表2　効果判定規準（International Neuroblastoma Response Criteria/INRC）

評価	原発巣	転移巣
CR (complete response)	腫瘍無し	腫瘍無し カテコールアミン代謝産物正常化
VGPR (very good partial response)	90%—99%縮小	腫瘍無し カテコールアミン代謝産物正常化，骨シンチでの集積は残存していてもよい (MIBG シンチは陰性化していなければならない)
PR (partial response)	50%以上縮小	測定可能病変が50%以上縮小 骨転移の病変数が50%以上減少 骨髄転移の病変数は0—1か所 (MIBG シンチでの集積は残存していてもよい)
MR (mixed response)	新病変の出現なし 原発巣および転移巣の測定可能病変において50%以上縮小する病変を認める 同時に他の病変は50%未満の縮小や25%未満の増大を示す	
NR (no response)	新病変の出現なし 原発巣および転移巣の測定可能病変は，50%未満の縮小や25%未満の増大を示す	
PD (progressive disease)	新病変の出現 あるいは原発巣および転移巣の測定可能病変において25%以上の増大を示す病変を認める もしくは骨髄の転移病変の新たな出現	

CR, VGPR, PR, MR, NR につては定義に述べられた全ての用件を満たしていることが必要である．PD に関しては定義に述べられたいずれかの要件を満たした状態である．

Brodeur GM, et al.: J Clin Oncol 11(8)：1466-7, 1993 より引用

表3　COG リスク分類

低リスク
1. 患者の年齢を問わず INSS 1 期
2. 1 歳未満の INSS 2A 期及び 2B 期
3. 1 歳以上で，FHG の INSS 2A 期及び 2B 期
4. 1 歳以上で，*N-MYC* 増幅なしの INSS 2A 期及び 2B 期
5. 1 歳未満で，*N-MYC* 増幅なし，かつ FHG，かつ高二倍体 DNA である INSS4S 期

中間リスク
1. 1 歳未満で，*N-MYC* 増幅なしの INSS 3 期
2. 1 歳以上で，*N-MYC* 増幅なし，かつ FHG の INSS 3 期
3. 1 歳未満で，*N-MYC* 増幅なしの INSS 4 期
4. 1 歳未満で，*N-MYC* 増幅なし，かつ二倍体に近い DNA の INSS 4S 期
5. 1 歳未満で，*N-MYC* 増幅なし，かつ UFHG の INSS 4S 期

高リスク
1. 1 歳以上で，*N-MYC* 増幅あり，かつ UFHG の INSS 2A 期及び 2B 期
2. 患者の年齢を問わず，*N-MYC* 増幅ありの INSS 3 期
3. 1 歳以上で，UFHG の INSS 3 期
4. 1 歳未満で，*N-MYC* 増幅ありの INSS 4 期
5. 1 歳以上の INSS 4 期
6. 1 歳未満で，*N-MYC* 増幅ありの INSS 4S 期

FHG;INPC で favorable histology group，UFHG;INPC で unfavorable histology group

記載のない項目は不問である

Castleberry RP. Eur J Cancer. 33：1430-1437, 1997 より引用

いは分子生物学的検査を施行する必要がある．また腫瘍マーカーとしての血清カテコールアミンやVMA，HVAなどを確認することが有効である．NSEは小円形細胞腫瘍ではいずれの腫瘍でも上昇することがある．また原発部位が腎近傍の場合には腎芽腫などの腎原発腫瘍との鑑別も必要である[1),3)]．

Ⅷ 予　後

予後を規定する最も重要な因子は適切な治療法である．このためリスク分類を設定し，リスクごとに治療強度を規定し，より適切な治療法の開発が検討されている．日本ではリスク分類は未だ議論の途上ではあるが，今後治療成績の集積により改善していくものと考えられる．今のところ1歳半以上あるいはstage 4あるいは*MYCN*が増幅している神経芽腫の予後は，骨髄破壊的大量化学療法を行っても3年無増悪生存割合として20〜40%台に過ぎない．大量化学療法後の主な再発形式は骨あるいは骨髄再発である．現在アメリカではCOGのリスク分類に基づき3年全生存割合を低リスク群で90%超，中間リスク群で70〜90%，高リスク群で30%超と推測している[1)]．日本全体としての治療成績は明らかではない．

Ⅸ 治療方針

神経芽腫の治療は以前よりリスクに基づいた治療戦略が行われてきた．現在のCOGの治療方針は低リスク群では外科切除後経過観察，中間リスク群では外科切除と通常の化学療法，高リスク群では集学的治療が必要となり，外科切除に加え積極的な化学療法と大量化学療法＋自家造血細胞移植及び放射線療法である．日本においても以前よりほぼ同様の治療方針が採られている．

低リスクあるいは中間リスク例の治療は第1選択としては外科切除を行う．外科療法単独で可能な場合や，あるいは6〜12〜24週間の化学療法を併用する場合がある．化学療法による障害を抑えるために薬剤使用量を少なくする努力が行われている．

ダンベル症候群をきたし脊髄圧迫症状を起こしている例では，神経学的な症状を一刻も早く改善し神経症状が永続するのを回避するために緊急化学療法を行うべきである．神経圧迫症状の期間が短いほど神経学的回復が期待される．椎弓切除術あるいは放射線療法も同等の治療効果が期待できるが，いずれもその後の化学療法が必要になるため，速やかな化学療法の実施が望ましいと考えられている[1)]．

Ⅹ 進行例に対する治療方針

種々の研究により明らかにされた予後不良因子（遠隔転移・*MYCN*増幅・UH・DNA index≦1など）を持つ高リスク神経芽腫は，現在なおその3年無増悪生存割合で20〜40%であり，新たな有効な治療法の開発が切に待ち望まれている．

高リスク神経芽腫の進行は早く，発見後あるいは化学療法中にも急速に進展する例が多く見られ，原発部位の増大だけでなく骨・骨髄・リンパ節・肝・後腹膜・後縦隔などに急速に転移・浸潤する特徴がある．また，集学的治療が奏効し治療を終了できたとしても，治療後すぐから再発を認めることが多い．近年では原発巣局所に対する外科及び放射線療法が進歩してきており，以前よりも原発巣局所からの再発は減少して来たが，その代わりに骨や骨髄再発が主体となってきている．再発時期は骨髄破壊的大量化学療法施行直後から2年以内にほとんどの再発がみられ，その後も5年以上にわたり再発がみられる[1)]．

現在日米欧では，高リスク神経芽腫に対しては，診断時に原発巣を安全に全摘出できる症例がほとんどなく，骨・骨髄転移例が多く，速やかに全身化学療法を行わないと生命の危険性が高いと判断されることから，生検後に速やかに寛解導入化学療法(induction)が開始される．化学療法を数コース行った後，局所療法として外科切除術および局所放射線療法を組み合わせた治療を行い，その後強化した化学療法かあるいは骨髄破壊的大量化学療法による地固め療法(consolidation)を行うものが多い．化学療法としてはシスプラチン・カルボプラチン・シクロフォスファミド・イホスファミド・エトポシド・ビンクリスチン・ピラルビシ

