Clinical Significance of Tumor-Associated Inflammatory Cells in Metastatic Neuroblastoma

Asgharzadeh, et al

Supplementary Data

1. **Detailed Materials and Methods**
2. **Supplemental Figures 1-6, and Tables 1 and 2**

Detailed Materials and Methods

Patient Selection and Sample Preparation

Immunohistochemical Studies

Neuroblastoma tissue microarray (TMA) was obtained from the Children's Oncology Group's biopathology center (Columbus, OH) and sample tissues from patients treated at Children's Hospital Los Angeles (n=12). The Biopathology TMA (http://www.nationwidechildrens.org/biopathology-center-resources) contains 90 unique cases (0.6 mm cores in duplicates, 60% of cases or triplicates 40% of cases) with 16 Stage 1, 16 Stage 2, 15 Stage 3, and 29 Stage 4, 15 Stage 4S cases, and twenty control cores (from 3 ganglioneuromas and 3 tonsils). The samples were all stained at ChildLab (www.childlab.com, a Division of Nationwide Children's Hospital, Columbus, OH) based on a double staining procedure and using antibodies against CD163 (Biocare Medical) and AIF1 (Leica Microsystems) with appropriate negative controls and counterstains. All stained sections were scored by three reviewers (HS, MH, SA). Twenty eight TMA cores were not evaluable due to poor tissue preservation, post-treatment samples, minimal presence of adequate neuroblasts, or extensive necrosis/fibrosis. Additional staining was performed on 12 tumor tissues from Children's Hospital Los Angeles. In total, there were 71 samples included in the final analyses: stages 1-3 (n=29), stage 4S (n=11), and metastatic, stage 4 (n=31); see Supplemental Table 1. Samples with large discrepancies in scores (difference in scores between 2 reviewers greater than 2) were re-reviewed by all three reviewers to reach consensus. Each antibody was scored independently with scores ranging from 0 to 3 allowing increments of 0.5. The scoring system was proposed by HS and reflects percentage of macrophages that occupy the septal spaces between tumor cells (0 <25%, 1: 25-50%, 2: 50-75%, 3: >75%). All scores were re-scaled to integers (0-7) for subsequent analyses.

Gene Expression Studies

Similar to our previous strategy for identifying a microarray-based gene expression signature for classification NBL-NA subgroup[1], we hypothesized that the gene expression profiles of tumors from patients younger than 18 months of age when diagnosed with metastatic NBL-NA, who have generally excellent outcomes, could be used to help identify the least aggressive tumors in patients who were 18 months or older at diagnosis. Therapy for patients in the CCG cohort was based on risk categories defined by age at diagnosis. In CCG trials, patients diagnosed before 12 months of age (CCG intermediate-risk group) received multi-agent chemotherapy. Ninety-seven of the 133 samples used in this study were from CCG patients that were also studied in our previous gene expression profiling study. All CCG patients diagnosed after 12 months of age were classified as high-risk according to CCG-3891 clinical trial and received multi-agent induction chemotherapy followed in most cases by randomized assignment to consolidation with either conventional dose chemotherapy or with myeloablative chemoradiotherapy and autologous hematopoietic stem cell transplantation (AHSCT). The high-risk group from CCG-3891 was further randomized to receive 13 cis-retinoic acid or no further therapy. Note that current COG classification has expanded the criteria for classifying neuroblastoma patients at intermediate-risk for disease progression to include all patients who are younger than 18 months at diagnosis and have a tumor that is hyperdiploid with favorable histology according to the International Neuroblastoma Pathology Classification (INPC). This change was the result of reanalysis of the prognostic importance of age at diagnosis, which indicated that any cutpoint for age at diagnosis between 15 and 19 months was statistically significantly associated with good prognosis among patients diagnosed with metastatic neuroblastoma without MYCN gene amplification.

The majority of GPOH patients and all COG patients received myeloablative chemoradiotherapy consolidation supported with AHSCT. Patients enrolled in COG-A3973 either non-randomly received 13-cis-retinoic acid or were randomized to receive 13-cis-retinoic acid with or without anti-GD2 antibody immunotherapy after myeloablative chemoradiotherapy. All samples were evaluated histologically to confirm diagnosis and features used in the INPC such as assessment of neuroblastic differentiation, mitosis-karyorrhexis index (MKI) and assignment to favorable or unfavorable groups.

Disease progression was defined *a priori* as the development of any new lesion, a greater than 25% increase in the mass of any measurable tumor, or a previously negative bone marrow sample that became positive for tumor cells. Two patients in the CCG group had inadequate

documentation for evidence of relapse and presumed to have a non-disease related event, and two other patients had evidence of disease progression at autopsy. These patients were considered to have disease progression in our analyses.

