ORIGINAL ARTICLE

Colorectal Liver Metastases

Recurrence and Survival Following Hepatic Resection, Radiofrequency Ablation, and Combined Resection–Radiofrequency Ablation

Ana L. Gleisner, MD; Michael A. Choti, MD, MBA; Lia Assumpcao, MD;
Hari Nathan, MD; Richard D. Schulick, MD; Timothy M. Pawlik, MD, MPH

Hypothesis: Although radiofrequency ablation (RFA) is increasingly an accepted option for patients with colorectal liver metastases, patients treated with resection vs RFA may have different tumor biology profiles, which might confound the relationship between choice of liver-directed therapy and outcome.

Design: Retrospective review of a prospectively collected database.

Setting: Major hepatobiliary center.

Patients: Between January 1, 1999, and August 30, 2006, 258 patients with colorectal liver metastases underwent hepatic resection with or without RFA.

Main Outcome Measures: Evaluation of outcome following resection alone, combined resection-RFA, and RFA alone using 3 statistical methods (paired-match control, Cox proportional hazards multivariate model, and propensity index) to identify and adjust for potential confounding variables.

Results: The median number of hepatic lesions was 2, and the median size of the largest lesion was 3.0 cm. One hundred ninety-two patients (74.4%) underwent resection alone, 55 patients (21.3%) underwent resection-RFA, and 11 patients (4.3%) underwent RFA alone. Patients who underwent resection-RFA had significantly increased risk of extrahepatic failure at 1 year vs patients who underwent resection alone or RFA alone ($P < .05$). On matched control and multivariate analyses, patients who underwent RFA with or without resection had significantly worse disease-free and overall survival than patients who underwent resection alone. Propensity score methods revealed that the aggregate distribution of clinical risk factors for resection-RFA was markedly different from that for resection alone. This suggested a lack of comparability to allow for statistical comparisons in the assessment of causal inferences regarding the efficacy of RFA therapy.

Conclusion: Although results of matched control and multivariate analyses suggested that RFA with or without resection was associated with worse outcome, propensity score methods revealed that the resection-RFA and resection-alone groups were different with regard to baseline tumor and treatment-related factors, making causal inferences about the efficacy of RFA unreliable.

Arch Surg. 2008;143(12):1204-1212

COLORECTAL CANCER IS THE fourth most common type of cancer in the Western world and is the second leading cause of cancer-related deaths in the United States.[1] Approximately 35% to 55% of patients with colorectal cancer develop hepatic metastases during the course of their disease.[2] Surgical resection of colorectal liver metastases is associated with long-term survival rates of 35% to 58%.[3-9] Radiofrequency ablation (RFA) has been proposed as a means to increase the number of patients eligible for liver-directed therapy. In patients with extensive metastatic disease who would otherwise be ineligible for resection, ablative approaches are often used instead of or combined with hepatic resection.[3,10-12]

In a recent study comparing recurrence and outcome following hepatic resection alone, combined resection-RFA, and RFA alone for colorectal liver metastases, Abdalla et al[3] reported significantly worse disease-free and overall survival for patients treated with RFA. In contrast, other studies[13-15] have reported more encouraging results following RFA therapy, with 3- and 5-year survival rates in excess of 55% and 30%, respectively. However, survival following RFA is difficult to interpret because many patients who undergo this therapy are characterized by poor prognostic factors (eg, unresectable disease, multiple tumors, or bilobar location), making comparisons with

Author Affiliations: Department of Surgery, The Johns Hopkins University School of Medicine, Baltimore, Maryland.

patients who have undergone complete surgical resection difficult.

Some investigators have argued that the main determinant of survival may not be the mode of treatment but rather the patient's disease, as disease extent seems to determine the modality of treatment.[3] As such, retrospective studies that compare resection vs RFA have been criticized for comparing apples to oranges.[3] Patients treated with resection vs RFA may have different tumor biologic profiles that might confound the relationship between choice of liver-directed therapy and outcome. The objective of the present study was to evaluate outcome following resection alone, combined resection-RFA, and RFA alone. We sought to determine the incidence of systemic and intrahepatic recurrence following liver-directed therapy. In addition, disease-free and overall survival were assessed using 3 statistical methods (paired-match control, Cox proportional hazards multivariate model, and propensity index) in an attempt to identify and adjust for potential confounding variables.

METHODS

Five hundred six consecutive patients who underwent hepatic resection alone, resection-RFA, or RFA alone for colorectal liver metastases at The Johns Hopkins Hospital between March 1, 1986, and August 30, 2006, were identified from our prospective institutional database. Only patients with colorectal liver metastases who were operated on with curative intent were included in the study. In addition, only patients undergoing their first liver-directed therapy were included. Similarly, only RFA treatments that were performed at the time of open laparotomy were included; patients who underwent percutaneous or laparoscopic-assisted RFA were excluded. To account for a possible period effect, only patients who underwent resection from January 1, 1999, onward (ie, when RFA was introduced as a therapeutic modality at our institution) were included in the present study. In total, 258 patients met the inclusion and exclusion criteria and were considered for further analyses.

SURGICAL RESECTION AND RFA

Patients were deemed to have resectable hepatic disease only if it was anticipated that their metastases could be completely resected, at least 2 adjacent liver segments could be spared, vascular inflow and outflow could be preserved, and the volume of the liver remaining after resection would be adequate.[16] Resection was classified as less than hemihepatectomy (ie, segmentectomy or subsegmentectomy), hemihepatectomy, or extended hepatectomy ($\geq$5 liver segments).[17]

Radiofrequency ablation of hepatic lesions was performed at the time of laparotomy according to a standardized treatment algorithm.[10,18] Intraoperative ultrasonography was used to insert needles into the lesions to be treated by RFA. Radiofrequency ablation was administered using an RFA generator (RITA Model 1500X; Rita Medical Systems, Inc, Fremont, California) with an enhanced device (Starburst XL or XLi, Rita Medical Systems, Inc) wherever applicable. The Starburst XL was used to create ablations up to 5 cm, while the Starburst XLi was used to create ablations up to 7 cm. Under intraoperative ultrasonographic guidance, the electrode was optimally positioned to achieve complete destruction of the tumor and at least a 1-cm zone of normal liver parenchyma when possible.

