Identification of Patient Subgroups with Markedly Disparate Rates of *MYCN* Amplification in Neuroblastoma: A Report from the International Neuroblastoma Risk Group (INRG) Project

Daria Thompson, MD, MPH[1], **Kieuhoa T. Vo, MD**[1], **Wendy B. London, PhD**[2], **Matthias Fischer, MD**[3], **Peter F. Ambros, PhD**[4], **Akira Nakagawara, MD**[5], **Garrett M. Brodeur, MD**[6], **Katherine K. Matthay, MD**[1], and **Steven G. DuBois, MD,MS**[1]

[1]Department of Pediatrics, UCSF Benioff Children's Hospital, University of California, San Francisco School of Medicine, San Francisco, CA, USA

[2]Dana-Farber Children's Hospital Cancer and Blood Disorders Center, Harvard Medical School, Boston, MA, USA

[3]Department of Pediatric Oncology, Children's Hospital and Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany

[4]Children's Cancer Research Institute, St. Anne Kinderkrebsforschung, Vienna, Austria

[5]Department Division of Biochemistry, Chiba Cancer Center Research Institute and Chiba University, Chiba, Japan

[6]Department of Pediatrics, Children's Hospital of Philadelphia and Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Abstract

Purpose—*MYCN* gene amplification (MNA) is a hallmark of aggressive neuroblastoma. We sought to determine the univariate and multivariate predictors of tumor MNA.

Patients and Methods—Data from the International Neuroblastoma Risk Group (INRG) were analyzed from the subset of 7,102 patients with known *MYCN* status. We used chi-squared testing and logistic regression to identify univariate and multivariate predictors of *MYCN* status. Recursive partitioning was used to identify groups of patients with maximal difference in rates of MNA.

Results—All clinical [age >18 months; high ferritin; high lactate dehydrogenase (LDH); INSS stage 4; adrenal site] and pathology/biology [DNA index 1; high MKI, undifferentiated/poorly differentiated grade; unfavorable histology by International Neuroblastoma Pathology Classification (INPC); segmental chromosomal aberrations (SCA)] features were significantly associated with MNA. LDH (OR=8.4; p<0.001) and chromosomal 1p LOH (OR 19.8; p<0.001) respectively were the clinical and biologic variables most strongly associated with *MYCN*-Amp. In

Corresponding Author Current Address: Steven DuBois, MD, Pediatric Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Dana 3, Boston, MA 02215, Telephone: 617-632-5640, Facsimile: 617-632-4811, steven_dubois@dfci.harvard.edu.

logistic regression, all variables except chromosome 17q aberration and pooled SCAs were independently predictive of MNA. Recursive partitioning identified subgroups with disparate rates of MNA, including subgroups with 85.7% MNA [patients with high LDH who had poorly differentiated adrenal tumors with chromosome 1p deletion] and 0.6% MNA [localized tumors with hyperdiploidy, low MKI, and lacking chromosome 1p aberration].

Conclusion—MNA is strongly associated with other clinical and biologic variables in neuroblastoma. Recursive partitioning identifies subgroups of neuroblastoma patients with highly disparate rates of MNA. These findings can be used to inform investigations of molecular mechanisms of MNA.

Keywords

neuroblastoma; neuroblastoma/genetics; gene amplification; oncogene proteins/genetics; decision trees

Introduction

Neuroblastoma (NB) is a childhood tumor with marked clinical and biological heterogeneity.[1,2,3] Stage, age[4], histology[3,5], lactate dehydrogenase (LDH) and ferritin levels[3], *MYCN* status[6,7], ploidy[8,9], segmental chromosomal aberrations (SCA)[3], and primary site[10] have all been shown to be independently prognostic. In particular, *MYCN* amplification (MNA) is associated with rapid disease progression in patients of all ages or stages[2,6,11] and is one of the strongest independent adverse prognostic factors.[3]

MYCN status is also associated with other adverse prognostic factors. For example, MNA is associated with unfavorable histology[5], mitotic karyorrhectic index (MKI)[12], and diploid/tetraploid tumors.[13] MNA is associated with the presence of a range of SCAs, especially 1p deletion[1,14,15], though 11q aberration shows a strong inverse relationship.[16,17,18] The adverse prognostic effect of MNA can supersede otherwise favorable tumor genetics (e.g., near-triploid DNA content).[2] MNA is found in a higher proportion of adrenal primary tumors[10] and its incidence may also differ based upon the involvement of specific metastatic sites.[19,20] Tumor MNA also varies by patient age, though this has been a complex nonlinear relationship in prior reports.[21,22,23] To our knowledge, a comprehensive and definitive analysis of clinical and pathology/biology predictors of MNA in a large cohort of patients has not been reported. Such an analysis has the potential to inform to what extent variables associated with MNA interact, or are independent predictors, and therefore whether or not these associations are stronger in specific patient subgroups.

In the current study, we used the largest available international neuroblastoma dataset to define comprehensively the association of MNA with clinical and pathology/biology factors. We identified variables independently associated with MNA and patient subgroups with significantly different rates of MNA. These findings will provide a valuable resource for the field, will facilitate care of patients with neuroblastoma, and may lead to greater biological insight into the propensity for MNA. In particular, the use of recursive partitioning has the potential to reveal differences in the relative impact of predictors of MNA depending upon specific contexts defined by other variables.

Patients and Methods

Patients

The International Neuroblastoma Risk Group (INRG) database includes 8,800 individual patients <21 years of age with pathologically confirmed neuroblastoma who were diagnosed between January 1, 1990 and December 31, 2002.[3] All patients were consented and enrolled on a neuroblastoma study within INRG member countries (France, Germany, Japan, Italy, Spain, United Kingdom) or within a cooperative group (Children's Oncology Group, SIOPEN LNESG1 study). Only patients with known *MYCN* status were included.

Outcome Variable

MYCN status was the primary dichotomous outcome variable for this analysis. Within each INRG member group or country, MNA was determined by either interphase fluorescence in situ hydridization (I-FISH)[24], polymerase chain reaction (PCR)[25], array-based comparative genomic hybridization (aCGH)[2], or multiplex ligation-dependent probe amplification (MPLA)[2]. For I-FISH, MNA has been defined as a more than 4-fold increase in the *MYCN* signal number compared to the reference probe on chromosome 2q.[2] Data on number of copies of *MYCN* were not available for this analysis.