Total RNA from frozen tumor sections of the CCG samples was previously isolated for microarray analysis using TRIzol reagent at Children's Hospital Los Angeles. TRIzol-based RNA extraction of samples from the COG and GPOH cohorts was conducted at the Children's Oncology Group's Biopathology Center (Columbus, Ohio) or University of Cologne (Cologne, Germany), respectively. RNA quality for all samples was assessed at Children's Hospital Los Angeles by gel electrophoresis and RNA integrity number (RIN) using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA). Only the samples with RIN>6.4 were subjected to reverse transcription into cDNA using M-MLV reverse transcriptase with oligo-dT priming (Invitrogen, Carlsbad, CA). Quantitative reverse transcription polymerase chain reaction (qRT-PCR) was carried out on a custom-designed TaqMan® Low Density Array (TLDA), using an ABI 7900 Sequence Detection System (Applied Biosystems, Carlsbad, CA). Five hundred nanograms of sample cDNA were loaded into the TLDA ports per manufacturer's protocol.

TLDA Analysis and Gene Selection

The TLDA was constructed with genes related to tumor- and inflammation-related genes. Tumor-related genes were selected based on previously published microarray studies.[1-3] Inflammation-related genes were selected based on prior knowledge and gene set enrichment analysis of microarray data.[1] After performing qRT-PCR reactions using 48-gene TLDA system, the cycle threshold (CT) values for the 44 genes and 3 housekeeping genes (GAPDH/SDHA/HPRT1) with detectible expression (CT<40) in more than 95% of specimens were determined; one gene (IL17) failed to generate any detectible expression. The CT values for the 44 genes were normalized to the geometric mean of the three housekeeping genes (Supplemental Table 2). Cycle threshold (CT) value for each gene was determined as follows: (1) Raw fluorescence values for each PCR cycle were exported from the Applied Biosystems 7900HT Version 2.3 Sequence Detection Systems software; (2) For each gene within each sample, a baseline value was computed as the median fluorescence from cycles 3-15. To avoid overestimating the baseline for some high-expressing genes, the upper limit of this range was adjusted to a value that was a least 3 cycles lower than the computed CT value for the gene; (3) The baseline value was subtracted from the raw fluorescence values, and a LOESS smoothing function was fitted. A CT value was computed as the point where the smoothed function

intersected a fixed threshold value of 0.20. The assay was considered negative if the baseline-corrected function did not intersect the fixed threshold. In these cases, the CT was assigned the value 40 (this occurred in less than 0.3% of all reactions).

Preliminary analysis of data generated using a 96-gene TLDA array and 40 tumor samples from the CCG cohort led to a 48-gene TLDA assay that was then used with the training and validation samples. The criteria for selection of the 48 genes were based on the genes from the 96-gene TLDA being highly correlated to their respective microarray data ($r > 0.70$ comparing TLDA and microarray data) or if categorization of patients based on median level of expression of the TLDA gene demonstrated statistically significant prediction of PFS (univariate analysis $P < 0.01$).

Heatmap was generated by obtaining fold-change values of each tumor over the average expression of all inflammation-related genes (Supplemental Table 2) in children diagnosed <18 months with metastatic NBL-NA. The data were winsorized (at 10%ile and 90%ile) to generate the heatmap.

In building the multivariate model, the association between each individual gene and probability of failure (i.e., having disease progression or not) was first examined by univariate logistic regression analysis, adjusting for age at diagnosis (Supplemental Table 2). Genes with a *P* value of ≤0.25 were included in a multivariate logistic model, together with age at diagnosis as a variable. Backward selection was then used to exclude genes with a *P* value >0.10. All *P* values were obtained using likelihood ratio tests. For LOOCV analysis, each of the 133 CCG samples was iteratively excluded from the analysis cohort, and a prediction model was built with the remaining 132 samples as described above. The tumor-progression score for the excluded sample was then computed using the newly-built prediction model.

NBL-NA Signature Accuracy

The relative contribution to the accuracy of the NBL-NA signature of each of the feature sets of the signature, i.e., age at diagnosis, the nine tumor cell-related genes, and five inflammation-related genes were assessed by comparing average AUC derived from 5000 permuted datasets

To evaluate the contribution of inflammation-related genes to the accuracy of prediction, we permuted the expression values of the five inflammation-related genes as a vector, without permuting age values or tumor-related gene expression values. This kept the correlation structure of these inflammation-related genes intact. The data for tumor cell-related genes and

