increases in tumor marker levels. Thirdly, pre-treatment SUVmaxs were not taken into consideration in this study, although these may also be related to post-treatment SUVmaxs for recurrence. In contrast, the strength of this study is that SUVmaxs for recurrence and non-recurrence were found to be sharply divided and the overlap ranges were small.

Studies of ^{18}F-FDG PET SUVmaxs have unavoidable problems. The SUVmaxs are semi-quantitative in that these values are influenced by physiological and technical factors, including the contrast between a tumor and normal lung, tumor size, scintillator type, image reconstruction methodology, and image noise. Therefore, clinicians should take into account that the exact threshold used for SUVmax is not always applicable to other situations.

5. Conclusions

Local recurrences after SBRT were observed as consolidation or nodular opacities and usually appeared within RILO on CT images. However, among these opacities, local recurrence rates were low and CT results had a limited ability to detect local recurrences. By comparison, ^{18}F-FDG PET could assist in the detection of local recurrence more accurately and the SUVmax on an early image was sufficient to differentiate a local recurrence from RILO.

Conflict of interest statement

Atsuya Takeda is funded by Grant-in-Aid for Scientific Research (C) from the Japan Society for the Promotion of Science.

Hirofumi Fujii is funded by Grants-in-Aid for Scientific Research (B) from the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science and Technology, Health and Labor Sciences Research Grants for Third Term Comprehensive 10-year Strategy for Cancer Control from the Ministry of Health, Labor and Welfare.

References

[1] Chi A, Liao Z, Nguyen NP, Xu J, Stea B, Komaki R. Systemic review of the patterns of failure following stereotactic body radiation therapy in early-stage non-small-cell lung cancer: clinical implications. Radiother Oncol 2010;94:1–11.
[2] Ettinger DS, Akerley W, Borghaei H, Chang A, Cheney R, Chirieac LR, et al. NCCN clinical practice guideline in oncology web site. In: Non-small Cell Lung cancer: National Comprehensive Cancer Network; 2012.
[3] Takeda A, Kunieda E, Takeda T, Tanaka M, Sanuki N, Fujii H, et al. Possible misinterpretation of demarcated solid patterns of radiation fibrosis on CT scans as tumor recurrence in patients receiving hypofractionated stereotactic radiotherapy for lung cancer. Int J Radiat Oncol Biol Phys 2008;70:1057–65.
[4] Matsuo Y, Nagata Y, Mizowaki T, Takayama K, Sakamoto T, Sakamoto M, et al. Evaluation of mass-like consolidation after stereotactic body radiation therapy for lung tumors. Int J Clin Oncol 2007;12:356–62.
[5] Takeda T, Takeda A, Kunieda E, Ishizaka A, Takemasa K, Shimada K, et al. Radiation injury after hypofractionated stereotactic radiotherapy for peripheral small lung tumors: serial changes on CT. Am J Roentgenol 2004;182:1123–8.
[6] Takeda A, Kunieda E, Shigematsu N, Hossain DM, Kawase T, Ohashi T, et al. Small lung tumors: long-scan-time CT for planning of hypofractionated stereotactic radiation therapy – initial findings. Radiology 2005;237:295–300.
[7] Takeda A, Kunieda E, Sanuki N, Ohashi T, Oku Y, Sudo Y, et al. Dose distribution analysis in stereotactic body radiotherapy using dynamic conformal multiple arc therapy. Int J Radiat Oncol Biol Phys 2009;74:363–9.
[8] Neri S, Takahashi Y, Terashi T, Hamakawa H, Tomii K, Katakami N, et al. Surgical treatment of local recurrence after stereotactic body radiotherapy for primary and metastatic lung cancers. J Thorac Oncol 2010;5:2003–7.
[9] Bylund KC, Chen Y, Okunieff P, Philip A, Milano M. Reirradiation with a second course of stereotactic body radiotherapy for locally recurrent lung lesions. Int J Radiat Oncol Biol Phys 2010;78:S495–6.
[10] Podoloff DA, Advani RH, Allred C, Benson 3rd AB, Brown E, Burstein HJ, et al. NCCN task force report: positron emission tomography (PET)/computed tomography (CT) scanning in cancer. J Natl Compr Canc Netw 2007;5(Suppl 1):S1–22, quiz S23–22.
[11] Huang K, Dahele M, Senan S, Guckenberger M, Rodrigues GB, Ward A, et al. Radiographic changes after lung stereotactic ablative radiotherapy (SABR) – can we distinguish recurrence from fibrosis? A systematic review of the literature. Radiother Oncol 2012;102:335–42.
[12] Aoki T, Nagata Y, Negoro Y, Takayama K, Mizowaki T, Kokubo M, et al. Evaluation of lung injury after three-dimensional conformal stereotactic radiation therapy for solitary lung tumors: CT appearance. Radiology 2004;230:101–8.
[13] Keidar Z, Haim N, Guralnik L, Wollner M, Bar-Shalom R, Ben-Nun A, et al. PET/CT using 18F-FDG in suspected lung cancer recurrence: diagnostic value and impact on patient management. J Nucl Med 2004;45:1640–6.
[14] Matsuo Y, Nakamoto Y, Nagata Y, Shibuya K, Takayama K, Norihisa Y, et al. Characterization of FDG-PET images after stereotactic body radiation therapy for lung cancer. Radiother Oncol 2010;97:200–4.
[15] Vahdat S, Oermann EK, Collins SP, Yu X, Abedalthagafi M, Debrito P, et al. CyberKnife radiosurgery for inoperable stage IA non-small cell lung cancer: ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography serial tumor response assessment. J Hematol Oncol 2010;3:6.
[16] Hoopes DJ, Tann M, Fletcher JW, Forquer JA, Lin PF, Lo SS, et al. FDG-PET and stereotactic body radiotherapy (SBRT) for stage I non-small-cell lung cancer. Lung Cancer 2007;56:229–34.
[17] Patz Jr EF, Lowe VJ, Hoffman JM, Paine SS, Burrowes P, Coleman RE, et al. Focal pulmonary abnormalities: evaluation with F-18 fluorodeoxyglucose PET scanning. Radiology 1993;188:487–90.
[18] Xiu Y, Bhutani C, Dhurairaj T, Yu JQ, Dadparvar S, Reddy S, et al. Dual-time point FDG PET imaging in the evaluation of pulmonary nodules with minimally increased metabolic activity. Clin Nucl Med 2007;32:101–5.
[19] Barger Jr RL, Nandalur KR. Diagnostic performance of dual-time (18)F-FDG PET in the diagnosis of pulmonary nodules: a meta-analysis. Acad Radiol 2012;19:153–8.
[20] Grills IS, Hope AJ, Guckenberger M, Kestin LL, Werner-Wasik M, Yan D, et al. A multinational pooled analysis of 434 cases of stage I non-small cell lung cancer (NSCLC) treated with volumetrically image-guided (VIGRT) stereotactic lung radiotherapy (SBRT): results from the Elekta Collaborative Lung Research Group. J Clin Oncol 2010;28:S15.
[21] Verstegen NE, Lagerwaard FJ, Haasbeek CJ, Slotman BJ, Senan S. Outcomes of stereotactic ablative radiotherapy following a clinical diagnosis of stage I NSCLC: comparison with a contemporaneous cohort with pathologically proven disease. Radiother Oncol 2011;101:250–4.
[22] Takeda A, Kunieda E, Sanuki N, Aoki Y, Oku Y, Handa H. Stereotactic body radiotherapy (SBRT) for solitary pulmonary nodules clinically diagnosed as lung cancer with no pathological confirmation: comparison with non-small-cell lung cancer. Lung Cancer 2012;77:77–82.

Med Oncol (2013) 30:415
DOI 10.1007/s12032-012-0415-5

ORIGINAL PAPER

Impact of the first tumor response at eight weeks on overall survival in metastatic breast cancer patients treated with first-line combination chemotherapy

Chikako Suzuki · Lennart Blomqvist · Thomas Hatschek · Lena Carlsson · Zakaria Einbeigi · Barbro Linderholm · Birgitta Lindh · Niklas Loman · Martin Malmberg · Samuel Rotstein · Martin Söderberg · Marie Sundqvist · Thomas M. Walz · Gunnar Åström · Hirofumi Fujii · Hans Jacobsson · Bengt Glimelius

Received: 10 July 2012 / Accepted: 11 November 2012

Abstract The aim of this was to determine whether the change of size observed at the first response evaluation after initiation of first-line combination chemotherapy correlates with overall survival (OS) in patients with metastatic breast cancer (MBC). The change in size of tumors derived from measurements according to Response Evaluation Criteria In Solid Tumors (RECIST) at the first evaluation on computed tomography (CT) was obtained from a multicenter, randomized phase III trial ("TEX trial," $n = 287$) comparing treatment with a combination of epirubicin and paclitaxel alone or with capecitabine (TEX). Cox regression and Kaplan–Meier analyses were performed to evaluate the correlations between the first change in tumor size, response according to RECIST and OS. Data from CT evaluations of 233 patients were available. Appearance of new lesions or progression of non-target lesions (new/non-target) indicated short OS by univariable regression analysis (HR 3.76, 95 % CI 1.90–7.42, $p < 0.001$). A decrease by >30 % at this early time point was prognostic favorable (HR 0.69, 95 % CI 0.49–0.98, $p = 0.04$) and not significantly less than the best overall response according to RECIST. After adjustment for previous adjuvant treatment and the treatment given within the frame of the randomized trial, OS was still significantly shorter in patients with new/non-target lesions after a median 8 weeks of treatment (HR 4.41, 95 % CI 2.74–7.11, $p < 0.001$). Disease progression at the first evaluation correlates with OS in patients with MBC treated with first-line combination chemotherapy. The main reason for early disease progression was the appearance of new lesions or progression of non-target lesions. These patients had poor OS even though more lines of treatment were available. Thus, these factors should be focused on in the response evaluations besides tumor size changes.

C. Suzuki (✉) · L. Blomqvist · H. Jacobsson
Departments of Diagnostic Radiology, Molecular Medicine and Surgery, Karolinska University Hospital and Karolinska Institutet, 171-76 Stockholm, Sweden
e-mail: Chikako.Suzuki@ki.se

T. Hatschek · B. Glimelius
Department of Oncology and Pathology, Karolinska Institutet, Stockholm, Sweden

L. Carlsson
Department of Oncology, Sundsvall General Hospital, Sundsvall, Sweden

Z. Einbeigi · B. Linderholm
Department of Oncology, Sahlgrenska University Hospital, Gothenburg, Sweden

B. Linderholm
Science for Life Laboratory, Karolinska Institutet, Stockholm, Sweden

B. Lindh
Department of Oncology, Umeå University Hospital, Umeå, Sweden

N. Loman · M. Söderberg
Department of Oncology, Skåne University Hospital, Lund, Sweden

M. Malmberg
Department of Oncology, Helsingborg General Hospital, Helsingborg, Sweden

S. Rotstein
Department of Oncology, Karolinska University Hospital Dannderyd, Stockholm, Sweden

M. Sundqvist
Department of Oncology, Kalmar General Hospital, Kalmar, Sweden

Published online: 16 January 2013

Keywords Metastatic breast cancer · Treatment response evaluation · Overall survival · First-line chemotherapy · Computed tomography (CT) · RECIST

Introduction

Overall survival (OS) is the most important endpoint by which to assess the clinical benefit in patients with metastatic cancer, especially in phase III trials [1]. In metastatic breast cancer (MBC), several lines of treatment and a long post-recurrence survival affect the significance of OS as a primary endpoint. Therefore, this endpoint is frequently replaced by progression-free survival (PFS) or time to progression (TTP) [2, 3]. To determine response or disease progression in a standardized manner, the imaging-based response evaluation system RECIST is used to measure changes in tumor size following treatment. RECIST is regularly used in clinical trials, and it was recently updated to version 1.1 [4, 5].

Analysis of data from a phase III multicenter trial on metastatic colorectal cancer (mCRC) showed that changes in the size of metastatic lesions at the first point of evaluation by CT scan correlated well with OS [6]. In that study, the categorization of response as determined by RECIST was questioned, which is in line with clinical trials evaluating the efficacy of targeted agents [7–9].

In primary breast cancer, pathological complete response after neoadjuvant chemotherapy is a strong favorable prognostic factor, at least in hormone receptor-negative tumors [10–13]. In MBC, however, the prognostic impact of tumor response rate on the outcome is rather limited [2, 14]. The prognostic impact of measurable changes in tumor size found at the first evaluation of response after initiation of chemotherapy in relation to OS in MBC has, to our knowledge, not been evaluated previously.