Predictor Variables

The INRG collects data on the following predictor variables used in this analysis: age; International Neuroblastoma Staging System (INSS) stage[26,27]; ploidy; presence of SCAs (1p, 11q, or 17q); ferritin level; LDH; primary site; sites of metastasis; histology (favorable vs. unfavorable by International Neuroblastoma Pathology Classification [INPC] system[5]); tumor diagnostic category (neuroblastoma vs ganglioneuroblastoma [intermixed, maturing, well-differentiated, or nodular]); grade (differentiating vs. poorly or undifferentiated); and MKI. Each continuous predictor variable was transformed into a binary variable for the analysis. Ferritin was categorized as high (92 ng/mL) versus low (<92 ng/mL) and LDH was categorized as high (587 U/L) versus low (<587 U/L), based on the mean values of ferritin and LDH in the overall INRG cohort, as previously reported[3]. Tumor primary site was categorized as thoracic vs. non-thoracic and adrenal vs. non-adrenal primary site. These site categories were chosen based on our previous work demonstrating the greatest effects of these sites on rates of MNA and because they are the most common primary sites[10]. We created a binary variable to reflect the presence of any SCA, defined as having 1p, 11q, and/or 17q aberrations.

Statistical Analyses

We used chi-squared testing to perform univariate analyses comparing the frequency of MNA as a function of each categorical predictor variable. We used a t-test to compare the distribution of ages as a continuous variable between patients with and without MNA.

We evaluated appropriate variables for inclusion in multivariate testing (logistic regression and recursive partitioning). Most potential predictor variables had missing data for >10% of patients; predictor variables ranged from 0% missing values (age) to 95.2% (17q chromosomal aberration). Within each predictor variable, we performed an assessment to

determine if *MYCN* status was missing at random. The proportion of patients with MNA was calculated for the favorable category (decreased rates of MNA expected), the unfavorable category (increased rates of MNA expected), and the missing category. If *MYCN* status was missing at random, we expected the proportion of MNA in patients with missing data for the variable to fall between the proportions seen in the favorable and unfavorable groups. Tumor diagnosis category violated this pattern and was excluded from multivariate models.

INPC histology classification system incorporates age as well as grade and MKI. To avoid confounding of INPC histology and age in multivariate models, only the underlying components (grade, age, and MKI) were included.

We performed multivariate logistic regression to select variables independently associated with MNA. Excluding 17q chromosomal aberration, only 188/7102 patients had known data for all other predictors. To utilize the full sample size of the dataset without selection bias and allow inclusion of patients for whom some predictor variables had missing values, we created a series of "dummy" variables that fully describe each predictor variable as unfavorable (yes/no), favorable (yes/no), and missing (yes/no). The "dummy" variables for unfavorable and missing categories were included in the model, leaving the favorable category as the reference. This approach prevents bias that might occur if only the subset of patients with complete data was utilized in the model.[28,29] We used a backward selection approach with p<0.05 to enter the model and p<0.05 to remain in the final model. Three multivariate logistic regression models were built: Model A - testing all covariates; Model B – testing clinical variables only; and, Model C – testing pathology/biology variables only. Model A was repeated including a binary variable for vital status (alive vs. dead), to see which baseline variables of interest remain associated with MNA after adjustment for vital status.

We used the classification and regression tree (CART) analysis method to identify age cut points associated with lower and higher rates of MNA.[30] In this analysis, age as a continuous variable was the only predictor variable and MNA was the outcome variable. The first three (most highly predictive) age cut points from CART were eligible for selection in future models.

For multivariate recursive partitioning models, we used the univariate odds ratios for MNA to prioritize variables for selection. Each predictor variable was tested within the overall cohort and the statistically significant predictor with the largest odds ratio was selected manually to form the split. Each split created two nodes; odds ratios were recalculated and the process was repeated within each node. This process proceeded iteratively until either or both of the following pre-specified conditions were satisfied: 1) no remaining variables with a statistically significant (p<0.05) odds ratio; and/or 2) further split yielded a subgroup with fewer than 15 total patients.

We utilized CART in R (http://www.R-project.org/) for recursive partitioning using age as the only predictor variable. STATA version 12 (StataCorp, College Station, TX) was used for all other analyses.

Results

Patient Characteristics

The characteristics of the 7,102 patients with known *MYCN* status are shown in Table 1 and are similar to the full INRG cohort reported previously.[3] MNA was reported in 1,155 patients in our cohort (16.3%).

Incidence of MNA Varies with Age at Diagnosis

Patients with *MYCN* amplified tumors were older compared to patients with *MYCN* non-amplified tumors (mean age 28.1 versus 24.3 months, p<0.001). A higher percentage of patients 18 months had MNA compared to patients <18 months (24.7% versus 9.9%); in addition, 65% of patients with MNA were 18 months of age. To further characterize the relationship of MNA with age, we analyzed the proportion of total cases with and without MNA as a function of age (Figure 1A). Patients 18–20 months old contributed the highest proportion of cases with MNA (9.3% of all cases with MNA). In contrast, patients 0–2.9 months old contributed the highest proportion of cases without MNA (17.6% of all cases without MNA). We then plotted the percent of patients in a given age interval with MNA (Figure 1B). The 3-month interval with the highest incidence of MNA was in patients 21–23.9 months old (38.4%) and the lowest incidence was in patients 0–2.9 months old (3.6%).

We used CART to identify optimal age cut points associated with the most disparate rates of MNA (Supplemental Figure 1). The top three age cutoffs identified by recursive partitioning were 367 days (12 months), 110 days (3.5 months), and 1282 days (42 months). Patients <110 days had an incidence of MNA of 3.6%, whereas patients 516–1086 days (17–35.5 months) had an incidence of MNA of 33.5%.

Additional Clinical and Pathology/Biology Features Are Associated with MNA

All clinical and pathology/biology variables were significantly associated with *MYCN* status in univariate analyses (Table 1). The clinical factor most strongly associated with MNA was high LDH [odds ratio (OR) 8.4; p<0.001], though multiple other variables had OR>4. The pathology/biology variable most strongly associated with MNA was LOH at 1p [OR 19.8; p<0.001], though multiple other variables had OR>10.

We created three separate logistic regression models (Table 2). When building a model that tested all clinical and pathology/biology variables, we demonstrated that all variables except gain of 17q and pooled SCAs were independently associated with MNA (Model A). After the inclusion of vital status, the results were similar to Model A except age was no longer independently associated with MNA (data not shown). Model B tested only clinical variables, and demonstrated that all clinical variables were independently associated with MNA. Model C tested only pathology/biology variables, and demonstrated that all pathology/biology variables except gain of 17q and pooled SCAs were independent predictors of MNA. In Model C, the point estimate for the OR for 1p LOH was similar to the univariate OR (17.8 from Table 2 model C vs. 19.8 from Table 1).

Recursive Partitioning Identifies Patient Subgroups with Disparate Rates of MNA

In multivariate recursive partitioning using only clinical variables (Figure 2A), we identified two extreme patient subgroups. The first subgroup with low LDH level, age <3.5 months, and non-stage 4 were *MYCN* amplified in 1.3%. In contrast, patients with high LDH, non-thoracic sites, and age >12 months but <42 months were *MYCN* amplified in 50.7%. The tree continued to split beyond these groups, but we made a *post hoc* decision to truncate the model after 5 splits to increase the practicality of the model (full tree with all splits shown in Supplemental Figure 2).

