ($n = 0.67$, $m = 0.15$, $TD_{50}(1) = 45$ Gy [16], D = 106.8, $p = 0.99$).

Qualitatively, the observed complication rates for rank order patient subgroups of the entire group corresponded to the complication rates for those patients predicted by the LKB model using the revised parameters (Fig. 2).

Because an n of 1 in the LKB model suggests a large volume effect and a strong correlation of RILD with the mean liver dose, the mean liver dose was studied in more detail. As expected, the mean liver dose was associated with the predicted NTCP, based on the LKB model, in a sigmoidal relationship (Fig. 3). For all 203 patients studied, after a threshold mean liver dose of 30 Gy (below which no patient developed RILD), the LKB NTCP increased by approximately 4%/Gy increase in the mean dose. A mean liver dose

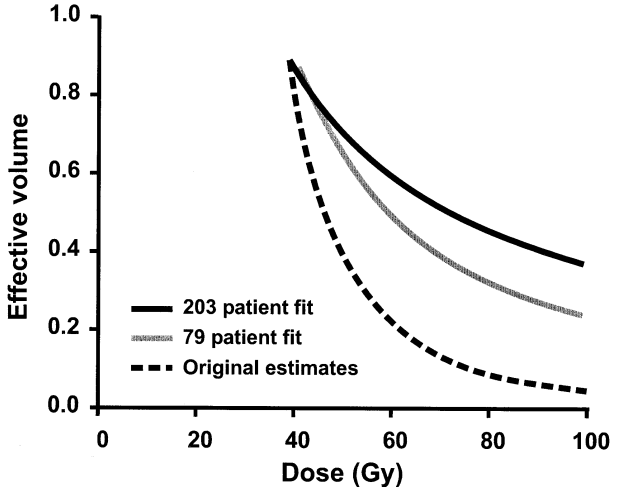

Fig. 1. Comparison of original estimates (25), previous fit (16), and newest fit [n = 1.1, m = 0.18, and $TD_{50}(1)$ = 43.3 Gy] for the LKB NTCP model. Ten percent iso-NTCPs are displayed in an effective volume (organ volume that if irradiated to the prescribed dose uniformly would be associated with the same NTCP as the nonuniform dose distribution) vs. dose (prescribed dose normalized to 1.5 Gy b.i.d.) graph.

of 31 and 43 Gy corresponded with a 5% and 50% probability of RILD, respectively.

Analysis of prognostic factors for RILD

The characteristics of patients who developed RILD were compared with those of the patients who did not develop RILD (Table 1). The median prescribed radiation dose to the tumor was not significantly different between those who developed RILD and those who did not (53 and 48 Gy, respectively). However, as expected, the median NTCP and mean liver dose were higher in the patients who developed RILD than in those who did not (NTCP: 0.17 vs. 0.04; mean liver dose: 37 vs. 31 Gy).

In addition to dosimetric factors, we investigated the influence of clinical and demographic factors on the development of RILD using a logistic regression model to complete a multivariate analysis. The mean liver dose ($p < 0.0001$), primary hepatobiliary carcinoma diagnosis ($p = 0.005$), use of BUdR chemotherapy ($p < 0.0001$), and male gender ($p = 0.002$) were statistically significant factors associated with the development of RILD, when the mean dose was put into the model before NTCP. Because the mean liver dose and NTCP correlated highly when NTCP was put into the regression model before the mean dose, NTCP replaced the mean liver dose as a significant factor associated with RILD ($p < 0.0001$). Age, use of whole liver irradiation, liver volume, treatment regimen, and prescribed dose were not independently associated with the development of RILD (Table 3).

Within the group of 169 patients treated with FUdR, the difference in risk of RILD between those with primary hepatobiliary cancer and those with metastatic liver cancer appeared to be largest in men. In that group of patients, men with primary hepatobiliary cancer had a significantly higher risk of RILD than did all other patients ($p = 0.007$).

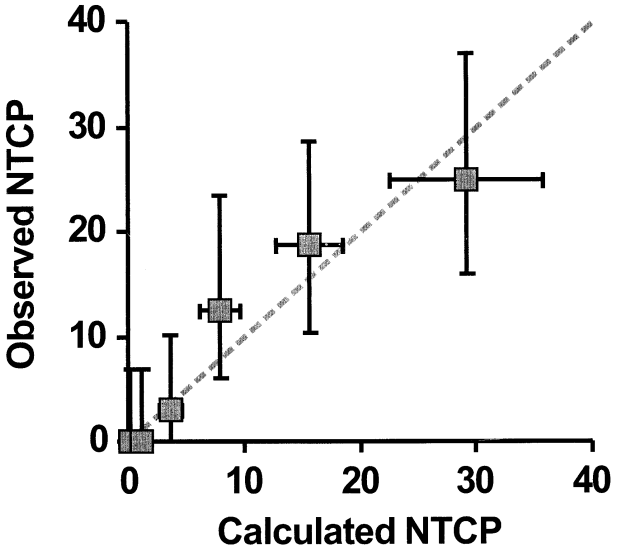

Fig. 2. Observed complication rates for all 203 patients studied, in rank order groups of 30 patients vs. predicted rates of complications according to the LKB NTCP model and revised parameters: n = 1.1, m = 0.18, and $TD_{50}(1)$ = 43.3 Gy.

