VOLUME 24 · NUMBER 31 · NOVEMBER 1 2006

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Customized Oligonucleotide Microarray Gene Expression–Based Classification of Neuroblastoma Patients Outperforms Current Clinical Risk Stratification

André Oberthuer, Frank Berthold, Patrick Warnat, Barbara Hero, Yvonne Kahlert, Rüdiger Spitz, Karen Ernestus, Rainer König, Stefan Haas, Roland Eils, Manfred Schwab, Benedikt Brors, Frank Westermann, and Matthias Fischer

From the Department of Pediatric Oncology and Hematology, Children's Hospital; the Center for Molecular Medicine; Department of Pathology, University of Cologne, Cologne; Departments of Tumor Genetics (B030) and Theoretical Bioinformatics (B080), German Cancer Research Center, Heidelberg; Max-Planck-Institute for Molecular Genetics, Berlin, Germany.

Submitted February 15, 2006; accepted August 31, 2006.

Supported by the Deutsche Krebshilfe (Grant No. 50-2719), the Bundesministerium für Bildung und Forschung through the National Genome Research Network 2 (NGFN2 Grants No. 01GS0456 and 01GR0450), the Competence Network Pediatric Oncology and Hematology, and the Fördergesellschaft Kinderkrebs-Neuroblastom-Forschung e.V.

B.B., F.W., and M.F. contributed equally to this work.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Address reprint requests to André Oberthuer, MD, Children's Hospital, Department of Pediatric Oncology and Hematology, University of Cologne, Kerpener Strasse 62, D-50924 Cologne, Germany; e-mail: andre.oberthuer@uk-koeln.de.

© 2006 by American Society of Clinical Oncology

0732-183X/06/2431-5070/$20.00

DOI: 10.1200/JCO.2006.06.1879

ABSTRACT

Purpose
To develop a gene expression–based classifier for neuroblastoma patients that reliably predicts courses of the disease.

Patients and Methods
Two hundred fifty-one neuroblastoma specimens were analyzed using a customized oligonucleotide microarray comprising 10,163 probes for transcripts with differential expression in clinical subgroups of the disease. Subsequently, the prediction analysis for microarrays (PAM) was applied to a first set of patients with maximally divergent clinical courses (n = 77). The classification accuracy was estimated by a complete 10-times-repeated 10-fold cross validation, and a 144-gene predictor was constructed from this set. This classifier's predictive power was evaluated in an independent second set (n = 174) by comparing results of the gene expression–based classification with those of risk stratification systems of current trials from Germany, Japan, and the United States.

Results
The first set of patients was accurately predicted by PAM (cross-validated accuracy, 99%). Within the second set, the PAM classifier significantly separated cohorts with distinct courses (3-year event-free survival [EFS] 0.86 ± 0.03 [favorable; n = 115] *v* 0.52 ± 0.07 [unfavorable; n = 59] and 3-year overall survival 0.99 ± 0.01 *v* 0.84 ± 0.05; both $P < .0001$) and separated risk groups of current neuroblastoma trials into subgroups with divergent outcome (NB2004: low-risk 3-year EFS 0.86 ± 0.04 *v* 0.25 ± 0.15, $P < .0001$; intermediate-risk 1.00 *v* 0.57 ± 0.19, $P = .018$; high-risk 0.81 ± 0.10 *v* 0.56 ± 0.08, $P = .06$). In a multivariate Cox regression model, the PAM predictor classified patients of the second set more accurately than risk stratification of current trials from Germany, Japan, and the United States ($P < .001$; hazard ratio, 4.756 [95% CI, 2.544 to 8.893]).

Conclusion
Integration of gene expression–based class prediction of neuroblastoma patients may improve risk estimation of current neuroblastoma trials.

J Clin Oncol 24:5070-5078. © 2006 by American Society of Clinical Oncology

INTRODUCTION

Neuroblastoma, a malignant pediatric tumor of the sympathetic nervous system, is the most common solid extracranial malignancy in children younger than 15 years[1] and is characterized by a remarkable heterogeneity of patients' courses. Whereas survival rates of older patients with metastatic disease (stage 4) have remained poor despite extensive efforts, spontaneous regression or maturation of the tumor is frequently found in younger patients with both localized (stage 1 to 3) and disseminated disease (stage 4S)[2] and result in an excellent outcome. Accordingly, therapeutic strategies vary from watch-and-wait approaches to intensive chemotherapeutic regimes. To choose the most appropriate form of first-line treatment, clinical trials stratify patients into groups of low (approximately 50% of patients), intermediate (approximately 10%) and high risk (approximately 40%) based on carefully developed combinations of markers with strong prognostic impact. These include tumor stage,[3] patients' age at diagnosis,[4] genomic amplification of the *MYCN* oncogene (MNA),[5] deletion or imbalance of chromosome 1p (del1p),[6] DNA ploidy[7] and a histopathologic classification proposed by Shimada.[8]

Downloaded from jco.ascopubs.org on April 15, 2008 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

However, despite elaborate risk stratification, current trials still fail to determine the best initial strategy for a substantial number of patients, resulting in over- and undertreatment of a yet unknown percentage. Therefore, other markers such as genetic alterations of the chromosomal regions 3p, 11q,[9,10] and 17q[11] or expression levels of candidate genes (eg, *TrkA*,[12] *FYN*[13]) have been proposed, but none of them is considered to provide enough additional value to be used in clinical practice, leaving an urgent need for novel risk estimation tools.

For several types of cancers, microarray-based gene expression analyses have been proposed to improve outcome prediction.[14,15] Regarding neuroblastoma, three recent studies predicted patients' outcome,[16-18] but either concentrated on small numbers of patients,[16] did not test their classifier in an independent set of patients,[18] or did not compare their classifier's performance with that of present risk stratification.[17] Thus, the clinical value of these classifiers still remains elusive.

In this study, we therefore aimed at constructing a robust gene expression–based classifier that reliably predicts neuroblastoma tumor behavior and may aid physicians in choosing the most appropriate form of first-line treatment.