ン・ドキソルビシンなどから2-4剤を選択し，それらを組み合わせて併用する種々の化学療法計画が考案され実施されている．また大量化学療法としては，現在主として行われているのはメルファラン＋エトポシド＋カルボプラチンの組み合わせによるMEC療法(CEM療法)やブスルファン＋メルファランの組み合わせによるBU/Mel療法などである．TBIについては議論の多い部分であるが，近年ではTBIによる腎障害，間質性肺炎，肝中心静脈閉塞症などの急性毒性と，成長障害，内分泌障害，白内障，二次がんなどの晩期毒性の重大性から，生存者の生活の質も重視したnon-TBIによる治療開発が世界的に進行している[1)]．

これらの高リスク神経芽腫に対する世界各国における標準的治療戦略による臨床試験成績はほぼ類似している．米国のCCG-A 3891臨床試験の大量化学療法群のデータでは，3年無増悪生存割合は34±4%で，非大量化学療法群では22±4%である[11)]．またドイツ及びスイスのNB97臨床試験の大量化学療法群のデータでは，3年無増悪生存割合は47%(95% CI38-55)で，非大量化学療法群では31%(95% CI 23-39)である[12)]．一方，日本から報告される3年無増悪生存割合は大体30～50%の範囲にある[13)]．

神経芽腫を含め発症数の少ない稀少疾患である小児固形腫瘍はいずれも未だ治療法開発段階であり，常に整備された前方視的な多施設共同臨床研究に従い治療が行われることが望ましいと考えられる．小児固形腫瘍は症例数が少なく，診断が不確実になることも多いため中央病理診断の必要性がある．各施設での個々の経験に基づいた治療では症例数が少なすぎて確かなデータが得られず，しかもデータ管理が不十分となりやすい．治療成績の向上には多施設臨床研究が不可欠である．

進行神経芽腫の再発例では再発後の予後はさらに悪い．現在でも非常に難治であり，再発後の長期生存例は少ない．新規薬剤や新規治療戦略の開発が望まれている．

現在世界的に期待されている新規薬剤・治療法には Topotecan, Irinotecan, Fenretinide, 131I-MIBG＋ASCT, Tandem Auto SCT, Allo SCT, Anti-GD2 monoclonal antibody 3F8＋GM－CSF, Demethylating agent (Decitabine) などがあげられる[1)]．

XI 治療終了後の長期的な問題点

治療終了後の長期的な問題点も最近クローズアップされるようになり，重大な課題として再認識されている．これらには副鼻腔炎，低身長，やせ，肥満，歯牙の崩出障害，甲状腺機能低下症，性腺機能障害，不妊症，毛髪，難聴，白内障，腎機能障害，心機能障害，呼吸器系の障害，学校に関連する問題，精神神経的問題，二次がんなどの種々の問題点がある[1)]．

XII 結　語

本稿では，神経芽腫についての基礎的知識の概略と，進行神経芽腫に対する化学療法についての知見を紹介した．

文　献

1）本稿全体を通じて次の文献を参考にした．Brodeur GM, Maris JM. Neuroblastoma. In Pizzo PA, Poplack DG, eds. Principle and Practice of Pediatric Oncology 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins, pp933-970, 2006

2）Rothenberg AB, et al: Neuroblastoma remembering the three physicians who described it a century ago: James Homer Wright, William Pepper, and Robert Hutchinson. Pediatr Radiol, 39 : 155-160, 2009

3）日本病理学会小児腫瘍組織分類委員会編集：小児腫瘍組織カラーアトラス第2巻．神経芽腫群腫瘍－国際分類INPCによる－第1版，金原出版，東京，pp3-33，2004

4）日本神経芽腫研究グループ運営委員会資料

5）澤田淳，他：神経芽腫スクリーニング．小児がんの臨床，初版，赤塚順一他編集，医薬ジャーナル，大阪，pp135-145，2000

6）Volchenboum SL, et al: Progress in defining and treating high-risk neuroblastoma: lessons from the bench and bedside. J Clin Oncol, 27 : 1003-1004, 2009

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

8) Cohn SL, et al: The international neuroblastoma risk group (INRG) classification system: an INRG task force report. J Clin Oncol, 27 : 289-297, 2008

9) Monclair T, et al: The international neuroblastoma risk group (INRG) staging system:an INRG task force report. J Clin Oncol, 27 : 298-303, 2008

10) Castleberry RP. Neuroblastoma. Eur J Cancer, 33 : 1430-1437, 1997

11) Matthay KK, et al: Treatment of high-risk neuroblastoma with intensive chemotherapy, radiotherapy, autologous bone marrow transplantation, and 13-cis-retinoic acid. N Engl J Med, 341 : 1165-1173, 1999

12) Berthold F, et al: Myeloablative megatherapy with autologous stem-cell rescue versus oral maintenance chemotherapy as consolidation treatment in patients with high-risk neuroblastoma: A randomized controlled trial. Lancet Oncol 6 : 649-658, 2005

13) Kaneko M, Tsuchida Y, Mugishima H, et al: Intensified chemotherapy increases the survival rate in patients with stage 4 neuroblastoma with *MYCN* amplification. J Pediatr Hematol Oncol; 24 : 613-621, 2002

Clinical Cancer Research

Accurate Outcome Prediction in Neuroblastoma across Independent Data Sets Using a Multigene Signature

Katleen De Preter[1], Joëlle Vermeulen[1], Benedikt Brors[2], Olivier Delattre[4], Angelika Eggert[5], Matthias Fischer[6], Isabelle Janoueix-Lerosey[4], Cinzia Lavarino[7], John M. Maris[8], Jaume Mora[7], Akira Nakagawara[9], André Oberthuer[6], Miki Ohira[9], Gudrun Schleiermacher[4], Alexander Schramm[5], Johannes H. Schulte[5], Qun Wang[8], Frank Westermann[3], Frank Speleman[1], and Jo Vandesompele[1]

Abstract

Purpose: Reliable prognostic stratification remains a challenge for cancer patients, especially for diseases with variable clinical course such as neuroblastoma. Although numerous studies have shown that outcome might be predicted using gene expression signatures, independent cross-platform validation is often lacking.

Experimental Design: Using eight independent studies comprising 933 neuroblastoma patients, a prognostic gene expression classifier was developed, trained, tested, and validated. The classifier was established based on reanalysis of four published studies with updated clinical information, reannotation of the probe sequences, common risk definition for training cases, and a single method for gene selection (prediction analysis of microarray) and classification (correlation analysis).