NTCP in patient subgroups

LKB model parameters were then fit for patient subgroups predicted to have different risks of RILD based on the multivariate analysis of clinical factors. The LKB model parameters for the 169 patients treated with FUdR chemotherapy were as follows: $n = 0.86$ (95% CI 0.63–1.68), $m = 0.11$ (95% CI 0.07–0.23), and $TD_{50}(1) = 42.7$ Gy (95% CI 40.4–49.8; D = 74.6, $p > 0.99$). The LKB model parameters for the 34 patients treated with BUdR chemotherapy were $n = 0.62$ (95% CI 0.55–0.71), $m < 0.001$ (95% CI <0.001–0.001), and $TD_{50}(1) = 33.8$ Gy (95% CI 32.2–37.3), indicating a substantially lower tolerance of the liver to radiation.

To better describe the different risk of RILD in patients with primary hepatobiliary malignancies and those with liver metastases, another fit was completed in which the LKB model parameters n and m were fit to the entire group of patients treated with FUdR (169 patients), but the $TD_{50}(1)$ was separately fit for patients with primary hepatobiliary cancer [$TD_{50}(1)_{HB}$; 84 patients] and liver metastases [$TD_{50}(1)_{LM}$; 85 patients]. The parameters were as follows: $n = 0.97$ (95% CI 0.69–2.3), $m = 0.12$ (95% CI 0.07–0.25), $TD_{50}(1)_{HB} = 39.8$ Gy (95% CI 38.8–41.1), and $TD_{50}(1)_{LM} = 45.8$ Gy (95% CI 43.4–50.4; D = 66.0, $p > 0.99$). The two $TD_{50}(1)$ values were significantly different ($p < 0.02$). This indicates a higher tolerance of the liver to radiation for patients with liver metastases compared with those with primary hepatobiliary malignancies. On the basis of this analysis, an opportunity exists for higher doses to be

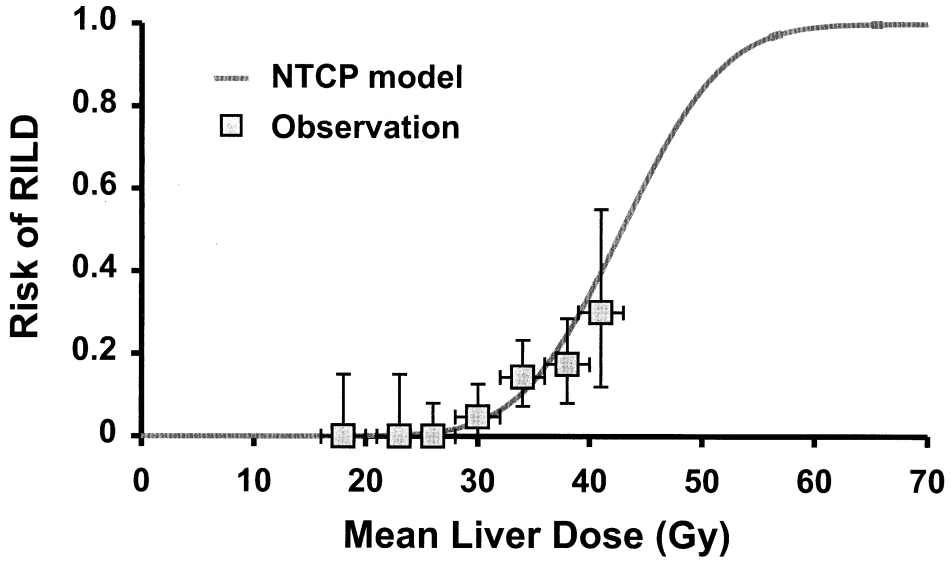

Fig. 3. Observed and predicted NTCP, according to the LKB NTCP model vs. mean liver dose (in 1.5 Gy b.i.d.). Observed NTCP calculated from patients grouped in 4-Gy bins, with 80% confidence intervals displayed. Predicted NTCP based on the LKB NTCP model, with n = 1.1, m = 0.18, and $TD_{50}(1)$ = 43.3 Gy.

delivered than previously estimated, especially for patients with liver metastases.

Figures 4 and 5 demonstrate the difference in the partial volume–dose-complication risk relationship in patients with primary liver cancer and liver metastases treated with FUdR. A threshold volume effect appears to be present, because the probability of complications is near 0 if the treated effective liver volume is less then one-third. The complication risk is also predicted to be low (<5%) if the whole effective liver volume treated is <32 Gy for primary liver cancer and <36 Gy for liver metastases (in 1.5-Gy fractions b.i.d.). The RILD risk is predicted to be 5% for an effective liver volume of two-thirds treated to 46 Gy for primary liver cancer and 54 Gy for liver metastases (in 1.5-Gy fractions b.i.d.). Other partial volume–dose-complication risks can be extrapolated from the curves in Figs. 4 and 5.

Table 3. Factors associated with RILD based on multivariate analysis

Prognostic factor	*p*	Odds ratio	95% CI
Mean liver dose*	<0.0001	1.6	1.4–2.1
BUdR chemotherapy	<0.0001	71	8.9–890
Primary hepatobiliary cancer diagnosis	0.005	6.4	1.7–30
Male gender	0.002	9.9	2.2–63
Age	NS		—
Normal liver volume	NS		—
Treatment regimen	NS		—
Use of whole liver radiation	NS		—

Mean liver dose was put into regression model before NTCP.