The aim of this study was to assess whether early effects on the size of metastases or appearance of new lesions assessed by CT correlates with OS in patients treated with first-line combination chemotherapy for MBC.

T. M. Walz
Division of Oncology, Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University, Linköping, Sweden

T. M. Walz
Department of Oncology, Karolinska University Hospital, Stockholm, Sweden

G. Åström · B. Glimelius
Department of Radiology, Oncology and Radiation Science, Uppsala University, Uppsala, Sweden

H. Fujii
Functional Imaging Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, Kashiwa, Japan

Patients and methods

A total of 287 patients were enrolled in a multicenter, randomized phase III trial comparing first-line chemotherapy with a combination of epirubicin and paclitaxel alone (ET) or with capecitabine (TEX). Patients with morphologically confirmed loco-regional inoperable or disseminated breast carcinoma were included unless they had received treatment with an anthracycline, a taxane, or 5-FU within 1 year before the study entry. Previous endocrine treatment for advanced disease in patients with hormone receptor-positive breast cancer was allowed. Patients with known brain metastases or other malignancies within the past 5 years were excluded. The clinical results of the trial have been reported previously [15].

Chemotherapy was repeated every 3 weeks until disease progression or severe toxicity occurred, or the treatment was terminated upon the patient's request or for medical reasons. Baseline tumor measurements were obtained by CT, MRI or, in superficial tumors, palpation within 4 weeks prior to randomization. The measurements were repeated after every three cycles of chemotherapy (7–9 weeks). Response evaluation was conducted according to RECIST [4]; complete response (CR) occurred when all tumor lesions disappeared; partial response (PR) when there was at least a 30 % decrease in the sum of the largest diameters of the target lesions, confirmed after at least 4 weeks; progressive disease (PD) when there was at least a 20 % increase in the sum of the largest diameters of the target lesions or the appearance of new lesions or unequivocal progression of non-target lesions; and stable disease (SD) when criteria for either PR or PD were not met.

Measurements of target lesions and best overall response (OR) according to RECIST were retrieved from the study database.

At the first response evaluation, changes observed between the sum of the largest diameters (LD) of target lesions at the baseline (baseline sum) and the sum of the LD of the same target lesions at the first evaluation (first sum) were calculated. The response at the first evaluation was calculated as the first response = [(sum of measurable lesions at 1st evaluation) − (sum of measurable lesions at baseline)]/(sum of measurable lesions at baseline) (%).

To allow comparison with RECIST and to be clinically relevant, patients were grouped on the basis of the first response value in 10 % steps ranging from a decrease in size by more than 50 % to an increase in size by more than 20 %. In cases with new lesions or progression of non-target lesions at the first evaluation, the first response value

was set to 100 %. If all metastases had disappeared at the first evaluation, the first response value was set to −100 %. Best OR rate was defined as the sum of CR + PR, whereas disease control also included patients with SD. OS was defined as the time from randomization to the date of death.

The landmark method was applied to compensate for longer guarantee time for responders for first response and OR according to RECIST using OS as endpoint [16, 17]. Eight weeks were considered as the guarantee time period for first response evaluation and 18 weeks were applied for OR confirmed. Patients alive at the last follow-up were censored.

Cox regression analysis was used to define prognostic variables and to estimate hazard ratios (HR) and 95 % confidence intervals (CI). Kaplan–Meier analyses were performed to evaluate correlations between first response and OS. The log-rank test was used for comparison in survival between groups.

To assess the reliability of the first response values, one independent radiologist (CS) obtained the baseline sum of measurable lesions and the sum at the first evaluation of 40 patients (17 %) in the same cohort whose CT examinations were accessible. The reviewer had no information about the original radiological response assessments or patient outcomes. The intraclass correlation coefficient (ICC) was estimated. The ICC can range from 0.00 to 1.00, where 1.00 represents the perfect accordance. The repeatability coefficient was also estimated based on ICC analysis.

A value of $p < 0.05$ was considered statistically significant. Statistical computations were performed with STATISTICA ver. 10 (StatSoft Inc., Tulsa, OK, USA).

Results

A total of 233 out of 287 patients were eligible for response evaluation. Cases with missing measurements in one of the two evaluations ($n = 26$) or only non-measurable lesions ($n = 28$) were excluded. Patient characteristics are shown in Table 1. Median OS of eligible patients was 28 months.

The ICC was 0.96 (95 % CI 0.92–0.98), which indicates very good reproducibility.

The median time from randomization to the first evaluation was 8.3 weeks (SD 1.6 weeks). The extent of the first change is shown in Fig. 1. Twenty-three (10 %) patients had a new lesion or progression of non-target lesions at this point, eight of these in combination with more than a 20 % increase in target lesions, the cutoff point for PD according to RECIST. No patient had more than a 20 % increase in target lesions only. Only one of these 23 patients had a survival that was longer than the median OS. Survival longer than the median OS (blue column in Fig. 1) was seen both in patients who had at least a 30 % decrease,

Table 1 Patient characteristics ($n = 233$)

Characteristic	
Age at entry	
Mean (SD)	54 (9.0)
Median	55
Range	29–74
No. of lesions/patient	
Mean (SD)	2.8 (1.7)
Median	3
Range	1–9
No. of target lesions evaluated	
Liver	262 (39.9 %)
Lymph node	235 (35.8 %)
Lung	86 (13.1 %)
Others	74 (11.3 %)
Total	657
Treatment arm	
TEX	117
ET	116
Previous adjuvant treatment	
Yes	116
No	117

fulfilling the requirement for being PR at the first evaluation, and among those who then had stationary disease (<20 % increase–<30 % decrease).

Univariate regression analyses showed that the group of patients with new lesion/progression of non-target lesion had a significantly increased hazard of death compared to the reference group with no change or decrease less than 10 % (HR 3.76, 95 % CI: 1.90–7.42, $p < 0.001$) (Table 2). The magnitude of tumor shrinkage for each 10 % change showed only weak correlations with OS without any statistical significance (Table 2). The OR rate according to RECIST correlated also with OS (Table 2). Non-responders (SD or PD) had poorer OS compared to responders (HR 1.97, 95 % CI: 1.43–2.71, $p < 0.001$); however, disease control (SD + PR + CR) showed an even stronger correlation with OS (HR 4.66, 95 % CI 2.90–7.47, $p < 0.001$).

Based on the results of the Cox regression model, patients were divided into three groups: (a) new/non-target, (b) 1st decrease ≤30 % to increase of 20 %, and (c) 1st decrease >30 %. There was a significant difference in OS among these three groups (Table 2 and log-rank $\chi^2 = 31.4$, $p < 0.001$; Fig. 2a). The hazard ratio (Table 3) and survival curves were similar to those according to RECIST (log-rank $\chi^2 = 35.1$, $p < 0.001$, Fig. 2b), indicating that the imaging findings already after 8 weeks give almost as much prognostic information as the summary findings included as RECIST criteria, defined as best response by confirmation.

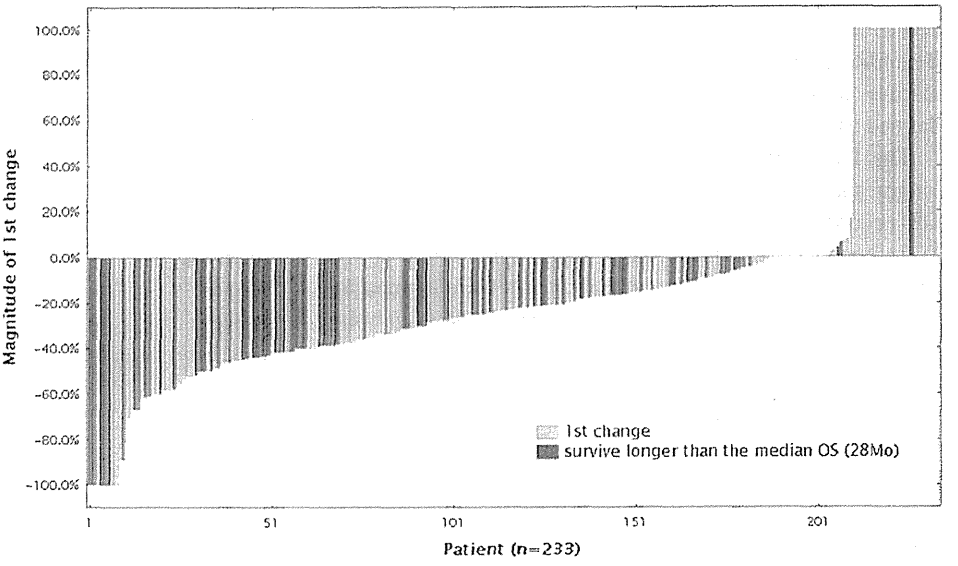

Fig. 1 Waterfall plot of the first change of tumor size based on all 233 patients. The change of size value was set to 100 % in case of new lesions or progression of non-target lesion observed at the first evaluation. The minus values imply size reduction. In case of disappearance of tumor, the change of size value was converted into −100 %. *Blue columns* indicate patients surviving longer than median OS (28 months)

Table 2 Univariable Cox regression analyses on overall survival (OS) in 233 patients with 158 events

	No. (censored)	HR	95 % CI	*p* value
Change of size at the first response evaluation				
New/non-target[a]	23 (1)	3.76	1.90–7.42	<0.001
Increase <20 %	23 (7)	0.86	0.42–1.76	0.68
No change–decrease 10 %	20 (5)	1	–	–
Decrease >10–20 %	36 (10)	0.94	0.49–1.77	0.84
Decrease >20–30 %	39 (12)	0.95	0.50–1.80	0.87
Decrease >30–40 %	35 (16)	0.69	0.35–1.37	0.29
Decrease >40–50 %	27 (13)	0.48	0.23–1.00	0.05
Decrease >50 %	30 (11)	0.79	0.40–1.56	0.49
New/non-target (a)	23 (1)	4.00	2.46–6.53	<0.001
Decrease ≤30 %–Increase <20 %	118 (34)	1	–	–
Decrease >30 %	92 (40)	0.69	0.49–0.98	0.04
Best overall response				
PD	23 (1)	3.49	2.09–5.84	<0.001
SD	72 (21)	1	–	–
PR	127 (48)	0.66	0.46–0.94	0.02
CR	11 (5)	0.43	0.18–1.01	0.053
Non-responder versus responder		1.97	1.43–2.71	<0.001
PD versus disease control (SD + PR + CR)		4.66	2.90–7.47	< 0.001

[a] New/non-target: appearance of new lesion or progression of non-target lesion

Clinical data from the TEX trial indicated that previous adjuvant treatment reduced the chance to respond to the study treatment [15]. Therefore, adjuvant treatment was included as a covariable in the multivariable Cox regression model together with the study treatment, ET versus TEX. Adjusted for these variables, new/non-target lesions were found to be an independent prognostic factor for OS (HR 4.41, 95 % CI: 2.74–7.11, $p < 0.001$, Table 4).

