Ewout W. Steyerberg, PhD
H. Jan Keizer, MD, PhD
Dirk T. Sleijfer, MD, PhD
Sophie D. Fosså, MD, PhD
Dean F. Bajorin, MD, PhD
Arthur Gerl, MD
Ronald de Wit, MD, PhD
Wim J. Kirkels, MD, PhD
Heimen Schraffordt Koops,
    MD, PhD
J. Dik F. Habbema, PhD

Retroperitoneal Metastases in Testicular Cancer: Role of CT Measurements of Residual Masses in Decision Making for Resection after Chemotherapy[1]

PURPOSE: To determine the relative importance of computed tomographic (CT) measurements for the prediction of histologic findings in residual masses in patients with nonseminomatous testicular cancer.

MATERIALS AND METHODS: Measurements of the maximum transverse size of retroperitoneal metastases before and after chemotherapy were available in 641 patients who underwent resection after chemotherapy while their levels of tumor markers were normal. Radiologic measurements of mass size and clinical characteristics (histologic findings in primary tumor and levels of α-fetoprotein, human chorionic gonadotropin, and lactate dehydrogenase before chemotherapy) were related to histologic findings in the residual mass with logistic regression analysis.

RESULTS: At resection, 302 patients had benign tissue, and 339 had residual tumor (mature teratomas or cancer). Tumor was more frequent in larger masses after chemotherapy but was unrelated to mass size before chemotherapy. Inclusion of the reduction in size significantly improved the logistic regression model, which included mass size after chemotherapy. This model was further improved with the addition of clinical characteristics. Areas under the receiver operating characteristic curves increased from 0.74 to 0.77 and 0.83 with these models.

CONCLUSION: A small retroperitoneal mass after chemotherapy is an important predictor of benign histologic findings in residual masses in patients with nonseminomatous testicular cancer. However, better predictions can be made when the reduction in size and clinical characteristics are considered as well. Decisions regarding resection should be based on the combination of these characteristics rather than on only mass size after chemotherapy.

Computed tomographic (CT) evaluations are important in patients with testicular cancer. At presentation, abdominal and thoracic imaging is performed to determine the extent of disease. Patients with metastatic disease are candidates for combination chemotherapy with cisplatin, which leads to a long-term survival of about 80% (1,2). Repeat CT scans are obtained to monitor the effect of treatment, which is apparent from a reduction in mass size (3). After chemotherapy, the presence of a residual mass on a CT scan may guide the decision to perform resection. The most frequent procedures are laparotomy with retroperitoneal lymph node dissection for residual retroperitoneal metastases and thoracotomy with wedge resection for residual lung nodules (2).

When retroperitoneal metastases are not detectable as residual masses on abdominal CT scans obtained after chemotherapy, a resection is usually not performed (2,4). This decision is based on the notion that residual tumor (mature teratoma or viable cancer) is rare in very small (eg, <10-mm) remnants of initial disease and that resection of only necrotic and/or fibrotic remnants has no therapeutic benefit (5).

In addition to the radiologic characteristics, clinical characteristics have been related to the presence of malignancy. These include histologic findings in the primary tumor

Characteristic	n (%)
Hospital	
Klinikum Grosshadern (Munich, Germany)	98 (15)
Memorial Sloan-Kettering Cancer Center (New York, NY)	88 (14)
Norwegian Radium Hospital (Oslo, Norway)	150 (23)
University hospitals* (the Netherlands)	151 (24)
University hospital (Groningen, the Netherlands)	154 (24)
Year of treatment	
Before 1985	256 (40)
1985–1989	226 (35)
1990 or after	159 (25)
Histologic findings at retroperitoneal resection	
Necrosis	302 (47)
Mature teratoma	268 (42)
Cancer	71 (11)

Note.—All patients were treated with chemotherapy for metastatic nonseminomatous testicular cancer and had normal levels of tumor markers AFP and HCG at the time of retroperitoneal lymph node resection.

* Amsterdam Free University ($n = 18$), Leiden ($n = 50$), Nijmegen ($n = 33$), Rotterdam Cancer Institute ($n = 37$), Rotterdam Dijkzigt Hospital ($n = 13$).

a.   b.

Figure 1. CT images show a left paraaortic mass (arrows) with a maximum transverse diameter of **(a)** 50 mm before chemotherapy and **(b)** 15 mm after chemotherapy. The residual mass contained fibrosis and some remnants of mature teratoma.

TABLE 2
Radiologic Measurements and Histologic Findings at Resection

Characteristic	Necrosis ($n = 302$)	Teratoma ($n = 268$)	Cancer ($n = 71$)	Total ($N = 641$)
Size before chemotherapy (mm)				
0–29	75 (52)	60 (41)	10 (7)	145
30–49	90 (51)	73 (42)	13 (7)	176
50–99	93 (44)	89 (42)	32 (15)	214
≥100	44 (42)	46 (43)	16 (15)	106
Size after chemotherapy (mm)				
0–9	90 (73)	28 (23)	6 (5)	124
10–19	99 (63)	53 (34)	6 (4)	158
20–29	60 (50)	46 (38)	15 (12)	121
30–49	29 (28)	53 (51)	23 (22)	105
≥50	24 (18)	88 (66)	21 (16)	133
Reduction in size (%)				
≥85	75 (80)	16 (17)	3 (3)	94
70–84	67 (64)	27 (26)	11 (11)	105
50–69	99 (54)	62 (34)	23 (13)	184
30–49	36 (43)	37 (45)	10 (12)	83
0–29	25 (20)	82 (66)	17 (14)	124
<0*	0 (0)	44 (86)	7 (14)	51

Note.—Numbers in parentheses are percentages.
* Increase in size.