*Because mean liver dose and NTCP are highly correlated, when NTCP was put into the regression model before mean liver dose, NTCP was substituted for mean liver dose as a significant factor, with similar results (NTCP p <0.0001, OR for logit (NTCP) = 6.7 (95% CI [3.0–19]).

Abbreviations: RILD = radiation-induced liver disease; CI = confidence interval; BUdR = bromodeoxyuridine; NS = not significant; NTCP = normal tissue complication probability; OR = odds ratio.

When LKB model parameter fits were separately completed for patients with primary hepatobiliary malignancies and liver metastases treated with FUdR (allowing all three parameters to vary in these two groups), similar results were obtained, although the confidence intervals were larger. The LKB parameters, 95% CIs, and deviance values for 84 patients with primary hepatobiliary cancer treated with FUdR were n = 0.90 (95% CI 0.66–2.0), m = 0.09 (95% CI 0.05–0.23), and $TD_{50}(1)_{HB}$ = 39.6 Gy (95% CI 37.1–43.7; D = 42.6, p >0.99). The LKB parameters, 95% CIs, and deviance values for the 85 patients with liver metastases treated with FUdR were n = 1.27 (95% CI 0.70 to >3), m = 0.18 (95% CI 0.10–0.4), and $TD_{50}(1)_{LM}$ = 50.8 Gy (95% CI 45.4–99.0; D = 22.4, p >0.99).

The mean liver dose was also evaluated further in the patient subgroups. The mean liver dose was not statistically different in patients with liver metastases and those with primary hepatobiliary cancer (average mean liver dose 31.6 and 30.6 Gy, respectively). Five percent of patients with liver metastases developed RILD, and 13% of patients with primary hepatobiliary cancer developed RILD. In patients treated with hepatic arterial FUdR, the mean liver dose associated with a 5% risk of RILD was 32 Gy for primary hepatobiliary cancer and 37 Gy for liver metastases (in 1.5 Gy per fraction). The mean liver dose associated with a 50% risk was 40 Gy for primary hepatobiliary cancer and 47 Gy for liver metastases.

