


# Individualized risk assessment in neuroblastoma: does the tumoral metabolic activity on $^{123}\text{I}$ -MIBG SPECT predict the outcome?

Julian M. M. Rogasch<sup>1</sup>  · Patrick Hundsdoerfer<sup>2,3</sup> · Christian Furth<sup>1</sup> · Florian Wedel<sup>1</sup> · Frank Hofheinz<sup>4</sup> · Paul-Christian Krüger<sup>5</sup> · Holger Lode<sup>6</sup> · Winfried Brenner<sup>1</sup> · Angelika Eggert<sup>2</sup> · Holger Amthauer<sup>1</sup> · Imke Schatka<sup>1</sup>

Received: 8 April 2017 / Accepted: 19 July 2017 / Published online: 14 August 2017  
© Springer-Verlag GmbH Germany 2017

## Abstract

**Purpose** Risk-adapted treatment in children with neuroblastoma (NB) is based on clinical and genetic factors. This study evaluated the metabolic tumour volume (MTV) and its asphericity (ASP) in pretherapeutic  $^{123}\text{I}$ -MIBG SPECT for individualized image-based prediction of outcome.

**Methods** This retrospective study included 23 children (11 girls, 12 boys; median age 1.8 years, range 0.3–6.8 years) with newly diagnosed NB consecutively examined with pretherapeutic  $^{123}\text{I}$ -MIBG SPECT. Primary tumour MTV and ASP were defined using semiautomatic thresholds. Cox regression analysis, receiver operating characteristic analysis (cut-off determination) and Kaplan-Meier analysis with the log-rank test for event-free survival (EFS) were performed for ASP, MTV, laboratory parameters (including urinary homovanillic acid-to-creatinine ratio, HVA/C), and clinical

(age, stage) and genetic factors. Predictive accuracy of the optimal multifactorial model was determined in terms of Harrell's *C* and likelihood ratio  $\chi^2$ .

**Results** Median follow-up was 36 months (range 7–107 months; eight patients showed disease progression/relapse, four patients died). The only significant predictors of EFS in the univariate Cox regression analysis were ASP ( $p = 0.029$ ; hazard ratio, HR, 1.032 for a one unit increase), MTV ( $p = 0.038$ ; HR 1.012) and MYCN amplification status ( $p = 0.047$ ; HR 4.67). The mean EFS in patients with high ASP (>32.0%) and low ASP were 21 and 88 months, respectively ( $p = 0.013$ ), and in those with high MTV (>46.7 ml) and low MTV were 22 and 87 months, respectively ( $p = 0.023$ ). A combined risk model of either high ASP and high HVA/C or high MTV and high HVA/C best predicted EFS.

**Conclusions** In this exploratory study, pretherapeutic image-derived and laboratory markers of tumoral metabolic activity in NB (ASP, MTV, urinary HVA/C) allowed the identification of children with a high and low risk of progression/relapse under current therapy.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00259-017-3786-1) contains supplementary material, which is available to authorized users.

✉ Julian M. M. Rogasch  
julian.rogasch@charite.de

<sup>1</sup> Department of Nuclear Medicine, Charité - Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany

<sup>2</sup> Department of Pediatric Oncology/Hematology, Charité - Universitätsmedizin Berlin, Berlin, Germany

<sup>3</sup> Berlin Institute of Health (BIH), 10178 Berlin, Germany

<sup>4</sup> PET Center, Helmholtz Zentrum Dresden-Rossendorf, Institute of Radiopharmaceutical Cancer Research, Dresden, Germany

<sup>5</sup> Institute for Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany

<sup>6</sup> Department of Pediatric Oncology and Hematology, University Medicine Greifswald, Greifswald, Germany

**Keywords** Neuroblastoma · Prognosis ·  $^{123}\text{I}$ -MIBG · Asphericity · Metabolic tumour volume

## Introduction

Neuroblastoma (NB) is the most frequent solid extracranial malignant tumour in children [1]. Based on a risk-adapted therapy, the prognosis in these patients has continuously improved during recent decades. However, NB is still the cause of 15% of all cancer-related deaths in childhood [2]. In Germany, treatment and follow-up are performed in accordance with the NB2004 trial protocol

[3]. This allows refined risk stratification into three treatment groups based on clinical factors (age, tumour stage) and genetic determinants (MYCN gene amplification, chromosome 1p36 abnormalities). The high-risk group includes all patients with MYCN amplification as well as patients older than 1 year with stage 4 disease. Despite a highly intensive treatment regimen, including high-dose chemotherapy, external beam irradiation,  $^{123}\text{I}$ -metaiodobenzylguanidine (MIBG) therapy and immunotherapy, 5-year event-free survival (EFS) and overall survival (OS) of high-risk patients are only about 40% and 50%, respectively [4, 5]. In contrast, long-term survivors may suffer from therapy-associated morbidity and mortality (e.g. hearing loss, infertility or secondary malignancies) [6, 7]. Furthermore, low-risk patients (no genetic abnormalities, tumour stage 1–3 or 4S) who are primarily treated in the observation group may experience progression to localized or disseminated disease and thus require adapted treatment. Early identification of this subgroup of patients might allow a reduction in intensity of therapy, and therefore of toxicity and complications, by enabling early chemotherapy or primary resection of initially small tumours. However, this would require accurate risk stratification.

MIBG scintigraphy with single-photon emission computed tomography (SPECT) is an integral part of the diagnostic work-up regarding staging and restaging in NB. Furthermore, the Curie and SIOPEN scores which represent the number of body segments affected by MIBG-avid skeletal metastases in patients with stage 4 disease correlate with EFS and OS [8, 9]. Nevertheless, there is currently no quantitative parameter derived from MIBG SPECT that can predict a patient's outcome independently of the tumour stage. Among other parameters, the metabolic tumour volume (MTV) obtained by positron emission tomography (PET) has shown prognostic value in different entities [10–12]. Furthermore, the asphericity (ASP) of the distribution of metabolic activity within the tumour can be quantified and serves as a surrogate marker for the aggressiveness of the tumour biology. It can predict progression-free survival and OS [13, 14]. Recently, the concept of ASP has been successfully transferred to SPECT imaging [15].