(presence of teratomatous elements) and levels of tumor markers before chemotherapy, namely, elevated levels of α-fetoprotein (AFP), elevated levels of human chorionic gonadotropin (HCG), and normal levels of lactate dehydrogenase (LDH) (6–10).

In this study, we analyzed the relative importance of the radiologic and clinical characteristics for the prediction of the histologic findings at retroperitoneal resection. We combined radiologic and clinical characteristics in multivariable statistical models. We further examined how these multivariable combinations may serve to improve decision making in the treatment of individual patients.

MATERIALS AND METHODS

Patients

We included patients who were treated with chemotherapy for metastatic nonseminomatous testicular cancer and who underwent resection of retroperitoneal lymph nodes after induction chemotherapy. To obtain a more homogeneous group of patients, we excluded patients who underwent resection after relapse and patients with elevated levels of tumor markers AFP or HCG at the time of surgery, extragonadal primary tumors, or histologically pure seminoma without elevated levels of serum tumor markers before chemotherapy.

Data were available in 716 patients. This data set consisted of data in 544 patients who were previously examined to derive a prediction model for the histologic findings in the residual masses (11) and data in 172 patients to validate this model (12). For the present analysis, we excluded 75 patients for whom measurements of mass size before ($n = 52$) or after ($n = 65$) chemotherapy were missing;

therefore, data in 641 patients were available for analysis (Table 1). Patients were treated at the Klinikum Grosshadern, Munich, Germany ($n = 98$ [13]); Memorial Sloan-Kettering Cancer Center, New York, NY ($n = 88$ [8]); Norwegian Radium Hospital, Oslo ($n = 150$ [9]); and a number of university hospitals in the Netherlands (Groningen, $n = 154$; other centers, $n = 151$ [7,14,15]).

Most patients were treated in the 1980s (median year 1986; year range, 1979–1996). Chemotherapeutic treatment was usually in the context of randomized clinical trials; details have been reported elsewhere (8,9,12–15). Data collection was
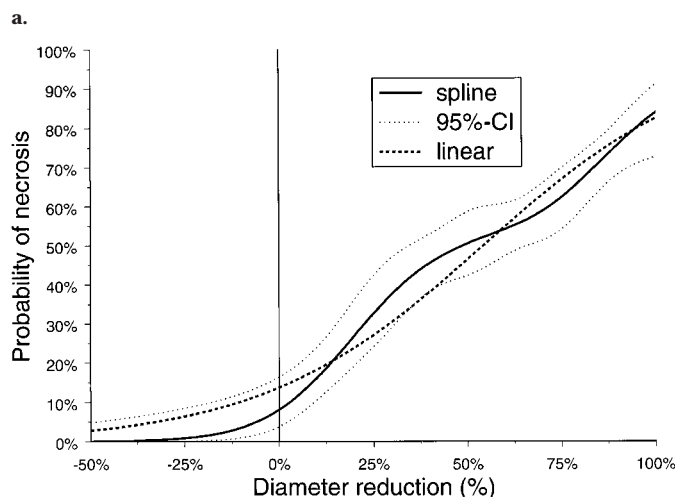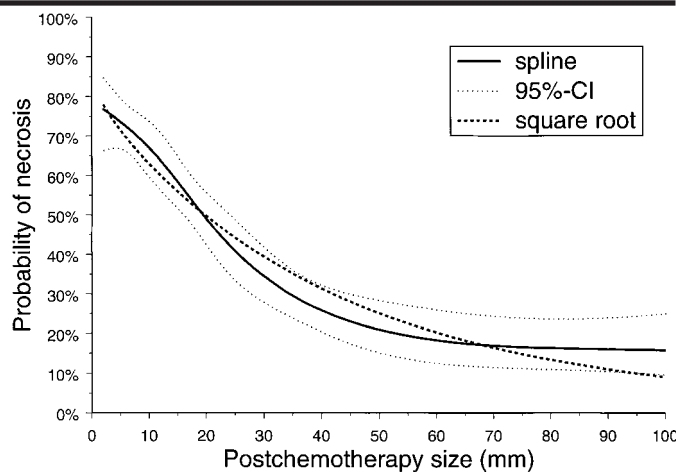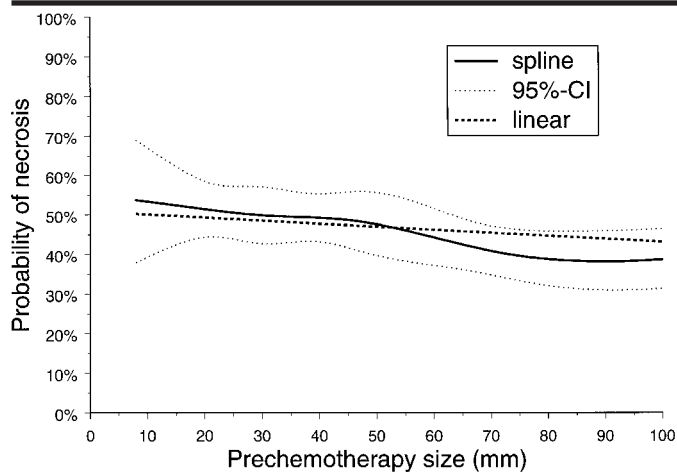
a.

b.

c.