Patients were treated with RFA alone or in combination with resection when at least 1 hepatic tumor was considered unresectable because of location of the disease, inadequate liver remnant, proximity of tumor to major vascular structures, or the presence of medical comorbidities that precluded major hepatic resection. Tumors were considered for RFA if near a major hepatic vein branch but not if adjacent to major biliary structures near the liver hilum.

DATA COLLECTION

The following data were collected for each patient: disease status, demographics, laboratory data, operative details, date of last follow-up, type of chemotherapy, response to neoadjuvant chemotherapy, administration and timing of chemotherapy, and date of death, as well as tumor number, size, and location. Tumor size and number were defined by the resection specimen or by intraoperative ultrasonographic measurement. Based on the main agent used in the regimen, chemotherapy was classified as follows: (1) no preoperative chemotherapy, (2) fluoropyrimidine-based chemotherapy with fluorouracil and leucovorin calcium, (3) irinotecan hydrochloride–based chemotherapy, or (4) oxaliplatin-based chemotherapy. Response to neoadjuvant chemotherapy was defined as at least a 25% reduction in tumor bidimensional measurements demonstrated on preoperative computed tomography or magnetic resonance imaging. Data on recurrence were categorized as referable to any site, intrahepatic only (eg, any intrahepatic site), or extrahepatic. The objective of the present study was not to evaluate the local efficacy of RFA; as such, true local recurrence on a per-lesion ablation basis was not evaluated.

STATISTICAL ANALYSIS

All statistical analyses were performed using commercially available software (SPSS version 11.5; SPSS Inc, Chicago, Illinois; and STATA, StataCorp LP, College Station, Texas). Summary statistics were obtained using established methods and were presented as percentages, means, and medians with their respective standard deviation and interquartile range (IQR). t Test or analysis of variance was used for comparison of variables that were normally distributed, while Kruskal-Wallis test was used to compare skewed continuous variables. χ^2 Statistic was used to compare frequencies of categorical variables among groups. Time to first recurrence and survival outcome were estimated using Kaplan-Meier nonparametric product-limit method.[19] Unadjusted differences in survival were examined using log-rank test.

Three statistical methods were used to control for possible confounding variables among the resection-alone, resection-RFA, and RFA-alone groups. First, patients who underwent RFA with or without resection (ie, cases) were matched 1:1 with patients who underwent resection alone (ie, controls). Patients were matched based on their metastatic lesion size and number and their primary T and N stages. Patterns of recurrence, disease-free survival, and overall survival were compared among the case and control groups. Second, multivariate Cox proportional hazards regression analyses were performed. Variables that were significant on univariate analysis or variables that were unbalanced among the treatment groups were included in the final multivariate model. Hazard ratios and 95% confidence intervals were estimated, and $P < .05$ was considered statistically significant. Third, propensity score methods were used to control for any systematic differences in the background characteristics between patients undergoing RFA vs resection. The propensity score method is used to determine the probability of an individual patient having received a certain treatment (herein RFA) as a function of several confounding covariates that are collapsed into a single predictor.[20] The propensity index provides a single composite

Table 1. Clinicopathologic Characteristics of the Study Patients

Characteristic	Value (N=258)
Patients	
Age, median (IQR), y	61 (53.0-69.5)
Male sex, No. (%)	169 (65.5)
White race/ethnicity, No. (%)	224 (86.8)
Primary Tumor	
Preoperative CEA level, median (IQR), ng/mL	8.0 (3.2-30.8)
T stage, No. (%)	
T1 or T2	37 (14.3)
T3 or T4	221 (85.7)
Positive lymph nodes, No. (%)	171 (66.3)
Synchronous, No. (%)	71 (27.5)
Hepatic Metastases	
Size of largest lesion, median (IQR), cm	3.0 (2.0-5.0)
Patients having solitary metastasis, No. (%)	119 (46.1)
No. of treated metastases, median (IQR)	2 (1-3)
Preoperative chemotherapy, No. (%)	**115** (44.6)
Fluorouracil alone	16 (6.2)
Oxaliplatin	35 (13.6)
Irinotecan hydrochloride	46 (17.8)
Unknown	18 (6.9)
Response to chemotherapy, No. (%) (n = 97)	
Responsive	61 (62.9)
Stable	20 (20.6)
Progression	10 (10.3)
Unknown	6 (6.2)

Abbreviations: CEA, carcinoembryonic antigen; IQR, interquartile range.

SI conversion factor: To convert CEA level to micrograms per liter, multiply by 0.1.

Table 2. Details of Surgical Procedures Stratified by Treatment Group

Surgical Procedure	Resection Alone (n=192)	Resection-RFA (n=55)	RFA Alone (n=11)	P Value
Resection, No. (%)				
Wedge or segmentectomy	100 (52.1)	47 (85.5)	. . .	
Hemihepatectomy	78 (40.6)	7 (12.7)	. . .	
Extended hepatectomy	14 (7.3)	1 (1.8)	. . .	<.001[a]
RFA				
No. of treated metastases, median	. . .	2	1	.03[b]
Lesion size, median, cm	. . .	2.4	3.0	.29

Abbreviations: RFA, radiofrequency ablation; ellipses, not applicable.

[a] Resection alone vs combined resection-RFA.