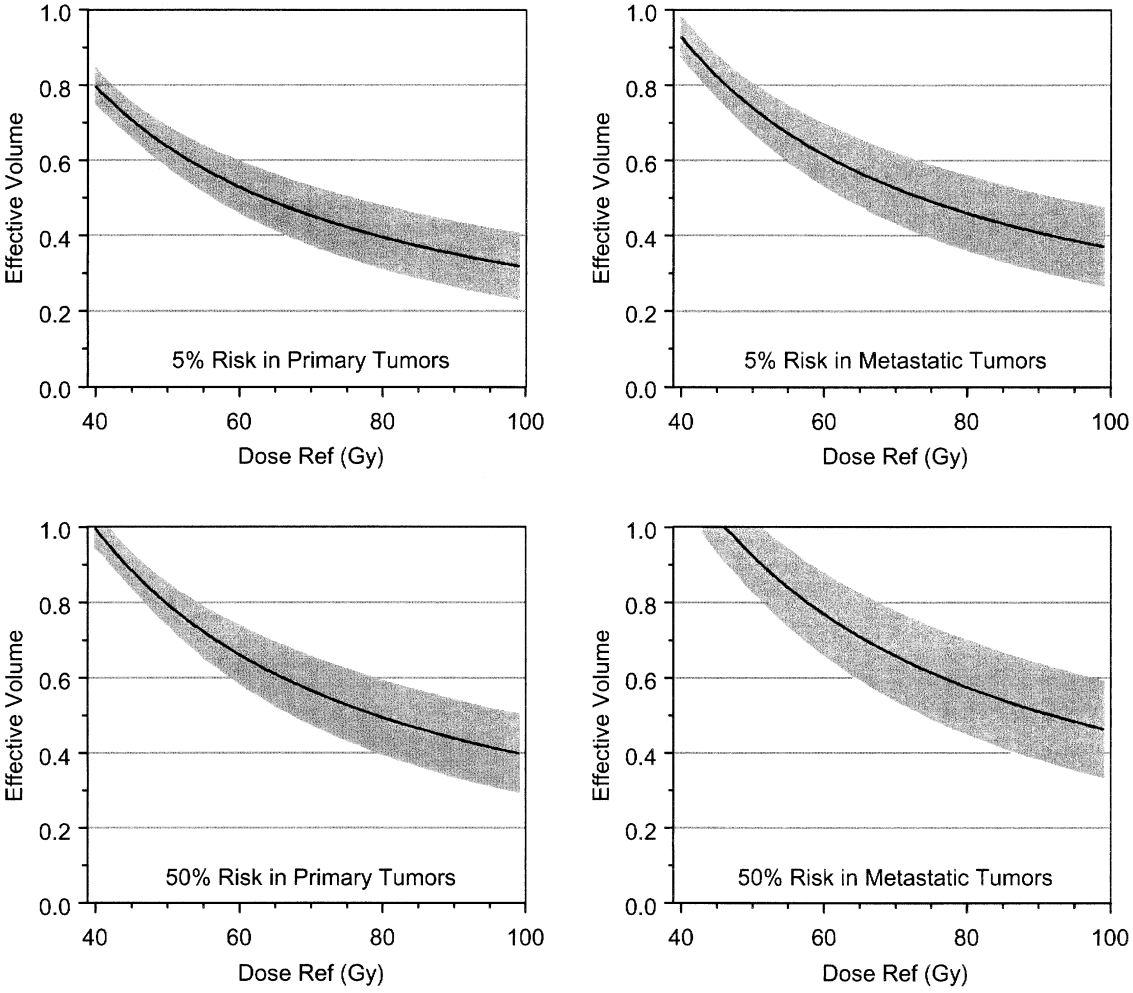

Fig. 4. Five and fifty percent iso-NTCP curves and 80% confidence limits for patients with primary hepatobiliary carcinoma and liver metastases treated with hepatic arterial FUdR using the LKB NTCP model [n = 0.97, m = 0.12, $TD_{50}(1)_{HB}$ = 39.8 Gy, and $TD_{50}(1)_{LM}$ = 45.8 Gy]. Effective volume (normal liver volume that, if irradiated uniformly, would be associated with the same NTCP as the nonuniform dose distribution actually delivered) vs. reference dose (prescribed dose normalized to 1.5 Gy b.i.d.).

DISCUSSION

Inherent biologic uncertainties are present in all NTCP models, and some authors have challenged their utility (25). Thus, clinical correlation with the NTCP predictions is needed. This report describes the largest series of patients with intrahepatic cancer treated with prospective dose–volume analyses and evaluation for RILD. We used the LKB NTCP model to study the dose–volume tolerance for RILD. Using this model with revised parameters, we were better able to describe the risk of RILD, compare rival radiation plans, and assign radiation doses for an individual patient on the basis of the predicted NTCP. The revised models suggest that patients with primary hepatobiliary malignancies have a lower tolerance to liver radiation than do patients with liver metastases. In addition, patients with a smaller volume of liver irradiated may be able to receive higher doses than previously estimated.

In this study, a diagnosis of primary hepatocellular carcinoma was associated with a significantly increased risk of RILD compared with a diagnosis of liver metastases. Although no patient had liver disease that substantially altered synthetic liver function (as assessed by prothrombin time), most patients with primary hepatobiliary carcinoma have preexisting cirrhosis or hepatitis, which may decrease the liver tolerance to radiation. It is not possible to comment on the development of RILD in patients with more advanced preexisting liver disease, because these patients were excluded from our studies. Not only may a different dose–volume relationship exist for RILD in patients with abnormal liver function, different types of radiation-associated liver injury may also occur, including exacerbation of preexisting hepatitis. Male gender was also associated with an increased risk of RILD, and this effect was most substantial among patients with a diagnosis of primary hepatobiliary cancer. BUdR hepatic arterial chemotherapy was associated