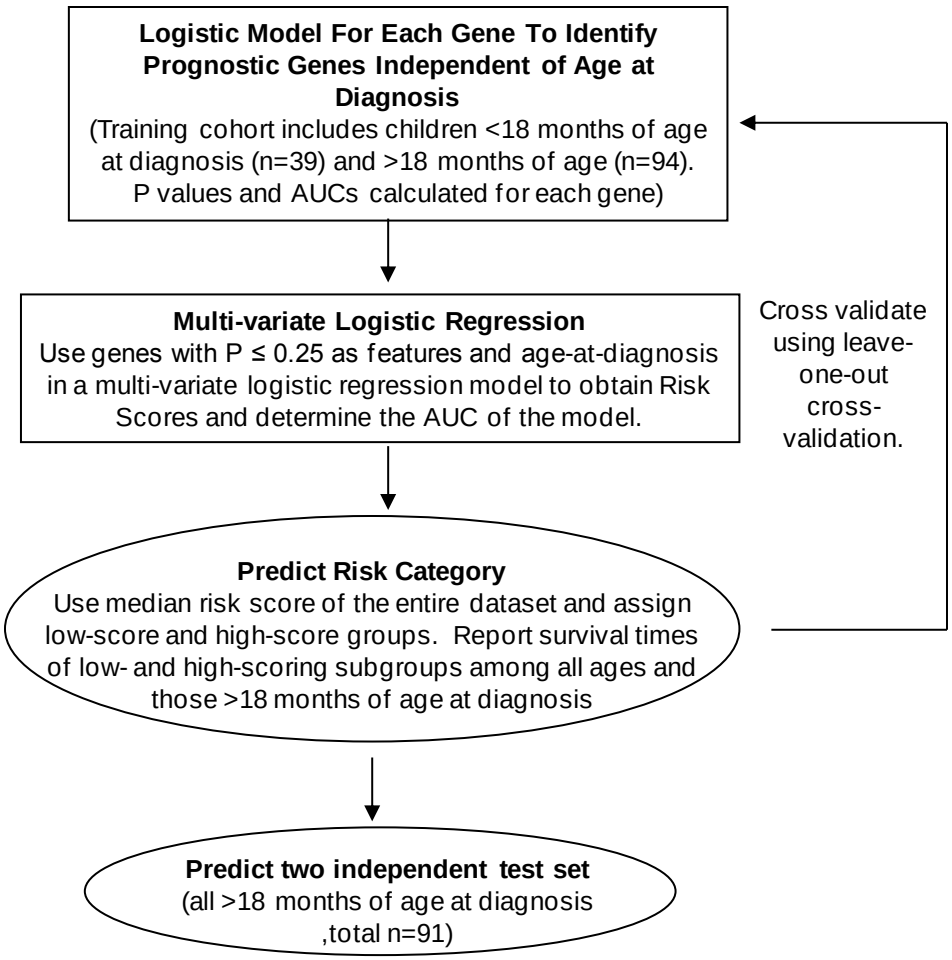
age at diagnosis were not changed. For each permutation, we applied the original regression coefficients from the final prediction model to the permutated data, calculated the predicted logit scores, and then calculated the AUC for the permuted dataset. To evaluate the contribution of tumor cell-related genes to the accuracy of prediction, we permutated the nine tumor genes as a vector without changing age and inflammation-related genes coefficients. We then calculated the average AUC using the 5000 permuted datasets. To evaluate the contribution of tumor cell-related genes, we permuted the nine tumor gene expression values as a vector without permuting age values or inflammation-related gene expression values, and again calculated average AUC using the 5000 permuted datasets. We performed the analogous analysis by permuting age at diagnosis without permuting gene expression values to evaluate the contribution of age towards the prediction in our NBL-NA model.

The original accuracy of the prediction model for the 133 CCG samples as estimated by AUC was 0.9634. The average AUC of permuted age at diagnosis, permuted inflammation-related genes, and permuted tumor cell-related genes were 0.09070, 0.8418, 0.6677, respectively. We also permuted all the features (age at diagnosis, tumor cell related and inflammation related genes), which gave us an average AUC of 0.4993 (~0.50), as expected.

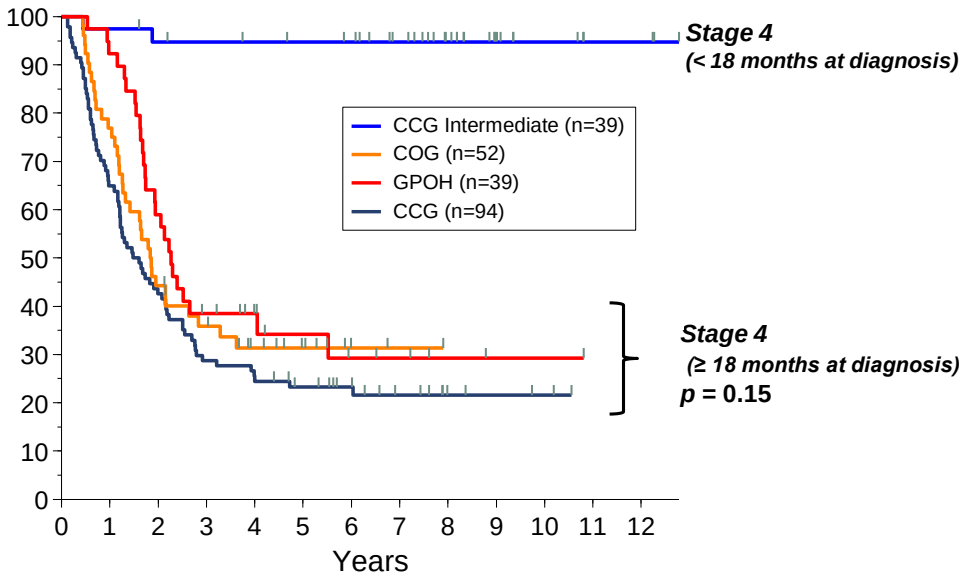
The overall AUC contribution from age at diagnosis, tumor cell-related, and inflammation-related genes together was 0.9634 – 0.5 = 0.4634. The relative contribution of a feature to the accuracy of the model was then computed as the percentage of the difference in the average permuted AUC for a given feature compared to the overall contribution of all features. Thus, age at diagnosis contributed (0.9634 – 0.9070)/0.4634 x 100% = 12.2%, contribution of the inflammation genes was (0.9634-0.8418)/0.4634 x 100% = 26.2%, and contribution of tumor-cell genes was (0.9634-0.6677)/0.4634 x 100% = 63.8%. These values were rescaled to generate the final contributions so that on average 63%, 25%, and 12% of the accuracy was attributable to the tumor cell-related genes, inflammation-related genes, and age at diagnosis, respectively.