Discussion

Commonly applied surrogate endpoints, such as OR, TTP, and PFS, have been used to assess the benefit of therapies for patients with MBC [3, 18, 19]. However, the ability of these endpoints to accurately predict OS has been limited even when information from multiple trials has been analyzed [2, 14, 20]. This study indicates that MBC patients

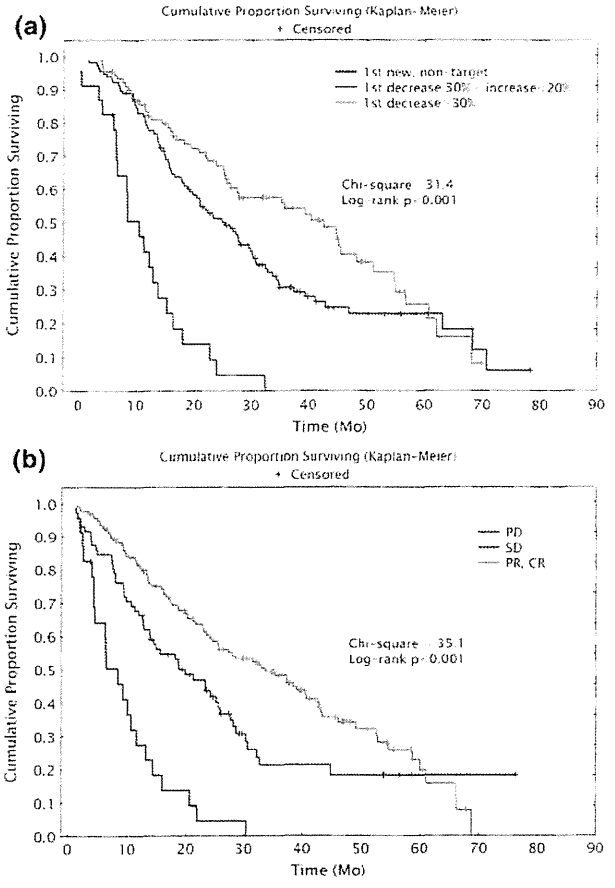

Fig. 2 Kaplan–Meier curves of overall survival according to the change of size at the first evaluation (**a**) and best overall response according to RECIST (**b**)

Table 3 Multivariate Cox regression model adjusting a decrease of 30 % or more at the first examination and responder (being PR and CR by RECIST) in terms of overall survival (OS) including 210 patients with 136 events

	HR	95 % CI	*p*
Decrease >30 %	0.77	0.51–1.17	0.22
PR, CR	0.73	0.49–1.11	0.14

Patients with PD were excluded

Table 4 Multivariate Cox regression model adjusting the first change of size to previous adjuvant treatment and study treatment arms in terms of overall survival (OS) including 233 patients with 158 events

	HR	95 % CI	*p*
New/non-target[a]	4.41	2.74–7.11	< 0.001
Previous adjuvant treatment (None = 1)	1.41	1.02–1.93	0.04
Study treatment (ET = 1.00)	0.92	0.67–1.26	0.59

[a] New/non-target: appearance of new lesion or progression of non-target lesion

with signs of progression at 8 weeks after the initiation of first-line combination chemotherapy have considerably shorter OS. The significant impact of the appearance of new lesions or progression of non-target lesions is concordant with previous results during chemotherapy for mCRC [6].

Clear signs of progression during chemotherapy for metastatic disease indicate that the current therapy has to be replaced by alternate treatments. The very strong prognostic impact of the appearance of new lesions or unequivocal progression of non-target lesions, found here, supports this strategy. This early progression to a combination of two (or three) drugs with known efficacy in breast cancer, resulting in poor OS, indicates that these tumors have an aggressive behavior and rarely respond to other cytostatics. Actually, only one of the 23 patients had an OS longer than median OS in the trial.

In a clinical trial setting, the concept of initial response may provide a rationale for modifying established procedures. PFS was the primary endpoint in the TEX trial, OR rates (CR + PR) were commonly used as an endpoint in many clinical trials in metastatic disease [3], and it was a secondary endpoint of the original TEX study [15]. Due to poor correlation with OS, OR rate was later replaced by PFS as the primary endpoint in most metastatic trials [3]. A reason for the limited association between OR rate and OS can be the arbitrary limit between PR and SD not separating those with and without favorable life-prolonging effect of the antitumor treatment. The disease control rate, that is, CR + PR + SD, showed a stronger correlation with OS than the OS rate in this study as well as in other studies [21]. Changes within ± 10 % may be a result of variability alone as shown in a study on the reproducibility of measurement of lung tumors [22]. The present study also questions why patients need to be followed until reaching best response or tumor progression; the first response could give almost the same prognostic information as the evaluation according to RECIST.

First-line treatment of breast cancer generally leads to response and delay of disease progression, in the present trial, between 10.8 and 12.4 months [15]. However, failure of obvious tumor shrinkage and the development of new lesions at the first evaluation may imply biologically aggressive therapy-resistant clones that require more intensive treatment. Since there was a complete overlap between patients with more than a 20 % increase and the development of new lesions or progression of non-target lesions, we were not able to find differences in OS between PD with and without development of new lesions and progression of non-target lesions. In our previous study on mCRC, a marked difference was found between those two groups [6].

There are several limitations of this study. The data come from a relatively limited number of patients. Since there was only one patient who had between 10–20 % increase and no patient with more than a 20 % increase and without new lesions, it was impossible to investigate the importance of the extent of progression on OS.

The population of patients with tumor shrinkage, a tendency to correlate with the extent of regression, was shown, yet it failed to reach statistical significance. Even though we could have had sufficient number of patients, it could still be difficult to determine the optimal cutoff point in case of tumor shrinkage, since the reduction should be a continuous variable related to the outcome. Furthermore, the relationship between the first tumor shrinkage and OS is diluted by those who respond to subsequent lines of treatment. On the other hand, this study highlights the poor prognosis of patients who had new lesions or progression of non-target lesions at the first evaluation even though several additional lines of treatment exist; resistance to one combination chemotherapy often mean resistance to other combination chemotherapies as well, although many exceptions are seen.

In this study, any optimal time point for readout of the first change could not be explored other than 8 weeks, as this was defined by the trial protocol. In studies of NSCLC and mCRC, the first response at 8 and 12 weeks, respectively, after initiation of chemotherapy has been suggested for use to predict prognosis [21, 23, 24].

In conclusion, although mainly restricted to those with the appearance of new lesions or progression of non-target lesions, in patients with MBC treated with combination chemotherapy, the first response approximately 8 weeks after initiation of chemotherapy correlates with OS.

Acknowledgments The authors thank Elisabeth Berg, the Medical Statistics Unit, Department of Learning, Informatics, Management and Ethics (LIME), Karolinska Institutet, Stockholm, Sweden, for professional assistance with the statistical analyses. The authors also thank Barbara L. Clough for her outstanding linguistic review. Financial supports were provided through the regional agreement on medical training and clinical research (ALF) between the Stockholm country council and the Karolinska Institutet, Pfizer, Roche and Sanofi-aventis.

Conflict of interest Although all authors completed the disclosure declaration of potential conflicts of interest, the following author indicated a financial or other interest that is relevant to the subject matter under consideration in this article. The sponsors had no influence on the design or performance of this study. Consultant/advisory role: Thomas Hatschek, Pfizer, Funding: Thomas Hatschek, Roche, Sanofi-aventis, Remuneration: None, Stock Ownership: None, Honoraria: None.

References

1. Johnson JR, Williams G, Pazdur R. End points and United States food and drug administration approval of oncology drugs. J Clin Oncol. 2003;21(7):1404–11.
2. Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. J Clin Oncol. 2008;26(12):1987–92.
3. Verma S, McLeod D, Batist G, et al. In the end what matters most? A review of clinical endpoints in advanced breast cancer. Oncologist. 2011;16(1):25–35.
4. Therasse P, Arbuck SG, Eisenhauer EA et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst. 2000;92(3):205–16.
5. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228–47.
6. Suzuki C, Blomqvist L, Sundin A, et al. The initial change in tumor size predicts response and survival in patients with metastatic colorectal cancer treated with combination chemotherapy. Ann Oncol. 2012;23(4):948–54.
7. Graf W, Pahlman L, Bergstrom R, et al. The relationship between an objective response to chemotherapy and survival in advanced colorectal cancer. Br J Cancer. 1994;70(3):559–63.
8. De Roock W, Piessevaux H, De Schutter J, et al. KRAS wild-type state predicts survival and is associated to early radiological response in metastatic colorectal cancer treated with cetuximab. Ann Oncol. 2008;19(3):508–15.
9. Piessevaux H, Buyse M, De Roock W, et al. Radiological tumor size decrease at week 6 is a potent predictor of outcome in chemorefractory metastatic colorectal cancer treated with cetuximab (BOND trial). Ann Oncol. 2009;20(8):1375–82.
10. Symmans WF, Peintinger F, Hatzis C, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol. 2007;25(28):4414–22.
11. von Minckwitz G, Sinn HP, Raab G, et al. Clinical response after two cycles compared to HER2, Ki-67, p53, and bcl-2 in independently predicting a pathological complete response after preoperative chemotherapy in patients with operable carcinoma of the breast. Breast Cancer Res. 2008;10(2):R30.
12. Akazawa K, Tamaki Y, Taguchi T, et al. Potential of reduction in total tumor volume measured with 3D-MRI as a prognostic factor for locally-advanced breast cancer patients treated with primary chemotherapy. Breast J. 2008;14(6):523–31.
13. Klauber-DeMore N, Ollila DW, Moore DT, et al. Size of residual lymph node metastasis after neoadjuvant chemotherapy in locally advanced breast cancer patients is prognostic. Ann Surg Oncol. 2006;13(5):685–91.
14. Brufsky AM, Hurvitz S, Perez E, et al. RIBBON-2: a randomized, double-blind, placebo-controlled, phase III trial evaluating the efficacy and safety of bevacizumab in combination with chemotherapy for second-line treatment of human epidermal growth factor receptor 2-negative metastatic breast cancer. J Clin Oncol. 2011;29(32):4286–93.
15. Hatschek T, Carlsson L, Einbeigi Z, et al. Individually tailored treatment with epirubicin and paclitaxel with or without capecitabine as first-line chemotherapy in metastatic breast cancer: a randomized multicenter trial. Breast Cancer Res Treat. 2012;131(3):939–47.
16. Buyse M, Piedbois P. On the relationship between response to treatment and survival time. Stat Med. 1996;15(24):2797–812.
17. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response and other comparisons of time-to-event by outcome variables. J Clin Oncol. 2008;26(24):3913–5.
18. Hackshaw A, Knight A, Barrett-Lee P, et al. Surrogate markers and survival in women receiving first-line combination anthracycline chemotherapy for advanced breast cancer. Br J Cancer. 2005;93(11):1215–21.

19. Bruzzi P, Del Mastro L, Sormani MP, et al. Objective response to chemotherapy as a potential surrogate end point of survival in metastatic breast cancer patients. J Clin Oncol. 2005;23(22): 5117–25.
20. Wilcken N, Dear R. Chemotherapy in metastatic breast cancer: a summary of all randomised trials reported 2000–2007. Eur J Cancer. 2008;44(15):2218–25.
21. Lara PN Jr, Redman MW, Kelly K, et al. Disease control rate at 8 weeks predicts clinical benefit in advanced non-small-cell lung cancer: results from Southwest Oncology Group randomized trials. J Clin Oncol. 2008;26(3):463–7.
22. Oxnard GR, Zhao B, Sima CS, et al. Variability of lung tumor measurements on repeat computed tomography scans taken within 15 minutes. J Clin Oncol. 2011;29(23):3114–9.
23. Yamamoto N, Nambu Y, Fujimoto T, et al. A landmark point analysis with cytotoxic agents for advanced NSCLC. J Thorac Oncol. 2009;4(6):697–701.
24. Heun JM, Grothey A, Branda ME, et al. Tumor status at 12 weeks predicts survival in advanced colorectal cancer: findings from NCCTG N9741. Oncologist. 2011;16(6):859–67.

Evaluation of Cartilage Invasion by Laryngeal and Hypopharyngeal Squamous Cell Carcinoma with Dual-Energy CT[1]

Hirofumi Kuno, MD
Hiroaki Onaya, MD, PhD
Ryoko Iwata, MD
Tatsushi Kobayashi, MD
Satoshi Fujii, MD, PhD
Ryuichi Hayashi, MD
Katharina Otani, PhD
Hiroya Ojiri, MD, PhD
Takeharu Yamanaka, PhD
Mitsuo Satake, MD, PhD

Purpose: To evaluate the clinical usefulness of dual-energy computed tomography (CT) with weighted-average (WA) images and iodine overlay (IO) images in the evaluation of laryngeal cartilage invasion in patients with laryngeal and hypopharyngeal squamous cell carcinoma (SCC).

Materials and Methods: The institutional review board approved this retrospective study, and written comprehensive consent was obtained from all patients. Seventy-two consecutive patients underwent 128-section dual-source dual-energy CT to stage laryngeal (n = 27) or hypopharyngeal (n = 45) cancer. Three observers who were blinded to the patients' clinical histories and histopathologic findings evaluated cartilage invasion on WA images alone or in combination with IO images (nonossified cartilages were selectively evaluated on IO images) by using a five-point scale. Thirty of the 72 patients (42%) underwent surgery, and findings from histopathologic examination in those patients were used as the standard of reference for the evaluation of diagnostic performance with receiver operating characteristic (ROC) curve analysis and in terms of sensitivity and specificity. Interobserver reproducibility was calculated with κ statistics.

Results: For thyroid cartilage, the area under the ROC curve (AUC) of the WA plus IO images was marginally larger than that for WA images alone (AUC = 0.957 vs 0.870, respectively; P = .075). The specificity of WA plus IO images was significantly superior to that of WA images alone (96% vs 70%, respectively; P = .031), with no compromise to the sensitivity (86% for both). For thyroid and cricoid cartilages, the interobserver reproducibility was higher for diagnoses made with WA plus IO images (κ = 0.68–0.72 and 0.64–0.79, respectively) than for those made with WA images alone (κ = 0.29–0.56 and 0.20–0.64, respectively).