[b] Resection-RFA vs RFA alone.

score that appropriately summarizes the collection of background characteristics for each patient cohort, allowing straightforward assessment of whether groups overlap sufficiently with respect to baseline characteristics to allow comparison between groups in a given data set.[20] In the present study, the propensity score was determined by predicting treatment group membership using logistic regression analysis, where the outcome was ablation and the covariates were the following: age, sex, disease-free interval, chemotherapy response, liver tumor number and size, primary T and N stages, history of preoperative or postoperative chemotherapy, and preoperative carcinoembryonic antigen (CEA) level exceeding 100 ng/mL (to convert CEA level to micrograms per liter, multiply by 0.1). Wherever appropriate, the propensity score was divided into quartiles and entered into Cox proportional hazards regression models along with the treatment group covariate (resection alone, resection-RFA, or RFA alone) to assess disease-free and overall survival.

Figure 1. For 11 patients who underwent radiofrequency ablation (RFA) alone, a total of 14 tumors were ablated. Five patients had a solitary lesion ablated that abutted the confluence of the hepatic veins (as shown) and was unamenable to resection.

RESULTS

CLINICOPATHOLOGIC CHARACTERISTICS

Table 1 gives the clinicopathologic features of 258 patients in the study. At the time of operation, surgical treatment was resection alone in 192 patients (74.4%), resection-RFA in 55 patients (21.3%), and RFA alone in 11 patients (4.3%) (**Table 2**). Among 192 patients who underwent resection alone, the extent of hepatic resection was less than hemihepatectomy in 100 patients (52.1%), hemihepatectomy in 78 patients (40.6%), and extended hepatectomy in 14 patients (7.3%). In contrast, among 55 who underwent resection-RFA, patients were less likely to undergo hemihepatectomy (7 patients [12.7%]) or extended hepatectomy (1 patient [1.8%]) ($P < .001$ for both).

For 11 patients who underwent RFA alone, a total of 14 tumors were ablated. Five patients had a solitary lesion ablated that abutted the confluence of the hepatic veins and was deemed unamenable to resection (**Figure 1**). The other 6 patients underwent RFA rather than resection because of medical comorbidities or underlying hepatic parenchymal injury that, in the surgeon's opinion, precluded safe performance of a major hepatic resection.

Although the distributions of certain risk factors (eg, disease-free interval, median preoperative CEA level, and primary tumor characteristics) were similar among the 3 study groups, univariate analyses revealed several tumor and treatment-related differences (**Table 3**). Patients who underwent resection alone were more likely to have larger tumors (median size, 3.5 cm; IQR, 2.0-5.0 cm) vs patients who underwent resection-RFA (me-

Table 3. Clinicopathologic Characteristics and Treatment-Relative Variables of Patients Stratified by Treatment Group

Characteristic	Resection Alone (n=192)	Resection-RFA (n=55)	RFA Alone (n=11)	*P* Value
Patients				
Age, median, y	61	61	60	.96
Male sex, No. (%)	121 (63.0)	40 (72.7)	8 (72.7)	.36
White race/ethnicity, No. (%)	167 (87.0)	50 (90.9)	7 (63.6)	.21
Primary Tumor				
Preoperative CEA level, median, ng/mL	4.3	6.3	6	.39
T stage				
T1 or T2	28 (14.6)	9 (16.4)	0	
T3 or T4	164 (85.4)	46 (83.6)	11 (100.0)	.72
Positive lymph nodes, No. (%)	122 (63.5)	42 (76.4)	7 (63.6)	.14
Synchronous, No. (%)	56 (29.2)	13 (23.6)	3 (27.3)	.16
Hepatic Metastases				
Size of largest lesion, median, cm	3.5	2.5	2.5	.02
Patients having solitary metastasis, No. (%)	112 (58.3)	0	7 (63.6)	<.001
No. of treated metastases, median	1	5	1	<.001
Preoperative chemotherapy, No. (%)	73 (38.0)	36 (65.5)	6 (54.5)	.001
Response to chemotherapy, No. (%)[a]				
Responsive	37 (50.7)	22 (61.1)	2 (33.3)	
Stable	13 (17.8)	7 (19.4)	0	
Progression	5 (6.8)	2 (5.6)	3 (50.0)	
Unknown	18 (24.7)	5 (13.9)	1 (16.7)	.16
Adjuvant hepatic arterial infusion therapy, No. (%)	15 (7.8)	11 (20.0)	2 (18.2)	.02
Positive microscopic surgical margins, No. (%)	7 (3.6)	9 (16.4)	0	.001

Abbreviation: CEA, carcinoembryonic antigen; RFA, radiofrequency ablation.
SI conversion factor: To convert CEA level to micrograms per liter, multiply by 0.1.
[a] Responses to chemotherapy for each column are 73 for resection alone, 36 for resection-RFA, and 6 for RFA alone.

dian size, 2.5 cm; IQR, 1.9-4.0 cm) (*P*=.02). In contrast, patients who underwent resection alone had fewer hepatic metastases (median, 1 metastasis; IQR, 1-2 metastases) than patients who underwent resection-RFA (median, 5 metastases; IQR, 3-6 metastases) (*P*<.001). Among patients who underwent resection alone, 58.3% had solitary tumors (*P*<.001). Preoperative systemic chemotherapy was less commonly administered to patients before resection alone (38.0%) vs before resection-RFA (65.5%) (*P*<.001). On final pathologic analysis, patients who underwent resection alone (3.6%) were less likely to have a positive microscopic surgical margin (R1) than patients who underwent resection-RFA (16.4%) (*P*=.001).

PATTERNS OF RECURRENCE

Recurrence at any site most often occurred after resection-RFA or after RFA alone compared with after resection alone (**Figure 2**A). The 1-year risks of any site recurrence were 24.4% for resection alone, 60.5% for resection-RFA, and 65.9% for RFA alone (*P*<.001). Differences in recurrence patterns were observed among the treatment groups. The proportion of patients in each group who developed distant metastases as a component of extrahepatic failure was markedly different among the 3 groups (Figure 2B). Patients who underwent resection-RFA had more than a 3-fold increased risk of extrahepatic failure at 1 year compared with patients who underwent resection alone (40.6% vs 12.8%; *P*<.001). In contrast, patients treated with RFA alone had an extrahepatic failure rate of 21.2%. Recurrence anywhere in the liver was also more common at 1 year in patients treated with resection-RFA (50.9%) or with RFA alone (62.5%) vs with resection alone (14.8%) (Figure 2C). Liver-only recurrence (ie, recurrence anywhere in the liver, without evidence of extrahepatic disease) at 1 year was 2% in the resection-alone group compared with 10.3% in the resection-RFA group and 41.3% in the RFA-alone group (*P*<.001).