PATIENTS AND METHODS

Patients

The study comprised 251 patients of the German Neuroblastoma Trials NB90-NB2004, diagnosed between 1989 and 2004. Informed consent was obtained from all patients before this study. Patients' age at diagnosis ranged from 0 to 296 months (median age, 15 months). Median follow-up for patients without fatal events was 4.5 years (range, 0.8 to 15.6 years). Stage was classified according to the International Neuroblastoma Staging System (INSS)[3]; response to treatment was defined according to the revised criteria of the International Neuroblastoma Response Criteria (INRC).[3] Analysis of chromosomal alterations was performed by fluorescence in situ hybridization as described previously[10] and aberrations were defined according to the guidelines of the European Neuroblastoma Quality Assessment Group.[19]

For the construction of a gene expression–based classifier, 77 patients with maximum divergent clinical courses were utilized. This first set comprised all patients who, by the end of February 2004 (beginning of the analysis), had died of disease despite cytotoxic treatment and for whom adequate tumor material was available from the German neuroblastoma tumor tissue bank (n = 1 stage 2, n = 2 stage 3 [n = 2 MNA], n = 19 stage 4 [n = 6 MNA], n = 1 stage 4S). Median time to relapse of disease in this subgroup was 1.3 years (range, 0.02 to 2.7 years) and median overall survival (OS) was 1.7 years (range, 0.02 to 5.84 years). Opposed to these, 54 patients who survived event free more than 1,000 days after diagnosis without treatment were randomly selected from the German neuroblastoma tumor bank (n = 28 stage 1, n = 13 stage 2, n = 1 stage 3, n = 12 stage 4S). Median follow-up time for this subgroup was 6.4 years (range, 2.9 to 11.8 years).

To validate the predictive power of the classifier in an independent set, another 174 samples were analyzed (n = 40 stage 1 (n = 2 MNA), n = 32 stage 2 [n = 1 MNA], n = 36 stage 3 [n = 9 MNA], n = 48 stage 4 [n = 9 MNA], n = 18 stage 4S disease [n = 2 MNA]). Three-year event-free survival (EFS) of this second set was 0.75 ± 0.04 and median duration of follow-up of the patients still alive was 3.8 years (range, 0.84 to 15.6 years).

Sample Preparation

Tumor samples were checked by a pathologist for at least 60% tumor content. Total RNA was isolated from 30 to 60 mg of snap-frozen tissue obtained before cytotoxic treatment using the FastPrep FP120 cell disruptor (Qbiogene-Inc, Carlsbad, CA) and the TRIzol reagent (Invitrogen, Karlsruhe, Germany). RNA integrity was assessed using the 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) considering only samples with an RNA Integrity Number of at least 7.5.

Gene Expression Analyses

Gene expression profiles were generated as dye-flipped dual-color replicates using customized 11K oligonucleotide-microarrays (Appendix, online only). For each sample, 1 μg of linearly amplified Cy3- and Cy5-labeled cRNA, respectively, was hybridized with 1 μg of reverse-color Cy-labeled cRNA of a total RNA pool of 100 neuroblastoma tumor samples using Agilent's Low-RNA-Input-LinearAmp-Kit and InSitu-Hyb-Kit-Plus. After washing and scanning, raw microarray data were preprocessed using software packages from the *R*-project (www.r-project.org)[20] and Bioconductor (www.bioconductor.org).[21] Quality control was performed utilizing the package ArrayMagic.[22] Samples were normalized using the variance stabilization algorithm (vsn)[23] and data from dye-flipped chip pairs were averaged. All raw and normalized microarray data are available at the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress; Accession: E-TABM-38).

Supervised Class Prediction Analysis

A detailed description of the supervised class prediction analysis is given in the Appendix (online only). In brief, the nearest shrunken centroids method (prediction analysis for microarrays [PAM][24]) was applied to the 77 patients with maximum distinct clinical courses (first set) and the classification performance was evaluated by a 10-times-repeated 10-fold cross validation as described previously.[25] As recommended by Li et al,[26] genes frequently selected (≥ 65%) in the classifier training phases of the cross-validation procedure were considered for a final signature and combined with the PAM algorithm to predict the class of the 174 patients of the second set.

Statistical Analysis

Kaplan-Meier estimates for EFS were calculated and compared by log-rank test. Recurrence, progression and death from disease were considered as events. For multivariate analysis, Cox's proportional hazards regression model based on EFS was applied. The factors age (continuous), *MYCN* (normal *v* amplified), tumor stage (1, 2, 3, and 4S *v* 4), status of chromosome 1p (normal *v* deletion/imbalance), Shimada classification (favorable *v* unfavorable), risk stratification by trials from Germany (NB2004), Japan,[27] and the United States (Children's Oncology Group [COG])[28] and the PAM classifier were fitted into a stepwise-backward selection. The likelihood-ratio test *P* value for inclusion was less than .05 and for exclusion more than .10.

RESULTS

Construction and Validation of a Gene Expression–Based Classifier

A comprehensive, neuroblastoma-related oligonucleotide microarray comprising 10,163 oligonucleotide probes (11K) was constructed based on extensive neuroblastoma transcriptome information from different whole-genome expression analyses (Fig 1A). Utilizing this chip, we generated 502 gene expression profiles from 251 neuroblastoma tumors and defined a gene expression–based classifier of patients' courses by applying the PAM algorithm[24] to unfiltered gene expression information of 77 neuroblastoma samples from patients with maximally divergent clinical outcome (Fig 1B). As estimated in a complete, 10-times-repeated 10-fold cross validation, the classification accuracy was high (99%) and comparable to the German neuroblastoma trial NB97 (97%), according to which these patients had been stratified. Prediction accuracies of current risk markers (stage, age, MNA, del1p, and Shimada) for these patients are depicted in Table 1. By considering all genes included in at least 65 of 100 training phases of the cross validation, a predictive signature comprising 144 genes was constructed and combined with the PAM algorithm to a prognostic classifier for neuroblastoma patients.