Figure 2. Graphs depict the relationships between the probability of finding necrosis at resection and the radiologic characteristics of **(a)** mass size before chemotherapy, **(b)** mass size after chemotherapy, and **(c)** reduction in mass diameter. Solid and dotted lines indicate the point estimates and 95% CIs, respectively, with a flexible nonlinear function (a spline). Dashed line indicates use of the linear term (size before chemotherapy and reduction) or the square root of a transformed variable (square root of the size after chemotherapy) in the logistic regression model, which provided an adequate fit.

268), or necrosis ($n = 302$). Immature teratoma was classified as cancer because of its more severe prognostic consequence compared with that of mature teratoma (16).

Statistical Analysis

Logistic regression analysis was used to relate the finding of necrosis or tumor (mature teratoma or viable cancer) at resection to the radiologic and clinical characteristics. We first performed univariable analyses of the following radiologic characteristics: size before chemotherapy, size after chemotherapy, and reduction in size. Reduction in size was calculated as follows: (size before chemotherapy − size after chemotherapy) ÷ size before chemotherapy. These characteristics were categorized for descriptive analyses by using clinically relevant cutoff values.

For further statistical analysis, these characteristics were considered as continuous variables. Since the relationship between these variables and the histologic findings in the residual mass did not necessarily have to be linear, restricted cubic spline functions with five knots were used in the logistic regression models (17,18). These flexible and smooth functions can take many forms, while the continuous character of the variable is maintained. Further, spline functions are used to test for nonlinearity on the log-

prospective at all participating centers. Before and after chemotherapy, abdominal CT scans were obtained and interpreted by the local radiologist at each participating institution. At all centers, patients were candidates for resection of retroperitoneal lymph nodes if residual masses were depicted on CT scans obtained after chemotherapy; this criterion implied resection of masses equal to or larger than approximately 10 mm. At the University Hospital Groningen, resection was also performed when the primary tumor contained teratomatous elements, even if no residual masses were detectable (7). At the Norwegian Radium Hospital, resection was routinely performed regardless of the size of the mass after chemotherapy (9).

Radiologic and Clinical Characteristics

Radiologic characteristics considered in this study were the maximum transverse

diameter (in millimeters) of the mass before and after chemotherapy. In principle, these measurements were read from the CT scans (Fig 1). When a CT scan was considered to be completely normal (no mass depicted), a mass size of 2 mm was assumed for statistical analysis. This procedure improved the statistical distribution of mass size.

Clinical characteristics included histologic findings in the primary tumor (teratoma-positive or teratoma-negative) and levels of tumor markers AFP, HCG, and LDH before chemotherapy. These characteristics were considered on the basis of their predictive value for the histologic content of the residual mass in previous studies. These four characteristics were combined in one predictor, simple score, in which all four characteristics were given equal weight.

The histologic outcome of retroperitoneal resection was classified according to the worst histologic finding present, that is, cancer ($n = 71$), mature teratoma ($n =$

odds scale. When statistically significant nonlinearity was observed ($P < .05$), a simple transformation (square, square root, or logarithm) was performed with the continuous variable, and the nonlinearity of the transformed variable was tested again.

Three logistic regression models were constructed and compared by using receiver operating characteristic (ROC) analysis (19). The first model included only measurements of mass size after chemotherapy. The second model included information about size before and after chemotherapy. The third model included the clinical characteristics in addition to radiologic measurements to evaluate their predictive contribution. Interaction terms between covariables were studied to test for violation of the additivity assumption (18).

In general, an evaluation of model performance may be too optimistic when the performance is determined with the same data set used to estimate the model (18,20). To correct for this bias, resampling methods have been advocated, such as the bootstrap method (21). We obtained areas under the ROC curves with 200 bootstrap samples consisting of 641 patients that were drawn with replacement from the original data set. Goodness of fit of the models was evaluated with the Hosmer-Lemeshow test, which we used to compare the observed and expected frequencies of the outcome (ie, necrosis vs tumor) over the whole range of predictions (22). Calculations were performed by using the following computer programs: nonparametric ROC analysis version 2.5 (23), SPSS for Windows (SPSS, Chicago, Ill), and SAS (SAS, Cary, NC) design library (FE Harrell, University of Virginia, Charlottesville; available at *http://lib.stat.cmu.edu/S/Harrell/*) for S-plus (Mathsoft, Seattle, Wash).

RESULTS

Radiologic Characteristics

Table 2 shows the relationship of mass size to the histologic findings at retroperitoneal resection. The size before chemotherapy, in contrast to the size after chemotherapy, was not clearly related to the histologic findings in the residual mass. Small residual masses contained necrosis more often and contained cancer less often when compared with larger masses. Nearly half of the patients with very small masses (0–9 mm, $n = 124$) were from Norway (Norwegian Radium Hospital, $n = 49$). The reduction in size proved to be a very strong predictor of necrosis at

resection: Seventy-five (80%) of 94 patients with a reduction of 85% or more had only necrosis, whereas 99 (54%) of 184 patients with a decrease of 50%–69% had necrosis. None of those with an increased mass size had necrosis (Table 2). Viable cancer was found in only three of 94 patients with a reduction in size of 85% or more. These three patients had normal scans after chemotherapy; before chemotherapy, their masses were 50, 48 and 25 mm.