DISEASE-FREE AND OVERALL SURVIVAL

Unadjusted Analysis

Disease-free survival for the 3 treatment groups is shown in **Figure 3**A. No significant difference was noted in disease-free survival for patients treated with resection-RFA vs those who underwent RFA alone (3-year actuarial disease-free survival rates, 34.1% vs 7.4%; *P*=.20); in contrast, the 3-year disease-free survival for the resection-alone group (39.8%) was longer than that in the resection-RFA group (*P*=.01). On univariate analyses, several other factors were associated with disease-free survival. The following patient characteristics were associated with increased risk of recurrence and shorter disease-free survival (*P*<.05 for all): preoperative CEA level exceeding 100 ng/mL (hazard ratio [HR], 2.11), size of tumors ablated exceeding 3 cm (3.29), disease-free interval shorter than 12 months (1.57), lack of response to preoperative chemotherapy (2.33), number of tumors ablated (3.05 for 2-3 lesions and 3.23 for >3 lesions), and multiple hepatic metastases (1.83 for 2-5 lesions and 2.80 for >5 lesions).

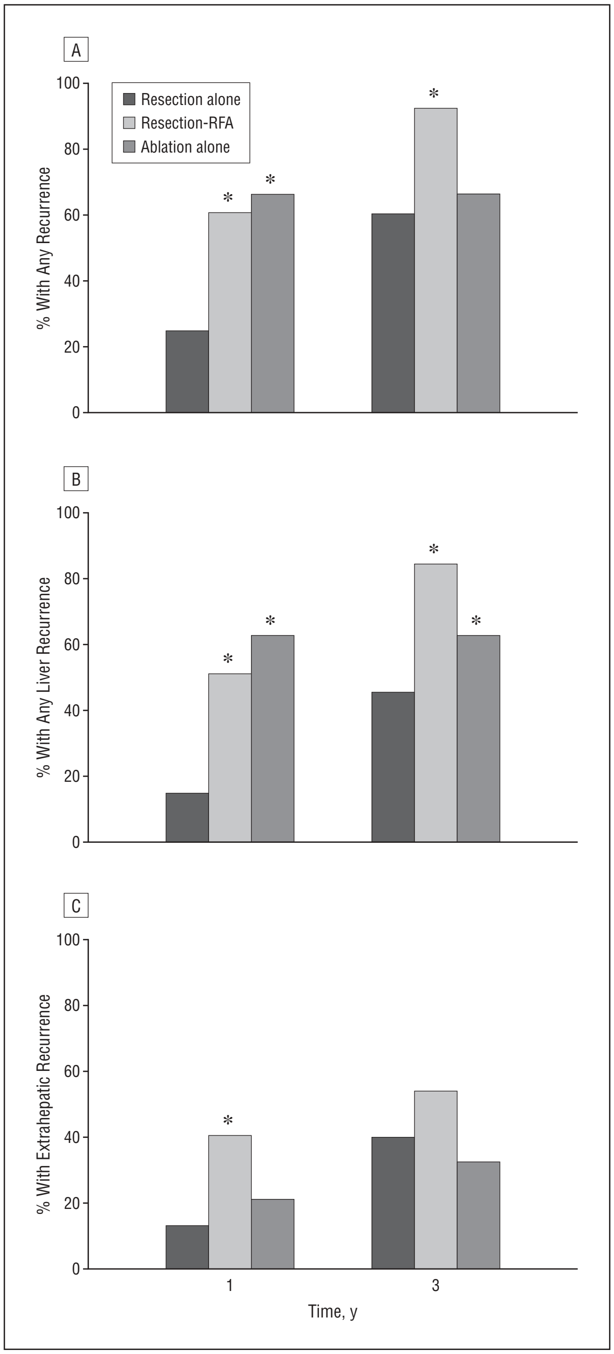

Figure 2. Any recurrence, any liver recurrence, and extrahepatic recurrence in the study groups. A, Recurrence at any site most often occurred after combined resection–radiofrequency ablation (RFA) or after RFA alone compared with after resection alone. B, Differences in recurrence patterns were observed among the treatment groups. Recurrence anywhere in the liver was more common at 1 year and at 3 years in patients treated with resection-RFA or with RFA alone vs with resection alone. C, The proportion of patients in each group who developed distant metastases as a component of extrahepatic failure was also markedly different. Patients who underwent resection-RFA had increased risk of extrahepatic failure at 1 year compared with patients who underwent resection alone. *$P<.05$.