Therefore, the aim of the current study was to evaluate the prognostic significance of quantitative measures derived from pretherapeutic MIBG SPECT images as well as laboratory parameters with respect to EFS. The goal was to identify prognostic parameters that might help differentiate and improve the existing risk assessment and identify patients with a high risk of impaired EFS despite having already received risk-adapted treatment. Finally, multifactorial predictive models were calculated to utilize the additional and complementary value of different parameters and their combinations for individualized risk assessment.

## Materials and methods

### Patients

This retrospective study included 23 consecutively examined children (11 girls, 12 boys; age median 1.8 years, range 0.3 to 6.8 years) newly diagnosed with NB and referred to the two participating departments (Charité – Universitätsmedizin Berlin, Germany, and University Medicine Greifswald, Germany) for pretherapeutic  $^{123}\text{I}$ -MIBG scintigraphy/SPECT between 2008 and 2016. All patients fulfilled the following criteria: the scintigraphic examination included SPECT or SPECT/CT, the primary tumour was MIBG-avid, and MIBG SPECT was performed before initiation of any antineoplastic treatment (18 patients) or the latest 1 week after the start of chemotherapy (5 patients). All patients were part of the German therapy optimization trial NB2004; i.e. risk stratification, treatment and follow-up visits were performed according to the NB2004 protocol [3]. An event was defined as disease progression, relapse or death from any cause.

### Laboratory parameters and genetic analysis

Laboratory parameters were obtained at the time of diagnosis. Urinary catecholamine metabolites were extracted from samples of spontaneously voided urine and levels of excreted vanillylmandelic acid (VMA) and homovanillic acid (HVA) were calculated relative to that of creatinine. Normal values for the VMA-to-creatinine ratio (VMA/C) were 95  $\mu\text{mol/g}$  in infants and declined to 42  $\mu\text{mol/g}$  in children >4 years of age. Maximum normal values for the HVA-to-creatinine ratio (HVA/C) also declined with increasing age and ranged from 70 to 179  $\mu\text{mol/g}$ . Normal maximum values for serum neuron-specific enolase (NSE) ranged from 20 to 25  $\mu\text{g/l}$  while the upper limits for serum lactate dehydrogenase (LDH) declined from 600 U/l in newborns to 300 U/l in children between 2 and 15 years of age. MYCN amplification status as well as chromosomal aberrations in 1p36 were assessed by fluorescence in situ hybridization (FISH).

### $^{123}\text{I}$ -MIBG SPECT acquisition

The activity of  $^{123}\text{I}$ -MIBG administered was weight-adapted based on the recommendations of the European Association of Nuclear Medicine (EANM) [16]. SPECT imaging was performed 24 h after tracer injection in addition to planar whole-body projections. In 9 of 23 patients (2008 to 2011), MIBG SPECT imaging was performed without attenuation correction using a GE Millennium VG Hawkeye SPECT/CT system (GE Medical Systems, Milwaukee, WI, USA) equipped with a low-energy high-resolution (LEHR) collimator and iterative reconstruction (2D OSEM algorithm, matrix  $128 \times 128 \times 128$ ,

voxel size  $4.42 \times 4.42 \times 4.42$  mm). In 13 of 23 patients (2012 to 2016), MIBG SPECT was performed using a Siemens Symbia T6 SPECT/CT (Siemens Healthcare, Erlangen, Germany). Images were acquired with a LEHR collimator and reconstructed iteratively (Flash 3D OSEM algorithm, matrix  $128 \times 128 \times 81$ , voxel size  $4.8 \times 4.8 \times 4.8$  mm). In 10 of these 13 patients, combined SPECT/low-dose CT (X-ray tube current 24 mA, voltage 130 kV, slice thickness 1.25 to 5 mm, convolution kernel B08s) was performed instead of SPECT alone, and CT-based attenuation correction was included in image reconstruction.

In one patient, SPECT images were acquired with a GE DST XL dual-head camera equipped with a LEHR collimator (GE Medical Systems, Milwaukee, WI, USA). Reconstruction was performed iteratively (2D OSEM algorithm, matrix  $128 \times 128 \times 128$ , voxel size  $3.95 \times 3.95 \times 3.95$  mm).

### <sup>123</sup>I-MIBG SPECT quantification

SPECT data were quantified with dedicated software (ROVER, version 3.0.23; ABX advanced biochemical compounds GmbH, Radeberg, Germany). The primary tumour was delineated using a semiautomatic, background-adapted algorithm [17], and subsequent manual correction to exclude nonmalignant tissue if required. The maximum count rate ( $T_{\max}$ ), average count rate ( $T_{\text{mean}}$ ), MTV and ASP of the primary tumour were calculated. ASP was defined as the relative deviation of the MTV surface from that of an isovolumetric sphere (as described by Apostolova et al. [14]). The average hepatic count rate ( $L_{\text{mean}}$ ) was defined based on a 30-ml sphere placed in physiological liver tissue to calculate tumour-to-liver count rate ratios (TLCRR;  $T_{\max}/L_{\text{mean}}$ ,  $T_{\text{mean}}/L_{\text{mean}}$ ).

### Statistical analysis

Statistical analysis was performed using SPSS 22 (IBM Corp., Armonk, NY, USA) and R 3.3.2 (Foundation for Statistical Computing, Vienna, Austria, 2016; <http://www.R-project.org>). Descriptive parameters are expressed as medians, interquartile range (IQR) and range, unless otherwise specified. The associations between the EFS and clinical parameters (sex, age, tumour stage), genetic factors (MYCN amplification status, 1p36), laboratory parameters (VMA/C, HVA/C, NSE, LDH), histology (Hughes grade [18]) and all quantitative SPECT parameters (ASP, MTV,  $T_{\max}/L_{\text{mean}}$ ,  $T_{\text{mean}}/L_{\text{mean}}$ ) were analysed using univariate Cox proportional hazards regression. Additionally, Kaplan-Meier curves with respective log-rank tests were calculated following binarization of metric parameters using cut-off values defined by receiver operating characteristic (ROC) curves (Supplementary material). The optimal cut-off value was defined as the point on

the ROC curve with the shortest distance  $d$  to the point (0,1) calculated as follows:

$$d = \sqrt{(1 - \text{Sensitivity})^2 + (1 - \text{Specificity})^2}$$

All variables with  $p \leq 0.1$  in the univariate Cox regression were also subjected to step-wise inclusion and exclusion into a predictive model for EFS based on their contribution to the likelihood ratio (LR)  $\chi^2$  and Harrell's  $C$  index [19] of the final model. The single parameter with the highest  $\chi^2$  was chosen as the starting parameter. Metric variables were included after binarization as described above. The contribution of single parameters to the overall predictive accuracy of the model was evaluated with the LR  $\chi^2$  test using the *survival* and *rms* package of R. Statistical significance was assumed at a  $p$  value of  $\leq 0.05$ .