References:

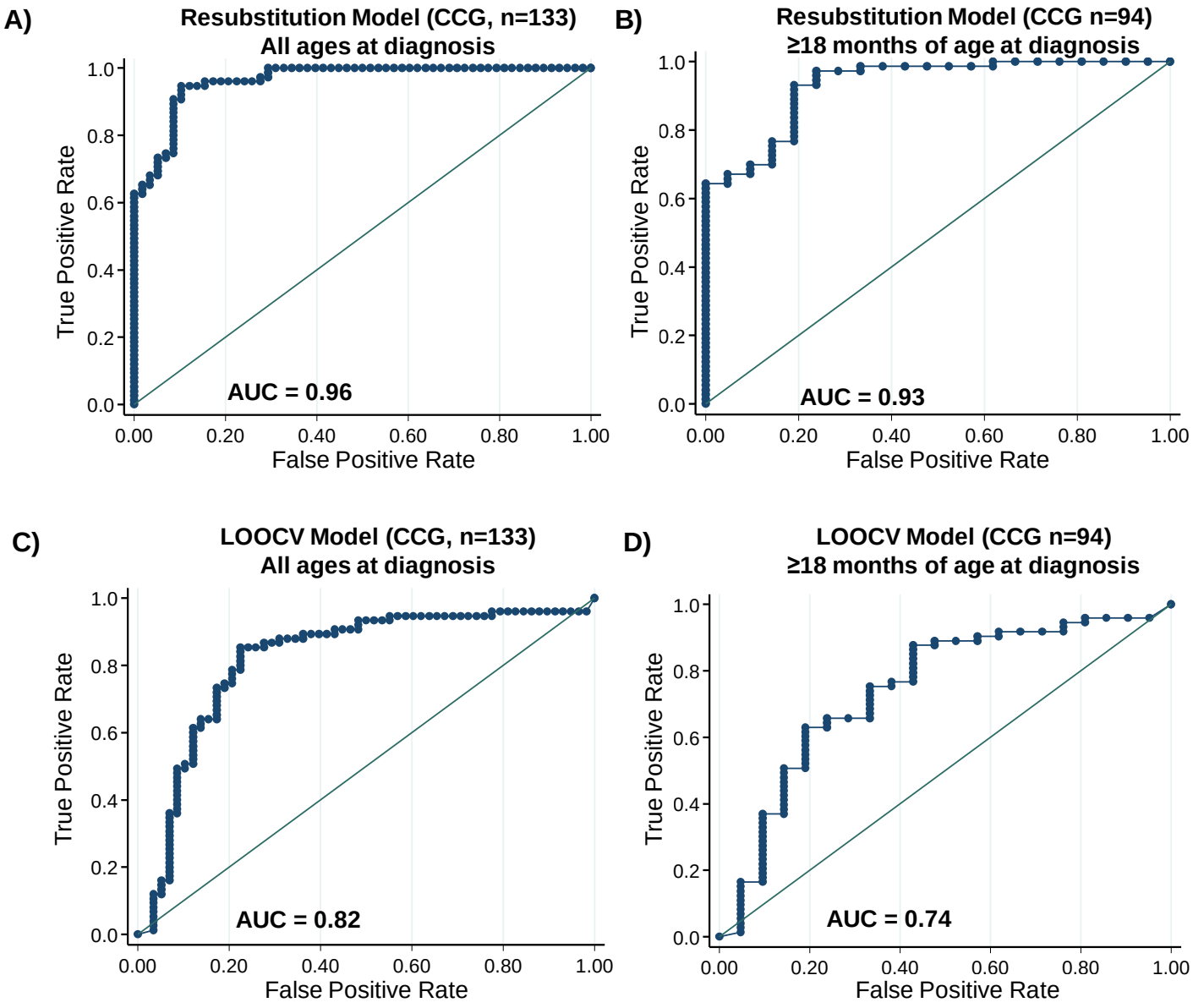
1. Asgharzadeh S, Pique-Regi R, Sposto R, et al: Prognostic significance of gene expression profiles of metastatic neuroblastomas lacking MYCN gene amplification. J.Natl.Cancer Inst. 98:1193-1203, 2006
2. Wei JS, Greer BT, Westermann F, et al: Prediction of clinical outcome using gene expression profiling and artificial neural networks for patients with neuroblastoma. Cancer Res. 64:6883-6891, 2004
3. 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