Our second tree utilized only pathology/biology variables (Figure 2B) and two extreme subgroups were identified. Tumors without 1p LOH and low MKI had a rate of MNA of 3.2%. In contrast, tumors with 1p LOH that were poorly differentiated, lacked 11q aberration and were diploid were *MYCN* amplified in 87.5%.

In our final recursive partitioning tree, all clinical and pathology/biology variables were available for selection (Figure 2C). This tree showed that non-stage 4 patients with low MKI, hyperdiploid tumors lacking 1p LOH were *MYCN* amplified in 0.6%. In contrast, patients with high LDH and adrenal primary tumors that were poorly differentiated and had 1p LOH were *MYCN* amplified in 85.7%. In this final tree with all variables available for selection, age was not chosen as a significant variable.

Metastatic Sites Correlate with *MYCN* Status

We repeated univariate analyses among the 2,176 patients with stage 4 disease and at least one known specific metastatic site to assess whether metastatic sites are associated with MNA (Table 3). Lung metastasis was the site most strongly associated with MNA (OR 3.0; p<0.001). Bone marrow and bone metastases were also significantly associated with MNA (OR 1.4 and 1.3, respectively; p<0.01), while skin metastases had lower likelihood of MNA (OR 0.4; p=0.01).

In this same group, a multivariate logistic regression model, testing all of the potential clinical and pathology/biology variables previously evaluated, demonstrated that LOH at 1p, high LDH, absence of 11q aberration, high MKI, non-thoracic site, lung metastases, diploid/hypodiploid tumors, absence of skin metastases, and adrenal site were independently significantly associated with MNA (Table 4).

Discussion

In this comprehensive analysis of predictors of *MYCN* amplified status, we demonstrate in a comprehensive and definitive manner the complex interaction between MNA and other features of this disease. The presence of each prognostic factor was statistically associated with *MYCN* status, with most remaining significant on multivariate testing. This novel finding demonstrating the independent association between MNA and almost all variables was unexpected. The result of our multivariable logistic regression model adjusting for vital status demonstrates that our findings cannot be fully explained by an association between MNA and prognosis. Pathology/biology variables demonstrated the strongest associations with MNA. Tumors with 1p LOH were almost 20 times as likely to have MNA compared to

tumors without 1p LOH, with a similar association even after controlling for other biological predictors. This is consistent with prior observations[1,14,15], and in context with our other findings, highlights the critical association between MNA and tumor genetic features compared to clinical features. For example, of the clinical variables, elevated LDH was the strongest predictor with a univariate odds ratio of 8.4 compared to 19.8 for 1p LOH.

Our novel recursive partitioning approach revealed dramatic differences in the rates of MNA between identified subgroups of patients, including two trees that yielded >80% absolute differences in rates of MNA between subgroups. In addition to highlighting the importance of SCA in prediction of MNA, this approach also enables us to illustrate the relative importance of some variables over others in predicting MNA as well as the importance of context of other predictor variables in the impact of a specific variable. For example, in the absence of LOH at 1p, the presence of high MKI is associated with a maximum incidence of MNA of 25% whereas, in the presence of LOH at 1p (along with other features), the presence of high MKI is associated with a maximum incidence of MNA of 71.4%. Moreover, among poorly differentiated tumors with LOH at 1p, aberration at chromosome 11q is associated with MNA, though the difference in maximum incidence of MNA based upon presence or absence of this aberration is not large (maximum possible incidence for those without 11q aberration of 87.5% compared to 71.4% for those with 11q aberration).

Along these same lines, it is noteworthy that age was a key predictor of differential rates of MNA in models relying solely on clinical variables, but age was not selected as a predictive variable in the recursive partitioning model that included all clinical and biologic/pathologic variables as potential predictors. This finding, along with the results of our multivariate logistic regression models, suggest that other variables associated with age account in part for some of the association between *MYCN* status and age, which fits with the concept that age behaves as a surrogate for the effects of other clinical and biological variables. Graphical representations of the relationship of age with MNA show a complex relationship. London and colleagues showed that any age cutoff between 15–20 months appropriately separates neuroblastoma patients into high and low risk groups for treatment decisions, with 18 months selected for future risk stratification.[23] We complement these findings by showing that MNA peaks in incidence at approximately this same age. Our results demonstrate the nonlinear relationship between age and MNA and should motivate additional studies into the developmental and molecular pathways involved in this association.

Our findings extend previous observations demonstrating positive associations between MNA and 1p LOH and gain of 17q, as well as a negative association between MNA and 11q aberration.[2,14–16,31,32] Our multivariate logistic regression model focused solely on pathology/biology variables showed that the 1p and 11q associations with *MYCN* status were independent of each other, while the 17q association was not. This finding may be due to the fact that 17q and 1p aberrations are themselves correlated,[2] or more likely reflect small sample size with available 17q data. The composite variable that included 1p, 11q, and/or 17q segmental aberrations was also not independently associated with *MYCN* status, likely due to opposing effects of 1p and 11q aberrations on the incidence of MNA. The overall pattern observed in this and other studies[33] is noteworthy as it indicates that specific

loci rather than non-specific SCAs are associated with the presence or absence of MNA, though the underlying biologic mechanisms for these associations are yet unknown.

Our results address the association between MNA and sites of primary and metastatic disease. We previously reported the association of primary site with *MYCN* status.[10] We have extended that finding to demonstrate that adrenal tumors are more likely and thoracic tumors are less likely to be *MYCN* amplified even after controlling for other predictors of *MYCN* status. Likewise, we have previously demonstrated that patients with lung metastases are more likely to have *MYCN* amplified tumors.[20] We now demonstrate that this association is independent of other predictors of *MYCN* status, though the underlying mechanisms for this association remain obscure. Unlike previous analyses, we did not find an association between MNA and central nervous system metastases, though this may be due to the rarity of this metastatic site at diagnosis.[34] We also demonstrate that skin metastases are predictive of being *MYCN* non-amplified. This site of disease is more common in stage 4S disease and also in young infants, groups which are less likely to have *MYCN* amplified tumors.[35]