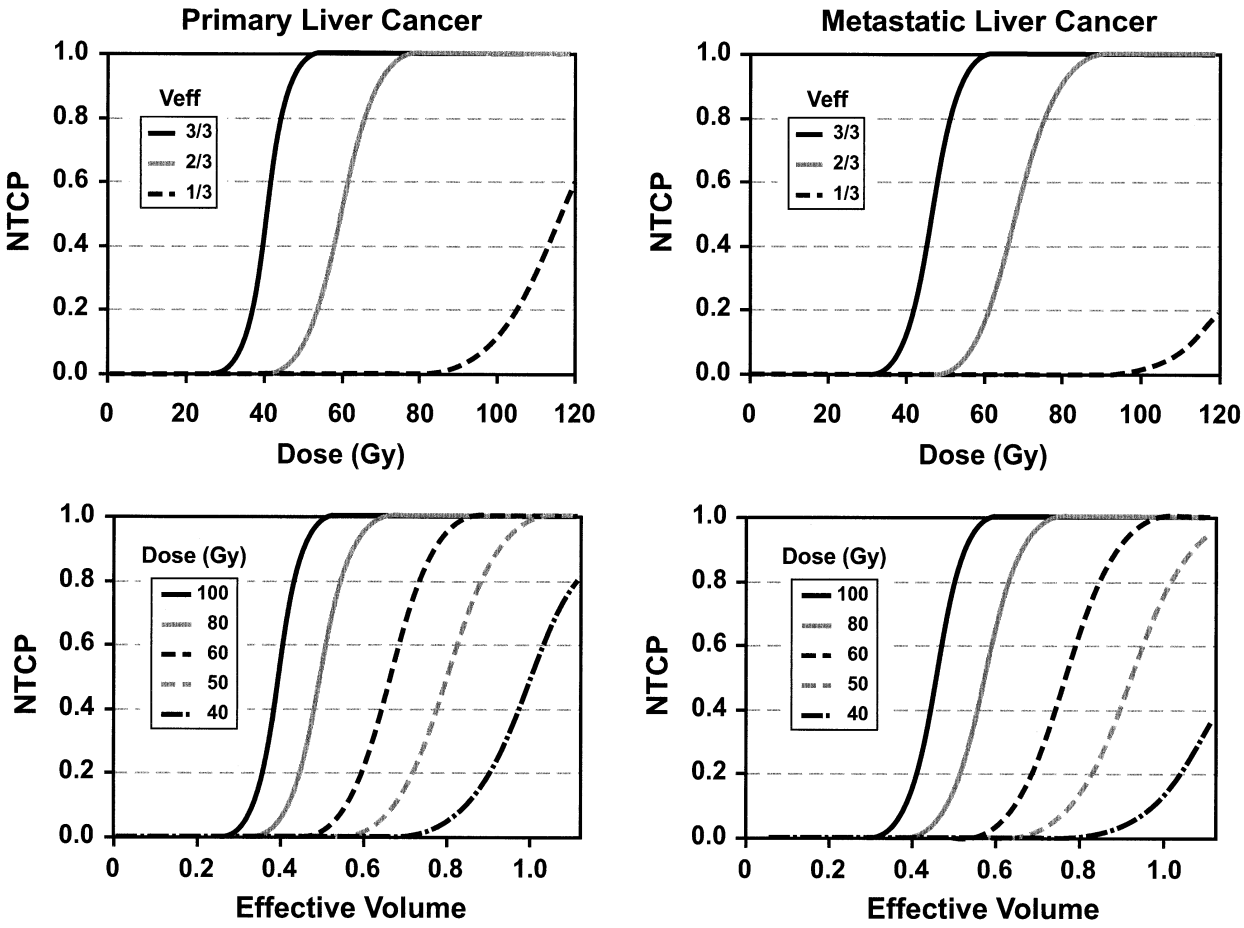

Fig. 5. NTCP vs. reference dose as a function of the volume irradiated uniformly (V_{eff}) (top row) and NTCP vs. V_{eff}, as a function of reference dose (prescribed dose normalized to 1.5 Gy b.i.d.) (bottom row), based on LKB NTCP model [n = 0.97, m = 0.12, $TD_{50}(1)_{HB}$ = 39.8 Gy, and $TD_{50}(1)_{LM}$ = 45.8 Gy], for patients with primary hepatobiliary cancer (left) and colorectal liver metastases (right).

with an increased risk of RILD compared with FUdR hepatic artery chemotherapy. Many of the patients who received hepatic artery BUdR also received a component of whole liver irradiation (24–31 Gy). Because the whole liver tolerance to radiation with hepatic arterial FUdR in this study was no different than the whole liver tolerance without chemotherapy previously reported (26), we hypothesize that hepatic arterial FUdR does not substantially alter the partial liver tolerance to radiation and the NTCP model parameter estimates. However, the use of other concurrent chemotherapeutic agents has been associated with a lower whole liver tolerance to radiation.

On the basis of our analysis using the LKB NTCP model, a larger volume effect for RILD than originally described (24) has been substantiated. Although the newest parameter estimates described our data well, our previous parameters estimates also predicted a large volume effect, and a clear distinction between these two parameter sets for all 203 patients was not observed. This was probably due to the lack of complications at low effective volumes/high doses (no complications observed with tumor dose >60 Gy), which would have a large influence on the exact value of the volume–effect parameter "n." However, the strong influence of the nondosimetric prognostic factors, including liver cancer type (primary vs. metastatic liver cancer), likely obscured the ability to make more refined estimates for the group as a whole.

Thus, in accordance with the large volume effect, the mean liver dose was found to be strongly associated with the development of RILD. The mean liver dose was related to NTCP (based on the LKB model) in a sigmoidal relationship, with a threshold for RILD at a mean liver dose of 30 Gy, and a 5% and 50% probability of RILD associated with a mean liver dose of 31 and 43 Gy, respectively (for the group of 203 patients studied). Of note, these liver tolerances are similar to the whole liver tolerances described by Emami *et al.* (27) in 1991 (5% and 50% risk of RILD with whole liver irradiation of 30 and 40 Gy, respectively). The mean liver dose is a relatively simple parameter that may be used to help rank plans and estimate the risk of RILD. The importance of the mean dose in predicting NTCP has been previously observed in the liver (16) and has also been

described for conformal therapy related to other organs, including the parotid salivary glands and lungs (28, 29).

Although the NTCP and mean liver dose may be useful in estimating the risk of RILD, there are limitations to using these factors as predictors. The LKB NTCP model and mean liver dose do not take into consideration volume thresholds for RILD, and they continually penalize (increase the NTCP prediction) for even very small volumes irradiated to high doses. As previously described by Jackson *et al*. (17), when the partial liver volume irradiated is kept below a threshold volume, the risk of RILD is estimated to be near 0, regardless of the radiation dose delivered. Our data are consistent with a threshold effect. For small volumes of normal liver treated (approximately one-third of the whole liver), doses as high as 100 Gy are predicted to be associated with little or no risk of toxicity for each of the patient subgroups studied. Finally, different radiation techniques or fractionation schemes used in other patient populations may be associated with different RILD partial organ tolerances, and our results may not be widely applicable.

A limitation of using one CT scan to obtain the DVHs on which the NTCP parameters are based is that the NTCP estimates can differ when ventilatory liver motion and setup uncertainty are accounted for (30). Although patients in the present study treated since 1996 were scanned in a breath-hold position, previous patients were scanned during free breathing. We are currently studying the magnitude of the effect of organ positional uncertainties on NTCP parameter determination for RILD. As we develop techniques for control of respiratory organ motion during radiation (31), the DVHs will better reflect the treatment delivered.