A Kaplan-Meier subanalysis separated patients based on the original risk group according to the NB2004 trial. The medium-risk group included one patient only and was therefore not considered in the subgroup analysis. Calculation of  $p$  values was omitted due to the low sample sizes in the subgroups. Cut-off values for metric variables calculated for the total population were retained unchanged for this subanalysis.

## Results

Table 1 provides an overview of all patients. Of the 23 patients, 12 had been assigned to the observation (low-risk) group, 1 patient to the medium-risk group and 10 patients to the high-risk group according to the NB2004 protocol [3]. The median follow-up duration of the survivors was 36.0 months (range 12 to 107 months; IQR 29 to 65 months). Events occurred in 8 of the 23 patients (34.8%; progression in 4; relapse in 3; death without progression/relapse in 1), and 4 patients (17.4%) died during follow-up. All patients in the medium-risk and high-risk groups showed a partial response (10 patients) or a complete response (1 patient) to induction chemotherapy. Both patients in the observation group treated primarily with short-term chemotherapy (patients 12 and 18) also showed a partial response (Table 1).

### Prognostic relevance of single parameters

In the univariate Cox regression analysis (Table 2), ASP ( $p = 0.029$ ; hazard ratio, HR, 1.032 for a one-unit increase; 95% CI 1.0 to 1.06), MTV ( $p = 0.038$ ; HR 1.012; 95% CI 1.0 to 1.024) and MYCN amplification status ( $p = 0.047$ ; HR 4.7; 95% CI 1.02 to 21.4) were significant predictors of EFS. Patient age ( $p = 0.09$ ; HR 0.6; 95% CI 0.3 to 1.1), chromosome 1p36 status ( $p = 0.09$ ; HR 4.1; 95% CI 0.8 to 21.4) and

**Table 1** Overview of all patients with all event data

Patient no.	Age (years)	Stage	MYCN amplification status	1p36	Risk group <sup>b</sup>	Curie score <sup>c</sup>	Response <sup>d</sup>	HVA/C (μmol/g)	HVA/C high	ASP (%)	ASP high	MTV (ml)	MTV high	Metabolic risk factors <sup>e</sup>	Event	Event-free survival (months)	Therapy after event
5	0.7	4S	-	-	1	-	-	174	-	21.6	-	40.8	-	0	-	96.5	
9	0.5	4S	-	+	1	-	-	201	+	4.1	-	19.2	-	1	Progression (M)	8.7	Surgery, chemotherapy, ASCT, mAb
18	0.4	4S	-	n.a.	1	-	Partial response	459	+	20.3	-	16.5	-	1	Progression (L)	20.7	Surgery
3	0.3	1	n.a.	-	1	-	-	712	+	8.7	-	47.0	+	1	Progression (L)	1.5	Chemotherapy
4	2.0	1	-	-	1	-	-	170	-	8.0	-	4.7	-	0	-	60.8	
17	1.1	1	-	-	1	-	-	159	-	4.6	-	35.2	-	0	-	28.7	
2	1.8	1	-	-	1	-	-	213	+	11.1	-	14.4	-	1	-	103.0	
15	1.9	2	-	-	1	-	-	188	+	86.8	+	58.4	+	2	Progression (L)	6.7	Surgery
23	0.7	2	-	n.a.	1	-	-	173	-	0.0	-	9.7	-	0	-	11.6	
12	0.5	3	-	-	1	-	Partial response	106	-	10.5	-	51.2	+	0	-	33.9	
10	1.9	3	-	-	1	-	-	378	+	13.7	-	26.4	-	1	-	36.0	
8	1.6	3	-	-	1	-	-	635	+	28.9	-	46.3	-	1	-	36.0	
20	5.8	3	-	-	2	-	Partial response	126	-	64.7	+	32.5	-	1	-	27.9	
11	3.1	4	-	+	3	1	Partial response	217	+	75.7	+	40.7	-	2	Relapse (M)	14.8	Chemotherapy, mAb, radiation therapy, surgery
14	1.7	4	-	+	3	13	Partial response	684	+	34.9	+	76.0	+	2	-	41.8	
21	2.8	4	-	-	3	14	Partial response	687	+	28.7	-	6.5	-	1	-	26.2	
1	5.1	4	-	+	3	13	Partial response	133	-	17.6	-	16.0	-	0	-	106.6	
19	4.7	4	-	n.a.	3	5	Partial response	229	+	15.2	-	21.5	-	1	-	28.6	
22	5.1	4	-	-	3	12	Partial response	355	+	30.3	-	82.1	+	1	-	20.8	
16	0.9	3	+	+	3	-	Partial response	366	+	50.6	+	106.1	+	2	Death from any cause	7.2	none
7	6.8	4	+	+	3	9	Partial response	133	-	8.5	-	41.6	-	0	-	65.1	
6	2.3	4	+	+	3	2	Partial response	904	+	62.9	+	88.9	+	2	Relapse (O)	20.4	Chemotherapy
13	1.2	4	+	+	3	5	Complete response	1,428	+	34.3	+	204.9	+	2	Relapse (O)	13.3	Chemotherapy

n.a. not available, (L) localized disease after the event, (M) multifocal/disseminated metastatic disease (Curie score >2), (O) oligometastatic disease (Curie score ≤2), ASCT autologous stem-cell transplantation, mAb monoclonal antibodies

<sup>a</sup> Patients numbered based on the date of <sup>123</sup>I-MIBG imaging and sorted by risk group according to the NB2004 protocol [3] and MYCN amplification status

<sup>b</sup> According to the NB2004 protocol

<sup>c</sup> At initial diagnosis [20]

<sup>d</sup> Response to induction chemotherapy (if performed)