Downloaded from jco.ascopubs.org on April 15, 2008 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

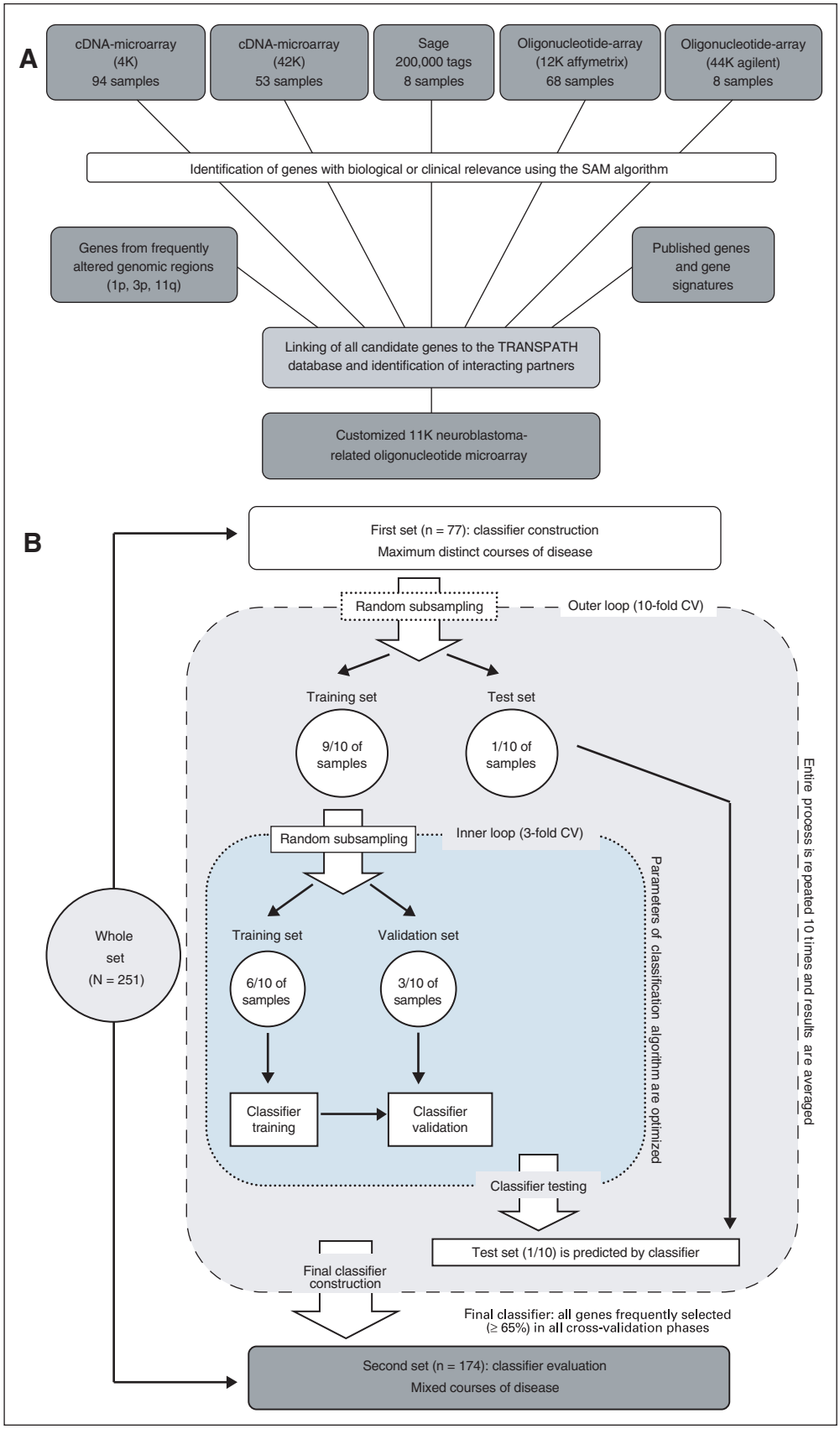

Fig 1. (A) Schematic illustration of the procedure for constructing the neuroblastoma-chip by combining genes with differential expression patterns between clinically relevant neuroblastoma subgroups from whole-genome expression information from multiple platforms (Appendix, online only). (B) Schematic overview of the experimental strategy for the generation and validation of the gene-expression classifier (Appendix). SAGE, serial analysis of gene expression; SAM, significance analysis for microarrays.

Downloaded from jco.ascopubs.org on April 15, 2008 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

Table 1. Classification Parameters for the Training Set

Classifier	No.	%				
		Accuracy	Sensitivity	Specificity	Negative Predictive Value	Positive Predictive Value
PAM (UF *v* F)	77	99	100	98	100	96
Age (< 1 *v* >1 year)	77	43	87	20	81	28
Stage 1, 2, and 4S *v* 3, 4	77	96	91	98	96	92
Stage 1, 2, 3, and 4S *v* 4	77	95	82	100	93	100
MYCN (amplified *v* normal)	77	80	34	100	82	100
Deletion 1p	77	81	57	98	84	93
Shimada (F *v* UF)	67	88	86	89	93	79
NB97 (LR *v* HR)	77	97	91	100	96	100

Abbreviations: PAM, prediction analysis for microarrays; F, favorable; UF, unfavorable; LR, low risk; HR, high risk.

Predictive Power of the Classifier in an Independent Set

The predictive power of this PAM classifier was compared with risk stratification systems of current international neuroblastoma trials[27,28] in an independent set of 174 patients reflecting the full spectrum of the disease's courses. According to the criteria of neuroblastoma trials from Germany (NB2004), the United States (COG),[28] and Japan,[27] 171, 167, and 173 patients could be stratified and all three classification systems separated patient subgroups with different EFS and OS (NB2004: low-risk [n = 100], 3-year EFS 0.80 ± 0.04 and OS 1.00; intermediate-risk [n = 13], 3-year EFS 0.76 ± 0.12 and OS 1.00; high-risk [n = 58], 3-year EFS 0.65 ± 0.07 and OS 0.82 ± 0.05; Figs 2A and 2B; similar results for the COG and Japanese risk stratification are not shown).

In comparison, classification by our PAM predictor also separated these 174 patients into subgroups with divergent outcome (favorable [n = 115], 3-year EFS 0.86 ± 0.03 and OS 0.99 ± 0.01; unfavorable [n = 59], 3-year EFS 0.52 ± 0.07 and OS 0.84 ± 0.05; both $P < .0001$; Figs 2C and 2D), but differed from risk group prediction of the clinical trials in a substantial number of patients (Fig 3). In a subset of 54 of 174 patients of the second set who were characterized by maximally divergent outcome as defined by the criteria for the training set, the PAM predictor achieved a classification accuracy of 93% (50 of 54), which was comparable to the accuracy of other current markers (Table 2).