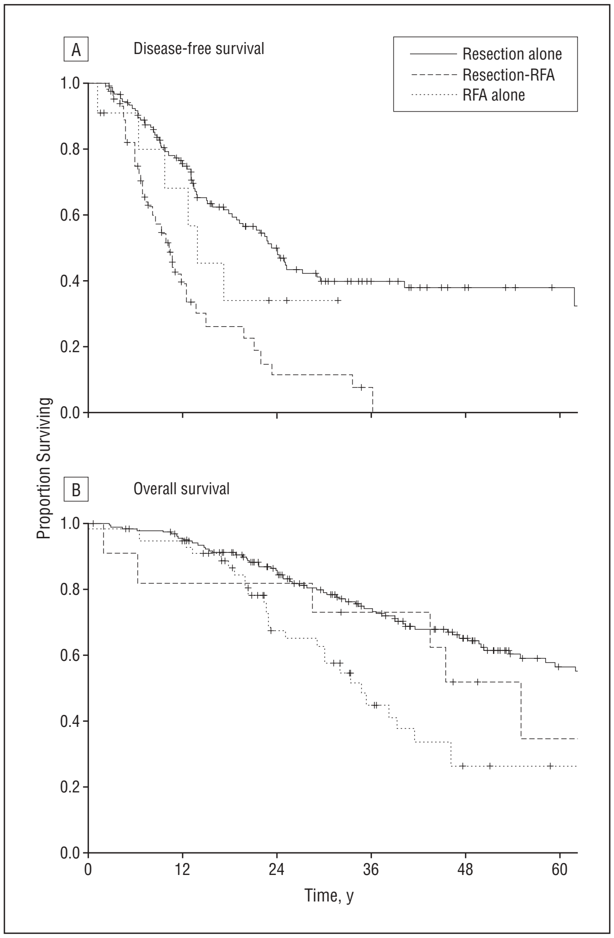

Figure 3. Survival in the study groups. A, Although there was no significant difference in disease-free survival for patients treated with combined resection–radiofrequency ablation (RFA) vs with RFA alone, disease-free survival was longest for patients treated with resection alone. B, On log-rank analysis, there was no difference in overall survival between patients who underwent resection alone vs RFA alone. In contrast, patients who were treated with resection-RFA had significantly shorter overall survival.

The median overall survival for the entire cohort of 258 patients was 59.0 months, while the 1-year, 3-year, and 5-year actuarial overall survival rates were 94.9%, 68.0%, and 49.6%, respectively (Figure 3B). On univariate log-rank analysis, there was no difference in 3-year overall survival between patients who underwent resection alone vs RFA alone (74.1% vs 72.7%; $P=.50$). In contrast, patients who were treated with resection-RFA had significantly worse overall 3-year survival (44.9%) ($P<.001$). Statistical analyses revealed several other factors that were associated with worse overall survival, including multiple hepatic metastases (HR, 1.50), size of tumors ablated exceeding 3 cm (2.97), and number of tumors ablated (2.77 for 2-3 lesions and 2.66 for >3 lesions) ($P<.05$ for all).

Adjusted Analysis

To adjust for relative intergroup differences in known risk factors for disease-free and overall survival, a matched-control analysis was performed. Patients who underwent RFA with or without resection (ie, cases) were matched 1:1 with patients who underwent resection alone (ie, controls). Matching was moderately successful in identifying cohorts of patients with comparable age, sex, primary tumor characteristics, and metastatic levels of hepatic disease burden (ie, similar number and size of liver lesions). On matched analysis, patients who underwent

Table 4. Matched Control Analysis of Adjusted Disease-Free and Overall Survival

Surgical Procedure	Survival, Median, mo	% Surviving 1 y	3 y	5 y	*P* Value
Disease-free survival					
Resection alone	19.5	76.0	41.3	41.3	
RFA with or without resection	10.2	43.3	8.9	0.0	<.001
Overall survival					
Resection alone	73.4	96.0	72.0	57.4	
RFA with or without resection	38.1	92.3	51.2	28.3	<.001

Abbreviation: RFA, radiofrequency ablation.

Table 5. Multivariate Cox Proportional Hazards Regression Analysis of Prognostic Factors Affecting Disease-Free and Overall Survival

Prognostic Factor	Hazard Ratio (95% Confidence Interval) Disease-Free Survival	Overall Survival
Primary tumor		
Disease-free interval >12 mo	0.77 (0.49-1.19)	. . .
Positive lymph nodes	0.97 (0.89-1.06)	. . .
Preoperative CEA level >100 ng/mL	1.99 (1.09-3.65)	1.03 (0.79-2.15)
Hepatic metastases		
Multiple lesions	1.46 (0.91-2.33)	1.41 (0.89-2.24)
Tumor size	1.00 (0.91-1.09)	1.13 (1.04-1.22)
Preoperative chemotherapy	1.51 (0.96-2.39)	1.00 (0.62-1.62)
Surgical procedure		
Resection alone		
Combined resection-RFA	2.09 (1.28-3.42)	2.82 (1.64-4.85)
RFA alone	1.41 (0.59-3.35)	1.77 (0.75-4.21)

Abbreviations: CEA, carcinoembryonic antigen; RFA, radiofrequency ablation; ellipses, not applicable.

SI conversion factor: To convert CEA level to micrograms per liter, multiply by 0.1.

RFA with or without resection still had worse disease-free and overall survival compared with patients who underwent resection alone (**Table 4**).

Because not all prognostic factors that were variable among the treatment groups were able to be matched, a mathematical model (ie, Cox proportional hazards regression model) was used to better control for potential confounding (**Table 5**). On multivariate analysis, only preoperative CEA level exceeding 100 ng/mL (HR, 1.99) and treatment with resection-RFA (2.09) were associated with increased risk of recurrence ($P < .05$ for both). Similarly, multivariate regression analysis revealed that only tumor size (HR, 1.13) and treatment with resection-RFA (2.82) were associated with increased risk of disease-specific death.

Standard model–based methods can potentially lead to inappropriate inferences in the situation where multiple confounding covariates do not overlap among treatment groups.[20] Therefore, propensity score methods were used.