Conclusion: Combined analysis of WA and IO images obtained with dual-energy CT improves the diagnostic performance and interobserver reproducibility of evaluations of laryngeal cartilage invasion by SCC.

[1] From the Diagnostic Radiology Division (H.K., R.I., T.K., M.S.), Pathology Division, Research Center for Innovative Oncology (S.F.), Head and Neck Surgery Division (R.H.), and Research Center for Innovative Oncology (T.Y.), National Cancer Center Hospital East, 6-5-1 Kashiwanoha, Kashiwa, Chiba 277-8577, Japan; Diagnostic Radiology Section, Clinical Trials and Practice Support Division, Center for Cancer Control and Information Services, National Cancer Center, Chuo-ku, Tokyo, Japan (H. Onaya); Imaging & Therapy Systems Division, Healthcare Sector, Siemens Japan K. K., Tokyo, Japan (K.O.); and Department of Radiology, Jikei University School of Medicine, Tokyo, Japan (H. Ojiri). Received August 15, 2011; revision requested October 27; revision received April 27, 2012; accepted May 9; final version accepted May 18. Supported in part by the National Cancer Center Research and Development Fund (23-A-35). **Address correspondence to** H.K. (e-mail: *hkuno@east.ncc.go.jp*).

The accurate evaluation of cartilage invasion is essential for determining the appropriate treatment strategies for laryngeal and hypopharyngeal cancer. Early stages without cartilage invasion are initially treated with the intention of preserving the larynx. For stages with limited local cartilage invasion, function-preserving partial laryngeotomy or chemoradiotherapy have been recently introduced (1–3); however, advanced stages with apparent cartilage invasion still necessitate total laryngectomy (1,4,5).

Computed tomography (CT) is a well-established method for evaluating cartilage invasion but nevertheless presents challenges. The CT appearance of laryngeal cartilage varies widely according to the different proportions of hyaline cartilage (which ossifies with aging), cortical bone, and fatty marrow. Moreover, nonossified cartilage and tumors show similar CT values of approximately 100 HU (6,7), making them almost indistinguishable—especially when the tumor is located adjacent to nonossified cartilage. Acceptable sensitivities (71%) and specificities (83%) have been reported by head and neck radiologists applying appropriate evaluation criteria to single-section spiral CT (7). The introduction of multisection CT has resulted in little progress in image interpretation, and cartilage invasion is thus still sometimes overestimated, resulting in unnecessary total laryngectomies in some patients (8,9).

In dual-energy CT, two data sets acquired with different tube voltages can be fused to generate weighted-average (WA) images that have a similar appearance to conventional CT images obtained at 120 kV (10) in addition to generating images of the distribution of iodinated contrast medium alone. For these applications, the material-specific x-ray energy dependence of the absorption coefficient is used in image postprocessing to mathematically extract iodine and separately calculate color-coded iodine images and virtual unenhanced images. Currently, dual-energy CT is being investigated in several clinical fields (11–15), including the differentiation of renal lesions, which show similar CT values to those on conventional CT scans (16,17).

Considering that an iodine staining distribution is seen in tumor tissues but not in normal cartilage, we assumed that the distinction of tumor and cartilage tissue may be possible by using dual-energy CT. Therefore, this modality may be applicable in the evaluation of cartilage invasion by laryngeal and hypopharyngeal squamous cell carcinoma (SCC).

The purpose of this study was to evaluate the clinical usefulness of dual-energy CT with WA images and iodine overlay (IO) images in the evaluation of laryngeal cartilage invasion in patients with laryngeal and hypopharyngeal SCC.

Advances in Knowledge

- Noncalcified cartilage with tumor invasion, which shows iodine distribution, is easier to visually differentiate from healthy cartilage on dual-energy CT iodine overlay (IO) images than on weighted-average (WA) images because iodine is selectively color coded and enhanced on IO images.
- The combined analysis of WA and IO images obtained with dual-energy CT can improve the specificity for thyroid cartilage ($P = .031$) without compromising sensitivity.
- Interobserver reproducibility in the detection of thyroid and cricoid cartilage invasion was higher for diagnoses made with WA plus IO images ($\kappa = 0.68–0.72$ and 0.64–0.79, respectively) than for those made with WA images alone ($\kappa = 0.29–0.56$ and 0.20–0.64, respectively).

Implication for Patient Care

- Preliminary evidence suggests that dual-energy CT has the potential to increase specificity without compromising sensitivity in the evaluation of thyroid cartilage invasion; this is particularly important when choosing a treatment strategy, especially when function-preserving therapy is being considered.

Materials and Methods

All work included in the manuscript was performed at the National Cancer Center Hospital East, Kashiwa, Chiba, Japan. One of the authors (K.O.) is an employee of Siemens Japan. Authors who are not employees of Siemens Japan monitored and had control of inclusion of any data and information submitted for publication.

Study Population

Our institutional review board approved this retrospective study. Previous written comprehensive consent from all patients was substituted for informed consent. Seventy-three consecutive patients with laryngeal or hypopharyngeal cancer were scheduled to undergo contrast material–enhanced CT for cancer staging between May 2010 and March 2011. One patient was excluded because of an iodine allergy. The remaining 72 patients (64 men, eight women; age range, 41–86 years; mean age, 67 years) were included in this study. Twenty-seven of the 72 patients (38%) had laryngeal cancer (13 patients had supraglottic carcinoma, 13 patients had glottic carcinoma, and

Published online before print
10.1148/radiol.12111719 **Content code:** HN

Radiology 2012; 265:488–496

Abbreviations:
AUC = area under the ROC curve
IO = iodine overlay
ROC = receiver operating characteristic
SCC = squamous cell carcinoma
WA = weighted average

Author contributions:
Guarantors of integrity of entire study, H.K., H. Onaya, M.S.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; literature research, H.K., H. Onaya, K.O., H. Ojiri; clinical studies, H.K., H. Onaya, T.K., S.F., R.H., M.S.; experimental studies, R.I.; statistical analysis, H.K., H. Onaya, T.Y.; and manuscript editing, H.K., H. Onaya, K.O., T.Y.

Conflicts of interest are listed at the end of this article.

one patient had subglottic carcinoma) and 45 (63%) had hypopharyngeal cancer (32 patients had piriform sinus carcinoma, seven had postcricoid carcinoma, and six had carcinoma of the pharyngeal posterior wall). SCC was histopathologically confirmed with endoscopic biopsy in all patients.

Thirty of the 72 patients (42%; 10 patients with laryngeal cancer and 20 with hypopharyngeal cancer) underwent surgical resection (nine patients underwent total laryngectomy, 17 underwent total pharyngolaryngectomy, and four underwent partial pharyngolaryngectomy), 26 (36%) underwent radical radiation therapy, and 16 (22%) underwent chemoradiotherapy. The median interval between CT diagnosis and surgery was 12 days (range, 1–34 days). In the 30 patients who underwent surgery, the CT findings were compared with the histopathologic findings, which served as the standard of reference. In the patients who underwent partial pharyngolaryngectomy, the involved side of the thyroid cartilage was removed. As a result, the cartilage was histologically evaluated in 108 preparations (30 [28%] thyroid cartilages, 26 [24%] cricoid cartilages, and 52 [48%] arytenoid cartilages).

CT Protocol

Scans were obtained by using a 128-section, dual-source CT unit (Somatom Definition Flash; Siemens Healthcare, Forchheim, Germany) in the dual-energy CT mode (18) with the following parameters: tube voltages of 100 and Sn140 kV (where Sn indicates the use of a 0.4-mm tin filter), 200 and 200 effective mAs, 0.33-second rotation time, 32 × 0.6-mm collimation with a z-flying focal spot, and pitch of 0.6. The average CT dose index was 14.35 mGy. The voltage combination of 100 and Sn140 kV was chosen to minimize noise while maximizing the separation of the x-ray tube's energy spectra.

Patients received a 100-mL injection of iodinated contrast material (Ioverin; Taiyo Yakuhin, Nagoya, Japan, 300 mg of iodine per milliliter) at a rate of 2.5 mL/sec. Scanning was started 70 seconds after the start of the injection and proceeded in a craniocaudal direction. The scanning range was set from the base of the skull to the bottom of the neck.

Image Reconstruction and Postprocessing

Two image sets (100 kV and Sn140 kV) were separately reconstructed from the acquired dual-energy data with 1.0-mm-thick sections in 0.7-mm increments by applying D30f kernels. A third image set (WA images) was calculated pixel by pixel from the 100-kV and Sn140-kV data sets, reconstructed by applying a B30f kernel. The pixels of the data sets (p_{100} and p_{140}, respectively) were linearly weighted and fused at the default ratio of $w = 0.5$ according to the following formula: $[w \times p_{100}] + [(1 - w) \times p_{140}]$ (19,20).

Next, all images were transferred to a commercial workstation (MMWP, Siemens Healthcare) and a three-material decomposition analysis (Syngo Dual Energy, Brain Hemorrhage; Siemens Healthcare) was used to subtract the iodine images from the contrast-enhanced dual-energy images, thereby generating virtual unenhanced images (12,13,19). The algorithm was applied at the default parameters for differentiating iodine from hemorrhage and brain parenchyma, which have similar CT values to those of soft tissue and cartilage. "Organ contour enhancement" and "resolution enhancement" were deselected to ensure a quantitative analysis of the small cartilage structures. Finally, the virtual unenhanced and iodine images were linearly weighted and fused at a ratio of 0.5 to create virtual unenhanced images with a color-coded iodine distribution (IO images).

Parallel ranges with 2.0-mm-thick sections and a 16-cm field of view were generated from the WA images, the virtual unenhanced images, and the IO images as follows: axial and coronal sections parallel and vertical to the vocal cords, respectively (in soft tissue and bone windows), from 1.0 cm above the hyoid bone to the inferior margin of the cricoid cartilage.

Image Interpretation

Three radiologists (H. Onaya, R.I., and T.K., with 22, 22, and 18 years of experience in diagnostic oncologic radiology, respectively), who were blinded to the patients' clinical histories and histopathologic findings, retrospectively and independently evaluated all images twice by using an NCC-CIR viewer (IBM, Armonk, NY). WA images were evaluated first, followed 4 weeks later by evaluation of the WA plus IO images. All images were presented in random order. The final diagnosis was determined by consensus and was used to compare the CT findings with the pathologic results.

The invasion of each laryngeal cartilage—thyroid, cricoid, and bilateral arytenoid cartilages—and the extralaryngeal tumor spread were evaluated according to diagnostic criteria by using a five-point scale, as follows: 1 = definitely negative, 2 = probably negative, 3 = possibly positive, 4 = probably positive, and 5 = definitely positive. These criteria were first applied to the WA images to evaluate tumor location, tumor extension, and the presence of invasion. Then, the IO images were read in addition to the WA images to identify the iodine distribution corresponding to the area noted on WA images. Areas of sclerosis and ossified cartilage were excluded from the evaluation of IO images. After a lesion had been classified as positive on WA images, the iodine distribution on the IO images was examined to derive the final classification as either positive or negative.

Thyroid cartilage.—For the thyroid cartilage, cartilage invasion was considered to be present when erosion (score, 3), lysis (score, 4), or extralaryngeal tumor spread through the cartilage (score, 5) was present (Fig 1). Cartilages with asymmetric sclerotic changes (ossification) without erosion or lysis were classified as negative on WA images because these changes may represent reactive changes (7,21). IO images were not used for additional information in these cases.

Cricoid and arytenoid cartilages.—For the cricoid and arytenoid cartilages, cartilage invasion was considered to be present when erosion or lysis was observed. Asymmetric sclerosis of the cartilage in contact with the tumor was also considered indicative of cartilage invasion on WA images according to previously reported criteria (6,7,22–24).