Although we utilized the largest available patient database, with 7,102 patients with known *MYCN* status, our work has certain limitations. Our overall rate of MNA of 16% is lower than widely cited estimates of 20–25%.[1] It is possible that the comprehensive nature of the INRG database provides a more accurate estimate of the rate of MNA across neuroblastoma. We note that our estimate is identical to that reported by a large analysis from the COG[16], though acknowledge that data from the COG are included within the INRG database and may bias the estimate towards those previously reported by the COG. It is also possible that the contribution of low-risk patients identified from national neuroblastoma screening efforts and/or higher rates of *MYCN* testing in localized patients enriched the population with biologically favorable tumors thereby reducing the proportion with MNA. Although our primary outcome, *MYCN* status, was not determined by identical techniques in each INRG member country, all those used are standard validated methods[2,24,25]. We chose to dichotomize LDH and ferritin based upon the median of the initial INRG cohort to maintain consistency with prior INRG analyses, though acknowledge that other cutpoints could be explored in future analysis for their association with *MYCN* status. The multivariate logistic regression analyses were limited by missing data. We attempted to circumvent this issue by utilizing dummy variables to allow the full data set to be used, thereby reducing the risk of selection bias. This approach could not be used as part of our recursive partitioning models and therefore certain variables with limited data, such as gain of 17q, may have been less likely to be selected, thus potentially decreasing the generalizability of our findings. Although the differences in patient subgroups revealed by recursive partitioning are innovative and can both inform insights into the biological and pathological factors that lead to *MYCN* amplification and lead to further research in the field, some of the resulting subgroups are small and therefore one must use caution in interpreting these subgroups for clinical decisions.

Despite these limitations, the current study is the largest analysis of predictors of MNA. Our work highlights the importance of obtaining adequate tissue for detailed molecular testing as part of the diagnostic process. While we identified groups of patients with very low

probabilities of MNA, no subgroups that completely lack MNA could be identified. Moreover, other pathology/biology variables, particularly SCAs, were key to identifying groups of patients with maximally different rates of MNA and improved our stratification beyond one restricted to the use of only clinical variables. Our findings should stimulate additional laboratory studies into the mechanisms of MNA as a function of age, as well as the genetics or epigenetics that may be orchestrating the complex interaction between the described molecular findings. Finally, additional studies exploring the ways that molecular changes associated with MNA interact with clinical features associated with MNA will be critical.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Brodeur GM, Bagatell R. Mechanisms of neuroblastoma regression. Nat Rev Clin Oncol. 2014; 11(12):704–713. [PubMed: 25331179]

2. Ambros PF, Ambros IM, Brodeur GM, et al. International consensus for neuroblastoma molecular diagnostics: report from the International Neuroblastoma Risk Group (INRG) Biology Committee. BMJ Cancer. 2009; 100:1471–1482.

3. 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; 27:289–297. [PubMed: 19047291]

4. Carlsen NL, Christensen IJ, Schroeder H, et al. Prognostic factors in neuroblastomas treated in Denmark from 1943 to 1980. A statistical estimate of prognosis based 253 cases. Cancer. 1986; 58:2726–2735. [PubMed: 3779618]

5. Shimada H, Ambros IM, Dehner LP, et al. The international neuroblastoma pathology classification (the Shimada system). Cancer. 1999; 86:364–372. [PubMed: 10421273]

6. Brodeur GM, Seeger RC, Schwab M, et al. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. Science. 1984; 224:1121–1124. [PubMed: 6719137]

7. George RE, Variend S, Cullinane C, et al. Relationship between histopathological features, MYCN amplification, and prognosis: A UKCCSG study. Med Pediatr Oncol. 2001; 36:169–176. [PubMed: 11464876]

8. Oppedal BR, Storm-Mathison I, Lie SO, et al. Prognostic factors in neuroblastoma: Clinical, histopathologic, and immunohistochemical features and DNA ploidy in relation to prognosis. Cancer. 1988; 62(4):772–780. [PubMed: 3395959]

9. Look AT, Hayes FA, Shuster JJ, et al. Clinical relevance of tumor cell ploidy and N-myc gene amplification in childhood neuroblastomas: a Pediatric Oncology Group study. J Clin Oncol. 1991; 9:581–591. [PubMed: 2066755]

10. Vo K, Matthay KK, Neuhaus J, et al. Clinical, biological, and prognostic differences based upon primary tumor site in neuroblastoma: a report from the International Neuroblastoma Risk Group (INRG) project. J Clin Oncol. 2014; 32:3169–3176. [PubMed: 25154816]

11. Seeger RC, Brodeur GM, Sather H, et al. Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. N Engl J Med. 1985; 313:1111–1116. [PubMed: 4047115]

12. Teshiba R, Kawano S, Wang L, et al. Age-dependent prognostic effect by mitosis-karyorrhexis index in neuroblastoma: A report from the children's oncology group. Pediatr Dev Pathol. 2014; 17(6):441–449. [PubMed: 25207821]

13. Bowman LC, Castleberry RP, Cantor A, et al. Genetic staging of unresectable or metastatic neuroblastoma in infants: a Pediatric Oncology Group study. J Natl Cancer Inst. 1997; 89(5):373–380. [PubMed: 9060959]

14. Fong CT, Dracopoli NC, White PS, et al. Loss of heterozygosity for the short arm of chromosome 1 in human neuroblastomas: correlation with N-myc amplification. Proc Natl Acad Sci U S A. 1989; 86(10):3753–3757. [PubMed: 2566996]

15. Maris JM, Guo C, Blake D, et al. Comprehensive analysis of chromosome 1p deletions in neuroblastoma. Med Pediatr Oncol. 2001; 36(1):32–36. [PubMed: 11464900]

16. Attiyeh EF, London WB, Mosse YP, et al. Chromosome 1p and 11q deletions and outcome in neuroblastoma. N Engl J Med. 2005; 353(21):2243–2253. [PubMed: 16306521]

17. Plantaz D, Vandesompele J, Van Roy N, et al. Comparative genomic hybridization (CGH) analysis of stage 4 neuroblastoma reveals high frequency of 11q deletion in tumors lacking MYCN amplification. Int J Cancer. 2001; 91(5):680–686. [PubMed: 11267980]

18. Spitz R, Hero B, Simon T, et al. Loss in chromosome 11q identifies tumors with increased risk for metastatic relapses in localized and 4S neuroblastoma. Clin Cancer Res. 2006; 12:3368–3373. [PubMed: 16740759]

19. Morgenstern DA, London WB, Stephens D, et al. Metastatic neuroblastoma confined to distant lymph nodes (stage 4N) predicts outcome in patients with stage 4 disease: a study from the International Neuroblastoma Risk Group database. J Clin Oncol. 2014; 32(12):1228–1235. [PubMed: 24663047]

20. DuBois SG, London WB, Zhang Y, et al. Lung metastases in neuroblastoma at initial diagnosis: a report from the international neuroblastoma risk group (INRG) project. Pediatr Blood Cancer. 2008; 51:589–592. [PubMed: 18649370]

21. Mosse YP, Deyell RJ, Berthold F, et al. Neuroblastoma in older children, adolescents and young adults: a report from the International Neuroblastoma Risk Group project. Pediatr Blood Cancer. 2014; 61:627–635. [PubMed: 24038992]