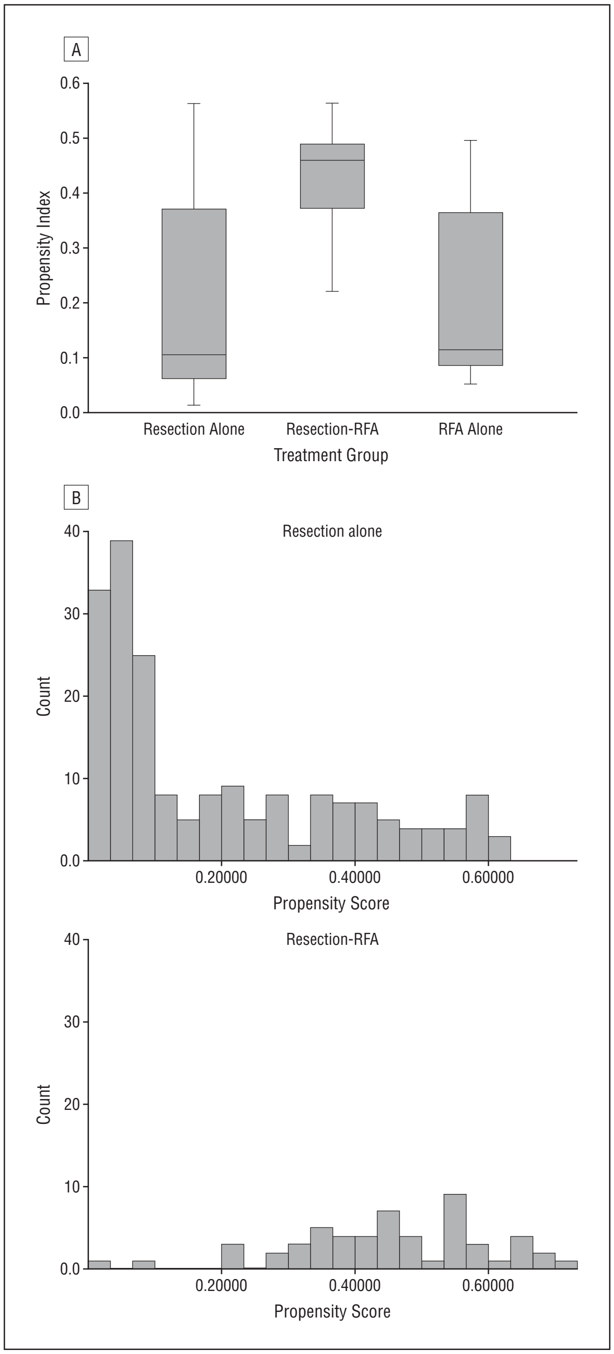

Figure 4. Use of the propensity index in the study groups. A, The overall balance in the propensity score distributions among the 3 groups was poor. Although the propensity score distributions between the resection-alone group vs the radiofrequency ablation (RFA)–alone group were similar, the distribution of propensity scores for the combined resection-RFA group was markedly different. B, Distribution of propensity scores comparing the resection-alone group vs the resection-RFA group. There is minimal overlap in the covariate profile for the groups, suggesting that comparison of the 2 groups would not be appropriate.

The distribution of propensity scores for the 3 treatment groups is shown in **Figure 4**A. The overall balance in the propensity score distribution among the 3 groups was poor. Although the propensity score distributions between the resection-alone group vs the RFA-alone group were similar, the distribution of propensity scores for the resection-RFA group was markedly different. Propensity score adjustment for the resection-alone group vs the RFA-alone group yielded no difference in risk of recurrence (HR, 1.41; *P*=.44) or in worse overall survival (1.77; *P*=.20). In contrast, propensity scores for the resection-alone group vs the resection-RFA group did not overlap sufficiently to allow subclassification to adjust for the covariates between the 2 groups (Figure 4B). In effect, because the covariate distributions failed to overlap in the resection-alone group vs the resection-RFA group, no causal inferences could be drawn to assess the efficacy of therapy between these 2 groups.

COMMENT

Alone or in combination with resection, RFA has been proposed as a therapeutic strategy for patients with colorectal liver metastases. Although RFA initially was used for local control of unresectable metastatic liver tumors, more recently the use of RFA has expanded to curative intent.[13,14] Data from several groups have shown that RFA alone[13] or in combination with hepatic resection[10,21] is feasible and safe and results in acceptable local recurrence rates. Several groups[10,13,14] reported 3-year and 5-year survival rates following RFA with or without concurrent hepatic resection at about 50% and 30%, respectively. However, other investigators reported that RFA with or without resection results in significantly higher rates of liver-only recurrence[3] and worse overall survival even in patients with small solitary colorectal liver metastasis.[22] These divergent results led some authors to call for randomized trials to investigate RFA compared with resection for resectable colorectal metastases[14]; however, other investigators argued that there is no equipoise between RFA and resection and that a randomized study to investigate RFA vs resection would be unethical.[23] While some evidence implicates RFA as a worse treatment,[23] other findings suggest that the main determinant of survival is not necessarily the mode of treatment but rather the patient's disease, as disease extent may determine the modality of treatment.[3] As such, whether studies looking at patients treated with resection vs RFA with or without resection are comparing apples to apples or apples to oranges has been a persistent controversy in the field. The present study is important because it sought to identify potential clinical, pathologic, and treatment-related factors that were distinct in each treatment group. Perhaps most important, unlike previous studies, we also used 3 statistical methods (paired-matched control, Cox proportional hazards multivariate model, and propensity index) in an attempt to adjust for potential confounding variables and to assess more quantitatively whether patients undergoing resection alone vs RFA with or without resection are comparable to allow for inferences about treatment efficacy.

Similar to previous findings,[3] patients treated with resection-RFA had significantly worse disease-free and overall survival compared with patients who underwent resection alone (Figure 3). Patients treated with RFA were also at significantly higher risk of intrahepatic recurrence (Figure 2). To ensure that differences in outcome were attributable to treatment allocation rather than to differences in baseline characteristics, analyses were performed to identify potential differences in the study populations. Patients treated with resection alone vs RFA with or without resection differed across several important clinical, pathologic, and treatment-related variables (Table 3). Patients treated with resection alone had more favorable prognostic characteristics, including fewer metastatic lesions and more R0 resections. In fact, patients undergoing resection alone had a median of 1 tumor compared with a median of 5 tumors among patients undergoing resection-RFA. In addition, patients treated with resection alone were significantly less likely to have received preoperative chemotherapy or adjuvant hepatic arterial infusion therapy, which suggests that the burden of disease may have been deemed less in this cohort of patients. The proportion of patients in each group who developed distant metastases as a component of extrahepatic failure was also markedly different. Patients who underwent resection-RFA had more than 3-fold increased risk of extrahepatic failure at 1 year compared with patients who underwent resection alone. In aggregate, these data suggest that patients undergoing resection-RFA (ie, patients with more tumors, increased R1 resections, extensive preoperative chemotherapy, and higher risk of distant metastases within 1 year of surgery) had different tumor biologic characteristics than patients undergoing resection alone.