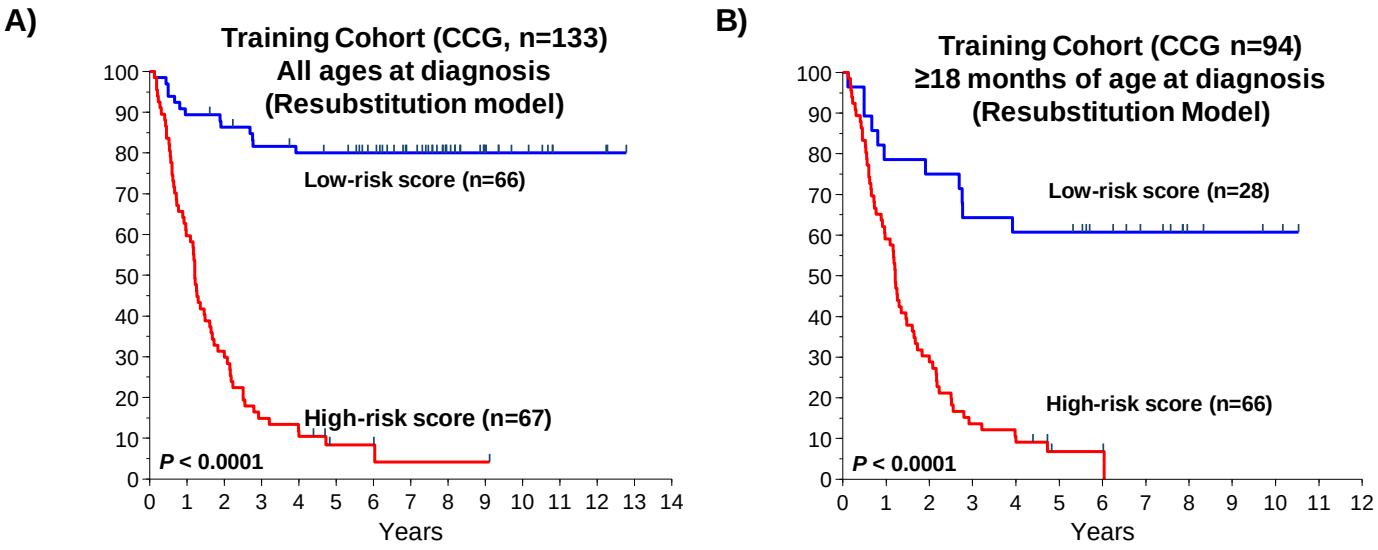
Supplemental Figure 1. Overview of the strategy used to develop and validate the NBL-NA Gene Signature. Logistic regression model for each gene and age-at-diagnosis was carried out to identify genes that are predictive of outcome and independent of age at diagnosis. Genes with P ≤ 0.25 were then used in a multi-variate logistic regression model to obtain a risk score for whether progression of disease will or will not occur. Risk groups (low-score or high-score groups) were simply based on the median score of the training set (n=133). This entire strategy was cross-validated using leave-one-out cross validation. The final model was tested on two independent cohorts of patients.