We now treat patients with unresectable intrahepatic malignancies with hepatic arterial FUdR and escalated focal liver irradiation, using the revised LKB NTCP model (169 patients/FUdR/4-parameter fit) to facilitate dose escalation, up to a maximal dose of 90 Gy (13). The use of separate models [different $TD_{50}(1)$ values] for patients with primary hepatobiliary malignancies and patients with liver metastases reflects the lower tolerance of the liver to radiation in primary hepatobiliary malignancies. In this approach, the iso-NTCP curve displaying the dose–volume pairs that subject a patient to an acceptable risk is chosen (15). After treatment planning has been performed to minimize the volume of liver treated, the dose prescribed is that which subjects each patient to a predetermined risk of RILD. As new technology that decreases the volume of normal liver irradiated is introduced, higher doses of radiation may be delivered by moving up the NTCP curve with the same risk of RILD. An alternative option is to use the mean liver dose to describe the risk of RILD. Using the revised LKB parameters, patients with small volumes of their liver irradiated may be able to receive much higher radiation doses to intrahepatic tumors than previously possible. Thus, it is possible that extrahepatic organs such as the small bowel may become dose limiting in the RT for some intrahepatic malignancies.

In an effort to allow even more patients with unresectable intrahepatic malignancies to receive higher doses safely, we use active breathing control (32) to suspend breathing during RT. By treating patients with their respiration suspended using active breathing control, we are able to decrease the amount of normal liver that has to be treated to account for organ motion due to breathing, producing a lower risk of RILD (33). Other potential methods of increasing the radiation dose safely include the use of radioprotective agents directed to the normal liver. Preclinical studies have demonstrated that amifostine can protect the normal liver without compromising tumor cell kill (34). Finally, it may be possible to select patients at low risk of RILD for higher dose escalation on the basis of serum transforming growth factor-β levels or other as yet unknown serum markers that may predict RILD risk before its clinical development. These interventions, along with a better understanding of the dose–volume tolerance of the liver to radiation, may allow additional dose escalation to tumors and avoidance of radiation injury to normal organs, leading to improved outcomes for patients with unresectable intrahepatic malignancies.

REFERENCES

1. Lawrence TS, Robertson JM, Anscher MS, *et al*. Hepatic toxicity resulting from cancer treatment. *Int J Radiat Oncol Biol Phys* 1995;31:1237–1248.
2. Ingold J, Reed G, Kaplan H. Radiation hepatitis. *AJR Am J Roentgenol* 1965;93:200–208.
3. Poussin-Rosillo H, Nisce LZ, D'Angio GJ. Hepatic radiation tolerance in Hodgkin's disease patients. *Radiology* 1976;121: 461–464.
4. Russell AH, Clyde C, Wasserman TH, *et al*. Accelerated hyperfractionated hepatic irradiation in the management of patients with liver metastases: Results of the RTOG dose escalating protocol. *Int J Radiat Oncol Biol Phys* 1993;27: 117–123.
5. Austin-Seymour MM, Chen GT, Castro JR, *et al*. Dose volume histogram analysis of liver radiation tolerance. *Int J Radiat Oncol Biol Phys* 1986;12:31–35.
6. Dawson L, McGinn C, Normolle D, *et al*. Escalated focal liver radiation and concurrent hepatic artery fluorodeoxyuridine for unresectable intrahepatic malignancies. *J Clin Oncol* 2000;18: 2210–2218.
7. Robertson JM, Lawrence TS, Andrews JC, *et al*. Long-term results of hepatic artery fluorodeoxyuridine, and conformal radiation therapy for primary hepatobiliary cancers. *Int J Radiat Oncol Biol Phys* 1997;37:325–330.
8. Mohiuddin M, Chen E, Ahmad N. Combined liver radiation and chemotherapy for palliation of hepatic metastases from colorectal cancer. *J Clin Oncol* 1996;14:722–728.
9. Lyman JT. Complication probability as assessed from dose-volume histograms. *Radiat Res Suppl* 1985;8:S13–S19.
10. Jackson A, Kutcher GJ, Yorke ED. Probability of radiation induced complications for normal tissues with parallel architecture subject to non-uniform irradiation. *Med Phys* 1993;20: 613–625.
11. Niemierko A, Goitein M. Modeling of normal tissue response