Analysis of data to assess the treatment effect of a given therapeutic modality that is based on patients who have been nonrandomly assigned a specific treatment is fraught with difficulty. Given that treatment assignment (ie, resection alone, resection-RFA, or RFA alone) is nonrandom, important clinical and prognostic characteristics may differ among the various treatment groups. To control or adjust for these differences when comparing treatment effects, one can use several statistical approaches. In the present study, we first used a matched-control analysis. Individuals were matched on a finite number of variables that were deemed to be important prognostically (ie, metastatic lesion size and number and primary T and N stages). On matched analysis, patients who underwent resection-RFA had worse disease-free and overall survival compared with patients who underwent resection alone. However, matched analyses had several disadvantages. Matching was limited because not all prognostic factors that were variable among the treatment groups were able to be matched; partly because of this, only 100 patients were included in the matched analysis. Another problem with matching was that it required tailoring in the selection of the study groups to make them as comparable as possible; however, this increased internal comparability may have resulted in a lack of representativeness.[24] In addition, matching was conducted according to categorical definitions of continuous variables (eg, tumor size <5 vs ≥ 5 cm and solitary vs multiple lesions); as such, residual differences between cases and controls may remain.[24] For these reasons, matched analyses between resection alone vs RFA with or without resection may have been associated with residual confounding, resulting in misleading conclusions about treatment efficacy.

Our second form of analysis used mathematical multivariate Cox proportional hazards regression models, which are the most frequently used technique to better control for potential confounding when analyzing time-to-event data. A Cox proportional hazards regression model is essentially a multiple linear regression of the logarithm of the hazard on the variable, with the baseline hazard being an intercept term that varies with time, while the covariates (eg, tumor size, number, and history of chemotherapy) act multiplicatively on the hazard at any point in time.[25] Similar to other studies[3,26] that used multivariate modeling, we noted that RFA with or without resection was associated with increased risk of recurrence and worse overall survival on multivariate analysis (Table 5). However, Rubin[20,27] has cautioned that initial distributional differences in the covariates among treated groups may be too substantial to allow for reliance on linear regression assessment. Rubin[20] also warned that multivariate estimation of causal treatment effects (herein the effect of resection alone vs resection-RFA vs RFA alone) may be misleading because of possible reliance on unwarranted assumptions and extrapolations without any warning. Although standard modeling software can automatically handle many regressor variables and produce results, these results can be misleading.[20] As such, unlike most (if not all) investigations that relied solely on multivariate analyses to investigate the efficacy of resection vs RFA, the present study also used propensity score technology[28] to further address estimation of causal effects relative to the confounding variables.

Our third model of analysis, propensity score methods, determined the propensity of a patient to have received a certain treatment (eg, RFA and resection) as a function of all confounding covariates collapsed into a single predictor. Because the propensity method provides a single composite score that appropriately summarizes the collection of background characteristics for each patient cohort, it allows for straightforward assessment of whether groups overlap sufficiently with respect to baseline characteristics to allow robust comparison between groups in a given data set.[20] As Rubin noted, "one critical advantage of propensity score methods is that they can warn the investigator that, because of inadequately overlapping covariate distributions, a particular database cannot address the causal question at hand without relying on untrustworthy model-dependent extrapolation."[20(p763)] In the present study, when propensity score methods were used, the covariates were imbalanced between the resection-alone group vs the resection-RFA group (Figure 4A). Propensity scores for the resection-alone group vs the resection-RFA group had minimal overlap (Figure 4B), suggesting that the data could not support any causal conclusions about the differential effects of the treatments. In effect, propensity score indexing revealed that the RFA-alone group and the resection-RFA group were so different that essentially no information could be drawn about the treatment effect.

The present study had several limitations. Only 11 patients underwent RFA alone, making any evidence-based conclusions regarding the effect of RFA alone difficult. However, patients who underwent RFA alone had overall survival similar to that of patients who underwent resection alone but had a higher recurrence rate. The higher incidence of intrahepatic recurrence may have been related to the fact that 5 of 11 patients in the RFA-alone group underwent RFA of a lesion immediately adjacent to the hepatic veins. Local recurrence adjacent to major venous structures has been well documented because of a "heat-sink" effect.[29] Another potential shortcoming of the present study was that the propensity index, as well as the matched and multivariate analyses, accounted for only a finite number of known measurable confounders. When addressing causal questions from nonrandomized studies, propensity score methods can only adjust for observed confounding covariates and not for unobserved ones. This is an inherent limitation to any nonrandomized study compared with randomized studies, as randomization tends to balance the distribution of all covariates, observed and unobserved.[20]

Accepted for Publication: September 28, 2007.

Correspondence: Timothy M. Pawlik, MD, MPH, Department of Surgery, The Johns Hopkins University School of Medicine, 600 N Wolfe St, Halsted, Room 614, Baltimore, MD 22187-6681 (tpawlik1@jhmi.edu).

Author Contributions: *Study concept and design:* Choti, Assumpcao, and Pawlik. *Acquisition of data:* Gleisner, Choti, Assumpcao, and Schulick. *Analysis and interpretation of data:* Gleisner, Assumpcao, Nathan, and Pawlik. *Drafting of the manuscript:* Gleisner, Choti, Assumpcao, and Pawlik. *Critical revision of the manuscript for important intellectual content:* Gleisner, Choti, Assumpcao, Nathan, Schulick, and Pawlik. *Statistical analysis:* Gleisner, Assumpcao, Nathan, and Pawlik. *Obtained funding:* Choti. *Study supervision:* Choti, Schulick, Pawlik.