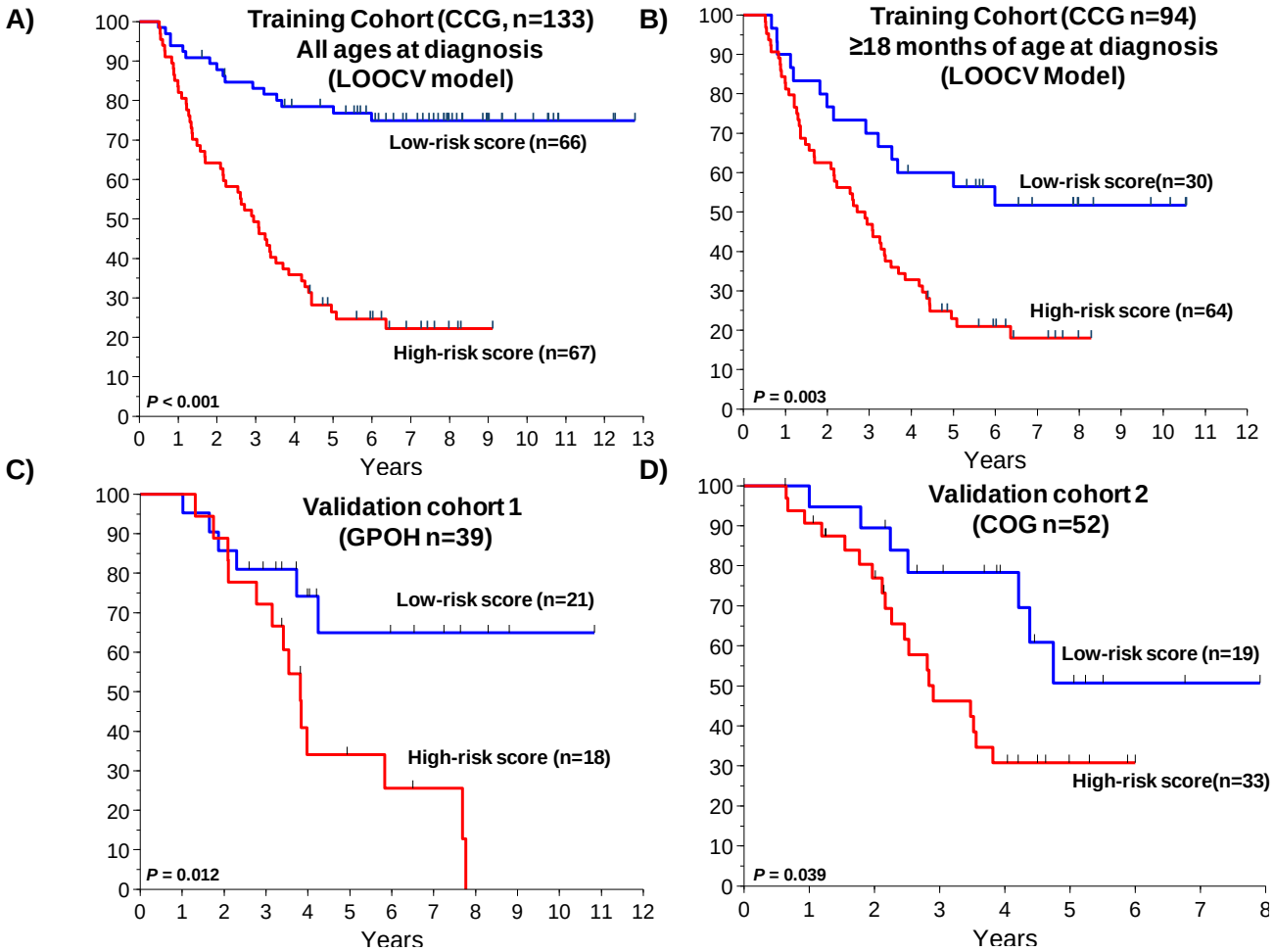
Supplemental Figure 2. Progression-Free Survival of Children Diagnosed with Metastatic Neuroblastoma Lacking MYCN Amplification by Age at Diagnosis. The graph shows no statistical difference in PFS for the patients clinically identified as high-risk (≥18 months of age at diagnosis) between the training (CCG) and validation (COG, GPOH) cohorts (*P* value = 0.15).