Results: Based on 250 training samples from four published microarray data sets, a correlation signature was built using 42 robust prognostic genes. The resulting classifier was validated on 351 patients from four independent and unpublished data sets and on 129 remaining test samples from the published studies. Patients with divergent outcome in the total cohort, as well as in the different risk groups, were accurately classified (log-rank $P < 0.001$ for overall and progression-free survival in the four independent data sets). Moreover, the 42-gene classifier was shown to be an independent predictor for survival (odds ratio, >5).

Conclusion: The strength of this 42-gene classifier is its small number of genes and its cross-platform validity in which it outperforms other published prognostic signatures. The robustness and accuracy of the classifier enables prospective assessment of neuroblastoma patient outcome. Most importantly, this gene selection procedure might be an example for development and validation of robust gene expression signatures in other cancer entities. *Clin Cancer Res; 16(5); 1532–41. ©2010 AACR.*

One of the main challenges in clinical cancer research remains accurate prediction of outcome, enabling better choice of risk-related therapy. This is particularly true for neuroblastoma, a pediatric tumor of the sympathetic nervous system, which is characterized by a remarkably heterogeneous clinical course. Tumors that are found in infants frequently regress spontaneously or show differentiation features on treatment, whereas tumors diagnosed in children >1 year of age often metastasize, causing accelerated cancer-related death despite intensive therapies. Accordingly, different therapeutic schemes exist ranging from watch-and-see approaches to multimodal therapies. Four major risk stratification systems are currently being used in various parts of the world (Europe, United States, Japan, and Germany) based on a combination of clinicopathologic and genetic parameters, such as age at diagnosis, tumor stage, *MYCN* gene status, histopathologic classification, ploidy, and chromosome 1p and 11q status (1–8). Clinical experience within these systems indicates that the stratification is useful, but misclassifications occur, resulting in overtreatment or undertreatment. Identification of more specific and sensitive markers for response to therapy and outcome prediction is clearly required and is expected to result in better choice of risk-related therapy.

As differences in outcome are considered to reflect underlying genetic and biological characteristics that have

Authors' Affiliations: [1]Center for Medical Genetics, Ghent University, Ghent University Hospital, Ghent, Belgium; Departments of [2]Theoretical Bioinformatics and [3]Tumour Genetics, German Cancer Research Center, Heidelberg, Germany; [4]Institut National de la Sante et de la Recherche Medicale U830, Institut Curie, Paris, France; [5]Division of Hematology and Oncology, University Children's Hospital Essen, Essen, Germany; [6]Children's Hospital of Cologne, Department of Pediatric Oncology, Cologne, Germany; [7]Developmental Tumor Biology Laboratory, Hospital Sant Joan de Déu, Barcelona, Spain; [8]Division of Oncology, Children's Hospital of Philadelphia, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; and [9]Division of Biochemistry, Chiba Cancer Center Research Institute, Chiba, Japan

Corresponding Author: Jo Vandesompele, Center for Medical Genetics, Ghent University Hospital, Medical Research Building, 2nd Floor, Room 120.055, De Pintelaan 185, B-9000 Ghent, Belgium. Phone: 32-9-332-5187; Fax: 32-9-332-6549; E-mail: Joke.Vandesompele@UGent.be.

doi: 10.1158/1078-0432.CCR-09-2607

©2010 American Association for Cancer Research.

Translational Relevance

Prognostic classification of heterogeneous diseases such as neuroblastoma remains challenging. In this study, a unique data-mining approach was applied for establishment of an accurate and robust gene expression classifier to predict clinical outcome of neuroblastoma patients. Using both published and unpublished microarray expression data of 933 primary neuroblastomas, a 42-gene classifier was developed and successfully validated. The powerful and independent prognostic value of the 42-gene classifier is shown using several lines of evidence. First, patients with divergent outcome were accurately classified. Second, multivariate analysis showed that the classifier is an independent prognostic factor, contributing to more accurate assessment of prognosis when considering the conventional risk factors alone. This was further confirmed within a subgroup of high-risk patients. Moreover, the excellent performance of the classifier across different expression platforms clearly shows its robustness. The presented gene selection procedure is suitable for the development of gene expression signatures in other cancer entities.

their effect on mRNA gene expression profiles, several microarray expression profiling studies have been undertaken to predict patient outcome in different cancer entities.

An important limitation of many published gene expression profiling studies is the lack of statistical power to identify markers and lack of independent validation. Typically, around 30,000 to 40,000 transcripts are tested, generating hundreds of thousands of data points for a relatively small subset of tumors (between 20 and 100). When such a high number of genes are evaluated as prognostic markers, there is a substantial chance that a random association between a gene and the prognostic classes is observed (9, 10). Consequently, many published studies do not classify patients better than chance due to lack of internal validation by repeated random sampling of training sets or external validation on independent samples. As such, there are a few inherent but often overlooked statistical issues, such as data overfitting, unstable gene lists, and lack of study power (11).

In this study, we established a prognostic 42-gene classifier for children with neuroblastoma by reanalysis of four published gene expression studies from four different microarray platforms comprising 582 patients in total (12–15). To facilitate data comparison across different platforms, probe annotations were updated with respect to the original publications. When available, clinical follow-up information was updated. All these aspects critically contribute to the success of our multigene signature. Successful validation of the multigene signature in four independent unpublished data sets shows its robust performance and platform independence.

Materials and Methods

Gene expression data sets. Four published studies were used for selecting the genes and building the prognostic classifier (phase 1 data sets), and four unpublished data sets were used as independent validation sets (phase 2 data sets).

The phase 1 data sets were downloaded either from the National Center for Biotechnology Information Gene Expression Omnibus (GSE2283 and GSE3960; refs. 14, 15) or from the European Bioinformatics Institute ArrayExpress database (E-TABM-38; ref. 13), or from the authors' Web site[10] (12).

A trained multigene correlation signature was validated on the four independent phase 2 data sets from which the 42 genes (when present) were extracted and standardized (per gene, the median value across the samples was subtracted followed by division by the SD of the gene): (*a*) hgu95av2 Affymetrix gene expression data from 106 neuroblastoma patients (validation set 1; 40 genes present), (*b*) hgu133plus2 Affymetrix gene expression data from 53 neuroblastoma patients (validation set 2; 40 genes present), (*c*) data set for 91 neuroblastoma patients obtained using an 11K custom Agilent oligonucleotide microarray (validation set 3; 41 genes present), and (*d*) Human Exon 1.0 ST Affymetrix expression data from 101 neuroblastoma patients (validation set 4; 42 genes present; standardized data of the 42-gene selection as well as clinical data are available in Supplementary Tables S1 and S2; Fig. 1).

For the remainder of the article, we will label the data sets according to the first author for the published phase 1 studies [Oberthuer (13), Wang (15), Berwanger (12), and Ohira (14)] and as validation sets 1, 2, 3, and 4 for the unpublished phase 2 studies.