Figure 2a–c illustrates the relationships between the radiologic characteristics and the probability of finding necrosis at resection. For size after chemotherapy, statistically significant nonlinearity was found ($P < .001$). This finding indicated, for example, that the likelihood of necrosis differed more between a 10-mm mass and a 20-mm mass than between a 50-mm mass and a 60-mm mass. Use of the square root of the size after chemotherapy proved adequate, since the remaining nonlinearity was nonsignificant ($P = .10$). Figure 2b, however, shows that this transformation led to some overestimation of the probability of necrosis for mass sizes of 25–60 mm. The reduction in size was approximately linearly related to the probability of necrosis in a logistic regression model; the nonlinearity, as depicted in Figure 2c, was statistically nonsignificant ($P = .11$).

Clinical Characteristics

Table 3 shows the relationship of the clinical characteristics to the histologic findings at retroperitoneal resection. Ne-

crosis was more frequent when the primary tumor was teratoma-negative, when the AFP or HCG levels were normal before chemotherapy, or when the LDH level was elevated before chemotherapy (all $P < .01$). These four characteristics were used to construct a simple score on the basis of the number of favorable clinical characteristics. This score was supported by the multivariable odds ratios, which were approximately 3.4, 3.1, 1.7, and 1.8. The simple score ranged from 0 to 4 and was clearly related to the probability of necrosis (Table 3). However, viable cancer was found in approximately 11% of the patients without a clear relationship with the score (Table 3). This implied that higher scores predicted, in particular, the absence of mature teratoma in the residual mass.

Combination of Radiologic and Clinical Characteristics

Radiologic characteristics and clinical characteristics were included in the logistic regression models to evaluate their ability to distinguish necrotic remnants from residual tumor at resection. Figure 3 shows that the area under the ROC curve was 0.737 (95% CI: 0.70, 0.78) when the size of the mass after chemotherapy was the only decision criterion. When the reduction in size was also considered, the ROC area increased to 0.771 (95% CI: 0.74, 0.80). This increase was especially due to a better true-positive rate (resection of tumor) in the lower left corner in Figure 3 where false-positive rates are low (resection of necrosis). In this area, the cutoff values for the residual mass size

TABLE 3
Clinical Characteristics and Histologic Findings at Resection

Characteristic	Necrosis ($n = 302$)	Teratoma ($n = 268$)	Cancer ($n = 71$)	Total ($N = 641$)
Histologic finding in primary tumor				
Negative for teratoma	194 (63)	80 (26)	35 (11)	309
Positive for teratoma	108 (33)	188 (57)	36 (11)	332
Levels of markers before chemotherapy				
AFP				
Normal	140 (66)	55 (26)	17 (8)	212
Elevated	162 (38)	213 (50)	54 (13)	429
HCG				
Normal	131 (58)	70 (31)	24 (11)	225
Elevated	171 (41)	198 (48)	47 (14)	416
LDH				
Normal	67 (38)	97 (55)	13 (7)	177
Elevated	235 (51)	171 (37)	58 (13)	464
Simple score for necrosis				
0	7 (15)	35 (75)	5 (11)	47
1	51 (27)	116 (62)	20 (11)	187
2	121 (50)	92 (38)	28 (12)	241
3	86 (69)	24 (19)	14 (11)	124
4	37 (88)	1 (2)	4 (10)	42

TABLE 4
Odds Ratios and 95% CIs of the Three Logistic Regression Models

Characteristic	Model 1		Model 2		Model 3	
	Odds Ratio	95% CI	Odds Ratio	95% CI	Odds Ratio	95% CI
Size after chemotherapy*	1.70	1.52, 1.91	1.21	1.05, 1.40	1.22	1.04, 1.42
Reduction in size†	...	...	1.31	1.21, 1.41	1.30	1.19, 1.41
Simple score‡	...	...	...	...	2.38	1.92, 2.96

* Square root transformation (comparison of a 10-mm mass with a 20-mm mass).
  † Linear covariate (coded per 10%).
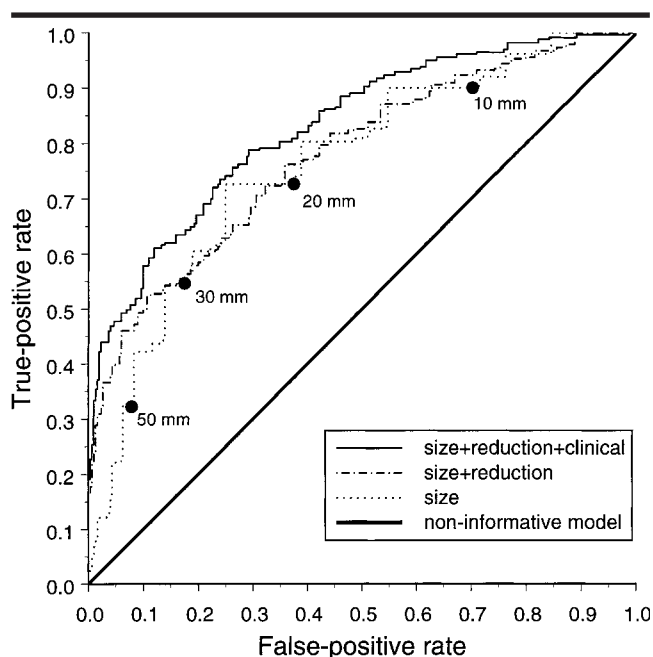   ‡ Linear covariate (coded per point).