Figure 1

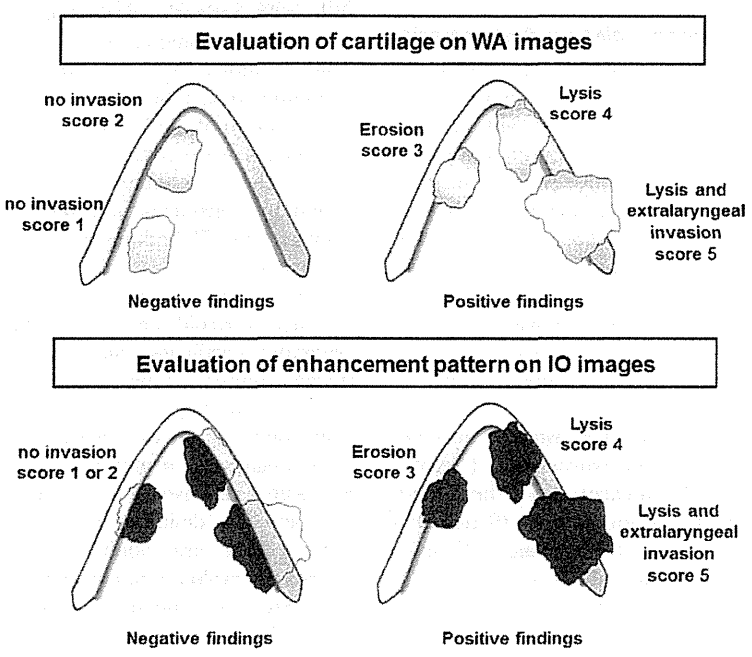

Figure 1: Diagrams show evaluation of thyroid cartilage invasion. Erosion (score, 3) was defined as invasion beyond inner cortex without reaching outer cortex (less than half the cartilage width), lysis (score, 4) was defined as almost reaching outer cortex but with preservation of that cortex, and extralaryngeal tumor spread (score, 5) was defined as all-layer invasion through both inner and outer cortex (penetration) of cartilage, including extralaryngeal soft tissues. For erosion and lysis, two findings needed to be confirmed on WA image: *(a)* disappearance of a thin hypoattenuating line between tumor and cartilage and *(b)* substitution of cartilage with tumor tissue. Conversely, preservation of a dark line between tumor and cartilage resulted in a negative finding. A fully defined line was scored as 1, and an almost continuously defined line was scored as 2.

Extralaryngeal tumor spread.—Extralaryngeal tumor spread was considered to be present when the primary tumor had expanded into the extralaryngeal soft tissues (eg, cervical soft tissues, infrahyoid muscles, thyroid gland, esophagus, trachea, or deep lingual muscle) with or without cartilage penetration.

Histologic Evaluation

Surgical specimens, including all cartilages around the tumor, were fixed in formalin and decalcified in advance and then cut into 3.0–4.0-mm-thick slices in the frontal direction, similar to the cross-sectional CT images, and evaluated by a pathologist (S.F., with 18 years of experience).

Statistical Analysis

Receiver operating characteristic (ROC) curves and sensitivity and specificity tables were used to evaluate the diagnostic performance of dual-energy CT with WA images alone and in combination with IO images. The histopathologic findings from the 30 patients who underwent surgery served as the standard of reference. Diagnostic performance was analyzed with ROC curves, and the area under the ROC curve (AUC) was calculated. ROC curves were compared to determine the diagnostic performance with WA images alone and in combination with IO images. Sensitivity and specificity were calculated on the basis of the assumption that a confidence level of 3 or higher was positive for the diagnosis of cartilage invasion and extralaryngeal tumor spread. The McNemar test was used to compare sensitivities and specificities between diagnoses made with WA images alone and in combination with IO images.

The interobserver reproducibility for the scores based on the WA images alone and in combination with IO images were calculated by using κ statistics (25). A 95% confidence interval was calculated for each κ (26).

All statistical tests were performed by using commercial software (IBM SPSS Statistics 18; SPSS, Chicago, Ill). $P < .05$ was considered indicative of a significant difference. Because the P values were exploratory in nature, Bonferroni correction was not applied.

Results

All WA and IO images generated with the dual-energy CT data were considered to be of diagnostic image quality.

Correlation between Dual-Energy CT Findings and Histopathologic Diagnosis

Pathologic evaluation of the 108 cartilages (30 thyroid cartilages, 26 cricoid cartilages, 52 arytenoid cartilages) from 30 patients revealed cartilage invasion in 14 of the 108 preparations (13%; seven thyroid cartilages, four cricoid cartilages, three arytenoid cartilages). On a per-patient basis, cartilage invasion was confirmed in eight of the 30 patients (27%). Extralaryngeal tumor spread was confirmed pathologically in 14 of 30 extralaryngeal soft-tissue preparations (47%). The 10 laryngeal tumors were classified as pT3 ($n = 5$) and pT4a ($n = 5$), and the 20 hypopharyngeal tumors were classified as pT1 ($n = 3$), pT2 ($n = 2$), pT3 ($n = 5$), and pT4a ($n = 10$).

Peritumoral inflammatory changes were found in both the invaded and noninvaded laryngeal cartilages surrounding the actual tumor borders. The overall pathologic evaluation revealed inflammatory changes in 27 of the 108 (25%) preparations (15 of 30 thyroid cartilages, seven of 26 cricoid cartilages, five of 52 arytenoid carti

lages). Inflammatory changes were found in many ossified cartilages, especially in fatty marrow infiltrated with calcified areas.

For the thyroid cartilage, use of WA images alone led to seven false-positive findings (erosion, $n = 4$; lysis, $n = 2$; highly asymmetric sclerosis, $n = 1$). With use of both WA and IO images, however, six of these seven cases were classified as negative because the iodine distribution on IO images did not match the shape of the suspected tumor on WA images. Figures 2 and 3 show cases with false-positive findings of erosion and lysis, respectively, on WA images. Figure 4 shows a true-positive case of invasion for comparison. One false-positive diagnosis remained even after examination of WA and IO images for a case with highly asymmetric sclerosis that was classified as possibly positive contrary to the criteria and despite the absence of any signs of erosion or lysis on IO images. False-positive findings on WA plus IO images caused by peritumoral inflammation were not encountered. One false-negative case was discovered incidentally on histologic preparations at locations other than those of cartilage invasion diagnosed on WA plus IO images. This case presented microinvasions of 2.0 mm or less to the ossifying thyroid cartilage.

For the arytenoid cartilage, diagnosis with WA images alone led to five false-positive findings, three of which remained false positive on WA plus IO images because they showed highly asymmetric sclerotic change. Peritumoral inflammatory changes were not found in these cases.

For extralaryngeal tumor spread, the sensitivity, specificity, and accuracy of the WA plus IO images were 100% (14 of 14 patients), 100% (16 of 16 patients), and 100% (30 of 30 patients), respectively. In one patient with pathologically proved extralaryngeal spread from a supraglottic SCC, the finding was missed on the WA images but was visible as a colored area on the IO images.

Statistical Results

The ROC curves for thyroid cartilage and the corresponding AUCs for diagnoses with WA images alone and in combination with IO images are shown in Figure 5. The AUC for diagnosis with the WA plus IO images (0.957) was larger than that for WA images alone (0.870), but the difference was not statistically significant ($P = .075$). The difference in the AUCs for the cricoid, right arytenoid, and left arytenoid cartilages were not significant ($P = .479$ and 0.292 for cricoid and right arytenoid cartilage, respectively; P value was not applicable to the left arytenoid cartilage).

The diagnostic performance of CT in the evaluation of cartilage invasion and extralaryngeal spread is summarized in Table 1. For the thyroid cartilage, evaluation of WA plus IO images resulted in a significantly higher specificity ($P = .031$) than did evaluation of WA images alone. The improvement in specificity was significant only for the thyroid cartilage ($P = .97$, .50, and .99, for cricoid, right arytenoid, and left arytenoid cartilage, respectively). The sensitivity for the thyroid, cricoid, and arytenoid cartilages did not differ between WA images alone and in combination with IO images.

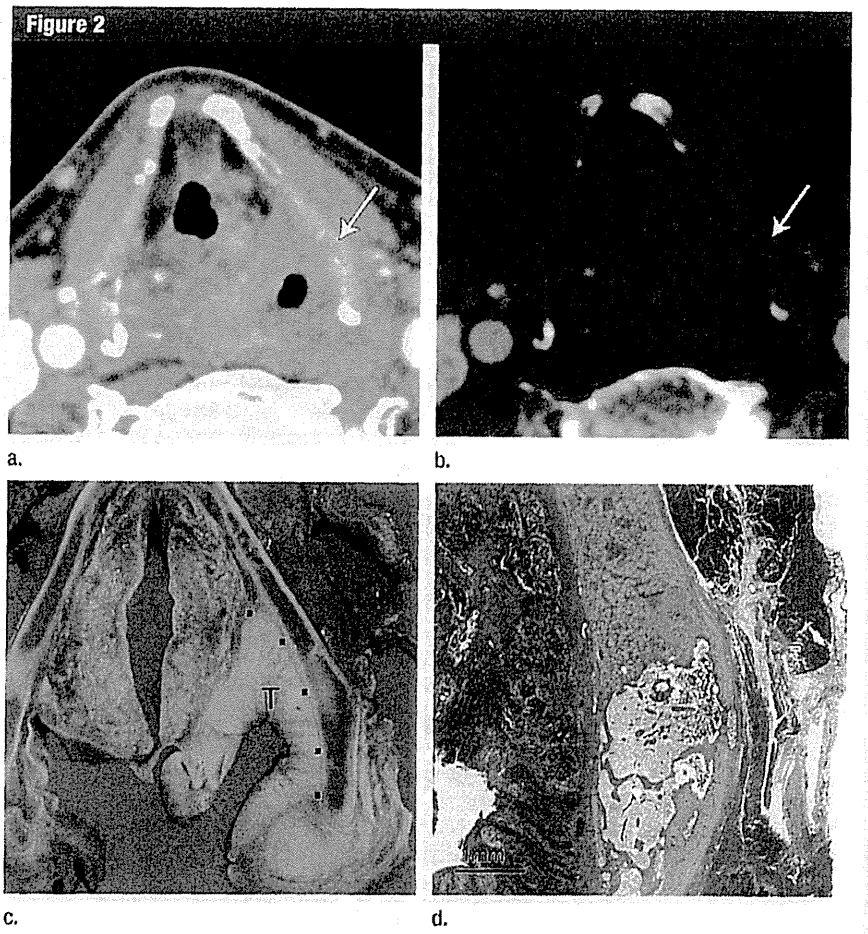

Figure 2: False-positive finding of erosion of thyroid cartilage with WA image alone in 69-year-old man with hypopharyngeal cancer. **(a)** WA image at supraglottic level shows tumor arising from left piriform sinus. Inner cortex of left thyroid cartilage shows focal erosion (arrow). **(b)** IO image does not show corresponding area of enhancement (arrow). **(c)** Corresponding slice from surgical specimen shows that left thyroid cartilage has not been invaded by tumor cells. ■ = macroscopic border of tumor (*T*). **(d)** Photomicrograph confirms that tumor cells have not invaded thyroid cartilage. (Hematoxylin-eosin stain; original magnification, ×5.)

Figure 3

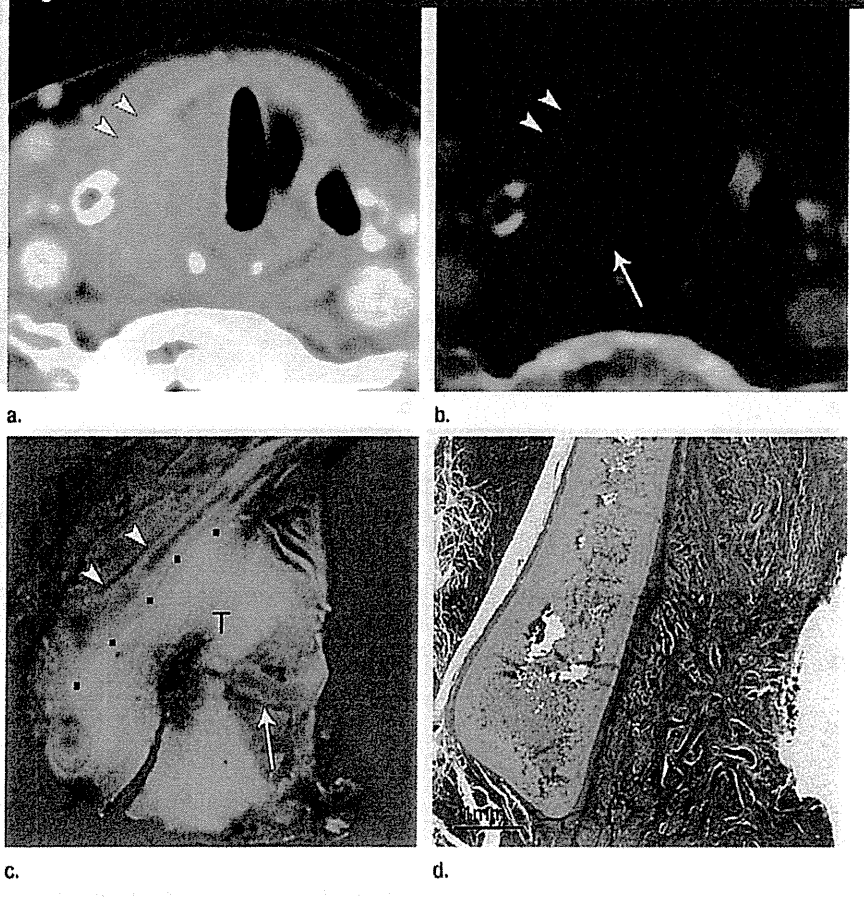

Figure 3: False-positive finding of lysis in thyroid cartilage on WA image alone in 70-year-old man with supraglottic cancer. **(a)** WA image at supraglottic level shows tumor invading the paraglottic space and piriform sinus. Nonossified cartilage of right thyroid lamina has been replaced by tumor (arrowheads). **(b)** IO image does not show a corresponding area of enhancement in thyroid cartilage (arrowheads), and right nonossified arytenoid cartilage behind tumor clearly shows no enhancement (arrow). **(c)** Corresponding slice from surgical specimen shows that right thyroid cartilage (arrowheads) and arytenoid cartilage (arrow) have not been invaded by tumor cells. ■ = macroscopic border of tumor (7). **(d)** Photomicrograph confirms that tumor cells have invaded thyroid and arytenoid cartilages. (Hematoxylin-eosin stain; original magnification, ×5.)