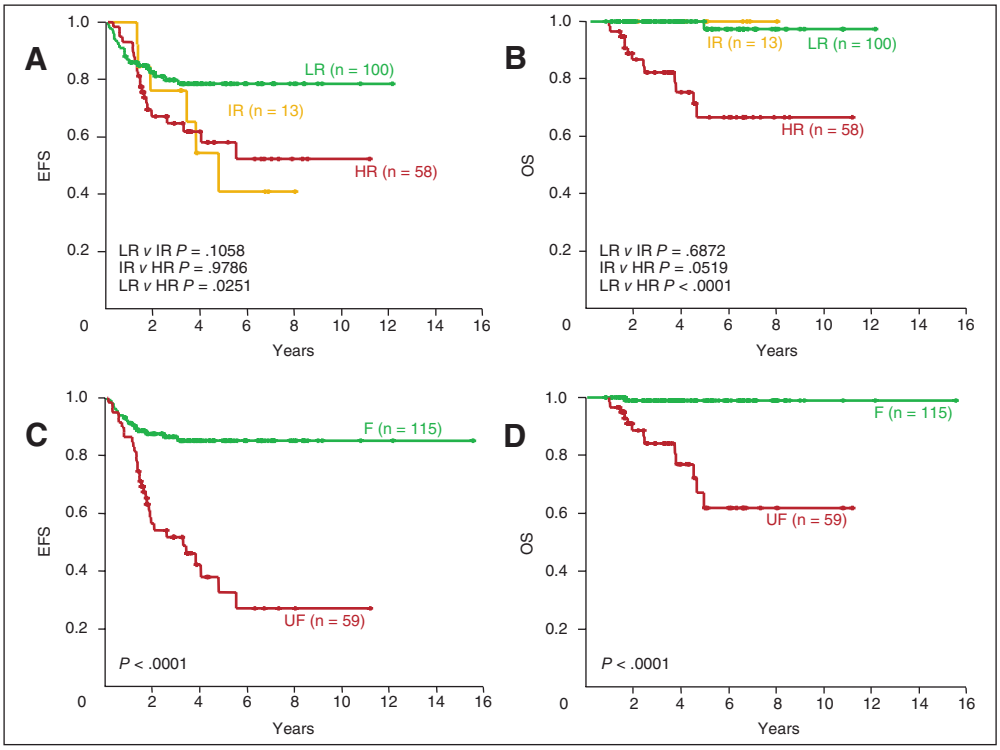

Fig 2. Comparison of neuroblastoma risk stratification according to the German neuroblastoma trial (NB2004) and the 144-gene prediction analysis for microarrays (PAM) classifier. Kaplan-Meier estimates for (A) event-free survival (EFS) and (B) overall survival (OS) for the 171 patients of the second set stratified according to the NB2004 trial. Kaplan-Meier estimates for (C) EFS and (D) OS of the 174 patients of the second set according to classification results of the 144-genes PAM classifier. LR, low risk (n = 100); IR, intermediate risk (n = 13); HR, high-risk (n = 58); F, favorable (n = 115); UF, unfavorable (n = 59).

Downloaded from jco.ascopubs.org on April 15, 2008 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