22. Cetinkaya C, Martinsson T, Sandgren J, et al. Age dependence of tumor genetics in unfavorable neuroblastoma: arrayCGH profiles of 34 consecutive cases, using a Swedish 25-year neuroblastoma cohort for validation. BMC Cancer. 2013; 13:231–242. [PubMed: 23656755]

23. London WB, Castleberry RP, Matthay KK, et al. Evidence for an age cutoff greater than 365 days for neuroblastoma risk stratification in the Children's Oncology Group. J Clin Oncol. 2005; 23:6459–6465. [PubMed: 16116153]

24. Shapiro DN, Valentine MB, Rowe ST, et al. Detection of N-myc gene amplification by fluorescence in situ hybridization: diagnostic utility for neuroblastoma. Am J Pathol. 1993; 142(5):1339–1346. [PubMed: 7684192]

25. Crabbe DC, Peters J, Seeger RC. Rapid detection of MYCN gene amplification in neuroblastoma using the polymerase chain reaction. Diagn Mol Pathol. 1992; 1(4):229–234. [PubMed: 1342970]

26. Brodeur GM, Seeger RC, Barrett A, et al. International criteria for diagnosis, staging and response to treatment in patients with neuroblastoma. J Clin Oncol. 1988; 6:1874–1881. [PubMed: 3199170]

27. 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; 11:1466–1477. [PubMed: 8336186]

28. Machin, D.; Cheung, Y-B.; Parmar, MKB. Survival Analysis: A Practical Approach. Chichester: Wiley; 2006.

29. Moroz V, Machin D, Faldum A, et al. Changes over three decades in outcome and the prognostic influence of age-at-diagnosis in young patients with neuroblastoma: A report from the International Neuroblastoma Risk Group Project. Eur J Cancer. 2011; 47(4):561–571. [PubMed: 21112770]

30. Cook EF, Goldman L. Empiric comparison of multivariate analytic techniques: advantages and disadvantages of recursive partitioning techniques. Journal of Chronic Dis. 1984; 37(9):721–731.

31. Maris JM, Weiss MJ, Guo C, et al. Loss of heterozygosity as chromosome 1p36 predicts for disease progression but no decreased overall survival probability in neuroblastoma patients: a Children's Cancer Group study. J Clin Oncol. 2000; 18(9):1888–1899. [PubMed: 10784629]

32. Guo C, White PS, Weiss MJ, et al. Allelic deletion at 11q23 is common in MYCN single copy neuroblastomas. Oncogene. 1999; 18(35):4948–4957. [PubMed: 10490829]

33. Mosse YP, Diskin SJ, Wasserman N, et al. Neuroblastomas have distinct genomic DNA profiles that predict clinical phenotype and regional gene expression. Genes Chromosomes Cancer. 2007; 46(10):936–949. [PubMed: 17647283]

34. Matthay KM, Brisse H, Couanet D, et al. Central nervous system metastases in neuroblastoma: radiologic, clinical, and biologic features in 23 patients. Cancer. 2003; 1(98):155–165. [PubMed: 12833468]

35. Taggart DR, London WB, Schmidt ML, et al. Prognostic value of the stage 4S metastatic pattern and tumor biology in patients with metastatic neuroblastoma diagnosed between birth and 18 months of age. J Clin Oncol. 2011; 29(33):4358–4364. [PubMed: 21969516]