Figure 3. Graph depicts the ROC curves for model 1 (size after chemotherapy, dotted line), model 2 (size after chemotherapy and reduction, dashed-and-dotted line) and model 3 (size after chemotherapy, reduction in diameter, and clinical characteristics; solid line). For model 1, combinations of true-positive (resection of tumor) and false-positive (resection of necrosis) rates are marked according to the cutoff values for the size of the residual mass. Best performance was achieved with model 3.

were large (>30 mm), which indicated that the reduction in size especially helped in the improved distinction of necrosis from malignancy in larger residual masses.

The area under the ROC curve increased to 0.826 (95% CI: 0.80, 0.86) when the simple clinical score was added to the model. The simple clinical score consisted of the following clinical characteristics: teratoma-negative primary tumor, normal prechemotherapy AFP level, normal prechemotherapy HCG level, and elevated prechemotherapy LDH level. This extension of the model led to an ROC curve that exceeded the previous model (size and reduction) over the whole

range of mass sizes and reductions, including small residual masses.

We noted that a true-positive rate of 100% was obtained only with a 100% false-positive rate; hence, any selection of patients for resection on the basis of the radiologic and/or clinical characteristics considered here missed some patients with residual tumor. The improvements from model 1 to model 2 and from model 2 to model 3 were statistically significant ($P < .01$). According to our bootstrap validation procedure, the overoptimism in the estimation of the areas under the ROC curves was trivial (eg, decrease from 0.826 to 0.823 for model 3). This finding

implied that the models were valid for similar patients.

Table 4 shows the odds ratios of the three logistic regression models. The odds ratio for size after chemotherapy was calculated so that it represented the comparison between a patient with a mass of 20 mm and a patient with a mass of 10 mm. The odds ratio in model 1 was 1.70, which indicated that the odds of necrosis was 1.70 times larger in the smaller mass. This odds ratio was considerably smaller in models 2 and 3, which included reduction in size, than in model 1 in which the size after chemotherapy was considered as the single predictor. This finding was caused by correlation: In general, smaller masses had a larger reduction (Pearson correlation coefficient $r = 0.63$). The simple score for necrosis had a low correlation with these radiologic characteristics ($r \approx 0.20$). The odds ratios of the radiologic characteristics were, hence, similar in models 2 and 3 (around 1.2 and 1.3).

In the third model, the clinical score was the most important predictor ($P < .001$), followed by the reduction in size ($P < .001$) and mass size after chemotherapy, which was the least important ($P = .01$). The following equation can be used to calculate the probability of necrosis: score $= -2.365 - (0.1514 \cdot$ square root of the size after chemotherapy) $+ (2.588 \cdot$ reduction) $+ (0.8680 \cdot$ simple score), where the size is in millimeters, the reduction is expressed as a fraction, and the simple score is expressed as points (0–4 points). The probability of necrosis was $1/(1 + e^{-\text{score}})$.

The goodness of fit of the model that included both radiologic and clinical characteristics is illustrated in Table 5. We focused on small residual masses (0–19 mm) in which resection decisions are most subject to debate. Within small masses, the reduction in size can be used to reasonably distinguish patients with a low risk of necrosis and patients with a high risk of necrosis; the simple score for necrosis further adds to this ability. The observed frequencies reasonably concur with the expected frequencies according to the model, although the small number of patients in some cells make comparisons difficult. Formal tests of goodness of fit were statistically nonsignificant.

Finally, we prepared graphs to show the predicted probabilities of necrosis in relation to radiologic and clinical characteristics (Fig 4). When the simple score equals 0, the probability of necrosis is predicted to be always less than 50%, even for the most favorable combinations of mass sizes before and after chemotherapy. For higher

scores (eg, scores of 3 or 4), the probability of necrosis may well exceed 80% in patients with a large reduction in mass size. Nearly straight lines are shown in Figure 4; this depiction is consistent with the finding that a reduction in size was a more important predictor of necrosis than the residual mass size ($P < .001$ vs $P = .01$). We marked the areas where a clear preference for resection (probability <60%) or observation (probability >90%) might be assumed.

DISCUSSION

These study findings show that simple, readily-available clinical characteristics significantly contribute to radiologic measurements in the prediction of histologic findings in a residual retroperitoneal mass after chemotherapeutic treatment for metastatic testicular cancer. Further, the response to chemotherapy, expressed in the reduction in size, is statistically more important than the size after chemotherapy. The combination of radiologic and clinical characteristics more closely approximates the ideal definition of the two subgroups—one with benign tissue only, for which resection can be avoided, and a second with residual cancer or mature teratoma, for which resection is beneficial.

The discriminative ability of the combination of reduction in mass size and size after chemotherapy was signifcantly higher than that of postchemotherapy size alone (areas under the ROC curves of 0.74 and 0.77). Adding the four clinical characteristics further improved the model (area under the ROC curve of 0.83). These characteristics included the histologic findings in the primary tumor (presence of teratoma elements) and the levels of tumor markers AFP, HCG, and LDH before chemotherapy, which were combined in a simple score based on the number of favorable values. The predictive value of these characteristics is well documented (6–10).