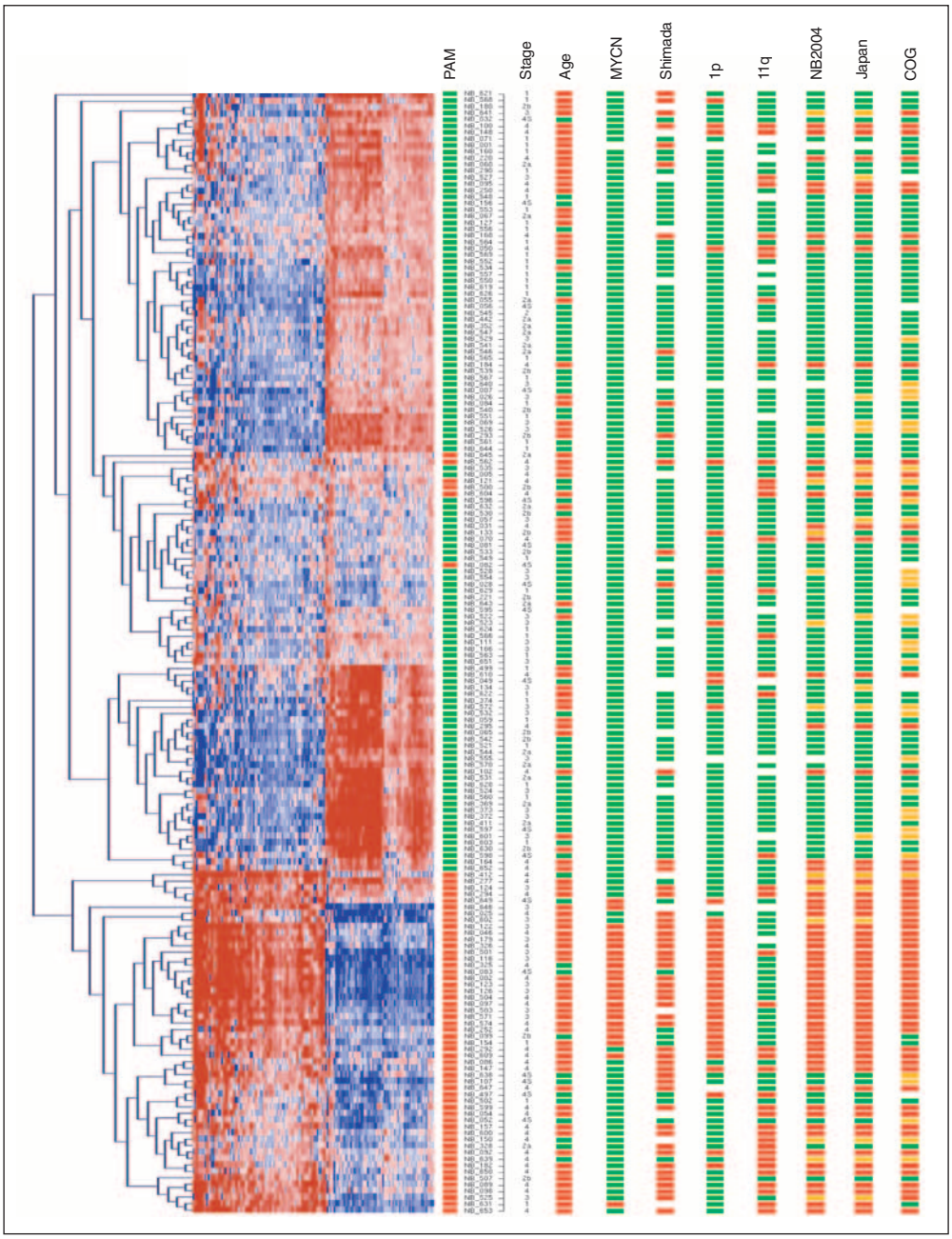

Fig 3. Hierarchical cluster analysis and clinical covariates of 174 patients using gene-expression data of the 144 prediction analysis for microarrays (PAM) signature genes. Lines represent patients, columns represent genes. Gene expression levels are visualized as log-values ranging from blue (+1.0) to red (−1.0). Indicated aside are both the results of the PAM prediction (green, favorable; red, unfavorable) and the clinical covariates for tumor stage, age (green < 1 year; red > 1 year), *MYCN* (green, nonamplified; red, amplified), chromosome 1p status (green, normal; red, deletion/imbalance), chromosome 11q status (green, normal; red, deletion/imbalance) and patients' risk groups as defined by the criteria from Germany, Japan, and the United States (green, low risk; orange, intermediate risk; red, high risk). COG, Children's Oncology Group.

Identification of Patients With Unfavorable Prognoses Among Patients Currently Considered Low Risk

Of 100 patients assigned to the low-risk group by the German risk stratification, PAM predicted 90 as favorable and 10 as unfavorable (3-year EFS, 0.86 ± 0.04 *v* 0.25 ± 0.15; $P < .0001$; Fig 4A). Apart from the significantly diverging EFS of these two subgroups, a remarkable difference in the nature of events was observed between low-risk patients with a favorable and an unfavorable PAM prediction. Whereas six of seven events observed in 10 low-risk patients with an unfavorable PAM classification were metastatic relapses or progressions to stage 4 disease (n = 1 stage 1, n = 3 stage 2, n = 2 stage 4S) resulting in immediate treatment of these patients according to the high-risk protocol, 13 events were noted in 90 low-risk patients with a favorable PAM prediction. Of them, six patients had small locoregional progressions, and another six had stage 4S-related progressions (patients < 1 year of age with localized disease who progressed to stage 4S disease or progressions of initial stage 4S skin metastases). After no (n = 3) or limited treatment (n = 5 surgery, n = 4 limited chemotherapy), these patients are currently in complete (n = 8) or incomplete (n = 4) remission for 29 to 2,589 days (median, 1,620 days). Only one of 90 NB2004 low-risk patients with a favorable PAM prediction had a metastatic relapse of disease (NB499), and therefore was misclassified.

Similar results with high statistical significance were observed using the COG (n = 75; 3-year EFS, 0.88 ± 0.04 *v* 0.25 ± 0.15;

Downloaded from jco.ascopubs.org on April 15, 2008 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

Table 2. Classification Parameters for 54 Patients of the Second Set With Maximally Divergent Courses of the Disease

Classifier	No.	%				
		Accuracy	Sensitivity	Specificity	Negative Predictive Value	Positive Predictive Value
PAM (UF *v* F)	54	93	93	93	97	81
Age (< 1 *v* > 1 year)	54	74	86	70	93	50
Stage 1, 2, and 4S *v* 3, 4	54	91	86	93	95	80
Stage 1, 2, 3, and 4S *v* 4	54	91	64	100	89	100
MYCN (amplified *v* normal)	54	89	64	98	89	90
Deletion 1p	53	89	69	95	91	82
Shimada (F *v* UF)	46	96	90	97	97	90

Abbreviations: PAM, prediction analysis for microarrays; F, favorable; UF, unfavorable.

$P < .0001$; Fig 4B) or the Japanese risk group system (n = 97; 3-year EFS, 0.85 ± 0.04 *v* 0.25 ± 0.15; $P < .0001$; Fig 4C).

Separation of Intermediate-Risk Group Patients Into Subgroups With Divergent Outcome

In both the German (n = 13) and the Japanese (n = 18) intermediate-risk group, no event was noted in patients with a favorable gene expression classification (Germany n = 6, Japan n = 11), whereas six events occurred in seven patients with an unfavorable PAM prediction each (3-year EFS NB2004, 1.00 *v* 0.57 ± 0.19; $P = .018$; 3-year EFS Japan, 1.00 *v* 0.57 ± 0.19; $P = .0005$; Figs 4D and 4F). In the intermediate-risk group as defined by the COG trial (n = 33), four events were observed in 26 patients with a favorable PAM classification (n = 1 progression of skin metastasis and n = 3 progression of primary tumor) and all had beneficial courses of the disease with no (n = 3) or moderate chemotherapeutic treatment (n = 1). In contrast, four of seven intermediate-risk patients with an unfavorable classification experienced events and were treated by surgery (n = 1) or intensive salvage chemotherapy (n = 3; 3-year EFS, 0.82 ± 0.08 *v* 0.71 ± 0.17; $P = .0266$; Fig 4E).

Identification of Event-Free Survivors Within Current High-Risk Groups

In 58 patients who were concordantly stratified as high risk by both the German and the Japanese risk-classification systems, a trend towards a significantly differing EFS between favorably (n = 16) and unfavorably predicted patients (n = 42) was observed (3-year EFS 0.81 ± 0.10 *v* 0.58 ± 0.08; $P = .0601$; Figs 4G and 4J). To date, 19 of 42 unfavorably predicted patients had a relapse/progression of disease, and 12 of them died as a result of disease. In contrast, three events occurred in the subgroup of 14 high-risk patients with a favorable PAM prediction (and one died as a result of disease). Similar results were observed applying risk classification according to the COG trial (n = 59; 3-year EFS, 0.82 ± 0.09 [n = 17] *v* 0.57 ± 0.08 [n = 42]; $P = .0386$; Fig 4H).

Multivariate Cox Regression Analysis

In a multivariate Cox proportional hazard regression model of single prognostic markers (age, stage, *MYCN*, status 1p, Shimada, and PAM) based on EFS, both the PAM classifier ($P = .004$; hazard ratio (HR), 3.318; 95% CI, 1.428 to 7.708) and the Shimada system ($P = .043$; HR, 2.284; 95% CI, 1.0002 to 5.205) were significant independent predictors of outcome for the second set of patients (Table 3). In a multivariate Cox regression analysis with the complete risk stratification strategies of current trials (NB2004, COG, and Japan), the PAM classifier was a superior, independent prognostic marker ($P < .0001$; HR, 4.756; 95% CI, 2.544 to 8.893; Table 2).