Data preprocessing. To make the data from the different microarray platforms maximally comparable, annotation information of the probes was updated using the MatchMiner tool (16) for the custom-made cDNA or oligonucleotide arrays (12–14) and using the latest version of the R packages hgu95av2 and hgu133plus2 for the Affymetrix array data (15). Probe identification numbers were converted into gene symbols to enable straightforward comparison of the gene lists between the different studies. Throughout the text, the number of unique gene symbols (represented by one or more array probes) in each study is indicated.

Updated clinical information with regard to progression-free survival (PFS) and overall survival (OS) times was obtained from the authors (14, 15) or was publicly available (13). For the Berwanger and Ohira studies and validation set 1, only OS data were available.

[10] http://www.imt.uni-marburg.de/microarray/download.html

Patients were divided in two clearly defined risk groups. The low-risk subgroup was defined by stage I, II, or IVS without *MYCN* amplification, and the high-risk subgroup comprised patients with age of diagnosis >1 y with stage IV tumors (irrespective of *MYCN* status) or with stage II and III tumors with *MYCN* amplification. To develop our classifier, as many patients as possible from the four phase 1 data sets were divided in the two risk groups with maximally divergent clinical course (Table 1), that is, low-risk patients with PFS time (or OS time for Berwanger and Ohira data sets) of at least 1,000 d and high-risk patients that died from the disease. The patients that did not belong to the above-mentioned low- or high-risk subgroups were used as independent test set.

Statistical analysis. Identification and validation of prognostic classifiers (for each single phase 1 data set) were done by prediction analysis of microarray (PAM) classification with 10-times repeated 10-fold cross-validation in

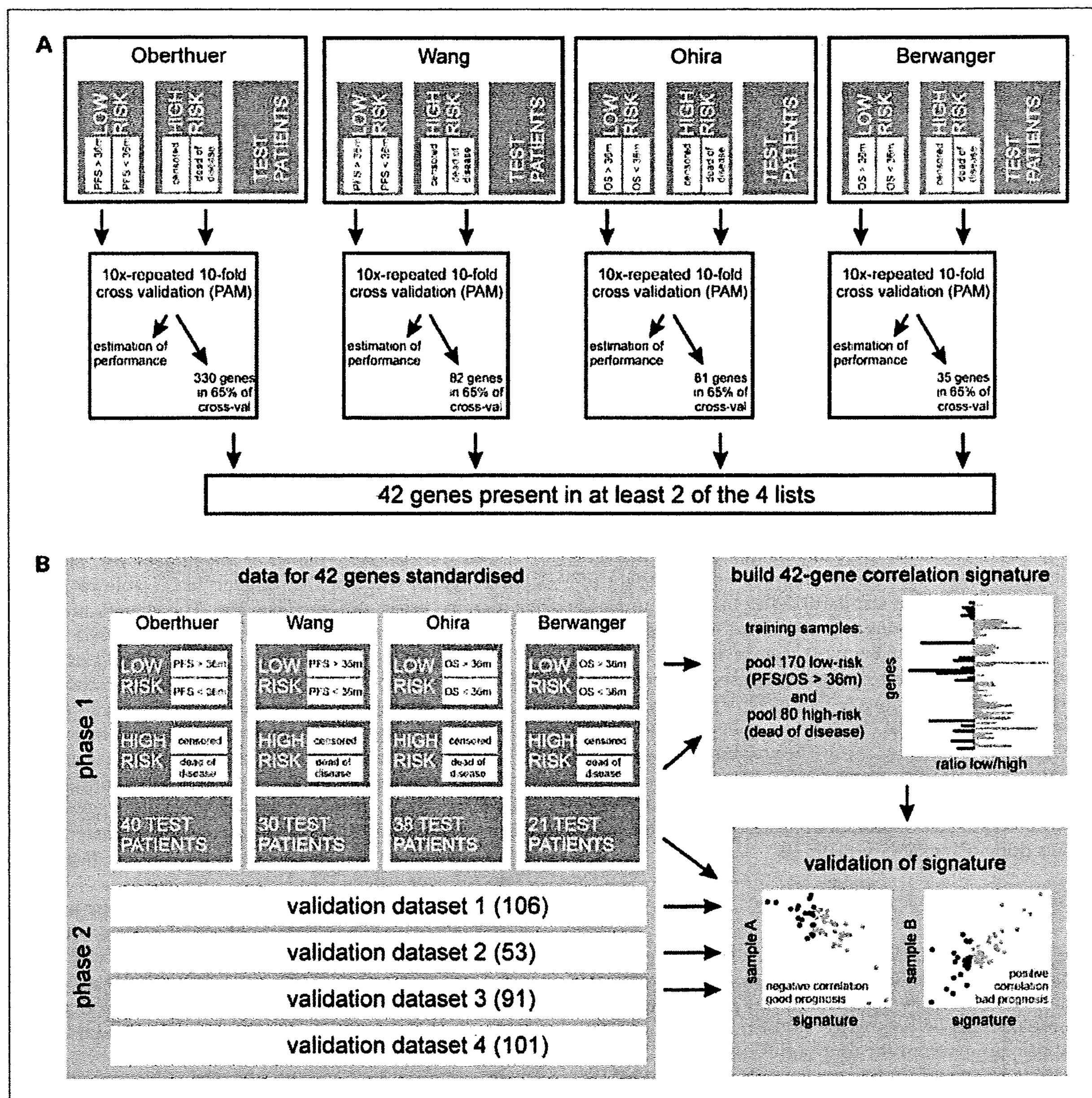

Fig. 1. Outline of the strategy used for prioritization of the 42 prognostic gene list (A) and construction of a 42-gene correlation signature and validation on independent test samples from phase 1 studies and phase 2 validation data sets (B). m, months.

Table 1. Published phase 1 studies used for training the classifier, with indication of the number of (training) samples, median OS or PFS (in months), and estimation of the performance of the study-specific PAM classifier for prediction of unfavorable outcome (OS)

	Berwanger	Oberthuer	Ohira	Wang
No. patients	94	251	136	101
No. low-risk training samples	22	87	43	18
No. high-risk training samples	13	25	20	22
Median OS/PFS (mo)	OS = 43	PFS = 55	OS = 46	PFS = 48
Specificity	0.955	0.977	0.814	1.000
Sensitivity	1.000	0.960	0.950	0.773
Negative predictive value	0.929	0.923	0.704	1.000
Positive predictive value	1.000	0.988	0.972	0.783
Accuracy	0.971	0.973	0.857	0.875
Performance (AUC)	0.977	0.969	0.882	0.886

the R statistical language using the Bioconductor package MCRestimate (Fig. 1A; refs. 13, 17). Forty-two genes were present in at least two of the four resulting gene lists.