<sup>e</sup> ASP, HVA/C

**Table 2** Univariate Cox regression analysis for all examined parameters

Variable	Hazard ratio <sup>a</sup>	95% confidence interval	<i>p</i> value
ASP	1.032	1.0–1.06	<b>0.029</b>
MTV	1.012	1.0–1.024	<b>0.038</b>
$T_{\max}/L_{\text{mean}}$	0.82	0.6–1.2	0.31
$T_{\text{mean}}/L_{\text{mean}}$	0.62	0.3–1.3	0.19
Stage			0.57
Stage (4 vs. <4)	0.78	0.2–3.3	0.73
Age	0.56	0.29–1.1	0.09
Sex (female)	0.98	0.2–3.9	0.98
Hughes grade			0.99
MYCN amplification status	4.7	1.02–21.4	<b>0.047</b>
1p36 status	4.1	0.79–21.4	0.09
VMA/C	1.0	1.0–1.0	0.67
HVA/C	1.002	1.0–1.003	0.071
NSE	1.0	1.0–1.0	0.6
LDH	1.0	1.0–1.0	0.14

Values in bold type are statistically significant ( $p < 0.05$ )

<sup>a</sup> Per one-unit increase in metric variables

elevated HVA/C ( $p = 0.071$ ; HR 1.002; 95% CI 1.0 to 1.0) showed a tendency to significance. All other parameters were not predictive of EFS.

ROC analysis for metric parameters identified ASP >32%, MTV >46.7 ml and HVA/C >181  $\mu\text{mol/g}$  as optimal cut-off values (Supplementary material). According to Kaplan-Meier analysis, the mean EFS in patients with high ASP (>32.0%) and low ASP were, respectively, 20.8 months (95% CI 10.5 to 31.2 months) and 88.1 months (95% CI 69.3 to 107.0 months; log-rank test,  $p = 0.013$ ; Figs. 1 and 2). The mean EFS in patients with high MTV (>46.7 ml) and low MTV were, respectively, 21.8 months (95% CI 10.5 to 33.1 months) and 87.3 months (95% CI 67.9 to 106.7 months;  $p = 0.023$ ). The mean EFS in patients with and without MYCN amplification were, respectively 26.5 months (95% CI 4.2 to 48.8 months) and 85.1 months (95% CI 66.5 to 103.6 months;  $p = 0.029$ ).

### Predictive models

The optimal model for predicting the EFS combined ASP and HVA/C and yielded a  $\chi^2$  of 11.5 and Harrell's *C* of 0.8. Only slightly higher predictive accuracies were obtained when MTV or MYCN amplification status was added to the model ( $\chi^2$  11.7, *C* 0.82, LR  $\chi^2$  test  $p = 0.64$ , for MTV, and  $\chi^2$  13.0, *C* 0.85,  $p = 0.47$ , for MYCN). In contrast, a slightly but significantly lower predictive value was obtained when MTV and HVA/C were combined in one model ( $\chi^2$  10.6, *C* 0.81,  $p < 0.001$ ).

Based on the predictive model including ASP and HVA/C, none of the seven patients without metabolic risk factors

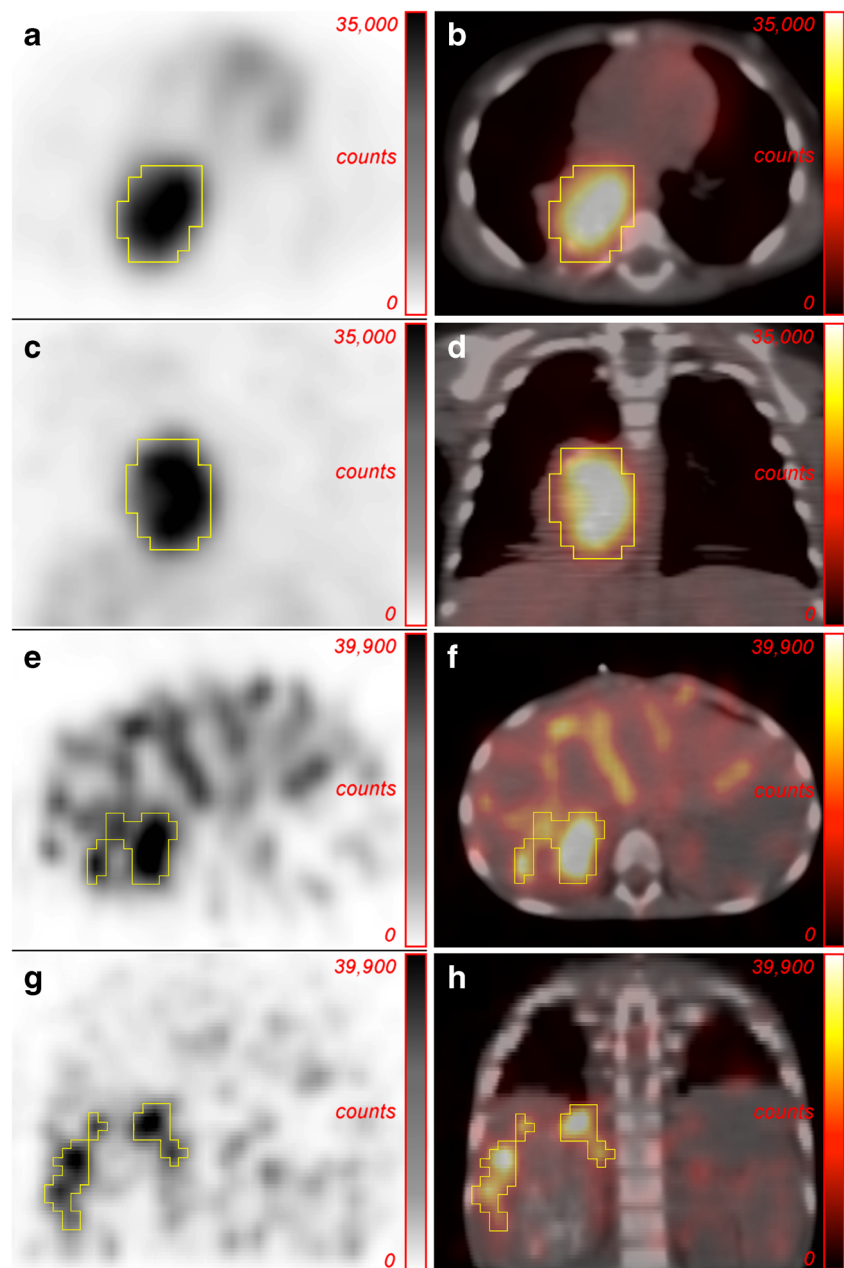
(mRF) suffered an event while three of ten patients (30%) with one mRF and five of six patients (83.3%) with two mRF suffered an event (log-rank test  $p = 0.005$ ; Fig. 3). In the risk model including MTV and HVA/C, none of the seven patients without mRF suffered an event, three of nine patients (33.3%) with one mRF and five of seven patients (71.4%) with two mRF suffered an event (log-rank test  $p = 0.01$ ).