Financial Disclosure: None reported.

Previous Presentation: This study was presented at the 7th Americas Congress of the American Hepato-Pancreato-Biliary Association; April 22, 2007; Las Vegas, Nevada.

REFERENCES

1. Jemal A, Murray T, Ward E, et al. Cancer statistics, 2005 [published correction appears in *CA Cancer J Clin.* 2005;55(4):259]. *CA Cancer J Clin.* 2005;55(1):10-30.
2. Steele G Jr, Ravikumar TS. Resection of hepatic metastases from colorectal cancer: biologic perspective. *Ann Surg.* 1989;210(2):127-138.
3. Abdalla EK, Vauthey JN, Ellis LM, et al. Recurrence and outcomes following hepatic resection, radiofrequency ablation, and combined resection/ablation for colorectal liver metastases. *Ann Surg.* 2004;239(6):818-827.
4. Scheele J, Stangl R, Altendorf-Hofmann A. Hepatic metastases from colorectal carcinoma: impact of surgical resection on the natural history. *Br J Surg.* 1990;77(11):1241-1246.
5. Choti MA, Sitzmann JV, Tiburi MF, et al. Trends in long-term survival following liver resection for hepatic colorectal metastases. *Ann Surg.* 2002;235(6):759-766.
6. Hughes KS, Rosenstein RB, Songhorabodi S, et al. Resection of the liver for colorectal carcinoma metastases: a multi-institutional study of long-term survivors. *Dis Colon Rectum.* 1988;31(1):1-4.
7. Fong Y, Fortner J, Sun RL, Brennan MF, Blumgart LH. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. *Ann Surg.* 1999;230(3):309-321.
8. Adson MA, van Heerden JA, Adson MH, Wagner JS, Ilstrup DM. Resection of hepatic metastases from colorectal cancer. *Arch Surg.* 1984;119(6):647-651.
9. Pawlik TM, Scoggins CR, Zorzi D, et al. Effect of surgical margin status on sur-

vival and site of recurrence after hepatic resection for colorectal metastases. *Ann Surg*. 2005;241(5):715-724.

10. Pawlik TM, Izzo F, Cohen DS, Morris JS, Curley SA. Combined resection and radiofrequency ablation for advanced hepatic malignancies: results in 172 patients. *Ann Surg Oncol*. 2003;10(9):1059-1069.
11. Evrard S, Becouarn Y, Fonck M, Brunet R, Mathoulin-Pelissier S, Picot V. Surgical treatment of liver metastases by radiofrequency ablation, resection, or in combination. *Eur J Surg Oncol*. 2004;30(4):399-406.
12. Heslin MJ, Medina-Franco H, Parker M, Vickers SM, Aldrete J, Urist MM. Colorectal hepatic metastases: resection, local ablation, and hepatic artery infusion pump are associated with prolonged survival [published correction appears in *Arch Surg*. 2001;136(7):809]. *Arch Surg*. 2001;136(3):318-323.
13. Abitabile P, Hartl U, Lange J, Maurer CA. Radiofrequency ablation permits an effective treatment for colorectal liver metastasis. *Eur J Surg Oncol*. 2007;33(1):67-71.
14. Machi J, Oishi AJ, Sumida K, et al. Long-term outcome of radiofrequency ablation for unresectable liver metastases from colorectal cancer: evaluation of prognostic factors and effectiveness in first- and second-line management. *Cancer J*. 2006;12(4):318-326.
15. Adam R, Delvart V, Pascal G, et al. Rescue surgery for unresectable colorectal liver metastases downstaged by chemotherapy: a model to predict long-term survival. *Ann Surg*. 2004;240(4):644-658.
16. Clavien PA, Emond J, Vauthey JN, Belghiti J, Chari RS, Strasberg SM. Protection of the liver during hepatic surgery. *J Gastrointest Surg*. 2004;8(3):313-327.
17. The Brisbane 2000 Terminology of Liver Anatomy and Resections: HPB 2000; 2:333-39 [letter and reply]. *HPB*. 2002;4(2):99-100.
18. Pearson AS, Izzo F, Fleming RY, et al. Intraoperative radiofrequency ablation or cryoablation for hepatic malignancies. *Am J Surg*. 1999;178(6):592-599.
19. Kaplan EL, Meier P. A non-parametric estimate from incomplete observations. *J Am Stat Assoc*. 1958;53:457-480.
20. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med*. 1997;127(8, pt 2):757-763.
21. Elias D, Goharin A, El Otmany A, et al. Usefulness of intraoperative radiofrequency thermoablation of liver tumours associated or not with hepatectomy. *Eur J Surg Oncol*. 2000;26(8):763-769.
22. Aloia TA, Vauthey JN, Loyer EM, et al. Solitary colorectal liver metastasis: resection determines outcome. *Arch Surg*. 2006;141(5):460-467.
23. Abdalla EK, Vauthey JN. Colorectal metastases: resect or ablate? *Ann Surg Oncol*. 2006;13(5):602-603.
24. Szklo M, Nieto FJ. Basic study designs in analytical epidemiology. In: Szklo M, Nieto FJ, eds. *Epidemiology Beyond the Basics*. Sudbury, MA: Jones & Barlett; 2004:45-47.
25. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis, part II: multivariate data analysis: an introduction to concepts and methods. *Br J Cancer*. 2003; 89(3):431-436.
26. de Baere T, Elias D, Dromain C, et al. Radiofrequency ablation of 100 hepatic metastases with a mean follow-up of more than 1 year. *AJR Am J Roentgenol*. 2000;175(6):1619-1625.
27. Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med*. 2007;26(1): 20-36.
28. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41-55.
29. Welp C, Siebers S, Ermert H, Werner J. Investigation of the influence of blood flow rate on large vessel cooling in hepatic radiofrequency ablation. *Biomed Tech (Berl)*. 2006;51(5-6):337-346.