A limitation of our analysis was that many patients with masses of less than 10 mm after chemotherapy (normal CT scan) did not undergo resection. Hence, the histologic findings in their residual masses remains unknown. Follow-up studies in these patients might be valuable, specifically studies in which retroperitoneal recurrence is examined, but these data are currently not available to us. This selection poses no problem for the estimation of regression coefficients once the size after chemotherapy is correctly incorpo-

rated into the regression model that is used to predict necrosis. Our analysis indicated that a square root transformation of size after chemotherapy was adequate. However, the selection causes verification bias, which has probably affected the estimates of sensitivity and specificity but not the estimates of the area under the ROC curve (24,25). The latter estimates proved to be very stable in an evaluation of internal validity with bootstrap techniques. A slightly more complicated model showed adequate external validity in a previous study (12). Therefore, these findings imply that mass size after chemotherapy should not be considered alone with regard to the likely histologic content of a residual mass, but it should be combined with the reduction in size (comparison with the size before chemotherapy) and clinical characteristics. On the other hand, better predictors of the histologic findings are needed to bring the area under the ROC curve closer to 1.

Several radiologic characteristics of re-

sidual masses have been related to the histologic findings in residual masses, in addition to mass size measurements. CT parameters, such as attenuation (26,27) and density (28), and magnetic resonance (MR) imaging parameters (29) have been studied but have been proved, in general, not to be very informative. Also, positron emission tomography (PET) has been evaluated. The results at Indiana University (30) were favorable: Although PET could not be used to distinguish between necrosis and mature teratoma, masses in four patients that contained cancer showed higher standardized uptake values. These findings, however, were not supported in another study (31). Therefore, measurements of mass size are currently the most relevant of radiologic characteristics for the distinction of necrosis from malignancy.

A further limitation of our analysis was the quality of the CT measurements available for statistical analysis. CT measurements were not standardized over all participating centers and merely reflected

TABLE 5
Goodness of Fit of Model 3 in 282 Patients with 0–19-mm Masses after Chemotherapy

Simple Score	Reduction in Size (%)				
	0–49	50–69	70–84	≥85	Total
0					
Expected frequency* (%)	8 (5, 14)	20 (14, 29)	32 (23, 43)	43 (31, 56)	21 (15, 30)
Observed frequency† (%)	··· ($n = 4$)	··· ($n = 6$)	··· ($n = 3$)	··· ($n = 3$)	28 ($n = 18$)
No. of patients with findings at resection‡	0/4/0	2/4/0	2/1/0	1/2/0	5/13/0
1					
Expected frequency* (%)	23 (17, 31)	39 (32, 46)	52 (44, 60)	65 (56, 73)	47 (40, 55)
Observed frequency† (%)	··· ($n = 8$)	48 ($n = 23$)	64 ($n = 22$)	63 ($n = 19$)	53 ($n = 73$)
No. of patients with findings at resection‡	2/5/1	11/11/1	14/8/0	12/6/1	39/31/3
2					
Expected frequency* (%)	38 (30, 47)	60 (54, 65)	71 (66, 77)	81 (75, 86)	65 (59, 71)
Observed frequency† (%)	57 ($n = 21$)	68 ($n = 28$)	58 ($n = 24$)	86 ($n = 35$)	69 ($n = 108$)
No. of patients with findings at resection‡	12/9/0	19/8/1	14/8/2	30/5/0	75/30/3
3					
Expected frequency* (%)	68 (60, 75)	78 (71, 83)	85 (80, 89)	92 (88, 95)	85 (80, 89)
Observed frequency† (%)	··· ($n = 2$)	88 ($n = 16$)	67 ($n = 15$)	86 ($n = 23$)	82 ($n = 55$)
No. of patients with findings at resection‡	2/0/0	14/2/0	10/3/2	19/2/1	45/7/3
4					
Expected frequency* (%)	70 (54, 82)	90 (84, 94)	94 (89, 96)	96 (93, 98)	92 (87, 95)
Observed frequency† (%)	··· ($n = 2$)	··· ($n = 7$)	··· ($n = 8$)	91 ($n = 11$)	89 ($n = 28$)
No. of patients with findings at resection‡	1/0/1	6/0/1	8/0/0	10/0/1	25/0/3
Total					
Expected frequency* (%)	35 (27, 43)	57 (51, 63)	69 (63, 75)	81 (75, 86)	64 (58, 70)
Observed frequency† (%)	46 ($n = 37$)	65 ($n = 80$)	67 ($n = 72$)	80 ($n = 90$)	67 ($N = 282$)
No. of patients with findings at resection‡	17/18/2	52/25/3	48/20/4	72/15/3	189/81/12