### Subanalysis of different risk groups according to NB2004

In patients in both the observation group and the high-risk group (according to NB2004 [3]), the mRF ASP, MTV and HVA/C as well as the predictive models were able to separate patients with a high risk from those with a low risk of progression/relapse (Fig. 4).

Patients in the observation group received treatment (chemotherapy) only if organ-threatening tumour growth or tumour progression occurred. Seven of ten patients in the observation group had at least one mRF (ASP HVA/C model). An event did occur in four of these seven patients (57.1%) and had to be treated subsequently (Table 1). Due to organ-threatening tumour growth, one of these four patients (patient 18) underwent limited chemotherapy at initial diagnosis, but experienced disease progression during follow-up necessitating tumour resection. None of the three patients with one mRF but a favourable outcome had been treated initially. No event occurred in any of the five patients in the observation group without a mRF. All patients in the high-risk group were treated with multidrug chemotherapy including high-dose chemotherapy followed by autologous stem-cell transplantation.

**Fig. 1** Imaging in two example patients with high and low ASP: transverse (a, e) and coronal (c, g) MIBG SPECT images and transverse (b, f) and coronal (d, h) fused SPECT/CT images in the two patients with the MTV delineated (yellow line). **a–d** A 1-year-old boy (patient 17) with NB of the right thoracic sympathetic trunk treated in observation group (tumour stage 1). ASP was very low (5%). No event occurred within 29 months of follow-up. **e–h** A 3-year-old girl in the high-risk group (patient 11) with NB originating from the right adrenal gland (stage 4, MYCN amplified). ASP was high (76%), mostly due to extensive central necrosis (red arrow). Relapse occurred after 15 months, and the patient died after 36 months

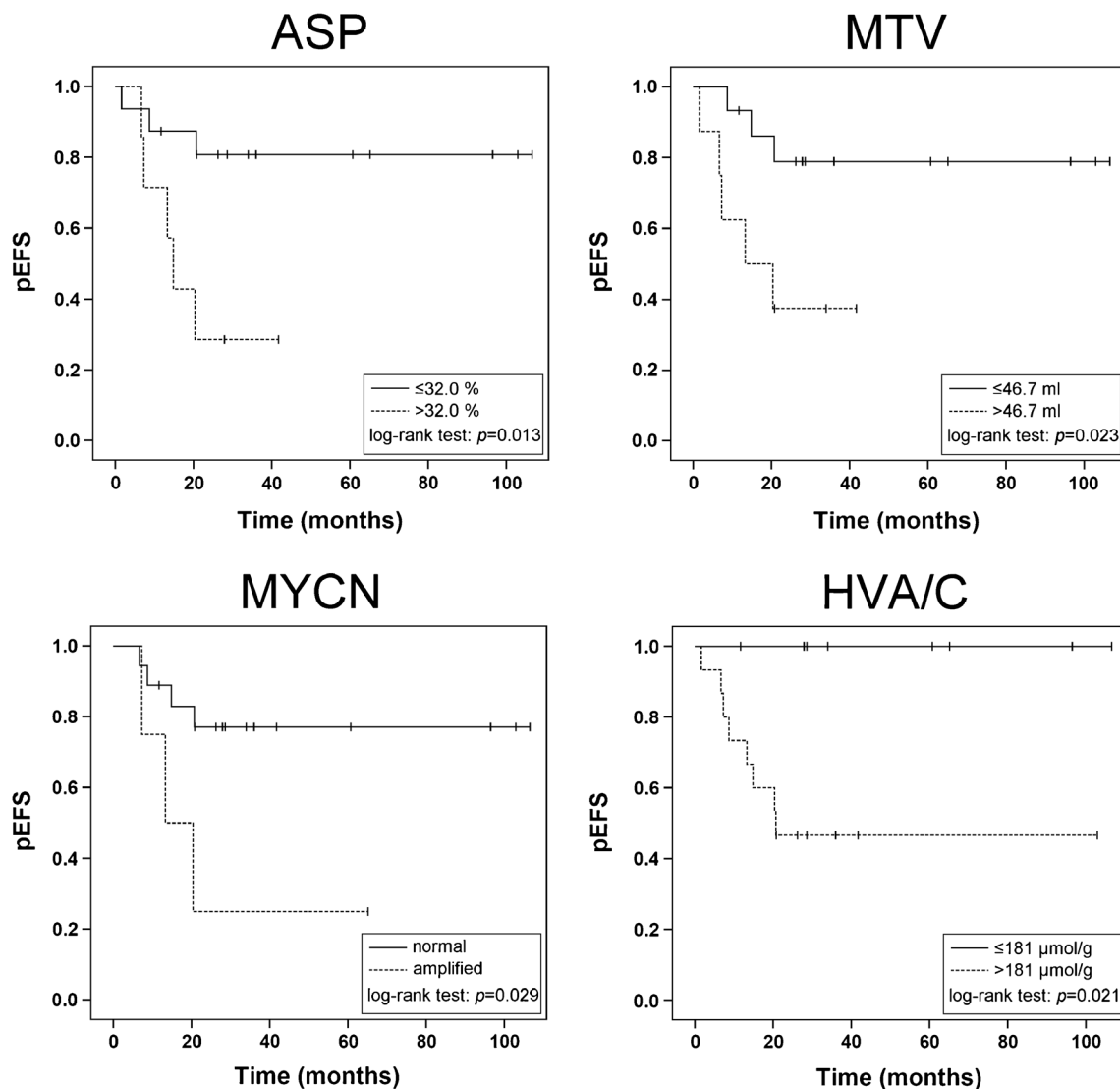


Five of ten patients in the high-risk group had both mRF, and an event occurred subsequently in four of these patients (80%) (progression/relapse). No event occurred in any of the five patients with only one or no mRF.