Interobserver Reproducibility

Table 2 shows the results of κ statistic analysis comparing diagnoses based on the WA images alone and in combination with IO images. For thyroid cartilage, the interobserver reproducibility of the invasion scores was higher for diagnoses made with WA and IO images, with substantial agreement (κ = 0.68–0.72) for WA and IO images and fair or moderate agreement (κ = 0.29–0.56) for WA images alone. For cricoid cartilage, the interobserver reproducibility tended to improve with the addition of IO images (from κ = 0.20–0.64 to κ = 0.64–0.79). However, the reproducibility with WA images alone did not differ from that with a combination of WA and IO images; there was slight to fair agreement for arytenoid cartilage (κ = 0.34–0.72 and 0.31–0.50, respectively) and substantial agreement for extralaryngeal tumor spread (κ = 0.66–0.80 and 0.63–0.85, respectively).

Discussion

In the current dual-energy CT study, we compared the diagnostic performance of combined WA plus IO images with that of WA images alone, which are equivalent to standard single-energy 120-kV CT images (10), in the evaluation of tumor invasion by laryngeal and hypopharyngeal SCC. Diagnosis with a combination of WA plus IO images showed significantly higher specificities, with no compromise in the sensitivity for thyroid cartilage. Furthermore, the interobserver reproducibility of diagnosis with a combination of WA plus IO images was higher than that with WA images alone (especially in the evaluation of thyroid and cricoid cartilage invasion) and that reported with single-section spiral CT (27,28). High specificity and reproducibility could be achieved because the IO images depicted areas of iodine distribution in nonossified cartilage and thus reduced the overestimation of invasion that occurred during diagnoses with WA images alone. This is particularly important for treatment decision making when function-preserving therapy is being considered because fewer false-positive diagnoses mean fewer unnecessary laryngectomies.

In the early 1990s, CT was considered a highly specific test for the evaluation of thyroid cartilage invasion, with specificities reaching 87%–97%. However, these values were obtained at the expense of sensitivity, which was reported to be 50%–67% (29–31). In an attempt to increase sensitivity, Becker et al (7) in 1997 redefined the diagnostic criteria for single-section CT and were able to reach an acceptable balance of 71% versus 83% between sensitivity and specificity by applying these criteria of erosion, lysis, and extralaryngeal tumor spread. Despite technologic advances in CT (eg, transition from single-section to multisection CT and the increase in spatial and temporal resolution), no major reports were published

Figure 4

a. b. c.

Figure 4: *True-positive finding of invasion through outer cortex in thyroid cartilage on WA plus IO image in 78-year-old man with glottic cancer.* **(a)** WA image at level of false vocal cords shows tumor invasion into left thyroid cartilage, spreading into extralaryngeal soft tissue (arrow). **(b)** IO image clearly shows extent of extralaryngeal tumor spread, which matches the tumor shape (white arrow). **(c)** Corresponding slice from surgical specimen shows extralaryngeal tumor (*T*) spreading through cartilage (white arrow), consistent with findings on IO image **(b)**. In addition, a nonenhanced islet is shown in anterior portion on IO image (black arrow in **b**), and pathologic findings confirmed absence of tumor (black arrow in **c**). Sclerosis of right thyroid cartilage is seen on WA image (arrowheads in **a**) but is not interpretable on IO image. Pathologic specimen reveals invasion by tumor cells at location of sclerosis (arrowheads in **c**).

Figure 5

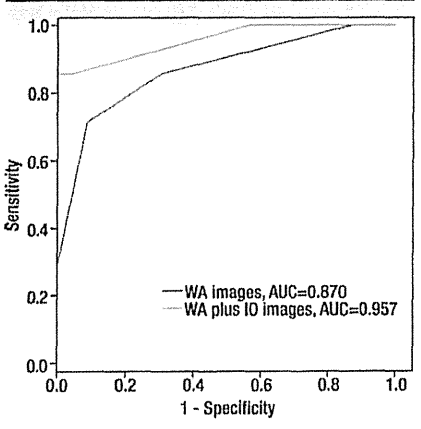

Figure 5: Graph shows ROC curves and corresponding AUCs in the prediction of thyroid cartilage invasion. The AUC for WA plus IO images was marginally significantly larger (AUC = 0.957, $P = .075$) than that for WA images alone (AUC = 0.870).

about the diagnostic performance of CT for thyroid cartilage invasion until 2007, when Li et al (8) reported a sensitivity of 85% and a specificity of 75% with 64-section CT.

The diagnostic performance with WA images alone in the present study was in line with that in previous reports (7,8). However, by adding information from IO images, we were able to improve the specificity of CT while maintaining a high sensitivity. The improvement in accuracy was possible because IO images revealed tumor tissue within (or adjacent to) the cartilage as red areas of iodine distribution, resulting in better visual differentiation between the tumor and noncalcified thyroid cartilage.

Magnetic resonance (MR) imaging seems to be more sensitive than CT for detecting cartilage invasion (sensitivity up to 96%) (32). However, peritumoral inflammation can mimic cartilage invasion even on MR images, leading to a number of false-positive findings. Because inflammatory changes are most common in the thyroid cartilage, the specificity of MR imaging for detecting invasion of the thyroid cartilage is only 56%–65% (30,32). Furthermore, MR imaging is often degraded by motion artifacts and lacks thin sections; both of these factors mean that MR imaging is not a satisfactory first choice for imaging laryngeal and hypopharyngeal SCC. Although we did not encounter false-positive findings on WA plus IO images caused by peritumoral inflammation in our present study, the possibility of problems similar to those with MR imaging occurring with IO images must be addressed in a future study.

The present study had some limitations, including one related to the technical aspects of dual-energy CT. First, IO imaging is unsuitable for the evaluation of areas of sclerotic change and previously ossified cartilage, as seen mostly in the results for the arytenoid cartilage. Because we used the three-material decomposition algorithm (12,13,19), bone and calcified structures were classified into iodine or soft tissues according to tissue density and could not be recognized on IO images. If lesions including these structures were evaluated on the basis of IO images alone, inappropriate overestimation of iodine distribution would have easily occurred. Therefore, areas of sclerosis and previously ossified parts of cartilage should be evaluated initially on WA images before the addition of IO images to avoid this inherent limitation. Second, the number of cases in this study was relatively small, and further studies involving a larger number of cases with histologic findings

Table 1

Diagnostic Performance in the Evaluation of Cartilage Invasion and Extralaryngeal Tumor Spread

Parameter	TP*	TN*	FN*	FP*	Sensitivity (%)†	Specificity (%)	PPV (%)	NPV (%)
Thyroid cartilage invasion (n = 30)								
WA images alone	6	16	1	7	86 (6/7)	70 (16/23)	46 (6/13)	94 (16/17)
WA plus IO images	6	22	1	1	86 (6/7)	96 (22/23)‡	86 (6/7)	96 (22/23)
Cricoid cartilage invasion (n = 26)								
WA images alone	4	21	0	1	100 (4/4)	95 (21/22)	80 (4/5)	100 (21/21)
WA plus IO images	4	22	0	0	100 (4/4)	100 (22/22)	100 (5/5)	100 (21/21)
Right arytenoid cartilage invasion (n = 26)								
WA images alone	2	19	1	4	67 (2/3)	83 (19/23)	33 (2/6)	95 (19/20)
WA plus IO images	2	21	1	2	67 (2/3)	91 (21/23)	50 (2/4)	95 (21/22)
Left arytenoid cartilage invasion (n = 26)								
WA images alone	0	25	0	1	NA	96 (25/26)	0 (0/1)	100 (25/25)
WA plus IO images	0	25	0	1	NA	96 (25/26)	0 (0/1)	100 (25/25)
Extralaryngeal tumor spread (n = 30)								
WA images alone	13	16	1	0	93 (13/14)	100 (16/16)	100 (13/13)	94 (16/17)
WA plus IO images	14	16	0	0	100 (14/14)	100 (16/16)	100 (14/14)	100 (16/16)

Note.—Numbers in parentheses are raw data. FN = false-negative findings, FP = false-positive findings, NPV = negative predictive value, PPV = positive predictive value, TN = true-negative findings, TP = true-positive findings.

* Data are numbers of findings.

† NA = not applicable.

‡ P = .031 for comparison between WA images and WA and IO images.

are needed to ensure a clinical effect. Third, the statistical analysis was potentially limited in that the more reasonable analyses, including the linear mixed model to account for clustering, did not work well and could not clearly justify improvement of the interobserver reproducibility owing to the small sample size. These problems remain to be solved in future studies. Furthermore, because differences among readers in the scoring of WA images remained relatively high, larger studies are warranted to clarify these issues.

In conclusion, combined analysis of WA and IO images generated with dual-energy CT has the potential to improve the diagnostic performance and interobserver reproducibility of CT in the evaluation of cartilage invasion by laryngeal and hypopharyngeal SCC.

Table 2

Interobserver Agreement in the Evaluation of Cartilage Invasion and Extralaryngeal Tumor Spread

Parameter and Reader	WA Images Alone	WA plus IO Images
Thyroid cartilage invasion (n = 72)		
A vs B	0.52 (0.29, 0.75)	0.72 (0.46, 0.98)
A vs C	0.46 (0.23, 0.70)	0.72 (0.46, 0.98)
B vs C	0.29 (0.02, 0.58)	0.68 (0.39, 0.97)
Cricoid cartilage invasion (n = 72)		
A vs B	0.64 (0.28, 1.01)	0.79 (0.49, 1.07)
A vs C	0.20 (−0.07, 0.46)	0.64 (0.32, 0.96)
B vs C	0.47 (0.18, 0.75)	0.68 (0.39, 0.97)
Arytenoid cartilage invasion (n = 144)		
A vs B	0.34 (0.07, 0.61)	0.50 (0.22, 0.78)
A vs C	0.40 (0.14, 0.65)	0.31 (0.02, 0.61)
B vs C	0.72 (0.53, 0.90)	0.45 (0.20, 0.71)
Extralaryngeal tumor spread (n = 72)		
A vs B	0.80 (0.64, 0.97)	0.85 (0.70, 0.99)
A vs C	0.76 (0.59, 0.92)	0.70 (0.52, 0.88)
B vs C	0.66 (0.46, 0.85)	0.63 (0.43, 0.82)

Note.—Data are κ values, with 95% confidence intervals in parentheses. κ values of less than 0 indicate poor agreement; 0–0.2, slight agreement; 0.21–0.4, fair agreement; 0.41–0.6, moderate agreement; 0.61–0.8, substantial agreement; and 0.81–1, almost perfect agreement.

Acknowledgments: We thank Keiichi Nomura, MS, National Cancer Center Hospital East, for his excellent technical assistance in image acquisition and processing and Bernhard Krauss, PhD, Siemens Healthcare, Forchheim, Germany, for technical advice on dual-energy CT.