* Frequency of necrosis. Numbers in parentheses are 95% CIs.
  † Frequency of necrosis. Data are shown if $n \geq 10$.
  ‡ Findings are no. of patients with necrosis/no. of patients with mature teratoma/no. of patients with cancer.
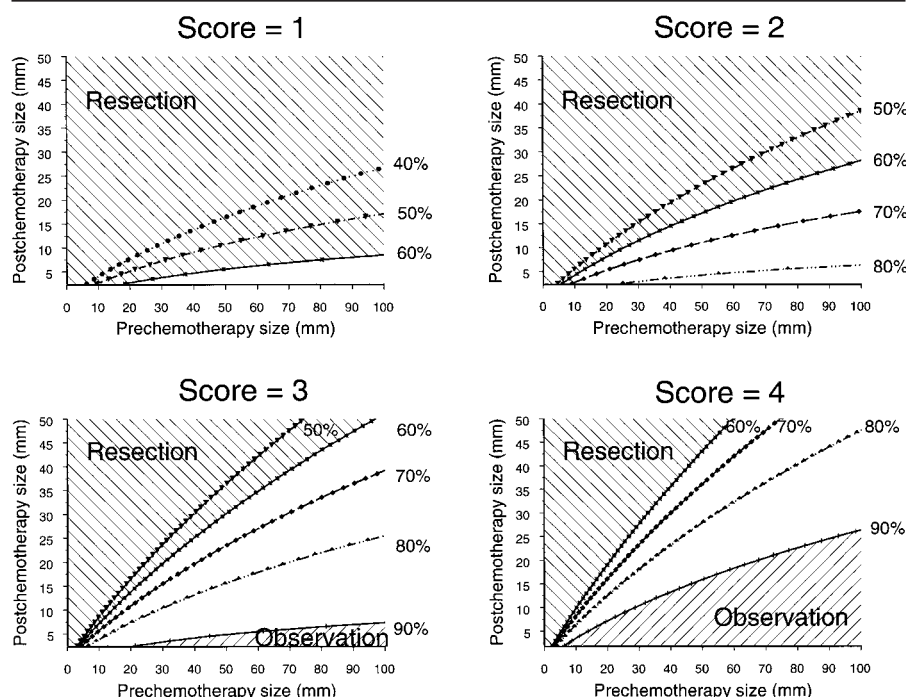
Figure 4. Graphs depict the predicted probabilities of necrosis in relation to the simple score, size before chemotherapy, and size after chemotherapy. Areas with probabilities less than 60% or greater than 90% (striped areas) depend on the combination of mass size before and after chemotherapy with the simple score of four clinical characteristics.

a routine-care situation. The level of expertise and, hence, the accuracy of the measurements may have differed among the radiologists. Also, current CT scanners provide measurements that are more precise than those possible with the equipment used in the 1980s (32). Therefore, the predictive ability of current CT measurements may be better than that shown in our evaluation.

When new radiologic characteristics are evaluated, we advocate that adjustments be made for mass size and clinical characteristics, as defined in the present study. This can be achieved, for example, with an evaluation of the new characteristic in a logistic regression model that includes the result from the formula for model 3 as a simple linear covariate. Further, we hope that new histologic, molecular, or genetic markers that can be used to predict the histologic findings in the residual mass will be identified. Again, evaluation of such markers should account for the simple predictive characteristics considered in the present study.

In the absence of stronger predictors, decisions regarding resection have to be made. Resection policies vary widely between centers, especially for small residual masses (6–9). Part of this variation may be justified by local circumstances,

such as the feasibility of frequent follow-up with high-quality CT scanning when resection is not performed (9). Other possible reasons for this variation include different appraisals of the consequences of leaving mature teratoma or cancer unresected and insufficient knowledge of the prevalence of residual malignancy in (very) small masses. The latter may be resolved by our "evidence graphs" (Fig 4), which may be useful in the communication between the radiologist and the treating physicians (urologic surgeon and medical oncologist).

In a previous evaluation, we compared several resection policies (33). The results of this comparison showed that resection is not considered to be indicated when the probability of necrosis exceeds 90%. For this subgroup, it may well be argued that the risks of short-term and long-term morbidity are not balanced by the potential benefits of resection, which include prevention of growth of mature teratoma, complete resection of viable tumor, and administration of additional chemotherapeutic treatments after resection of viable cancer (34).

These benefits likely outweigh the risks of resection in patients with a probability of less than 60%. This group contains patients with masses that enlarge despite

chemotherapy. We identified 51 of such patients in our data set: Forty-four (86%) had mature teratoma, and seven (14%) cancer. These findings confirmed those of earlier reports (35). The optimal treatment in patients with probabilities between 60% and 90% is debatable. A cutoff of 70% or 80% may be considered, but individual assessment of the patient's condition and his or her preferences are important. Also, in general, resection of small residual masses will not lead to long-term morbidity such as retrograde ejaculation; this consideration supports resection of small masses.

In conclusion, decisions regarding resection of residual masses should not be based on only the size of the mass after chemotherapy. The scan obtained before chemotherapy should be considered to determine the reduction in size, which should be combined with readily available clinical characteristics. This evidence-based approach will optimize the decisions in the patient's interest.

References

1. Einhorn LH. Treatment of testicular cancer: a new and improved model. J Clin Oncol 1990; 8:1777–1781.
2. Bosl GJ, Motzer RJ. Testicular germ-cell cancer. N Engl J Med 1997; 337:242–253.
3. Javadpour N, Anderson T, Doppman JL. Computed tomography in evolution of testicular cancer during intensive chemotherapy. J Urol 1979; 122:565–567.
4. Mead GM, Stenning SP, Parkinson MC, et al. The Second Medical Research Council study of prognostic factors in nonseminomatous germ cell tumors. J Clin Oncol 1992; 10:85–94.
5. Stomper PC, Jochelsen MS, Garnick MB, et al. Residual abdominal masses after chemotherapy for nonseminomatous testicular cancer: correlation of CT and histology. AJR Am J Roentgenol 1985; 145:743–746.
6. Donohue JP, Rowland RG, Kopecky K, et al. Correlation of computerized tomographic changes and histological findings in 80 patients having radical retroperitoneal lymph node dissection after chemotherapy for testis cancer. J Urol 1987; 137:1176–1179.
7. Gelderman WAH, Schraffordt Koops H, Sleijfer DTH, et al. Results of adjuvant surgery in patients with stage III and IV nonseminomatous testicular tumors after cisplatin-vinblastine-bleomycin chemotherapy. J Surg Oncol 1988; 38:227–232.
8. Toner GC, Panicek DM, Heelan RT, et al. Adjunctive surgery after chemotherapy for nonseminomatous germ cell tumors: recommendations for patient selection. J Clin Oncol 1990; 8:1683–1694.
9. Fosså SD, Qvist H, Stenwig AE, Lien HH, Ous S, Giercksky KE. Is postchemotherapy

retroperitoneal surgery necessary in patients with nonseminomatous testicular cancer and minimal residual tumor masses? J Clin Oncol 1992; 10:569–573.