## Discussion

This retrospective, explorative study evaluated the potential of MIBG SPECT-derived, clinical, genetic and laboratory parameters for individualized prediction of the risk of disease progression/relapse in 23 children with NB.

To the best of our knowledge, this is the first study demonstrating evidence that quantitative parameters obtained from pretherapeutic  $^{123}\text{I}$ -MIBG SPECT can help stratify patients with NB according to their risk of disease progression/relapse under current treatment regimens. ASP of the primary tumour was a significant predictor of EFS in the univariate Cox regression analysis. The HR of 1.032 per one-unit increase represents a 15-fold increased risk of progression/relapse in the patient with the highest measured ASP (87%) compared to the patient with the lowest ASP (0%). High ASP is regarded as a surrogate marker for aggressive tumour biology as it results from either intratumoral parts with low tracer avidity



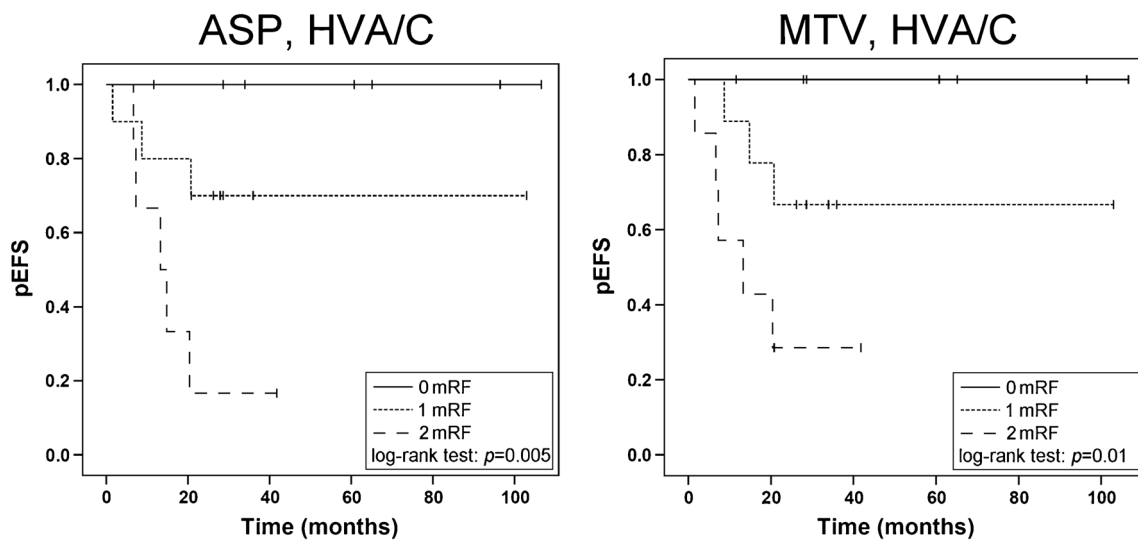
**Fig. 2** Kaplan-Meier curves with respective log-rank tests for EFS for selected parameters with  $p \leq 0.1$  in the univariate Cox regression ( $pEFS$  predicted EFS)

(e.g. low tumour perfusion with central necrosis; Fig. 1) or a heterogeneous, spiculated outer surface of the MTV. In patients with stage 4 disease, metastases were not delineated to ensure a uniform methodology throughout the cohort. Applying ASP for the first time in SPECT imaging, Wetz et al. recently showed that high ASP in somatostatin receptor imaging significantly predicts failure to respond to peptide receptor radionuclide therapy in patients with metastatic gastroenteropancreatic neuroendocrine neoplasms [15].

Also, the MTV of the primary tumour itself significantly predicted EFS. Notably, the size of the MTV and its ASP are independent parameters in principle because the ASP of the MTV is calculated after correction for its volume (according to the equation used to define ASP by Apostolova et al. [14]). Nevertheless, as well as large size in itself being a risk factor, large tumours are more likely to show higher ASP and

therefore less favourable tumour biology. The TLCRR did not predict EFS in our cohort. Fendler et al. found that a high TLCRR is helpful in differentiating neuroblastic tumours with an unfavourable histopathology (i.e. NB) from ganglioneuroma or ganglioneuroblastoma [21]. However, no dedicated evaluation of its prognostic value in NB has been performed.

In addition to these SPECT-derived parameters, common clinical and genetic risk factors in NB patients were evaluated. However, these established and proven prognostic factors were only included for the purpose of comparison; the aim was not to verify whether or not the factors are relevant. In this small study, although tumour stage was not predictive of EFS, patient age showed a tendency towards significance. It should be noted that all patients underwent risk-adapted treatment which aims to account for these two variables. Moreover, neither of these parameters was found to be a stand-alone



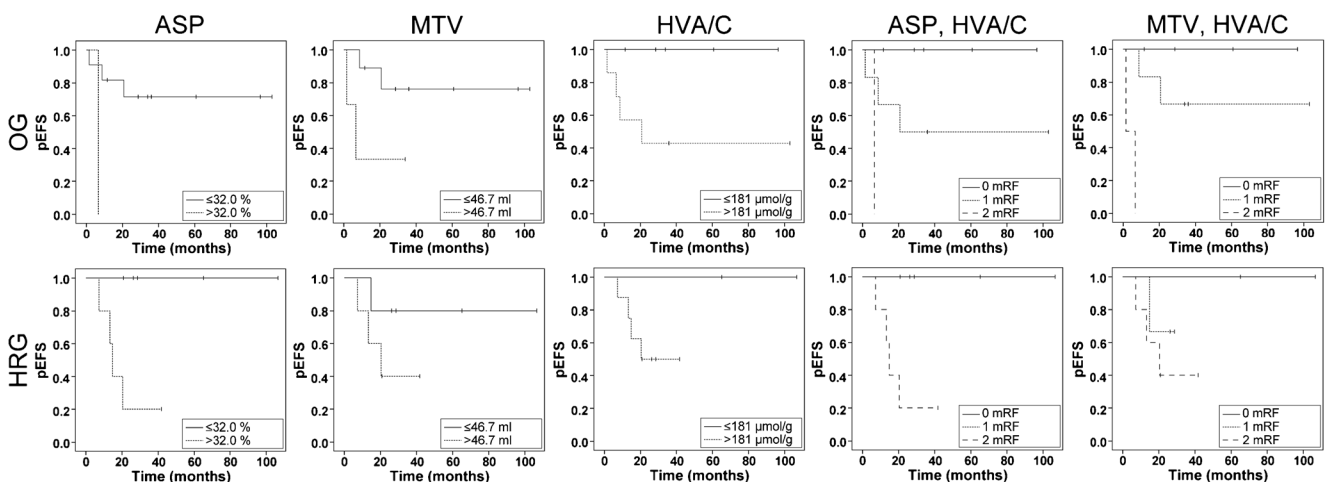
**Fig. 3** Kaplan-Meier curves with log-rank tests for EFS in patients with zero, one and two metabolic risk factors (*mRF*) according to the predictive models including ASP and HVA/C or MTV and HVA/C. The log-rank test was performed for the overall model (*pEFS* predicted EFS)