Disclosures of Conflicts of Interest: **H.K.** No relevant conflicts of interest to disclose. **H. Onaya** No relevant conflicts of interest to disclose. **R.I.** No relevant conflicts of interest to disclose. **T.K.** No relevant conflicts of interest to disclose. **S.F.** No relevant conflicts of interest to disclose. **R.H.** No relevant conflicts of interest to disclose. **K.O.** No relevant conflicts of interest to disclose. **H. Ojiri** No relevant conflicts of interest to disclose. **T.Y.** No relevant conflicts of interest to disclose. **M.S.** No relevant conflicts of interest to disclose.

References

1. American Society of Clinical Oncology, Pfister DG, Laurie SA, et al. American Society of Clinical Oncology clinical practice guideline for the use of larynx-preservation strategies in the treatment of laryngeal cancer. J Clin Oncol 2006;24(22):3693–3704.
2. Hoffman HT, Porter K, Karnell LH, et al. Laryngeal cancer in the United States: changes in demographics, patterns of care, and survival. Laryngoscope 2006;116(9 Pt 2 Suppl 111):1–13.
3. Forastiere AA, Goepfert H, Maor M, et al. Concurrent chemotherapy and radiotherapy for organ preservation in advanced laryngeal cancer. N Engl J Med 2003;349(22):2091–2098.
4. Rodriguez CP, Adelstein DJ, Rybicki LA, et al. Clinical predictors of larynx preservation after multiagent concurrent chemoradiotherapy. Head Neck 2008;30(12):1535–1542.
5. Castelijns JA, Becker M, Hermans R. Impact of cartilage invasion on treatment and prognosis of laryngeal cancer. Eur Radiol 1996;6(2):156–169.
6. Hermans R. Staging of laryngeal and hypopharyngeal cancer: value of imaging studies. Eur Radiol 2006;16(11):2386–2400.
7. Becker M, Zbären P, Delavelle J, et al. Neoplastic invasion of the laryngeal cartilage: reassessment of criteria for diagnosis at CT. Radiology 1997;203(2):521–532.
8. Li B, Bobinski M, Gandour-Edwards R, Farwell DG, Chen AM. Overstaging of cartilage invasion by multidetector CT scan for laryngeal cancer and its potential effect on the use of organ preservation with chemoradiation. Br J Radiol 2011;84(997):64–69.
9. Hartl DM, Landry G, Hans S, Marandas P, Brasnu DF. Organ preservation surgery for laryngeal squamous cell carcinoma: low incidence of thyroid cartilage invasion. Laryngoscope 2010;120(6):1173–1176.
10. Tawfik AM, Kerl JM, Razek AA, et al. Image quality and radiation dose of dual-energy CT of the head and neck compared with a standard 120-kVp acquisition. AJNR Am J Neuroradiol 2011;32(11):1994–1999.
11. Graser A, Johnson TR, Hecht EM, et al. Dual-energy CT in patients suspected of having renal masses: can virtual nonenhanced images replace true nonenhanced images? Radiology 2009;252(2):433–440.
12. Gupta R, Phan CM, Leidecker C, et al. Evaluation of dual-energy CT for differentiating intracerebral hemorrhage from iodinated contrast material staining. Radiology 2010;257(1):205–211.
13. Johnson TR, Krauss B, Sedlmair M, et al. Material differentiation by dual energy CT: initial experience. Eur Radiol 2007;17(6):1510–1517.
14. Kang MJ, Park CM, Lee CH, Goo JM, Lee HJ. Dual-energy CT: clinical applications in various pulmonary diseases. RadioGraphics 2010;30(3):685–698.
15. Leschka S, Stolzmann P, Baumüller S, et al. Performance of dual-energy CT with tin filter technology for the discrimination of renal cysts and enhancing masses. Acad Radiol 2010;17(4):526–534.
16. Brown CL, Hartman RP, Dzyubak OP, et al. Dual-energy CT iodine overlay technique for characterization of renal masses as cyst or solid: a phantom feasibility study. Eur Radiol 2009;19(5):1289–1295.
17. Karlo C, Lauber A, Götti RP, et al. Dual-energy CT with tin filter technology for the discrimination of renal lesion proxies containing blood, protein, and contrast-agent: an experimental phantom study. Eur Radiol 2011;21(2):385–392.
18. Flohr TG, McCollough CH, Bruder H, et al. First performance evaluation of a dual-source CT (DSCT) system. Eur Radiol 2006;16(2):256–268.
19. Krauss B, Schmidt B, Flohr TG. Dual source CT. In: Johnson TRC, Fink C, Schönberg SO, Reiser MF, eds. Dual energy CT in clinical practice. Heidelberg, Germany: Springer, 2011; 11–20.
20. Paul J, Bauer RW, Maentele W, Vogl TJ. Image fusion in dual energy computed tomography for detection of various anatomic structures: effect on contrast enhancement, contrast-to-noise ratio, signal-to-noise ratio and image quality. Eur J Radiol 2011;80(2):612–619.
21. Beitler JJ, Muller S, Grist WJ, et al. Prognostic accuracy of computed tomography findings for patients with laryngeal cancer undergoing laryngectomy. J Clin Oncol 2010;28(14):2318–2322.
22. Nix PA, Salvage D. Neoplastic invasion of laryngeal cartilage: the significance of cartilage sclerosis on computed tomography images. Clin Otolaryngol Allied Sci 2004;29(4):372–375.
23. Schmalfuss IM, Mancuso AA, Tart RP. Arytenoid cartilage sclerosis: normal variations and clinical significance. AJNR Am J Neuroradiol 1998;19(4):719–722.
24. Muñoz A, Ramos A, Ferrando J, et al. Laryngeal carcinoma: sclerotic appearance of the cricoid and arytenoid cartilage—CT-pathologic correlation. Radiology 1993;189(2):433–437.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–174.
26. Agresti A. Measuring agreement. In: Categorical data analysis. New York, NY: Wiley, 1990; 366–367.
27. Hermans R, Van der Goten A, Baert AL. Image interpretation in CT of laryngeal carcinoma: a study on intra- and interobserver reproducibility. Eur Radiol 1997;7(7):1086–1090.
28. Hoorweg JJ, Kruijt RH, Heijboer RJJ, Eijkemans MJC, Kerrebijn JDF. Reliability of interpretation of CT examination of the larynx in patients with glottic laryngeal carcinoma. Otolaryngol Head Neck Surg 2006;135(1):129–134.
29. Zbären P, Becker M, Läng H. Pretherapeutic staging of laryngeal carcinoma: clinical findings, computed tomography, and magnetic resonance imaging compared with histopathology. Cancer 1996;77(7):1263–1273.
30. Becker M, Zbären P, Laeng H, Stoupis C, Porcellini B, Vock P. Neoplastic invasion of the laryngeal cartilage: comparison of MR imaging and CT with histopathologic correlation. Radiology 1995;194(3):661–669.
31. Sulfaro S, Barzan L, Querin F, et al. T staging of the laryngohypopharyngeal carcinoma: a 7-year multidisciplinary experience. Arch Otolaryngol Head Neck Surg 1989;115(5):613–620.
32. Becker M, Zbären P, Casselman JW, Kohler R, Dulguerov P, Becker CD. Neoplastic invasion of laryngeal cartilage: reassessment of criteria for diagnosis at MR imaging. Radiology 2008;249(2):551–559.

Geminin Expression in Pancreatic Neuroendocrine Tumors

Possible New Marker of Malignancy

Masaki Aizawa, MD, Motohiro Kojima, MD, PhD,† Naoto Gotohda, MD, PhD,* Satoshi Fujii, MD, PhD,† Yuichiro Katoh, MD, PhD,* Takahiro Kinoshita, MD, PhD,* Shinichiro Takahashi, MD, PhD,* Masaru Konishi, MD, PhD,* Taira Kinoshita, MD, PhD,* and Atsushi Ochiai, MD, PhD†*

Objectives: We evaluated geminin labeling index (LI) as a new prognostic indicator of pancreatic neuroendocrine tumor.

Methods: Twenty-seven patients who underwent surgery were retrospectively referred. Labeling indices for geminin and Ki-67 were calculated and compared with clinicopathologic factors. Then, the concordance of positivity between 2 LIs was evaluated using the color difference quotation.

Results: The median (range) of LIs for geminin and Ki-67 was 1.0% (0.05%–14.9%) and 1.5% (0.02%–8.8%), respectively. When the high LI was defined as more than 2.0% according to the receiver operating characteristic curves determining the metastasis, both geminin LI (hazard ratio [HR], 31.382; 95% confidence interval [CI], 3.177–309.99; $P = 0.003$) and Ki-67 LI (HR, 6.182; 95% CI, 1.221–31.298; $P = 0.028$) were significant risk factors of recurrence in the univariate analysis. The Kaplan-Meier curves consistently exhibited the superiority of geminin LI (log rank, $P < 0.001$) to Ki-67 LI (log rank, $P = 0.041$) in predicting the disease-free survival. In the color difference quotation, the median ΔE of geminin stain (16.12; range, 5.8–41.9) was significantly larger than that of Ki-67 stain (13.17; range, 3.4–37.9).

Conclusions: The geminin LI was suggested to be more closely correlated with outcome and had more consistent positivity than the Ki-67 LI.

Key Words: geminin, neuroendocrine tumor, color difference quotation, prognostic factor, Ki-67

(*Pancreas* 2012;41: 512–517)

The annual incidence of pancreatic neuroendocrine tumor (PNET) is about 0.32 cases per population of 100,000 in the United States and 2.23 cases per population of 100,000 in Japan. Pancreatic neuroendocrine tumors are thought to represent 1%–2% of all pancreatic neoplasms. The apparent incidence and prevalence of PNET have increased substantially during the last 30 years, probably because of the rapid progress of innovative diagnostic techniques.[1] The best treatment for PNET is curative surgical resection, which has a disease-free survival rate of 82% after surgery.[2] Pancreatic neuroendocrine tumors have a wide spectrum of clinical presentations. Therefore, multiple studies have attempted to develop staging and grading systems to better define prognosis.[2–5] The 2000 World Health Organization (WHO) classification system used both stage-related criteria (size and presence of metastases) and grade-related criteria (mitotic rate, perineural invasion, angioinvasion, and Ki-67 proliferative index) to predict outcome. Though this approach included most well-accepted pathologic prognostic factors, the multiple grading parameters made it difficult to reproduce grades reliably among pathologists and institutions, and this grading system has since been replaced by the current WHO classification.[6] Immunohistochemistry for Ki-67 protein is commonly used to evaluate the proliferative activity of tumor cells, and numerous studies have shown that the labeling index (LI) of the Ki-67 protein is correlated with the clinical outcome of patients with a variety of malignant tumors, including PNET.[2,3,7–10] The Ki-67 protein is detected during all active cell cycle phases (G_1, synthesis, G_2, and mitosis and cytokinesis) but not in resting (G_0) cells, although its function remains uncertain.[11,12] Although histologic grade-based estimations of prognosis are extremely useful for interpreting biopsy samples, additional reliable markers are needed.

Geminin, a negative regulator of DNA replication, has recently been described as a novel marker of malignant potential.[13] During the G_1 phase, DNA replication is initiated by the recruitment of the origin recognition complex, composed of cell division cycle-6 and Cdt1, to specific points of replication origin in the genome; this recruitment, in turn, loads the minichromosome maintenance (Mcm) complex, which is composed of Mcm-2 to Mcm-7.[14] Geminin is specifically expressed during the S, G_2, and early M phases and interacts with Cdt1 to prevent the loading of the Mcm complex to points of origin that have already been initiated, thus ensuring a single replication per 1 cell cycle.[15,16] Geminin expression has been widely observed in various malignant neoplasms, and the number of geminin-positive cells is reportedly proportional to the cell proliferation index, as measured using Ki-67 expression.[13] High levels of geminin expression are reportedly correlated with a poorer clinical outcome in breast cancer,[17] renal cell carcinoma,[18] prostatic adenocarcinoma,[19] salivary gland carcinomas,[20] lung cancer,[21] and gastric hyperplasia.[22] However, the prognostic significance of geminin expression in PNET remains unknown.

The purpose of this study was to determine whether geminin expression defines the aggressiveness of PNET and to compare the clinical and diagnostic use of the geminin LI with that of the Ki-67 LI.

MATERIALS AND METHODS

Patients

Between 1994 and 2010, a total of 27 consecutive patients underwent primary surgical treatment for PNET at our institution. The medical records and surgical specimens of these patients were retrospectively examined in the present study.

From the *Division of Digestive Surgery and †Pathology Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, Kashiwa, Japan.

Received for publication February 2, 2011; accepted September 23, 2011.