10. Steyerberg EW, Keizer HJ, Stoter G, Habbema JDF. Predictors of residual mass histology following chemotherapy for metastatic non-seminomatous testicular cancer: a quantitative overview of 996 resections. Eur J Cancer 1994; 30A:1231–1239.

11. Steyerberg EW, Keizer HJ, Fosså SD, et al. Prediction of residual retroperitoneal mass histology following chemotherapy for metastatic nonseminomatous germ cell tumour: multivariate analysis of individual patient data from 6 study groups. J Clin Oncol 1995; 13:1177–1187.

12. Steyerberg EW, Gerl A, Fosså SD, et al. Validity of predictions of residual retroperitoneal mass histology in nonseminomatous testicular cancer. J Clin Oncol 1998; 16:269–274.

13. Gerl A, Clemm C, Schmeller N, et al. Outcome analysis after postchemotherapy surgery in patients with nonseminomatous germ cell tumours. Ann Oncol 1995; 6:483–488.

14. Steyerberg EW, Keizer HJ, Zwartendijk J, et al. Prognosis after resection of residual masses following chemotherapy for metastatic non-seminomatous testicular cancer: a multivariate analysis. Br J Cancer 1993; 68:195–200.

15. Gels ME, Nijboer AP, Hoekstra HJ, et al. Complications of the postchemotherapy resection of retroperitoneal residual tumor mass in patients with non-seminomatous testicular germ cell tumors. Br J Urol 1997; 79:263–268.

16. Terrier-Lacombe MJ, Martinez-Madrigal F, Porta W, Rahal J, Droz JP. Embryonal rhabdomyosarcoma arising in a mature teratoma of the testis: a case report. J Urol 1990; 143:1232–1234.

17. Harrell FE, Lee KL, Pollock BG. Regression models in clinical studies: determining relationships between predictors and response. J Natl Cancer Inst 1988; 80:1198–1202.

18. Harrell F, Lee K, Mark D. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996; 15:361–387.

19. Hanley JA, McNeil BJ. The meaning and use of the area under the receiver operating characteristic (ROC) curve. Radiology 1982; 143:29–36.

20. Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. Stat Med 1990; 9:1303–1325.

21. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Assoc 1983; 78:316–331.

22. Hosmer DW, Lemeshow S. Assessing the fit of the model. In: Applied logistic regression. New York, NY: Wiley, 1989; 135–175.

23. Vida S. A computer program for nonparametric receiver operating characteristic analysis. Comput Methods Programs Biomed 1993; 40:95–101.

24. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med 1978; 299:926–930.

25. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. Biometrics 1983; 39:207–215.

26. Husband JE, Hawkes DJ, Peckham MJ. CT estimations of mean attenuation values and volume in testicular tumors: a comparison with surgical and histologic findings. Radiology 1982; 144:553–558.

27. Stomper PC, Kalish LA, Garnick MB, Richie JP, Kantoff PW. CT and pathologic predictive features of residual mass histologic findings after chemotherapy for nonsemi-

nomatous germ cell tumors: can residual malignancy or teratoma be excluded? Radiology 1991; 180:711–714.

28. Sagalowsky AI, Ewalt DH, Molberg K, Peters PC. Predictors of residual mass histology after chemotherapy for advanced testis cancer. Urology 1990; 35:537–542.

29. Hogeboom WR, Hoekstra HJ, Mooyaart EL, et al. Magnetic resonance imaging of retroperitoneal lymph node metastases of non-seminomatous germ cell tumours of the testis. Eur J Surg Onc 1993; 19:429–437.

30. Stephens AW, Gonin R, Hutchins GD, Einhorn LH. Positron emission tomography evaluation of residual radiographic abnormalities in postchemotherapy germ cell tumor patients. J Clin Oncol 1996; 14:1637–1641.

31. Nuutinen JM, Leskinen S, Elomaa I, et al. Detection of residual tumours in postchemotherapy testicular cancer by FDG-PET. Eur J Cancer 1997; 33:1234–1241.

32. Heiken JP, Brink JA, Vannier MW. Spiral (helical) CT. Radiology 1993; 189:647–656.

33. Steyerberg EW, Keizer HJ, Fosså SD, et al. Resection of residual retroperitoneal masses in testicular cancer: evaluation and improvement of selection criteria. Br J Cancer 1996; 74:1492–1498.

34. Fox EP, Weathers TD, Williams SD, et al. Outcome analysis of patients with persistent nonteratomatous germ cell tumor in postchemotherapy retroperitoneal lymph node dissections. J Clin Oncol 1993; 11:1294–1299.

35. Panicek DM, Toner GC, Heelan RT, Bosl GJ. Nonseminomatous germ cell tumors: enlarging masses despite chemotherapy. Radiology 1990; 175:499–502.