to radiation: The critical volume model. *Int J Radiat Oncol Biol Phys* 1993;25:135–145.
12. Yorke ED, Jackson A, Fox RA, *et al*. Can current models explain the lack of liver complications in Y-90 microsphere therapy? *Clin Cancer Res* 1999;5:3024s–3030s.
13. McGinn CJ, Ten Haken RK, Ensminger WD, *et al*. Treatment of intrahepatic cancers with radiation doses based on a normal tissue complication probability model. *J Clin Oncol* 1998;16: 2246–2252.
14. Kutcher GJ, Burman C. Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method. *Int J Radiat Oncol Biol Phys* 1989; 16:1623–1630.
15. Ten Haken RK, Martel MK, Kessler ML, *et al*. Use of V_{eff} and iso-NTCP in the implementation of dose escalation protocols. *Int J Radiat Oncol Biol Phys* 1993;27:689–695.
16. Lawrence TS, Ten Haken RK, Kessler ML, *et al*. The use of 3-D dose volume analysis to predict radiation hepatitis. *Int J Radiat Oncol Biol Phys* 1992;23:781–788.
17. Jackson A, Ten Haken RK, Robertson JM, *et al*. Analysis of clinical complication data for radiation hepatitis using a parallel architecture model. *Int J Radiat Oncol Biol Phys* 1995; 31:883–891.
18. Ten Haken RK, Lawrence TS, McShan DL, *et al*. Technical considerations in the use of 3-D beam arrangements in the abdomen. *Radiother Oncol* 1991;22:19–28.
19. Balter JM, Lam KL, McGinn CJ, *et al*. Improvement of CT-based treatment-planning models of abdominal targets using static exhale imaging. *Int J Radiat Oncol Biol Phys* 1998;41:939–943.
20. Fraass B, McShan D. 3-D treatment planning: I. Overview of a clinical planning system. In: van Klefffens H, Wittkamper F, editors. The use of computers in radiation therapy. North-Holland: Elsevier Science Publishers BV; 1987. p. 273–276.
21. Roberts SA, Hendry JH. The delay before onset of accelerated tumour cell repopulation during radiotherapy: A direct maximum-likelihood analysis of a collection of worldwide tumour-control data. *Radiother Oncol* 1993;29:69–74.
22. Cox D, Snell E. Analysis of binary data. New York: Chapman and Hall; 1989. p. 179–184.
23. Agresti A. Categorical data analysis. New York: John Wiley & Sons; 1990.
24. Burman C, Kutcher GJ, Emami B, *et al*. Fitting of normal tissue tolerance data to an analytic function. *Int J Radiat Oncol Biol Phys* 1991;21:123–135.
25. Langer M, Morrill SS, Lane R. A test of the claim that plan rankings are determined by relative complication and tumor-control probabilities. *Int J Radiat Oncol Biol Phys* 1998;41: 451–457.
26. Russell AH, Clyde C, Wasserman TH, *et al*. Accelerated hyperfractionated hepatic irradiation in the management of patients with liver metastases: Results of the RTOG dose escalating protocol. *Int J Radiat Oncol Biol Phys* 1993;27: 117–123.
27. Emami B, Lyman J, Brown A, *et al*. Tolerance of normal tissue to therapeutic irradiation. *Int J Radiat Oncol Biol Phys* 1991;21:109–122.
28. Eisbruch A, Ten Haken RK, Kim HM, *et al*. Dose, volume, and function relationships in parotid salivary glands following conformal and intensity-modulated irradiation of head and neck cancer. *Int J Radiat Oncol Biol Phys* 1999;45:577–587.
29. Kwa SL, Lebesque JV, Theuws JC, *et al*. Radiation pneumonitis as a function of mean lung dose: An analysis of pooled data of 540 patients. *Int J Radiat Oncol Biol Phys* 1998;42: 1–9.
30. Lujan AE, Ten Haken RK, Larsen EW, *et al*. Quantization of setup uncertainties in 3-D dose calculations. *Med Phys* 1999; 26:2397–2402.
31. Dawson LA, Brock KK, Litzenberg D, *et al*. The reproducibility of organ position using active breathing control (ABC) during liver radiotherapy. *Int J Radiat Oncol Biol Phys* 2001; 51:1410–1421.
32. Wong JW, Sharpe MB, Jaffray DA, *et al*. The use of active breathing control (ABC) to reduce margin for breathing motion. *Int J Radiat Oncol Biol Phys* 1999;44:911–919.
33. Ten Haken RK, Balter JM, Marsh LH, *et al*. Potential benefits of eliminating planning target volume expansions for patient breathing in the treatment of liver tumors. *Int J Radiat Oncol Biol Phys* 1997;38:613–617.
34. Symon Z, Levi M, Ensminger WD, *et al*. Selective radioprotection of hepatocytes by systemic and portal vein infusions of amifostine in a rat liver tumor model. *Int J Radiat Oncol Biol Phys* 2001;50:473–478.
35. Billingsley P. Probability and measure. 2nd ed. New York: John Wiley & Sons; 1986.

APPENDIX 1

Maximal likelihood estimation

Normal liver DVHs and the occurrence or lack of occurrence of RILD from 203 patients comprised the input data for the determination of the NTCP model parameters using a maximal likelihood analysis. The parameters $TD_{50}(1)$, m, and n were adjusted to maximize the probability of predicting complications for those patients who experienced complications and to maximize the probability of predicting no complications for those patients who were complication free.

For each patient i, let the vectors d_i and v_i be the normalized dose (d_i)–volume (v_i) histogram, when the components of v_i sum to 1 and the components of d_i are divided by the reference dose (the prescribed target dose), D_i, so that the largest component of d_i equals, or is very close to, 1. Let the components of d_i and v_i be d_{ij} and v_{ij}, respectively, where "j" represents the number of dose–volume bins for each patient's DVH (range of j approximately 50–150). Let $R_i = 1$ if the patient i experienced RILD and $R_i = 0$ otherwise.

Using the LKB NTCP model, the effective volume (V_{eff}) of the liver is

$$V_{eff_i} = \sum_j v_{ij} d_{ij}^{1/n}$$

where n represents the volume effect parameter relating the tolerance dose of uniform whole organ irradiation to uniform partial organ irradiation.