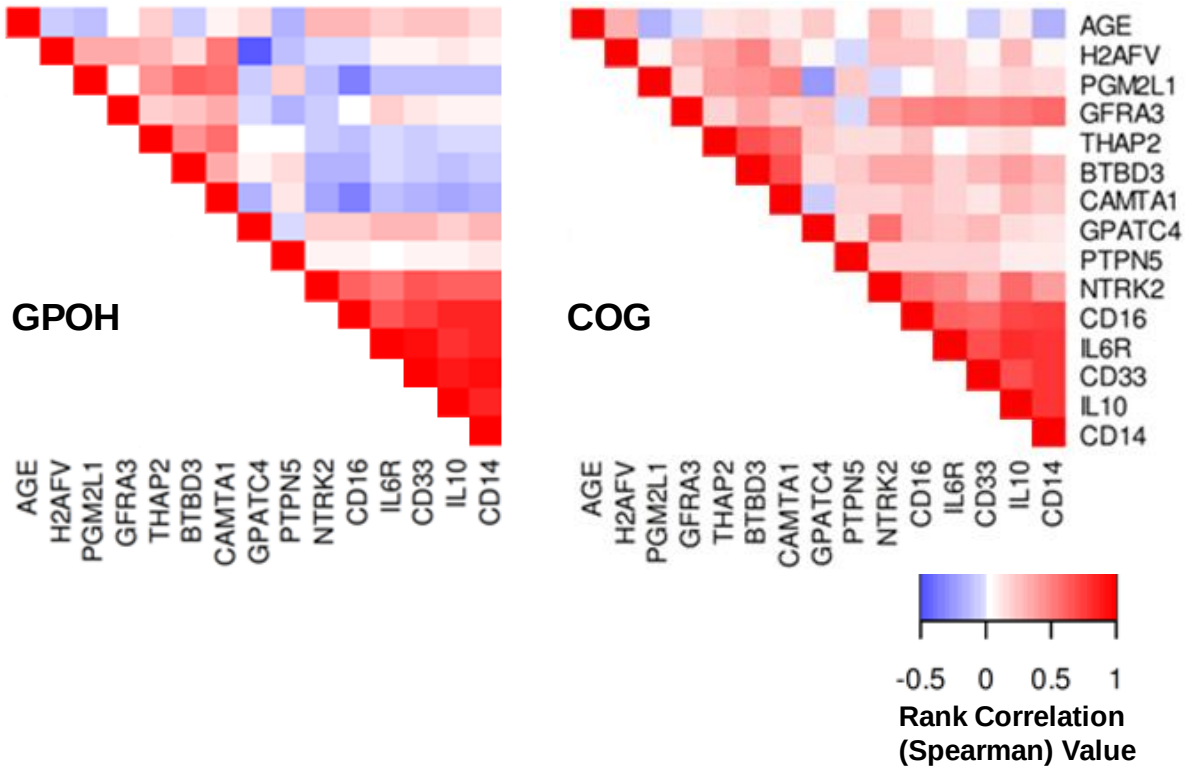
Supplemental Figure 3. Accuracy of Prediction and Progression-Free Survival based on Re-substitution and LOOCV Analysis in the Training Cohort According to the NBL-NA 14-Gene Signature. The accuracy of the NBL-NA 14-gene signature was obtained using the receiver-operating-characteristic curves (ROC) of this model based on its true positive rate (sensitivity) and false positive rate (1-specificity). The measured Area Under the Curve (AUC) value of the ROC curve using re-substitition and leave-one-out cross-validation (LOOCV) tumor-progression scores demonstrates high accuracy of the signature in **A & C)** the entire CCG cohort (n=133) and **B &D)** the CCG patients diagnosed at ≥18 months of age (clinically-defined high-risk group, n=94), respectively. The median re-substitution value of the tumor-progression score of the entire CCG cohort was used to categorize patients into low- (score less than the median value) and high-risk score (score equal or greater than the median value) groups. The re-substitution analysis is known to over-estimate the accuracy of a given model and this bias reflects the need for cross-validation.



Supplemental Figure 4. Progression-Free Survival based on Re-substitution Analysis in the Training Cohort According to the NBL-NA 14-Gene Signature. The graphs show PFS estimates using the 14-gene signature classification scores based on re-substitution analysis for **A)** the entire CCG cohort and **B)** the CCG patients diagnosed at ≥18 months of age (clinically-defined high-risk group, n=94). The re-substitution analysis is known to over-estimate the accuracy of a given model and this bias reflects the need for cross-validation.



Supplemental Figure 5. Overall Survival based on NBL-NA 14-Gene Disease Progression Signature. The median LOOCV tumor-progression score of the entire CCG cohort was used to categorize patients into low- (score less than the median value) and high-risk score (score equal or greater than the median value) groups. The graphs show Kaplan-Meier estimates of overall survival for patients with metastatic neuroblastoma lacking MYCN gene amplification according to 14-gene signature classification scores and treated on **A)** CCG intermediate and high-risk protocols (LOOCV analysis), **B)** the CCG patients diagnosed ≥18 months of age treated on high-risk protocols (LOOCV analysis), and validation cohorts treated on **C)** GPOH high-risk protocols and **D)** COG high-risk protocols.



Supplemental Figure 6. Heatmap of the Rank Correlation Matrix of the NBL-NA 14-Gene Signature. Pairwise rank correlation (Spearman rank correlation) analyses were performed for all genes in the NBL-NA 14-gene signature and age at diagnosis. The patterns of correlation using samples from CCG patients diagnosed ≥18 months of age (Figure 2) were similar to the pairwise rank correlations obtained using samples from the GPOH and COG validation cohorts. The red color represents positive rank correlation level above zero (white color) for a given gene pair and the blue color represents negative rank correlation level. The inflammation-related genes (CD16/FCGR3, CD33, CD14, IL6-R, IL-10) show high levels of correlation across all cohorts. NTRK2 has the strongest correlation of any tumor cell-related gene with inflammation-related genes and in particular with IL-6R expression (all groups combined: Spearman r = 0.58).

Supplemental Table 1. Characteristics of patients used in IHC analysis*

Characteristic		
Age at diagnosis	**<18 months (n=38)**	**≥18 months (n=33)**
Locoregional		
Stage 1	8	2
Stage 2	9	2
Stage 3	3	5
Metastatic (Stage 4)		
MYCN non-amplified	5	7
MYCN amplified	1	2
missing MYCN status	1	15
Stage 4S	11	0

*Analysis was conducted using tissue microarray developed at COG and its Biopathology center as well as cases obtained at Children's Hospital Los Angeles. TMA cores were evaluated by pathologist (HS) to ensure adequate tumor cell content. Ganglioneuromas and ganglioneuroblastomas present on the TMA were excluded from the analyses.