A.
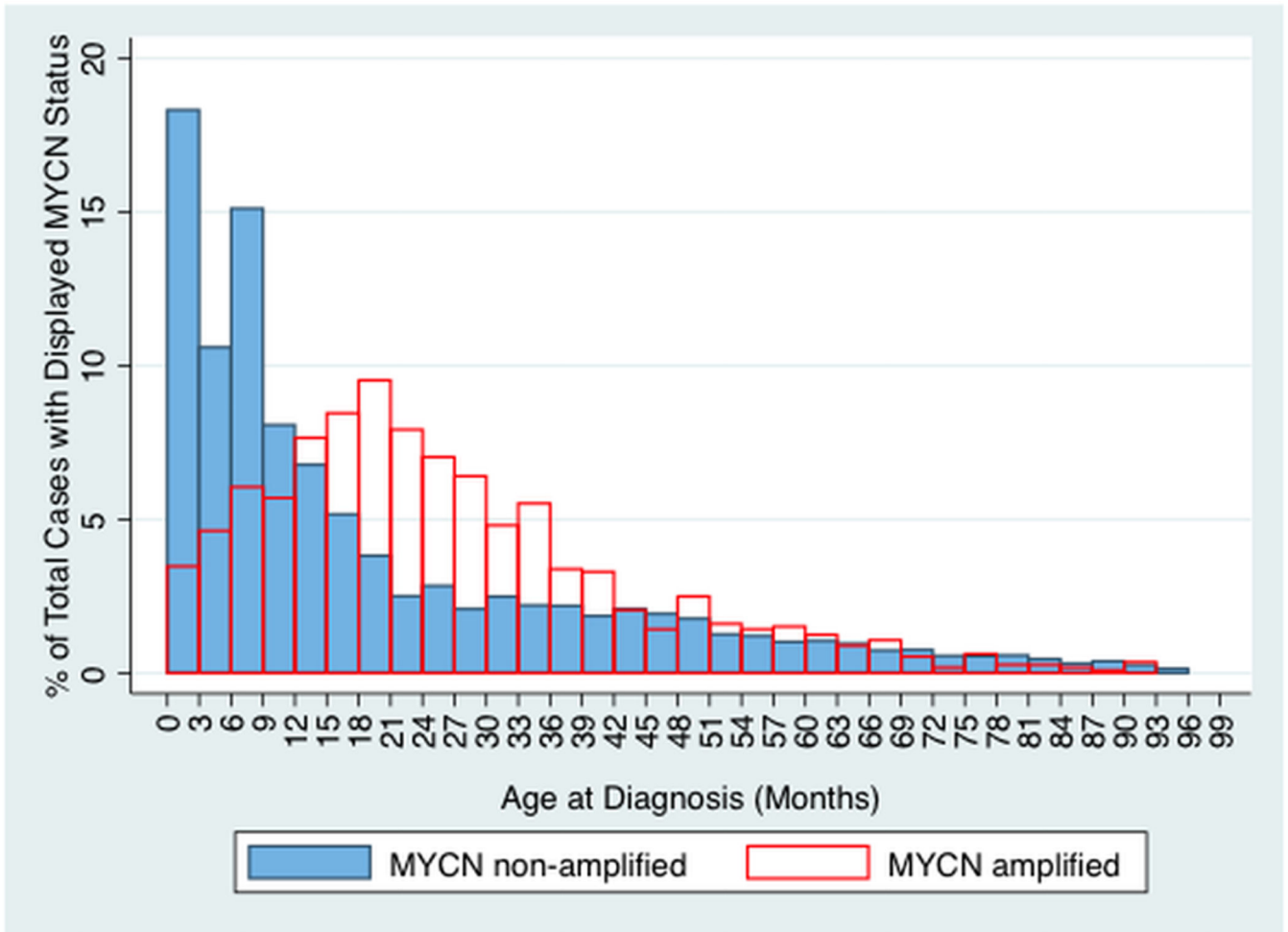
B.
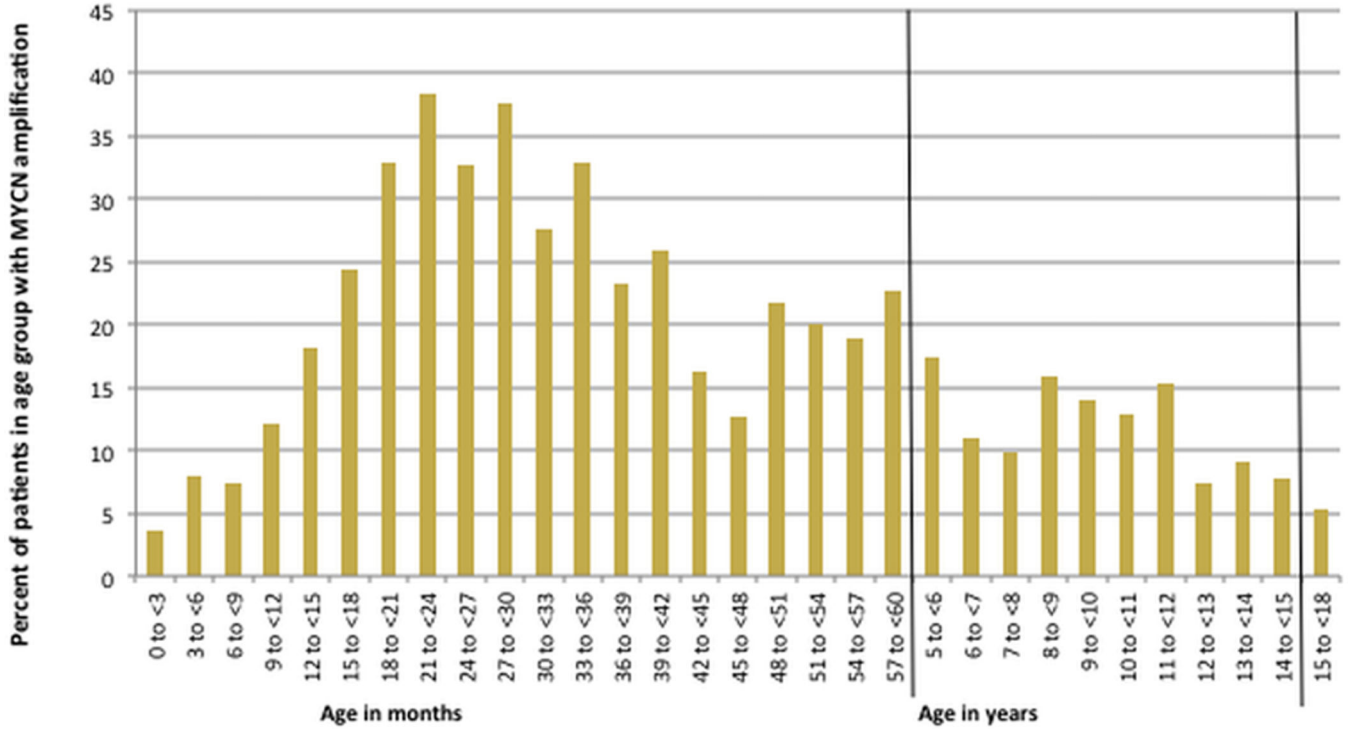
Figure 1.
A. Two-way histogram comparing the proportion of total cases with *MYCN* amplification to those without, by age (sum of bars for each histogram = 100%).
B. Histogram showing percent of patients within a given age group that have *MYCN*-amplified tumors.

A.

Total: 4,350
MYCN-A = 16.6%

LDH Low
N = 2,234
MYCN-A = 4.7%

OR: 8.4

LDH High
N = 2,116
MYCN-A = 29.1%

OR: 3.3

OR: 5.6

Age ≤ 3.5 mo
N = 411
MYCN-A = 1.7%

Age > 3.5 mo
N = 1,814
MYCN-A = 5.4%

Thoracic Primary
N = 232
MYCN-A = 5.6%

Non-Thoracic Primary
N = 1,872
MYCN-A = 32.1%

OR: 6.7

OR: 2.1

OR: 4.3

OR: 3.3

Not Stage 4
N = 384
MYCN-A = 1.3%

Stage 4
N = 25
MYCN-A = 8.0%

Ferritin Low
N = 921
MYCN-A = 3.8%

Ferritin High
N = 310
MYCN-A = 7.4%

Age ≤ 18 mo
N = 148
MYCN-A = 2.7%

Age > 18 mo
N = 84
MYCN-A = 10.7%

Age ≤ 3.5 mo
N = 244
MYCN-A = 9.8%

OR: 2.6

OR: 3.5

Not Stage 4
N = 733
MYCN-A = 3.0%

Stage 4
N = 173
MYCN-A = 7.5%

Age > 3.5 mo
N = 1,626
MYCN-A = 35.5%

Not Adrenal Primary
N = 95
MYCN-A = 4.2%

Adrenal Primary
N = 149
MYCN-A = 13.4%

OR: 4.6

OR: 3.8

Age ≤ 42 mo
N = 125
MYCN-A = 4.0%

Age > 42 mo
N = 49
MYCN-A = 16.3%

Age ≤ 12 mo
N = 508
MYCN-A = 17.1%

Age > 12 mo
N = 1,118
MYCN-A = 43.8%

OR: 6.8

OR: 2.5

Not Stage 4
N = 349
MYCN-A = 8.3%

Stage 4
N = 150
MYCN-A = 38.0%

Age ≤ 42 mo
N = 757
MYCN-A = 50.7%

Age > 42 mo
N = 360
MYCN-A = 29.4%

B.

```
                              Total: 2,029
                              MYCN-A = 16.1%

                           OR: 19.8

        No Chrom 1p LOH                        Chrom 1p LOH
        N = 1,581                              N = 448
        MYCN-A = 5.5%                          MYCN-A = 53.7%

     OR:                                    OR:
     10.0                                   10.0

  Low MKI          High MKI      Differentiating    Poorly Differentiated
  N = 906          N = 56        N = 24             N = 259
  MYCN-A = 3.2%    MYCN-A = 25.0% MYCN-A = 12.5%    MYCN-A = 58.9%

                                             OR:
                                             4.4

                          NoChrom 11q Ab              Chrom 11q Ab
                          N = 116                     N = 29
                          MYCN-A = 66.7%              MYCN-A = 31.0%

                       OR:                         OR:
                       10.5                        11.7

                  Hyperdiploid   Diploid      Low MKI         High MKI
                  N = 15         N = 24       N = 16          N = 7
                  MYCN-A = 40.0% MYCN-A = 87.5% MYCN-A = 17.7% MYCN-A = 71.4%
```