A cross-platform gene signature was built using standardized expression data of the 42 genes (if present on the respective arrays, see Supplementary Data 2) from four published phase 1 studies. The correlation method was used to build and test a cross-platform prognostic signature (Fig. 1B). Log-transformed data were merged in one file (if more than one probe was present for a certain gene, the probe with the highest expression value was selected), and for each of the 42 genes, the mean expression value in low-risk neuroblastoma patients with PFS of at least 1,000 d was subtracted from the mean expression value in high-risk neuroblastoma patients that died of disease. For classification, the Pearson's correlation coefficient of the signature with the standardized expression values of independent test patients was calculated. Patients with a correlation coefficient below 0 were predicted to have good prognosis, whereas the other patients were predicted to have bad prognosis [according to Liu et al. (18)].

Kaplan-Meier survival analysis was done with the R survival package (R version 2.6.1). The area under the receiver operating characteristic curve (AUC) was used as a measure for the accuracy of the classifiers (ROCR R-package). Multivariate forward conditional logistic regression analysis was done using SPSS version 16.

Results

Gene prioritization for inclusion in a robust prognostic classifier. A complete 10-times repeated 10-fold cross-validation using the PAM algorithm (13, 19) was done on the training patients belonging to one of the two clearly defined risk groups from the four published phase 1 studies separately to identify robust prognostic markers (Fig. 1). This process was accompanied by determination of the classification accuracy, providing a first estimation of the utility of the expression data to predict outcome (Table 1).

For each data set, we selected the probes that were included in at least 65 of the 100 cross-validation gene lists, as these genes are likely to be the ones with the highest prognostic value as determined by Oberthuer et al. (13). The resulting prognostic gene lists from the four studies showed significant overlap (Table 2; Supplementary Data 1). Two genes were in common between three lists (i.e., *MYCN* and *NTRK1*), whereas 40 genes were in common between two lists. Thirty-two were previously reported in at least 1 of 10 published prognostic gene lists, of which only 10 were found in 2 or more published prognostic lists (12–14, 20–26). The occurrence of the 42 genes in at least two of the four lists makes them robust, platform-independent, prognostic markers.

Classification performance of the 42-gene list. Next, we investigated whether the 42-gene list is able to predict prognosis across different data sets. The classification performance was estimated in the different phase 1 data sets using a complete 10-times repeated 10-fold cross-validation method using all patients from the two clearly defined risk groups. For this analysis, it is important to note that not all 42 genes are present on all platforms; hence, the performance test was inherently done with a different number of genes for the different data sets (Supplementary Data 2). As already indicated, the 10-times repeated 10-fold cross-validation provides a good estimate for the classification performance using the expression data of the selected gene list.

As a reference, the 35-, 330-, 81-, and 82-gene lists obtained through single PAM analysis of each of the four phase 1 data sets were evaluated in the same way as the 42-gene list. The classification performance was also tested for a subset of 11 genes (from the 42-gene list) that were present on all four platforms. This analysis showed that all performance parameters for the 42-gene list are best or second best for all studies compared with the other gene lists, whereby the overall accuracy is highest for the 42-gene list subset (AUC = 0.935; Supplementary Data 2). This analysis also shows that the performance of a classifier built for

a given data set is not always best, which indicates the power and utility of our meta-analysis for the identification of a prognostic gene list by using expression data of 250 training samples (170 low risk and 80 high risk). When only 11 genes of the 42-gene list were selected that are present on all four platforms, the overall accuracy was lower due to loss in sensitivity and positive predictive value. The 42-gene classifier was also compared with two published classifiers (13, 27) and showed that the 42-gene classifier performs best.

Table 2. Genes that are in common between the 42-gene list and the different individual classifier gene lists (number of common genes in list/total number of genes in list)

	Berwanger (10/35)	Oberthuer (38/330)	Ohira (12/81)	Wang (26/82)	published lists
AHCY		–	–		2
AKR1C1		+		+	1
ARHGEF7		+	+		2
BIRC5	–		–		1
CADM1		+		+	0
CAMTA2		+		+	0
CDCA5	–	–			2
CDKN3		–		–	2
CLSTN1		+		+	1
DDC		+	+		1
DPYSL3		+	+		1
ECEL1		+	+		0
EPB41L3		+		+	0
EPHA5	+	+			1
EPN2		+		+	0
FYN			+	+	1
GNB1		+	+		1
HIVEP2		+		+	1
INPP1	+			+	1
MAP7	+	+			1
MAPT		+	+		1
MCM2		–		–	0
MRPL3		–		–	1
MYCN	–	–	–		4
NCAN		–		–	0
NME1	–	–			2
NRCAM		+		+	2
NTRK1		+	+	+	4
ODC1	–			–	1
PAICS		–		–	1
PLAGL1	+	+			1
PMP22		+		+	1
PRKACB		+		+	2
PRKCZ		+		+	1
PTN		+		+	1
PTPRN2		+	+		0
SCG2		+		+	1
SLC25A5		–		–	1
SNAPC1		–		–	0
TYMS		–		–	1
ULK2		+		+	0
WSB1	+	+			4

NOTE: The number of published prognostic gene lists (other than the four reanalyzed studies) in which these genes are found is indicated in the last column. –, associated with poor outcome; +, associated with favorable outcome.

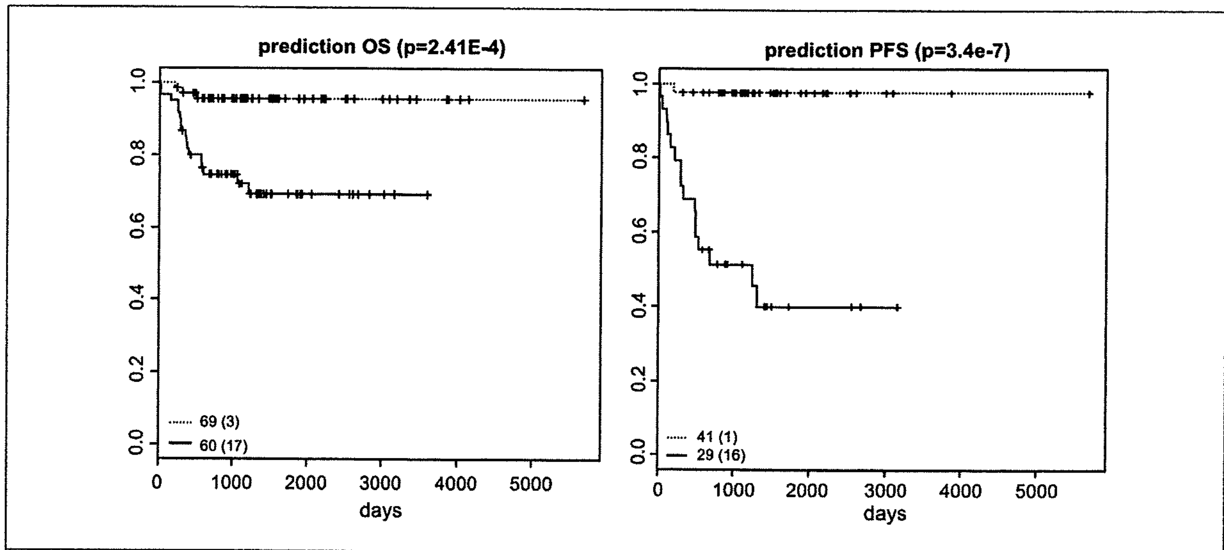

Fig. 2. Kaplan-Meier and log-rank analysis of 129 test patients (OS) and 70 test patients (PFS) from the four published phase 1 studies classified using the prognostic correlation signature. Legend, number of patients in predicted subgroups; between brackets, number of patients with event (relapse, progression, or death).