prognostic factor in the current risk assessment [3]. In contrast, the MYCN amplification status – the most important single conventional prognostic parameter – predicted EFS in the univariate Cox regression analysis even in patients under risk-adapted treatment.

A higher HVA/C ratio tended to predict EFS. The urinary VMA/HVA ratio has been studied more extensively. However, it was prognostic only in a subgroup of patients with localized disease and non-amplified MYCN [22, 23]. The current results on HVA/C are supported by the findings of two previous studies. Strenger et al. showed in 114 children with NB that an elevated urinary HVA (higher than 2.5 times the mean of healthy patients) predicted shorter OS ( $p = 0.015$ ) [24]. Zambrano and Reyes-Múgica found in 33 patients with NB significantly higher urinary HVA excretion in patients with stage 4 than in patients with stage 1–3 tumours [25]. A meta-analysis showed that urinary levels of VMA or HVA

alone – in contrast to serum LDH or NSE – are not significant predictors, although some of the available data were insufficient for analysis. The authors concluded that interstudy comparison is limited due to a lack of uniform methodology in data acquisition (study sample composition, laboratory sample collection and reporting, definition of “elevated levels”, and statistical parameters) [26]. Thus, the universal usage of any laboratory parameter for risk stratification would first require methodological standardization. In the current cohort, the predictive significance of all laboratory markers and their optimal cut-off values were calculated independently of patient age and age-dependent ranges of normal values. Normal values are only provided for interstudy comparison.

As no single parameter will be sufficient to predict the outcome in any individual patient, multifactorial risk models were designed and calculated. Both final models included the HVA/C ratio and one SPECT-derived parameter (ASP or



**Fig. 4** Kaplan-Meier curves for selected single parameters and the predictive models in patients in the observation group (*OG*) and the high-risk group (*HRG*) (*pEFS* predicted EFS)



MTV) and showed only slightly different predictive accuracies. Most importantly, both predictive models allowed the identification of a subgroup of patients in whom no events occurred. These patients would be ideal candidates for treatment de-escalation. In contrast, patients with two combined mRF were very likely to experience disease progression or relapse and could benefit from intensified treatment. Further characterization of patients with intermediate risk (one mRF) is pending – although even the presence of only one mRF in the patients in the observation group substantially increased the risk of an unfavourable EFS. Calculation of these risk models also underlined that the optimal combination was not necessarily defined by the two parameters with the highest single predictive accuracy but by those parameters that provided the highest supplementary value and thus the best combined accuracy (assessed in terms of Harrell's *C* and the LR  $\chi^2$  value).

The implications of improved initial risk stratification are substantially different between low-risk (observation group) and high-risk patients. While OS in patients with low-risk disease is very good, disease progression may still occur necessitating intensive treatment analogous to that in high-risk patients at worst [3] – which was true for patient 9 in our cohort (Table 1). Early identification of patients with a substantial risk of tumour progression may allow less intensive treatment compared to the situation when tumour progression has already occurred. On the other hand, identification of a subgroup of high-risk patients with a better prognosis than “ultra-high-risk” patients might allow a reduction of therapy intensity in these patients in the future. Kaplan-Meier analysis demonstrated that patients with a low risk and a high risk of tumour progression/relapse can be differentiated among those current “low-risk” and “high-risk” patients using ASP, MTV, HVA/C and the combined risk models (Fig. 4). However, due to the small number of patients, these first results must be interpreted with caution and require validation in a larger cohort. Moreover, these results do not imply that mRF should supplant any of the existing risk factors; as stated above, the aim was to identify parameters derived from previously potentially under-represented modalities to further differentiate the risk assessment in this heterogeneous tumour entity.

The limited sample size was the major limitation of the current study and may have resulted in either underestimation of the predictive relevance of some of the parameters examined (beta error, large 95% CI) or premature confirmation of the predictive value (alpha error) due to a selection bias. However, the latter was minimized by including all consecutive patients treated at both institutions who met basic inclusion criteria. In addition, the medium-risk group according to NB2004 [3] comprised only one patient, which must be taken into consideration, as no conclusions in this risk group can be drawn. Three different SPECT devices were used, and attenuation correction (CT-based) was performed in only ten patients.

However, post-hoc intraindividual comparison of SPECT parameters obtained from noncorrected and corrected data did not yield substantially different results in these patients.

## Conclusions

In order to improve the outcome of patients with NB a more sophisticated, individualized risk stratification is required. In this explorative study, imaging parameters relating to tumoral metabolic activity, ASP and MTV, derived from pretherapeutic  $^{123}\text{I}$ -MIBG SPECT in particular in combination with the level of urinary HVA excretion allowed identification of children with a high risk and a low risk of progression/relapse treated according to current therapy protocols. These results, however, need validation in a considerably larger patient cohort in a multicentre setting to verify the correct selection of predictive parameters as well as the optimal cut-off values.

**Acknowledgments** The authors thank Dr. Ingo Steffen for his advice on statistical methodology and Mr. René Höhne for his support in obtaining patient data.