DISCUSSION

As several studies have shown, many recent gene expression–based approaches lacked statistical stringency and therefore presented overly optimistic results[29] or classifiers that are overfitted to the particular set of patients used in a study.[30,31] Because such effects are in part caused by unspecific information generated by large gene expression profiles, we designed a disease-related chip covering a high percentage of transcripts related to neuroblastoma tumor behavior and a reduced fraction of unspecific probes. Subsequently, we applied a statistically rigid methodology to generate a robust and clinically applicable gene expression classifier from a first set of 77 tumors, representing maximally contrasting subtypes of neuroblastoma. Remarkably, gene expression–based classification of these patients was highly accurate (99%, as estimated by cross validation) and confirmed to be precise in 54 of 174 patients of an independent set that met the criteria of the training set (93%; 50 of 54). In comparison, comparable high cross-validated prediction accuracies were observed in two recent studies by Schramm et al (77% to 85%)[18] and Ohira et al (89%),[17] suggesting that gene expression analysis could be introduced into future neuroblastoma trials. It should be stressed though, that Ohira et al did not compare their classifier's performance with that of risk stratification of current neuroblastoma trials and that the clinical value of this predictor therefore remains elusive.

In contrast, we show in the present study that our PAM predictor separates patients of current low-, intermediate- and high-risk group (as defined by risk stratification according to current international neuroblastoma trials) into subgroups with divergent outcome. In addition, a remarkable difference in the nature of events was observed in patients with a favorable and an unfavorable PAM prediction, respectively. This effect was most prominent in patients currently considered low risk, in whom an unfavorable gene expression classification was closely associated with metastatic relapse/progression to stage 4 disease (six of seven patients with event), whereas a favorable one was related to stage 4S-related events (six of 13) or locoregional events (six of 13). The finding that 12 of 13 favorably predicted low-risk patients with event were curable by very limited or no treatment at

Downloaded from jco.ascopubs.org on April 15, 2008 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

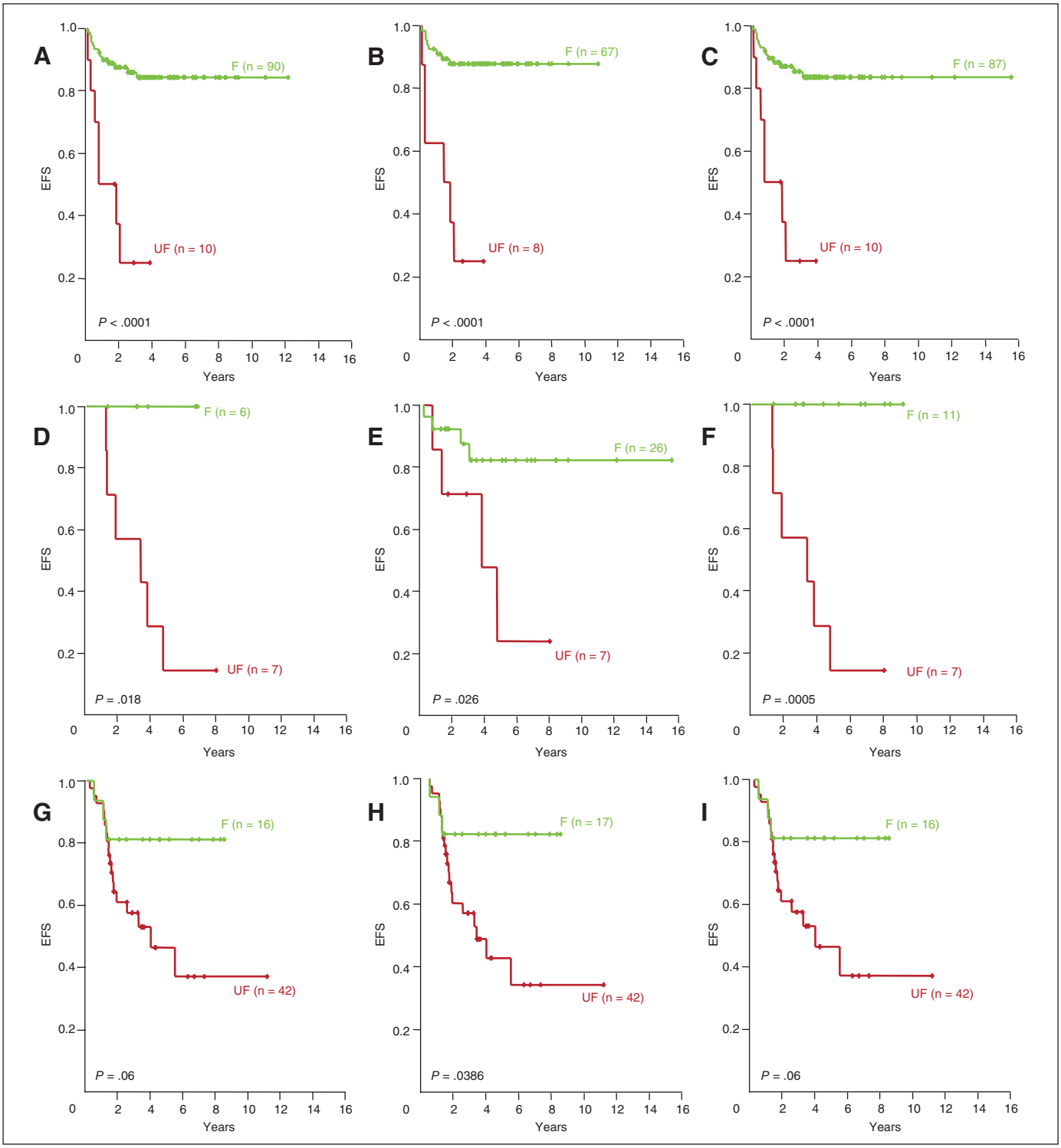

Fig 4. Kaplan-Meier estimates for event-free survival (EFS) for the 174 patients of the second set (subdivided into low-, intermediate-, and high-risk groups according to current neuroblastoma trials) according to their classification by the 144-gene prediction analysis for microarrays (PAM) predictor. Kaplan-Meier estimates for EFS for patients of the low-risk groups defined by (A) NB2004, (B) the Children's Oncology Group (COG), and (C) the Japanese trial; Kaplan-Meier estimates for EFS for patients of the intermediate-risk groups defined by (D) NB2004, (E) COG and the (F) Japanese trial; Kaplan-Meier estimates for EFS for patients of the high-risk groups defined by (G) NB2004, (H) COG, and (I) the Japanese trial. F, favorable; UF, unfavorable.