Validation of a cross-platform prognostic 42-gene correlation signature for neuroblastoma. A major disadvantage of the PAM classification method is the need for a training set of samples that are analyzed on the same gene expression measurement platform as the one used to evaluate the test samples. We therefore applied an alternative method to build a classifier based on the 42-gene list that can be used for completely independent data sets even on other platforms.

The prognostic signature is determined using 250 training samples from the four phase 1 studies. A 42-gene classification vector was created and tested using the correlation method (see Materials and Methods; Fig. 1).

First, the correlation signature was tested on the 129 test samples (patients not belonging to the low- and high-risk subgroup) from the four phase 1 studies and revealed a very high predictive power for OS (log-rank P = 2.41E−4) and PFS (log-rank P = 3.40E−7; Fig. 2).

Next, this correlation signature was evaluated on the four independent phase 2 data sets (351 patients), whereby the patients could be clearly separated into groups with significant differences in OS (log-rank P = 2.17E−23) and PFS (log-rank P = 2.03E−21; Fig. 3A). Kaplan-Meier analysis of patients stratified using known risk factors (i.e., age, stage, and *MYCN* gene status) showed that the correlation signature outperforms these risk factors ($P < 0.001$, except for *MYCN*-amplified samples; Supplementary Fig. S2). This was confirmed using multivariate logistic regression analysis evaluating age, stage, *MYCN* status, and the gene classifier, indicating that the 42-gene signature is an independent predictor for PFS and OS in the four phase 2 data sets as well as in the test samples of the phase 1 data sets (Table 3). Of note, whereas phase 2 data sets are representative of the general neuroblastoma population, test samples from the phase 1 data sets only represent intermediate risk patients.

As the different validation data sets include patients stratified using different risk stratification systems (Europe, United States, and Germany), we defined a common low- and high-risk group (Supplementary Data 3). As there was only 1 patient of 50 that died of disease within the common low-risk group of patients, we did not do Kaplan-Meier analysis. However, we could show that this single patient was classified in the high-molecular risk group using our classifier. Most interestingly, the correlation signature could partition patients within the common high-risk subgroup into groups with significant differences in OS and PFS (Fig. 3B) and was an independent prognostic marker (odds ratios, >4; Supplementary Table S4). To exclude that the significant survival differences in high-risk tumors is solely due to the effect of the *MYCN* amplification and related downstream *MYCN* signaling, we also tested the survival in high-risk tumors without *MYCN* amplification and could show that the classifier also significantly discriminates these patients with respect to outcome (Fig. 3C; Supplementary Table S4). In line with this, inspection of the 42-gene list indicated that not all 42-genes are related to *MYCN* amplification (Supplementary Data 4).

Discussion

In this study, we developed and validated a 42-gene prognostic classifier for children with neuroblastoma through a reanalysis strategy of published data complemented with gene expression data from 351 patients from