## Compliance with ethical standards

**Conflicts of interest** None.

**Ethical approval** All procedures were in accordance with the ethical standards of the respective institutional ethics commissions (vote, EA2/150/16; MV05/11) and with the principles of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

## References

1. Kaatsch P. Epidemiology of childhood cancer. *Cancer Treat Rev*. 2010;36:277–85.
2. Mueller S, Matthay KK. Neuroblastoma: biology and staging. *Curr Oncol Rep*. 2009;11:431–8.
3. Society for Paediatric Oncology and Haematology. NB2004 Trial Protocol for Risk Adapted Treatment of Children with Neuroblastoma. Version 1.00. Berlin: Society for Paediatric Oncology and Haematology; 2004.
4. Matthay KK, Reynolds CP, Seeger RC, Shimada H, Adkins ES, Haas-Kogan D, et al. Long-term results for children with high-risk neuroblastoma treated on a randomized trial of myeloablative therapy followed by 13-cis-retinoic acid: a Children's Oncology Group study. *J Clin Oncol*. 2009;27:1007–13.
5. Maris JM. Recent advances in neuroblastoma. *N Engl J Med*. 2010;362:2202–11.
6. Trahair TN, Vowels MR, Johnston K, Cohn RJ, Russell SJ, Neville KA, et al. Long-term outcomes in children with high-risk neuroblastoma treated with autologous stem cell transplantation. *Bone Marrow Transplant*. 2007;40:741–6.
7. Applebaum MA, Henderson TO, Lee SM, Pinto N, Volchenboum SL, Cohn SL. Second malignancies in patients with neuroblastoma:

- the effects of risk-based therapy. *Pediatr Blood Cancer*. 2015;62:128–33.
8. Decarolis B, Schneider C, Hero B, Simon T, Volland R, Roels F, et al. Iodine-123 metaiodobenzylguanidine scintigraphy scoring allows prediction of outcome in patients with stage 4 neuroblastoma: results of the Cologne Interscore Comparison study. *J Clin Oncol*. 2013;31:944–51.
  9. Yanik GA, Parisi MT, Shulkin BL, Naranjo A, Kreissman SG, London WB, et al. Semiquantitative mIBG scoring as a prognostic indicator in patients with stage 4 neuroblastoma: a report from the Children's Oncology Group. *J Nucl Med*. 2013;54:541–8.
  10. Li YJ, Dai YL, Cheng YS, Zhang WB, Tu CQ. Positron emission tomography (18)F-fluorodeoxyglucose uptake and prognosis in patients with bone and soft tissue sarcoma: a meta-analysis. *Eur J Surg Oncol*. 2016;42:1103–14.
  11. Marinelli B, Espinet-Col C, Ulaner GA, McArthur HL, Gonen M, Jochelson M, et al. Prognostic value of FDG PET/CT-based metabolic tumor volumes in metastatic triple negative breast cancer patients. *Am J Nucl Med Mol Imaging*. 2016;24:120–7.
  12. Tamandl D, Ta J, Schmid R, Preusser M, Paireder M, Schoppmann SF, et al. Prognostic value of volumetric PET parameters in unresectable and metastatic esophageal cancer. *Eur J Radiol*. 2016;85:540–5.
  13. Apostolova I, Steffen IG, Wedel F, Lougovski A, Mamitz S, Derlin T, et al. Asphericity of pretherapeutic tumour FDG uptake provides independent prognostic value in head-and-neck cancer. *Eur Radiol*. 2014;24:2077–87.
  14. Apostolova I, Rogasch J, Buchert R, Wertzel H, Achenbach HJ, Schreiber J, et al. Quantitative assessment of the asphericity of pretherapeutic FDG uptake as an independent predictor of outcome in NSCLC. *BMC Cancer*. 2014;14:896. doi:10.1186/1471-2407-14-896.
  15. Wetz C, Apostolova I, Steffen IG, Hofheinz F, Furth C, Kupitz D, et al. Predictive value of asphericity in pretherapeutic [<sup>111</sup>In]DTPA-octreotide SPECT/CT for response to peptide receptor radionuclide therapy with [<sup>177</sup>Lu]DOTATATE. *Mol Imaging Biol*. 2017;19:437–45.
  16. European Association of Nuclear Medicine. Dosage card (Version 5.7.2016). Vienna: European Association of Nuclear Medicine; 2016.
  17. Hofheinz F, Pöttsch C, Oehme L, Beuthien-Baumann B, Steinbach J, Kotzerke J, et al. Automatic volume delineation in oncological PET. Evaluation of a dedicated software tool and comparison with manual delineation in clinical data sets. *Nuklearmedizin*. 2012;51:9–16.
  18. Hughes M, Marsden HB, Palmer MK. Histologic patterns of neuroblastoma related to prognosis and clinical staging. *Cancer*. 1974;34:1706–11.
  19. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–87.
  20. Mattay KK, Edeline V, Lumbroso J, Tanguy ML, Asselain B, Zucker JM, et al. Correlation of early metastatic response by <sup>123</sup>I-metaiodobenzylguanidine scintigraphy with overall response and event-free survival in stage IV neuroblastoma. *J Clin Oncol*. 2003;21:2486–91.
  21. Fendler WP, Wenter V, Thornton HI, Ilhan H, von Schweinitz D, Coppentrath E, et al. Combined scintigraphy and tumor marker analysis predicts unfavorable histopathology of neuroblastic tumors with high accuracy. *PLoS One*. 2015;10:e0132809.
  22. Berthold F, Hunnemann DH, Harms D, Käser H, Zieschang J. Serum vanillylmandelic acid/homovanillic acid contributes to prognosis estimation in patients with localised but not with metastatic neuroblastoma. *Eur J Cancer*. 1992;28A:1950–4.
  23. Cangemi G, Reggiardo G, Barco S, Barbagallo L, Conte M, D'Angelo P, et al. Prognostic value of ferritin, neuron-specific enolase, lactate dehydrogenase, and urinary and plasmatic catecholamine metabolites in children with neuroblastoma. *Onco Targets Ther*. 2012;5:417–23.
  24. Strenger V, Kerbl R, Dornbusch HJ, Ladenstein R, Ambros PF, Ambros IM, et al. Diagnostic and prognostic impact of urinary catecholamines in neuroblastoma patients. *Pediatr Blood Cancer*. 2007;48:504–9.
  25. Zambrano E, Reyes-Múgica M. Hormonal activity may predict aggressive behavior in neuroblastoma. *Pediatr Dev Pathol*. 2002;5:190–9.
  26. Riley RD, Heney D, Jones DR, Sutton AJ, Lambert PC, Abrams KR, et al. A systematic review of molecular and biological tumor markers in neuroblastoma. *Clin Cancer Res*. 2004;10:4–12.