Reprints: Atsushi Ochiai, MD, PhD, Pathology Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, 6-5-1 Kashiwanoha, Kashiwa, Chiba 277-8577, Japan (e-mail: aochiai@east.ncc.go.jp).

Dr Aizawa states no source of financial support and no disclosure of funding received for this work.

The authors declare no conflict of interest.

Patients with recurrent tumors were excluded. Follow-up clinical information was obtained from the patients' medical records. The follow-up time was measured from the date of surgery until disease-caused death or the end of the follow-up period.

Clinicopathologic Parameters

The grading and staging of each tumor were performed according to the WHO classification, the cancer staging manual of the American Joint Committee on Cancer (AJCC), and the classification proposed by the European Neuroendocrine Tumor Society (ENETS).[23,24]

The prognostic values of the following clinicopathologic parameters were examined in the present study: tumor diameter (<2 vs ≥2 cm), lymphatic or blood vessel infiltration (absent vs present), perineural invasion (absent vs present), serosal or retroperitoneal invasion (absent vs present), tumor extension beyond the pancreas (absent vs present), mitotic index per 10 high-power fields (10 HPFs) (<2 vs ≥2), regional and distant metastasis (absent vs present), and pathological stage (AJCC stage ≥ IIA vs AJCC stage ≤ I, ENETS stage ≥ IIb vs ENETS stage ≤ IIa).

Histologic Examination and Immunohistochemistry

Surgical specimens were fixed in 10% formalin and embedded in paraffin. Two pathologists (M.A. and M.K.), who were unaware of the clinical data, reviewed all the hematoxylin-and-eosin–stained sections and reclassified and graded the specimens according to the histologic parameters.

Serial 4-μm sections were used for immunohistochemical staining. Deparaffinized and rehydrated sections were immersed in 0.3% hydrogen peroxide in methanol for 30 minutes to block endogenous peroxidase activity. Heat-induced antigen retrieval was performed for 20 minutes at 95°C with a 10-mM citrate buffer (pH 6.0). After the slides had cooled at room temperature for 1 hour, they were exposed to 2% normal swine serum in phosphate-buffered saline for 30 minutes, then allowed to react overnight at 4°C with the following mouse monoclonal antibodies: antihuman geminin (diluted 1:40, clone EM6; Novocastra, Newcastle, United Kingdom) or antihuman Ki-67 (diluted 1:100, clone MIB-1; Dako, Glostrup, Denmark). After washing with phosphate-buffered saline 3 times, the sections were reacted with EnVision plus (Dako) for 30 minutes at room temperature. The peroxidase reaction products were developed with 3,3′-diaminobenzidine, and the sections were counterstained with hematoxylin.

The LI of each marker was calculated by manually counting the number of brown-stained tumor cell nuclei among the total number of tumor cells in the most highly immunoreactive area at a magnification of 400-fold, with the aid of an eyepiece grid (5 × 5 squares). Indices were expressed as a percentage value corresponding to the number of positive cells among approximately 2000 tumor cells.

Evaluation of Color Difference Quotation

Immunohistochemically stained full-face sections from each case with geminin and Ki-67 overexpression were digitized using the Slide Path and the NanoZoomer Digital Pathology System (Hamamatsu, Welwyn Garden City, United Kingdom). Approximately 7 minutes was required to scan a slide at a resolution of 40×. Subsequently, 400-fold magnified images from highly immunoreactive areas were exported for analysis.

The image analysis was performed using Photoshop (version 7; Adobe Systems, San Jose, CA). First, the image mode was converted from RGB to Lab color mode. Two hundred fifty representative positive cells were selected in each of 4 cases with high geminin expression levels (LI ≥ 2.0%) and 7 cases with high Ki-67 expression levels (LI ≥ 2.0%). All the cells with a recognizable brown stain were measured. The L*a*b* value of the nuclear areas was outputted as a range from 0 to 255. The values represented a Lab color space composed of 3 axes in a spherical form: L^*, a^*, and b^*. The L^* axis was associated with the lightness of the color, whereas the a^* and b^* axes were associated with the red-green scale and the yellow-blue scale, respectively. After the conversion of the scale to the CIE LAB color system, the L^* value was described as a decimal scale from 0 to 100, and the a^* and b^* values ranged from −128 to 127. The difference between the average values for a positive cell and an adjacent negative cell was calculated as ΔL^*, Δa^*, and Δb^*. Then, the color difference ΔE was estimated using the formula $\Delta E = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2}$. ΔE was then compared between the geminin-stained and Ki-67–stained sections.

Statistical Analysis

The Spearman rank correlation test was used to determine associations between continuous variables. Receiver operating characteristic curves were plotted to calculate the sensitivity, specificity, positive predictive value, and negative predictive value for the presence of metastasis. The cutoff values for the geminin LI and the Ki-67 LI were chosen so as to obtain the best combination of predictive values. A univariate analysis using the Cox proportional hazards model was applied to estimate the associations of clinicopathologic factors, including the immunohistochemical results, with the disease-free survival period. Survival curves were drawn using the Kaplan-Meier method, and the differences were analyzed using a log-rank test. Differences in nonparametric data were estimated using the Mann-Whitney U test. All P values < 0.05 were considered statistically significant. The statistical analyses were performed using Dr. SPSS II for Windows (SPSS Japan, Tokyo, Japan).

RESULTS

Demographic Characteristics and Tumor-Related Factors

Among the 27 patients, the median age at the time of diagnosis was 56 years; 14 patients (51.9%) were men, and 13 patients (48.1%) were women. In all the patients, a curative resection was performed, followed by a histologic assessment of the tumor grading. The tumor-related factors are summarized in Table 1. The tumor was located in the pancreatic head in 11 cases (40.7%), in the body in 12 cases (44.4%), and in the tail in 4 cases (14.8%). Only 1 case of functional PNET (insulinoma) was included. The median and range of the maximum tumor diameter were 26 mm and 8 to 92 mm, respectively. The diameter was 2 cm or larger in 11 cases (40.7%). Local invasion was observed in 2 cases (7.4%). One of these cases exhibited an obstruction of the inferior common bile duct, and another case presented with invasion to both the splenic artery and vein. Lymph node metastasis was encountered in 10 cases (37.0%), and a solitary liver metastasis was resected in 1 case (3.7%). With respect to mitosis, fewer than 2 mitoses per 10 HPFs were present in 19 cases (70.4%), from 2 to 10 mitoses per 10 HPFs were encountered in 8 cases (29.6%), and none of the cases had more than 20 mitoses per 10 HPFs.

The tumor grade classifications according to the WHO system and tumor staging according to the AJCC or ENETS criteria are shown in Table 1. Nineteen cases (70.4%) were classified as grade 1 neuroendocrine tumor (NET), and 8 cases

TABLE 1. Patient Demographics and Tumor-Related Factors for 27 Patients

Characteristics		
Age, y	Median	56
	Range	31–76
Sex, n (%)	Men	14 (51.9)
	Women	13 (48.1)
Location, n (%)	Ph	11 (40.7)
	Pb	12 (44.4)
	Pt	4 (14.8)
Functional tumor, n (%)		1 (3.7)
Maximum diameter, mm	Median	26
	Range	8–92
Diameter, n (%)	<2 cm	16 (59.3)
	≥2 cm	11 (40.7)
Local invasion, n (%)		2 (7.4)
Metastasis, n (%)	—	17 (63.0)
	Lymph node	10 (37.0)
	Liver	1 (3.7)
Mitosis, n (%)	0–1 per 10 HPFs	19 (70.4)
	2–20 per 10 HPFs	8 (29.6)
	>20 per 10 HPFs	0 (0.0)
WHO grading, n (%)	NET G1	19 (70.4)
	NET G2	8 (29.6)
	NEC	0 (0.0)
AJCC stage, n (%)	IA	8 (29.7)
	IB	9 (33.3)
	IIA	0 (0.0)
	IIB	9 (33.3)
	III	0 (0.0)
	IV	1 (3.7)
ENETS stage, n (%)	I	8 (29.7)
	IIa	7 (25.9)
	IIb	2 (7.4)
	IIIa	0 (0.0)
	IIIb	9 (33.3)
	IV	1 (3.7)

NET G1 indicates grade 1 NET; NET G2, grade 2 NET; Pb, pancreatic body; Ph, pancreatic head; Pt, pancreatic tail.

(29.6%) were classified as grade 2. According to the AJCC staging, 17 cases (63.0%), 9 cases (33.3%), 0 cases (0%), and 1 case (3.7%) were classified as stages I, II, III, and IV, respectively. According to the classification proposed by ENETS, 8 cases (29.7%), 9 cases (33.3%), 9 cases (33.3%), and 1 case (3.7%) were classified as stages I, II, III, and IV, respectively.

The median and range of the observation period were 1704 days and 37 to 4206 days, respectively. Three patients died, one of whom had a tumor-related death; the other 2 patients had treatment-related deaths. Recurrence after surgery was observed in 6 patients (22.2%).

Geminin and Ki-67 Expression

The immunohistochemical analysis examined the expressions of geminin and Ki-67 protein in all the cases (Fig. 1). Immunoreactivity was observed exclusively in the nuclei of the tumor cells. Geminin was also immunoreactive in the perichromosomal cytoplasm of mitotic cells in a few cases. The median LIs for geminin and Ki-67 were 1.0% and 1.5%, respectively. The geminin LI was slightly but significantly lower than that of Ki-67. Figure 2 shows the positive correlation between the geminin LI and the Ki-67 LI (Spearman $R_s = 0.757$, $P < 0.001$).

The receiver operating characteristic curves for the geminin LI, the Ki-67 LI, and the mitosis count (all of which were continuous variables), which were used to predict the presence of metastatic lesions, are shown in Figure 3. The curves for the 2 LIs were similar. The area under the curve was calculated to be 0.829 (95% confidence interval [CI], 0.660–0.999) for the geminin LI, 0.776 (95% CI, 0.598–0.955) for the Ki-67 LI, and 0.594 (95% CI, 0.362–0.826) for the mitosis count. The geminin LI seemed to have a slightly superior ability to predict metastasis, compared with the Ki-67 LI. The sensitivity, specificity, positive predictive value, and negative predictive value of a geminin LI greater than 2.0% (n = 4) and of a Ki-67 LI greater than 2.0% (n = 7) for determining metastasis were 33.3%, 90.5%, 50.0%, and 82.6% and 50.0%, 80.0%, 42.9%, and 85.0%, respectively. We defined high-geminin expression cases as those with a geminin LI greater than 2% because this cutoff had the best discriminatory power for the predictive values. According to the current WHO classification, a Ki-67 LI of 2.0% can be used to discriminate G1 tumors, and this cutoff also had the best discriminatory power for the predictive values in the present analysis. Thus, we regarded a Ki-67 LI greater than 2.0% as indicating a high Ki-67 expression level.

Correlations of Geminin and Ki-67 LIs With Prognosis

Because there was only 1 tumor-related death in this series, we examined the predictive values of each LI for the disease-free survival period after surgery. The results of a univariate Cox regression analysis are shown in Table 2. A mitosis count of 2 or more per 10 HPFs (hazard ratio [HR], 10.204; 95% CI, 1.684–61.834; $P = 0.012$), a local invasion (HR, 18.762; 95%

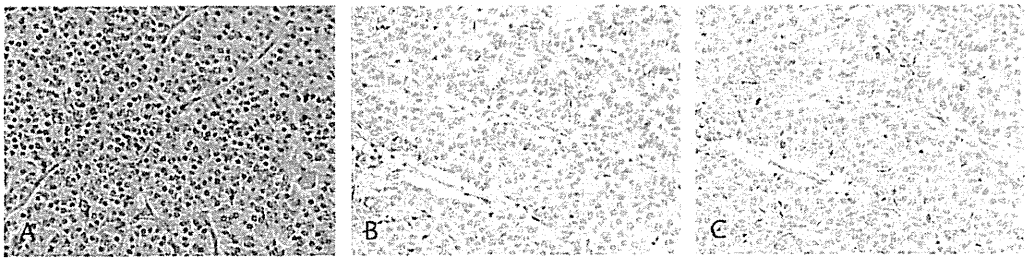

FIGURE 1. Representative photomicrograph of a PNET specimen. Hematoxylin and eosin staining (A) shows a typical trabecular arrangement of uniform tumor cells. The cells have eosinophilic cytoplasm and centrally located, round nuclei. Immunohistochemical staining for geminin (B) and Ki-67 (C) shows brown-stained tumor cell nuclei. The number of geminin-positive cells was smaller than the number of Ki-67–positive cells in most cases (original magnification ×400).