Downloaded from jco.ascopubs.org on April 15, 2008 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

Table 3. Multivariate Cox Regression Model

	P	Hazard Ratio	95% CI
Single marker			
Age (continuous)	.159		
Stage (1-3, 4S *v* 4)	.775		
MYCN (normal *v* amplified)	.299		
Status 1p (normal *v* deletion/imbalance)	.658		
Shimada (F *v* UF)	**.043**	2.284	1.0002-5.205
PAM classifier (F *v* UF)	**.004**	3.318	1.428-7.708
Trial			
NB2004 (LR *v* IR *v* HR)*	.646		
Japan (LR *v* IR *v* HR)†	.518		
COG (LR *v* IR *v* HR)‡	.350		
PAM (F *v* UF)	**< .0001**	4.756	2.544-8.893

NOTE. Boldfacing indicates statistical significance.
Abbreviations: F, favorable; UF, unfavorable; PAM, prediction analysis for microarrays; LR, low risk; IR, intermediate risk; HR, high risk; COG, Children's Oncology Group.
*Stage, age, *MYCN*, del1p.
†Stage, age, *MYCN*.
‡Stage, age, *MYCN*, ploidie, Shimada.

all raises the possibility that spontaneous regression or maturation would have occurred in all these patients, because a favorable vote by our classifier is based on a gene expression pattern resembling that of long-term survivors who did not receive cytotoxic treatment.

Regarding the significant discrimination of current intermediate-risk patients, it appears reasonable to assume that one potentially very important use of the classifier could be in the context of risk stratification for patients currently considered intermediate risk. However, one has to consider the possible effects of the cytotoxic treatment, because patients in the intermediate-risk group received considerable chemotherapeutic dose intensity. Thus, the question whether therapy could safely be reduced in current intermediate- or even high-risk patients with a favorable gene expression classification needs to be evaluated in a prospective setting. Nonetheless, a favorable PAM classification might reflect an intrinsic ability of neuroblastoma tumors to mature/regress either spontaneously or after limited treatment; such phenomena were observed in patient NB005, in whom extensive maturation of the tumor justified early withdrawal of further cytotoxic treatment.

Yet, apart from the question how to optimally adjust therapeutic regimens depending on the results of a gene expression classification, the important question is how gene expression–based classifiers can optimally be integrated into current risk estimation systems. On the one hand, the PAM predictor could be combined with other markers such as the Shimada score to form an optimal risk stratification system; on the other hand, the multivariate Cox's regression of the PAM predictor with the complete risk stratification systems of current trials also supports the implementation of the classifier as a stand-alone test at least for certain subgroups (eg, current low-risk patients). These important issues clearly have to be addressed in future prospective studies. In any case, both multivariate Cox regression analyses underline the high degree of prognostic information that is covered by the 144 classifying genes. Analysis of the biologic functions of these genes revealed that (1) many are major players in the process of chromosome segregation or cell cycle regulation (eg, *CCNB1*, *CENPA*, *MAP7*, *STK6* etc). Elevated expression levels of this functional group has also been found to be associated with poor outcome in both other tumor entities[32] and neuroblastoma patients with adverse outcome,[33] indicating that upregulation of these processes may describe neuroblastoma tumor progression. (2) A substantial number of genes are involved in apoptosis (eg, *NALP1* etc)[34] and/or in the promotion of neuronal differentiation (eg, *DST*, *MAPT*, *NXPH1* etc). Genes of the latter category have also been observed in all current microarray classifiers for neuroblastoma[16-18] and universally had higher expression levels in patients with beneficial courses, which could suggest ongoing regression or maturation in these tumors. (3) Several signature genes were reported to be of prognostic impact on their own (eg, *NTRK1* = *TrkA*[12], *SCG* = Chromogranin C[35]) and some were found to contribute to other gene expression classifiers for neuroblastoma (eg, *CNR1*,[16] *AHCY*[17]).

However, although similar functional groups of genes contribute to all current gene expression classifiers for neuroblastoma, only little overlap exists between the particular candidate genes for these expression signatures.[16-18] This could be explained by several factors, such as differences in the set of tumors, the microarray-platforms, or the classifying algorithms used in these studies. Yet, it also has to be considered that several differing sets of predictive genes yielded similar classification results for breast cancer patients.[36] Therefore, the application of rigid statistical methodologies to generate a predictor and the testing of sufficiently large independent sets of tumors appear to be most important factors to warrant higher reproducibility of gene expression–based classification results and to allow for the integration of gene expression–based prediction into clinical trials.

REFERENCES

1. Kaatsch PKU, Spix C, et al: Annual report 2004 (1980-2003), German Childhood Cancer Registry. Mainz: German Cancer Registry 88:11-12, 2004

2. Berthold F, Hero B: Neuroblastoma: Current drug therapy recommendations as part of the total treatment approach. Drugs 59:1261-1277, 2000

3. Brodeur GM, Pritchard J, Berthold F, et al: Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. J Clin Oncol 11:1466-1477, 1993

4. Evans AE, D'Angio GJ, Randolph J: A proposed staging for children with neuroblastoma: Children's Cancer Study Group A. Cancer 27:374-378, 1971

5. Brodeur GM, Seeger RC, Schwab M, et al: Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. Science 224:1121-1124, 1984

6. Brodeur GM, Fong CT, Morita M, et al: Molecular analysis and clinical significance of N-myc amplification and chromosome 1p monosomy in human neuroblastomas. Prog Clin Biol Res 271:3-15, 1988

7. Look AT, Hayes FA, Nitschke R, et al: Cellular DNA content as a predictor of response to chemotherapy in infants with unresectable neuroblastoma. N Engl J Med 311:231-235, 1984