The dose associated with a 50% risk of complication (D_{50i}), related to V_{eff} for patient i is then as follows:

$$D_{50_i} = TD_{50}(1) \cdot V_{eff_i}^{-n}$$

where $TD_{50}(1)$ is the dose associated with a 50% risk of complication for whole organ irradiation.

A probit model was assumed for the probability of RILD of patient *i*:

$$p_i = p_i\,(m, n, TD_{50}(1); D_i, d_i, v_i) = \phi\left(\frac{D_i - D_{50}}{m \cdot D_{50_i}}\right)$$

The log-likelihood for the entire data set,

$$L(m, n, TD_{50}(1) = \sum_i \log(p_i)^{Ri} + \log(1 - p_i)^{1-Ri} \quad (1)$$

was then maximized over all feasible values of $TD_{50}(1)$, m, and n using the Newton-Raphson method implemented in Statistical Analysis System PROC Mixed software (SAS Institute, Cary, NC).

Given the limited number of complications (19), estimates of all three parameters in arbitrary subsets of the 203 patients were not generally possible. Subset specific estimates of $TD_{50}(1)$ could be obtained while estimating the n and m parameters from the entire data set. For instance, the whole liver tolerance was estimated for patients with primary hepatobiliary cancer (HB) vs. metastatic liver (LM) cancer by maximizing the log-likelihood:

$$\begin{aligned} &L(m, n, TD_{50}(1)_{HB}, TD_{50}(1)_{LM}) \\ &= \sum_{\text{HB patients}} \log(p_i(m, n, TD_{50}(1)_{HB}, D_i, d_i, v_i))^{Ri} \\ &\quad + \log(1 - p_i(m, n, TD_{50}(1)_{HB}, D_i, d_i, v_i))^{1-Ri} \\ &\quad + \sum_{\text{LM patients}} \log(p_i(m, n, TD_{50}(1)_{LM}; D_i, d_i, v_i))^{Ri} \\ &\quad + \log(1 - p_i(m, n, TD_{50}(1)_{LM}; D_i, d_i, v_i))^{1-Ri} \end{aligned}$$

APPENDIX 2

Confidence intervals for the dose–V_{eff} iso-NTCP curves

The confidence intervals for the dose–V_{eff} curves were obtained using a method to determine the confidence intervals for nonlinear functions of parameters estimates from a nonlinear model, as follows.

The confidence intervals for a given risk level are constructed by determining a confidence band for V_{eff} for each dose. The NTCP model is given by

$$p = \Phi\left(\frac{D - TD_{50}(1) \cdot V_{eff}^{-n}}{m \cdot TD_{50}(1) \cdot V_{eff}^{-R}}\right)$$

For a given risk, p, solve this for V_{eff} in terms of D:

$$V_{eff}(D; n, m, TD_{50}(1)) = \left(\frac{D}{nm\Phi^{-1}(p) + 1/m}\right)^{-1/n}$$

substituting $TD_{50}(1)_{HB}$ or $TD_{50}(1)_{LM}$ in the above formula depending on whether the patient has primary (HB) or metastatic cancer (LM).

For a given D, this is a nonlinear function of the parameter estimates. The estimates themselves have an approximate variance–covariance matrix that is a function of the Hessian [3×3 second-derivative matrix for parameters n, m, and $TD_{50}(1)$] at the end of the iterative solution process.

Given the function V_{eff} [D, n, m, $TD_{50}(1)$], parameter estimates $\theta = [n\ m\ TD_{50}(1)]$ and estimated covariance matrix $\hat{\Sigma}$, the delta method (35) may be applied, which, in brief, states that the distribution of the function of parameters about the function applied to their expected value is asymptotically normal, with variance:

$$\mathrm{Var}(\sigma') = \frac{\partial Veff'}{\partial \theta} \sum \frac{\partial Veff}{\partial \theta}$$

Estimates for θ and Σ are substituted into the above formula to produce the estimated variance, and Wald confidence intervals are thereby calculated. This is available directly in Statistical Analysis System PROX NLMIXED software (SAS Institute).

APPENDIX 3

Goodness of fit

The deviance of a given set of parameters is related to the log-likelihood (Eq. 1) as follows:

$$D(m, n, TD_{50}(1) = -2\sum_i \log(p_i)^{Ri} + \log(1 - p_i)^{1-Ri}$$

The deviance has, asymptotically, an approximate χ^2 distribution with $N - 3$ degrees of freedom, where N is the number of observations (3 is the number of estimated parameters). Thus, the goodness of fit of a given set of parameters can be assessed by $1 - \chi^2\,\{D[m, n, TD_{50}(1)], N - 3\}$, where $\chi^2\,(\cdot,\cdot)$ is the cumulative central χ^2 distribution function, and values closer to 0 indicate a significant lack of fit. Although this approximation is poor when modeling binary data with continuous variables (23), the statistic was used as a general guide.

If two models are fit according to the maximal likelihood method in which one is a nested within the other (i.e., the models are of the same form but certain parameters in one model are constrained to equal 0), the difference of the

deviances has an approximate χ^2 distribution with degrees of freedom equal to the number of constrained parameters, and values close to 0 indicate that constraining the parameters significantly degrades the fit. This statistic converges to the χ^2 significantly faster than the deviances of the individual models. This statistic cannot be used to compare two non-nested models.

p values were used to compare across sample sizes. A large goodness of fit *p* value implies a better fit than a smaller *p* value.