C.

```
                              Total: 2,029
                              MYCN-A = 16.1%

                           OR: 19.8

        No Chrom 1p LOH                        Chrom 1p LOH
        N = 1,581                              N = 448
        MYCN-A = 5.5%                          MYCN-A = 53.7%

        OR: 10.0                                    OR:
                                                    9.9

  Low MKI          High MKI      Not Thoracic Primary    Thoracic Primary
  N = 906          N = 56        N = 417                 N = 26
  MYCN-A = 3.2%    MYCN-A = 25.0% MYCN-A = 56.5%         MYCN-A = 11.5%

     OR:                                  OR:
     3.9                                  10.2

  Not Stage 4      Stage 4       Differentiating     Poorly Differentiated
  N = 714          N = 166       N = 23              N = 245
  MYCN-A = 2.1%    MYCN-A = 7.8% MYCN-A = 13.0%      MYCN-A = 60.6%

     OR:                                          OR: 9.0
     9.0

  Hyperdiploid     Diploid            Low LDH            High LDH
  N = 333          N = 56             N = 34             N = 108
  MYCN-A = 0.6%    MYCN-A = 5.3%      MYCN-A = 23.5%     MYCN-A = 73.4%

                                                      OR:
                                                      5.7

                              Not Adrenal Primary    Adrenal Primary
                              N = 38                 N = 70
                              MYCN-A = 51.3%         MYCN-A = 85.7%
```

Figure 2.

A. Manual recursive partitioning using logistic regression to identify factors associated with *MYCN* amplification, testing only clinical variables (tree truncated at 5 levels of variables; see Supplemental Figure 2 for full tree). Odds Ratio (OR) listed refers to odds ratio for *MYCN* amplification for the variable that resulted in each node (e.g. OR 8.4 indicates patients with high LDH had 8.4 fold higher odds of *MYCN* amplification compared to patients with low LDH).

B. Manual recursive partitioning using logistic regression to identify factors associated with *MYCN* amplification, testing only pathology/biology variables. Variables tested and not included in the final tree were 17q gain and pooled segmental chromosomal aberrations.

C. Manual recursive partitioning using logistic regression to identify factors associated with *MYCN* amplification, testing all potential predictor variables. Variables tested and not included in the final tree were: age; ferritin; 17q gain; pooled segmental chromosomal aberrations; and 11q aberration.

Table 1

Clinical and pathology/biology features of 7,102 patients with neuroblastoma according to *MYCN* amplification status.

Clinical features	Sample size (N)[1]	All (%)	*MYCN* Non-amplified (%)[2]	*MYCN* Amplified (%)[2]	P-value[3]	Odds Ratio[4]
Mean age at diagnosis, months	7102	25.1	24.3	28.1	<0.001	N/A
Age at diagnosis	7102				<0.001	3.0
18 months		42.8	75.3	24.7		
< 18 months		57.2	90.1	9.9		
INSS Stage	7003	36.9	69.6	30.4	<0.001	5.1
4						
Not Stage 4		63.1	92.1	7.9		
INSS Stage	7003				<0.001	N/A
1		23.9	97.1	2.9		
2		14.1	96.0	4.0		
3		17.6	82.4	17.6		
4S		7.5	91.1	8.9		
4		36.9	69.6	30.4		
Primary tumor site	6815				<0.001	2.5
Adrenal[5]		45.9	77.1	22.9		
Non-Adrenal		54.1	89.2	10.8		
Primary tumor site	6815				<0.001	7.1
Non-Thoracic		84.7	81.2	18.8		
Thoracic[5]		15.3	96.9	3.1		
Lactate dehydrogenase (U/L)	4354				<0.001	8.4
587		48.6	70.9	29.1		
< 587		51.4	95.3	4.7		
Serum ferritin (ng/mL)	3602				<0.001	4.2

Clinical features	Sample size (N)[1]	All (%)	MYCN Non-amplified (%)[2]	MYCN Amplified (%)[2]	P-value[3]	Odds Ratio[4]
92		48.8	74.0	26.0		
<92		51.2	92.2	7.8		
Pathology/Biology features						
International Neuroblastoma Pathology Classification	3866				<0.001	14.8
Unfavorable		35.0	66.7	33.3		
Favorable		65.0	96.7	3.3		
Ploidy	3516				<0.001	3.9
1 (diploid, hypodiploid)		29.2	71.2	28.8		
>1 (hyperdiploid)		70.8	90.5	9.5		
Tumor diagnosis[7]	3349				<0.001	10.1
Neuroblastoma or ganglioneuroblastoma, nodular		95.3	83.6	16.4		
Ganglioneuroblastoma (except nodular) or ganglioneuroma (maturing subtype)		4.7	98.1	1.9		
Tumor diagnosis (N=3349)[7]	3349				<0.001	9.3
Neuroblastoma		89.1	82.7	17.3		
Ganglioneuroblastoma or ganglioneuroma (maturing subtype)		10.9	97.8	2.2		
Grade of differentiation	2871				<0.001	4.6
Undifferentiated/Poorly differentiated		83.8	82.1	17.9		
Differentiating		16.2	95.5	4.5		
MKI	2738				<0.001	16.8
High		11.8	39.6	60.4		
Low/Intermediate		88.2	91.7	8.3		
LOH at 1p	2026				<0.001	19.8

Clinical features	Sample size (N)[1]	All (%)	MYCN Non-amplified (%)[2]	MYCN Amplified (%)[2]	P-value[3]	Odds Ratio[4]
Present		22.2	46.3	53.7		
Absent		77.8	94.5	5.5		
Pooled segmental chromosomal aberration[6]	2060				<0.001	10.4
Present		31.7	61.5	38.5		
Absent		68.3	94.3	5.7		
Aberration at 11q	1033				0.001	0.4
Present		20.9	91.7	8.3		
Absent		79.1	82.1	17.9		
Gain of 17q	340				0.01	2.1
Present		47.7	73.5	26.5		
Absent		52.3	84.8	15.2		

[1] N represents total with known data for the given variable.

[2] Row percentages sum to 100%.

[3] P-value refers to a student's T-test (for continuous age variable) or a chi-squared test for all other variables.

[4] Odds ratio for univariate logistic regression with MYCN amplification as dependent variable.