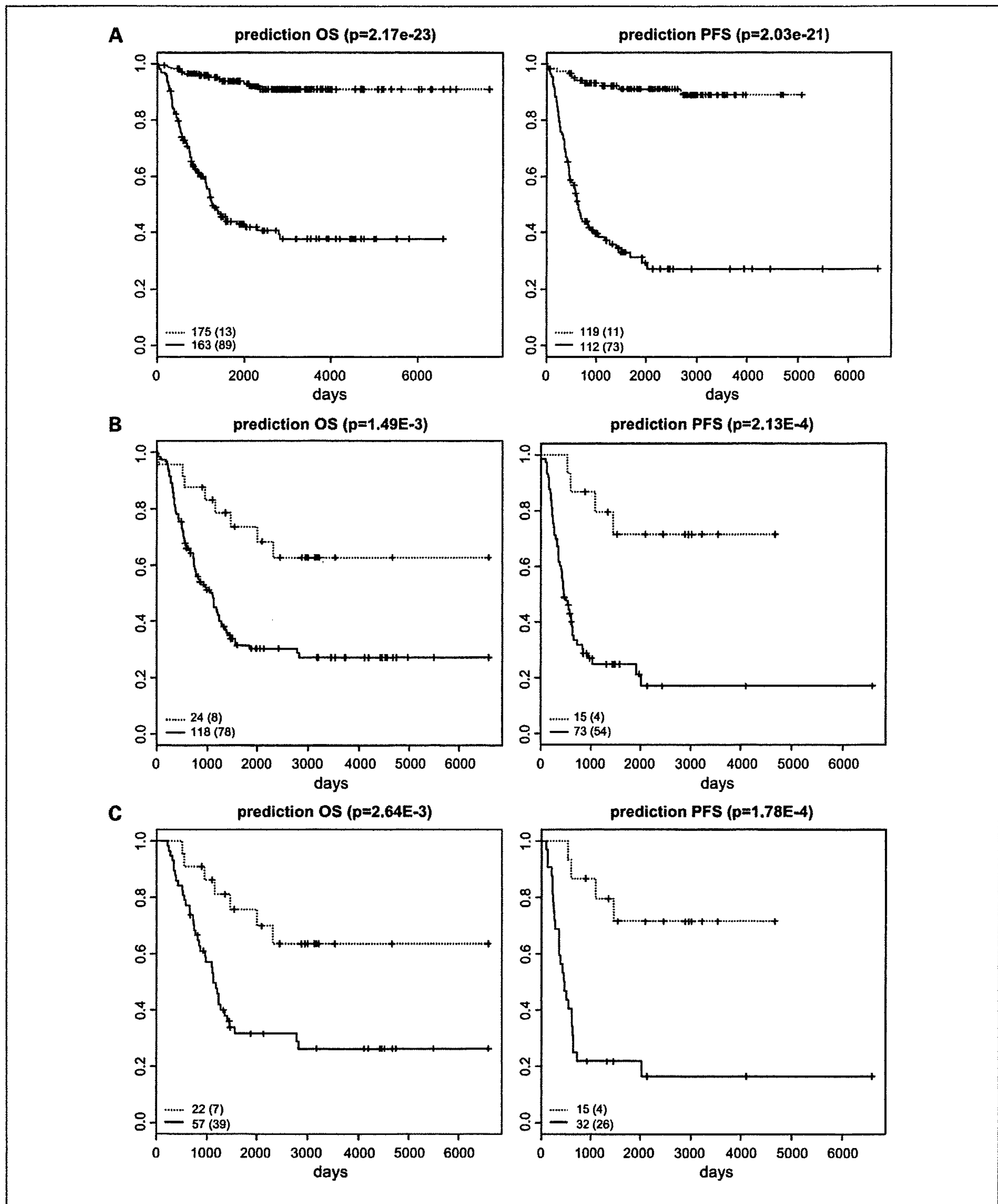

Fig. 3. Kaplan-Meier and log-rank analysis of the patients from four independent unpublished phase 2 validation data sets classified using the prognostic correlation signature for all patients together {5-y OS of 93.9% [95% confidence interval (95% CI), 90.2-97.6] for low molecular risk versus 43.1% (95% CI, 35.6-52.2) for high molecular risk and 5-y PFS of 91.1% (95% CI, 86.0-96.6) for low molecular risk versus 30.4% (95% CI, 22.1-41.8) for high molecular risk; A}, for the common high-risk subgroup (B), and for the common high-risk subgroup without MYCN amplification (C). Legend, number of patients in predicted subgroups; between brackets, number of patients with event (relapse, progression, or death).

Table 3. Multivariate logistic regression analysis (with correlation signature classification, MYCN status, International Neuroblastoma Staging System stage, and age at diagnosis; A) and sensitivity, specificity, and accuracy (AUC with 95% CI) results (follow-up time of at least 36 mo; B) for correlation signature prediction in the independent test samples from the phase 1 data sets and in the phase 2 validation data sets

A

		OS		PFS	
		P	OR (95% CI)	*P*	OR (95% CI)
Test samples from phase 1 data sets	Correlation signature	3.16E–2	5.11 (1.16-22.58)	3.12E–4	54.00 (6.17-472.41)
	MYCN amplification	7.80E–5	21.50 (4.69-98.54)	1.26E–1	—
	Stage (IV versus other)	1.80E–1	—	2.65E–1	—
	age (<1 or >1 y)	1.52E–1	—	8.65E–1	—
Phase 2 validation data sets	Correlation signature	9.07E–7	7.02 (3.23-15.28)	1.1E–14	16.45 (8.09-33.48)
	MYCN amplification	4.19E–2	2.23 (1.03-4.84)	3.13E–1	—
	Stage (IV versus other)	1.35E–2	2.50 (1.21-5.16)	2.16E–1	—
	age (<1 or >1 y)	1.45E–4	4.14 (1.99-3.66)	1.1E–4	4.18 (2.03-8.64)

B

	Test samples from phase 1 data sets	Phase 2 validation data sets
Sensitivity OS	17/20 = 0.85	89/102 = 0.87
Specificity OS	41/67 = 0.61	140/195 = 0.72
Performance, AUC (95% CI), OS	0.731 (0.612-0.850)	0.795 (0.742-0.849)
Sensitivity PFS	16/17 = 0.94	93/110 = 0.85
Specificity PFS	27/35 = 0.77	95/119 = 0.80
Performance, AUC (95% CI), PFS	0.856 (0.748-0.964)	0.822 (0.764-0.879)

NOTE: —, not analyzed.
Abbreviation: OR, odds ratio.

four unpublished data sets (Fig. 1). To accomplish this, four published microarray studies comprising >500 neuroblastoma patients were reanalyzed generating four new prognostic gene lists with a high overlap of genes between them. Comparison of the genes in the classifiers showed that 42 unique genes were present in at least two of the four lists. Not surprisingly, this set of 42 predictor genes contains numerous genes that have been reported in the context of neuroblastoma (e.g., *MYCN, NTRK1, NME1, CADM1, FYN, ODC1*, and *WSB1*). The finding of these genes in at least two independent studies indicates their robustness as prognostic markers. Comparison of the performance of the 42-gene list with the lists that were generated on the individual phase 1 studies and with two published prognostic gene lists (13, 27) showed that the classifier based on the 42-gene list has the highest overall accuracy while using the lowest number of genes. However, we have to keep in mind that this observed superiority of the 42-gene set might in part be due to the fact that, for some of the other gene lists, a large proportion of genes were not present on the platform (Supplementary Table S3).

The high prognostic classification performance of the 42-gene list is undoubtedly due to our unique reanalysis approach. First, annotations of the probes on the different platforms were updated according to the latest genome build. Second, a uniform risk definition was applied to select training patients across the different studies. Only patients with maximally divergent courses were used for training. Third, the same powerful algorithm with built-in cross-validation was used for identification of prognostic genes in four major published data sets, enabling the generation of relatively stable prognostic gene lists with high overlap.

This list of 42 prognostic genes was used to build a cross-platform classification signature. As the PAM algorithm is not suitable for cross-platform classification, we used a more intuitive, alternative method for building a 42-gene classifier. In this study, we generated a prognostic correlation signature based on expression data of the 42 genes in 250 training samples of the four phase 1 data sets. The signature was subsequently applied on independent test samples from the phase 1 data sets and on four independent and unpublished phase 2 data sets, generated on different expression profiling platforms, totaling 480 patients. The excellent prognostic performance of the 42-gene list (Table 3) further shows the validity of our meta-analysis approach and the utility of the recognized prognostic markers for neuroblastoma. The classifier predicts overall (OS) and PFS for the patients from the four phase 2 studies as well as for the test patients from the phase 1 studies (which could not be unequivocally classified in the low- or high-risk subgroups using known risk factors) with high sensitivity and specificity (Table 3). Importantly, the classifier was shown to be an independent predictor for both PFS and OS when stratifying for known risk factors such as age, stage, and *MYCN* status. Indeed, the 42-gene list does not only contain *MYCN*-regulated genes and, thus, not only reflects the *MYCN* copy number status of the samples. This is further substantiated by the excellent performance of the classifier in the high-risk neuroblastoma patients without *MYCN* amplification.

Thus far, this is the largest prognostic meta-analysis study in neuroblastoma, totaling >900 patients, including 351 patients from four independent and unpublished validation data sets. In contrast to other studies on neuroblastoma gene expression classifiers (13, 14, 21, 25, 27, 28), we could show an excellent performance of our classifier on these four independent data sets involving patients from different risk protocols from Germany, Europe, and United States by using a smaller gene set and a more intuitive classification method.

This survival classifier will definitely help to identify patients with increased risk in the current risk groups and to make a better choice of risk-related therapy. For example, low-risk patients with high molecular risk might benefit from more aggressive treatment protocols, whereas more intensive follow-up and new experimental therapies might be considered for high-risk patients with high molecular risk.