8. Shimada H, Ambros IM, Dehner LP, et al: The International Neuroblastoma Pathology Classification (the Shimada system). Cancer 86:364-372, 1999

9. Attiyeh EF, London WB, Mosse YP, et al: Chromosome 1p and 11q deletions and outcome in neuroblastoma. N Engl J Med 353:2243-2253, 2005

10. Spitz R, Hero B, Ernestus K, et al: Deletions in chromosome arms 3p and 11q are new prognostic markers in localized and 4s neuroblastoma. Clin Cancer Res 9:52-58, 2003

11. Bown N, Lastowska M, Cotterill S, et al: 17q gain in neuroblastoma predicts adverse clinical outcome: UK Cancer Cytogenetics Group and the UK Children's Cancer Study Group. Med Pediatr Oncol 36:14-19, 2001

12. Nakagawara A, Arima-Nakagawara M, Scavarda NJ, et al: Association between high levels of expression of the TRK gene and favorable outcome in human neuroblastoma. N Engl J Med 328:847-854, 1993

13. Berwanger B, Hartmann O, Bergmann E, et al: Loss of a FYN-regulated differentiation and growth arrest pathway in advanced stage neuroblastoma. Cancer Cell 2:377-386, 2002

Downloaded from jco.ascopubs.org on April 15, 2008 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

14. van de Vijver MJ, He YD, van't Veer LJ, et al: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347:1999-2009, 2002

15. Bullinger L, Dohner K, Bair E, et al: Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med 350:1605-1616, 2004

16. Wei JS, Greer BT, Westermann F, et al: Prediction of clinical outcome using gene expression profiling and artificial neural networks for patients with neuroblastoma. Cancer Res 64:6883-6891, 2004

17. Ohira M, Oba S, Nakamura Y, et al: Expression profiling using a tumor-specific cDNA microarray predicts the prognosis of intermediate risk neuroblastomas. Cancer Cell 7:337-350, 2005

18. Schramm A, Schulte JH, Klein-Hitpass L, et al: Prediction of clinical outcome and biological characterization of neuroblastoma by expression profiling. Oncogene 24:7902-7912, 2005

19. Ambros PF, Ambros IM: Pathology and biology guidelines for resectable and unresectable neuroblastic tumors and bone marrow examination guidelines. Med Pediatr Oncol 37:492-504, 2001

20. The R project for statistical computing. www.r-project.org

21. Bioconductor open source software for bioinformatics. www.bioconductor.org

22. Buness A, Huber W, Steiner K, et al: ArrayMagic: Two-colour cDNA microarray quality control and preprocessing. Bioinformatics 21:554-556, 2005

23. Huber W, von Heydebreck A, Sultmann H, et al: Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 18:S96-S104, 2002 (suppl)

24. Tibshirani R, Hastie T, Narasimhan B, et al: Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 99:6567-6572, 2002

25. Ruschhaupt M, Huber W, Poustka A, et al: A compendium to ensure computational reproducibility in high-dimensional classification tasks. Stat Appl Genet Mol Biol 3, 2004 (article 37)

26. Li X, Rao S, Wang Y, et al: Gene mining: A novel and powerful ensemble decision approach to hunting for disease genes using microarray expression profiling. Nucleic Acids Res 32:2685-2694, 2004

27. Cheung NK, Cohn S (eds): Neuroblastoma. Berlin, Germany, Springer, 2005, pp 78-79

28. Maris JM: The biologic basis for neuroblastoma heterogeneity and risk stratification. Curr Opin Pediatr 17:7-13, 2005

29. Michiels S, Koscielny S, Hill C: Prediction of cancer outcome with microarrays: A multiple random validation strategy. Lancet 365:488-492, 2005

30. Ransohoff DF: Rules of evidence for cancer molecular-marker discovery and validation. Nat Rev Cancer 4:309-314, 2004

31. Ntzani EE, Ioannidis JP: Predictive ability of DNA microarrays for cancer outcomes and correlates: An empirical assessment. Lancet 362:1439-1444, 2003

32. Nakamura N, Yamamoto H, Yao T, et al: Prognostic significance of expressions of cell-cycle regulatory proteins in gastrointestinal stromal tumor and the relevance of the risk grade. Hum Pathol 36:828-837, 2005

33. Krasnoselsky AL, Whiteford CC, Wei JS, et al: Altered expression of cell cycle genes distinguishes aggressive neuroblastoma. Oncogene 24:1533-1541, 2005

34. Liu F, Lo CF, Ning X, et al: Expression of NALP1 in cerebellar granule neurons stimulates apoptosis. Cell Signal 16:1013-1021, 2004

35. Pagani A, Forni M, Tonini GP, et al: Expression of members of the chromogranin family in primary neuroblastomas. Diagn Mol Pathol 1:16-24, 1992

36. Ein-Dor L, Kela I, Getz G, et al: Outcome signature genes in breast cancer: Is there a unique set? Bioinformatics 21:171-178, 2005

37. Tusher VG, Tibshirani R, Chu G: Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98:5116-5121, 2001

38. Schacherer F, Choi C, Gotze U, et al: The TRANSPATH signal transduction database: A knowledge base on signal transduction networks. Bioinformatics 17:1053-1107, 2001

Acknowledgment

We thank Dr Andreas Polten (Agilent Technologies, Germany) for his excellent collaboration.

Appendix

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

Authors' Disclosures of Potential Conflicts of Interest

The authors indicated no potential conflicts of interest.

Author Contributions

Conception and design: André Oberthuer, Patrick Warnat, Rüdiger Spitz, Manfred Schwab, Benedikt Brors, Frank Westermann, Matthias Fischer
Financial support: Frank Berthold, Frank Westermann, Matthias Fischer
Administrative support: Roland Eils, Manfred Schwab
Provision of study materials or patients: Frank Berthold, Rüdiger Spitz
Collection and assembly of data: André Oberthuer, Barbara Hero, Rainer König, Stefan Haas
Data analysis and interpretation: André Oberthuer, Patrick Warnat, Barbara Hero, Yvonne Kahlert, Karen Ernestus, Roland Eils, Benedikt Brors
Manuscript writing: André Oberthuer
Final approval of manuscript: André Oberthuer, Frank Berthold, Manfred Schwab, Benedikt Brors, Frank Westermann, Matthias Fischer

Downloaded from jco.ascopubs.org on April 15, 2008 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.