[5] Primary tumor site [adrenal or thoracic] versus all other primary tumor sites.

[6] LOH at 1p, gain of 17q, and/or aberration at 11q.

[7] International Neuroblastoma Pathology Classification tumor diagnosis category (includes neuroblastoma, ganglioneuroblastoma intermixed, ganglioneuroma maturing subtype or ganglioneuroblastoma well differentiated, ganglioneuroblastoma nodular, or unknown).

Table 2

Results of multivariate logistic regression models predicting *MYCN* amplification in 7,102 patients with neuroblastoma and known *MYCN* status.

Predictor variable (Unfavorable category shown / Favorable category is reference)	Number of patients with unfavorable variable / Number of patients with available data	Odds Ratio for *MYCN* Amplification in Presence of Unfavorable Predictor Variable (95% CI)	P-value
Model A (N=7,102) - All Variables[1]			
Age 18 Months	3041 / 7102	1.5 (1.3–1.8)	<0.001
Ferritin 92 (ng/mL)	1757 / 3602	1.6 (1.3–2.1)	<0.001
LDH 587 (U/L)	2114 / 4354	5.0 (4.0–6.4)	<0.001
INSS Stage 4	2585 / 7003	2.2 (1.9–2.6)	<0.001
Adrenal Primary Site	3129 / 6815	1.5 (1.3–1.8)	<0.001
Non-thoracic Primary Site	5771 / 6815	3.5 (2.4–5.1)	<0.001
Diploid or Hypodiploid	1025 / 3516	2.6 (2.1–3.2)	<0.001
High MKI	323 / 2738	9.9 (7.2–13.4)	<0.001
Poorly Differentiated or Undifferentiated	2407 / 2871	1.9 (1.1–3.1)	0.014
Chromosome 1p LOH Present	449 / 2026	11.4 (8.2–15.9)	<0.001
Chromosome 11q Aberration Present	216 / 1033	0.1 (0.06–0.2)	<0.001
Model B (N=7,102) - Clinical Variables Only[1]			
Age 18 Months	3041 / 7102	1.7 (1.5–2.0)	<0.001
Ferritin 92 (ng/mL)	1757 / 3602	1.7 (1.4–2.2)	<0.001
LDH 587 (U/L)	2114 / 4354	5.5 (4.4–6.9)	<0.001
INSS Stage 4	2585 / 7003	2.7 (2.3–3.1)	<0.001
Adrenal Primary Site	3129 / 6815	1.6 (1.4–1.9)	<0.001
Non-thoracic Primary Site	5771 / 6815	3.8 (2.6–5.6)	<0.001
Model C (N=7,102) - Pathology/Biology Variables Only[1]			
Diploid or Hypodiploid	1025 / 3516	3.3 (2.7–4.1)	<0.001
High MKI	323 / 2738	13.7 (10.2–18.4)	<0.001

Predictor variable (Unfavorable category shown / Favorable category is reference)	Number of patients with unfavorable variable / Number of patients with available data	Odds Ratio for *MYCN* Amplification in Presence of Unfavorable Predictor Variable (95% CI)	P-value
Poorly Differentiated or Undifferentiated	2407 / 2871	2.6 (1.6–4.3)	<0.001
Chromosome 1p LOH Present	449 / 2026	17.8 (13.1–24.3)	<0.001
Chromosome 11q Aberration Present	216 / 1033	0.2 (0.1–0.4)	<0.001

[1]Variables tested in the model and found non-significant were: A: Gain of 17q & pooled segmental chromosomal aberration; B: None; C: Gain of 17q & pooled segmental chromosomal aberration.

Table 3

Univariate analysis of metastatic sites as predictors of *MYCN* status in patients with stage 4 neuroblastoma.[1]

Metastatic site[1]	Sample size (N)[2]	Percent with metastatic site (%)	*MYCN*-Non amplified (%)[3]	*MYCN*-amplified (%)[3]	P-value[4]	Odds Ratio[5]
Bone marrow	2171				0.002	1.4
Yes		72.1	68.1	31.9		
No		27.9	74.9	25.1		
Bone	2164				0.006	1.3
Yes		58.8	67.9	32.2		
No		41.2	73.3	26.7		
Distant lymph node	2165				0.76	1.0
Yes		33.9	70.4	29.6		
No		66.1	69.8	30.2		
Liver	2172				0.74	1.0
Yes		18.1	69.4	30.6		
No		81.9	70.2	29.8		
Skin	1958				0.01	0.4
Yes		3.3	84.6	15.4		
No		96.7	70.1	29.9		
Lung	1923				<0.001	3.0
Yes		4.1	46.2	53.9		
No		95.9	71.7	28.3		
Central Nervous System	1923				0.6	1.2
Yes		2.6	67.4	32.7		
No		97.4	70.8	29.2		
Other	2128				<0.001	1.5
Yes		27.8	64.2	35.8		

Metastatic site[1]	Sample size (N)[2]	Percent with metastatic site (%)	*MYCN*-Non amplified (%)[3]	*MYCN*-amplified (%)[3]	P-value[4]	Odds Ratio[5]
No		72.2	72.5	27.5		

[1]At least one known metastatic site specified in the INRG database.

[2]N represents total with known data for the given variable.

[3]Row percentages sum to 100%.

[4]p-value refers to a chi-squared test.

[5]Odds ratio for univariate logistic regression with *MYCN* amplification as dependent variable.

Table 4

Multivariate predictors of *MYCN* status in 2,585 patients with stage 4 neuroblastoma.[1]

Predictor variable[2] (Unfavorable category shown / Favorable category is reference)	Number of patients with unfavorable variable / Number of patients with available data	Odds Ratio (95% CI)	P-value
Chromosome 1p LOH Present	263 / 650	8.6 (5.4–14.0)	<0.001
LDH > 587 (U/L)	1028 / 1504	6.8 (4.7–9.8)	<0.001
Chromosome 11q Aberration Absent	151 / 379	6.1 (2.7–13.9)	<0.001
High MKI	216 / 869	4.8 (3.1–7.5)	<0.001
Non-thoracic Primary Site	2300 / 2508	3.1 (1.8–5.6)	<0.001
Lung Metastasis	78 / 2270	2.4 (1.4–4.1)	0.001
No Skin Metastasis	2240 / 2305	2.1 (1.0–4.4)	0.05
Diploid or Hypodiploid	489 / 1166	1.8 (1.3–2.5)	<0.001
Adrenal Primary Site	1450 / 2508	1.6 (1.2–2.0)	0.001

[1] At least one known metastatic site specified in the INRG database.

[2] Variables tested in the model and found non-significant were: age; ferritin level; grade of differentiation; aberration at 11q; gain of 17q; pooled segmental chromosomal aberrations; and metastasis to bone, bone marrow, distant lymph node, liver, and central nervous system.