In conclusion, we applied a unique meta-analysis strategy for the identification of a robust set of 42 prognostic genes for outcome prediction in neuroblastoma. Furthermore, we propose a prognostic gene signature that is significantly associated with outcome prediction in neuroblastoma samples from independent studies using different technological platforms, making it a useful and practical classifier for risk stratification in neuroblastoma patients. The signature remains to be tested in a prospective clinical validation. The low number of genes makes this signature very well suited for cost-effective and fast PCR-based analysis, requiring only minimal amounts of tumor material, as exemplified by a recently published quantitative PCR study in which a 59-gene classifier containing the 42 genes from this study was trained, tested, and independently validated on a large cohort of patients (29). The outlined strategy for robust selection of prognostic markers and the use of a cross-platform correlation signature have wide application potential in other cancer entities.

In the search of an optimal prognostic classifier, it could prove useful to do an integrated analysis to determine the combined prognostic power of a mRNA gene expression signature along with gene copy number levels, microRNA gene expression patterns, and epigenetic modifications.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Vincent Detours for critical reading of the manuscript and helpful suggestions and the Children's Oncology group.

Grant Support

European Community under the FP6 (project: STREP: EET-pipeline, number: 037260). This article presents research results of the Belgian program of Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. K. De Preter is a postdoctoral researcher with the Fund for Scientific Research-Flanders. J. Vermeulen is a Ph.D. student with the Belgian Kid's Fund and the Fondation pour la recherche Nuovo-Soldati. This study was sponsored by the 'Kinderkankerfonds,' the 'Stichting tegen Kanker,' the Fund for Scientific Research and BOF-UGent, and the fondation Fournier Majoie pour l'innovation.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 09/27/2009; revised 11/27/2009; accepted 12/22/2009; published OnlineFirst 02/23/2010.

References

1. Brodeur GM, Fong CT, Morita M, Griffith R, Hayes FA, Seeger RC. Molecular analysis and clinical significance of N-myc amplification and chromosome 1p monosomy in human neuroblastomas. Prog Clin Biol Res 1988;271:3–15.
2. Brodeur GM, Pritchard J, Berthold F, et al. Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. J Clin Oncol 1993;11:1466–77.
3. Brodeur GM, Seeger RC, Schwab M, Varmus HE, Bishop JM. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. Science 1984;224:1121–4.
4. Evans AE, D'Angio GJ, Randolph J. A proposed staging for children with neuroblastoma. Children's cancer study group A. Cancer 1971; 27:374–8.
5. Look AT, Hayes FA, Nitschke R, McWilliams NB, Green AA. Cellular DNA content as a predictor of response to chemotherapy in infants with unresectable neuroblastoma. N Engl J Med 1984;311:231–5.

6. Maris JM. The biologic basis for neuroblastoma heterogeneity and risk stratification. Curr Opin Pediatr 2005;17:7–13.
7. Shimada H, Ambros IM, Dehner LP, et al. The International Neuroblastoma Pathology Classification (the Shimada system). Cancer 1999;86:364–72.
8. Cohn SL, Pearson AD, London WB, et al. The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. J Clin Oncol 2009;27:289–97.
9. Simon R. Diagnostic and prognostic prediction using gene expression profiles in high-dimensional microarray data. Br J Cancer 2003; 89:1599–604.
10. Tinker AV, Boussioutas A, Bowtell DD. The challenges of gene expression microarrays for the study of human cancer. Cancer Cell 2006;9:333–9.
11. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. Lancet 2005; 365:488–92.
12. Berwanger B, Hartmann O, Bergmann E, et al. Loss of a FYN-regulated differentiation and growth arrest pathway in advanced stage neuroblastoma. Cancer Cell 2002;2:377–86.
13. Oberthuer A, Berthold F, Warnat P, et al. Customized oligonucleotide microarray gene expression-based classification of neuroblastoma patients outperforms current clinical risk stratification. J Clin Oncol 2006;24:5070–8.
14. 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 2005;7:337–50.
15. Wang Q, Diskin S, Rappaport E, et al. Integrative genomics identifies distinct molecular classes of neuroblastoma and shows that multiple genes are targeted by regional alterations in DNA copy number. Cancer Res 2006;66:6050–62.
16. Bussey KJ, Kane D, Sunshine M, et al. MatchMiner: a tool for batch navigation among gene and gene product identifiers. Genome Biol 2003;4:R27.
17. Ruschhaupt M, Huber W, Poustka A, Mansmann U. A compendium to ensure computational reproducibility in high-dimensional classification tasks. Stat Appl Genet Mol Biol 2004;3:Article37.
18. Liu R, Wang X, Chen GY, et al. The prognostic role of a gene signature from tumorigenic breast-cancer cells. N Engl J Med 2007;356:217–26.
19. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 2002;99:6567–72.
20. Yamanaka Y, Hamazaki Y, Sato Y, et al. Maturational sequence of neuroblastoma revealed by molecular analysis on cDNA microarrays. Int J Oncol 2002;21:803–7.
21. 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 2004;64: 6883–91.
22. Takita J, Ishii M, Tsutsumi S, et al. Gene expression profiling and identification of novel prognostic marker genes in neuroblastoma. Genes Chromosomes Cancer 2004;40:120–32.
23. Ohira M, Morohashi A, Inuzuka H, et al. Expression profiling and characterization of 4200 genes cloned from primary neuroblastomas: identification of 305 genes differentially expressed between favorable and unfavorable subsets. Oncogene 2003;22: 5525–36.
24. Hiyama E, Hiyama K, Yamaoka H, Sueda T, Reynolds CP, Yokoyama T. Expression profiling of favorable and unfavorable neuroblastomas. Pediatr Surg Int 2004;20:33–8.
25. Asgharzadeh S, Pique-Regi R, Sposto R, et al. Prognostic significance of gene expression profiles of metastatic neuroblastomas lacking MYCN gene amplification. J Natl Cancer Inst 2006;98: 1193–203.
26. Abel F, Sjoberg RM, Nilsson S, Kogner P, Martinsson T. Imbalance of the mitochondrial pro- and anti-apoptotic mediators in neuroblastoma tumours with unfavourable biology. Eur J Cancer 2005; 41:635–46.
27. Chen QR, Song YK, Wei JS, et al. An integrated cross-platform prognosis study on neuroblastoma patients. Genomics 2008;92: 195–203.
28. Warnat P, Oberthuer A, Fischer M, Westermann F, Eils R, Brors B. Cross-study analysis of gene expression data for intermediate neuroblastoma identifies two biological subtypes. BMC Cancer 2007;7:89.
29. Vermeulen J, De Preter K, Naranjo A, et al. Predicting outcomes for children with neuroblastoma using a multigene-expression signature: a retrospective SIOPEN/COG/GPOH study. Lancet Oncol 2009;10: 663–71.

厚生労働科学研究費補助金
がん臨床研究事業

神経芽腫におけるリスク分類にもとづく標準的治療の
確立と均てん化および新規診断・治療法の開発研究

平成19年度～平成21年度　総合研究報告書

研究代表者　池田　均

印刷：㈱松井ピ・テ・オ・印刷