Supplemental Table 2. List of tested genes and their univariate significance of association and predictive accuracy of event, after adjusting for continuous age.

Rank	ABI Probe Name	Gene Symbol	*P* Value from Likelihood Ratio Test	AUC for Training cohort	AUC for Training cohort >= 18 mo at diagnosis
1	Hs00370894_m1	H2AFV	0.0001	0.8308	0.6275
2	Hs00737786_g1	GPATC4	0.0003	0.8262	0.6060
3	Hs00262161_s1	PTPN5	0.0006	0.8133	0.5636
4	Hs00328100_m1	PGM2L1	0.0017	0.8051	0.5127
5	Hs01053355_m1	GNAI1	0.0018	0.8117	0.4971
6	Hs00186495_m1	TMEFF1	0.003	0.8147	0.5258
7	Hs00415042_m1	IGKC;IGKV1	0.0048	0.8145	0.5486
8	Hs00275547_m1	FCGR3A;FCGR3B	0.0058	0.8156	0.5649
9	Hs00326433_m1	NXPH1	0.0212	0.8205	0.5225
10	Hs00275009_s1	CNR1	0.023	0.8062	0.4814
11	Hs00854282_g1	C5orf13	0.0265	0.8090	0.5258
12	Hs00171257_m1	TGFB1	0.0283	0.8051	0.5290
13	Hs00173925_m1	GPR85	0.036	0.8087	0.4873
14	Hs00233544_m1	CD33	0.0392	0.8048	0.5179
15	Hs00171690_m1	HOXC6	0.0487	0.8099	0.4905
16	Hs00181751_m1	GFRA3	0.0501	0.8044	0.4729
17	Hs00230167_m1	THAP2	0.0542	0.7993	0.4703
18	Hs00176787_m1	NTRK1	0.0612	0.7931	0.4442
19	Hs00705034_s1	SMARCE1	0.0617	0.8011	0.4814
20	Hs00174131_m1	IL6	0.0631	0.7966	0.4808
21	Hs00209118_m1	BTBD3	0.0634	0.7936	0.4514
22	Hs00930455_m1	CXCL12	0.0672	0.8023	0.5016
23	Hs00169842_m1	IL6R	0.083	0.8002	0.4977
24	Hs00391998_m1	CAMTA1	0.1034	0.7956	0.4494
25	Hs00383314_m1	C6orf168	0.1406	0.7998	0.4644
26	Hs02621496_s1	CD14	0.1498	0.8028	0.5036
27	Hs00299139_s1	YPEL1	0.201	0.8011	0.4664
28	Hs01016341_m1	ST7	0.2157	0.8016	0.4690
29	Hs00174086_m1	IL10	0.2193	0.7991	0.4814
30	Hs02786786_s1	NTRK2	0.2338	0.8025	0.4886
31	Hs00544818_m1	MS4A1	0.2883	0.7899	0.4553
32	Hs00392922_g1	MYT1L	0.2918	0.7920	0.4827
33	Hs00225656_m1	C1orf35	0.2941	0.7968	0.4710
34	Hs00220252_m1	PARP6	0.3671	0.7982	0.4599
35	Hs00234174_m1	STAT3	0.3716	0.7966	0.4853
36	Hs00174333_m1	CD19	0.3819	0.7908	0.4534
37	Hs00227602_m1	PRG2	0.3949	0.8030	0.4684
38	Hs00705213_s1	HRK	0.4326	0.7938	0.4449
39	Hs00379318_m1	PAK7	0.4394	0.7968	0.4508
40	Hs00365842_m1	CX3CR1	0.7708	0.7979	0.4579
41	Hs00194072_m1	APBA2	0.8091	0.7972	0.4547
42	Hs00289942_s1	LOC284244	0.85	0.7975	0.4540
43	Hs00268388_s1	SOX4	0.9115	0.7968	0.4534
44	Hs00366902_m1	SCN3A	0.9456	0.7966	0.4547

*To develop a clinically-applicable prognostic assay for patients diagnosed ≥18 months of age with metastatic NBL-NA, we used the TaqMan® Low Density Array (TLDA) to obtain qRT-PCR gene expression values of 44 genes in 133 samples from the CCG training cohort. Genes were selected based on available TLDA probes matching the genes identified in our previously published 55-gene microarray signature (24 matching genes with mean spearman correlation of r = 0.61 SD±0.2), previous studies in neuroblastoma, and inflammation-related genes selected based on prior knowledge and identification by gene set enrichment analysis of neuroblastoma microarray data. Thirty-two of the 44 (73%) genes related to tumor cells, while 12 of the 44 (23%) were inflammation-related genes.