NEUROTROPHIN RECEPTOR GENE EXPRESSION IN NEUROBLASTOMA RISK GROUPS

by

Caroline Wallace Koudelka

A THESIS

Presented to the Department of Public Health and Preventive Medicine

and the Oregon Health and Science University

School of Medicine

in partial fulfillment of

the requirements for the degree of

Master of Public Health

March 2005

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the Master of Public Health thesis of

Caroline W. Koudelka

has been approved

[redacted]

Jodi A. Lapidus, PhD
Chair of Thesis Committee

[redacted]

Motomi Mori, PhD
Committee Member

[redacted]

Srinivasa R. Nagalla, MD
Committee Member

[redacted]

Stephen Roberts, MD
Committee Member

TABLE OF CONTENTS

List of Figures

List of Tables

Acknowledgments

I am grateful to all organizations who contributed to the data used in this study, and who granted me the permission to analyze it. Participating members of the Children's Cancer Group recruited and followed subjects, and provided clinical assessments. Genetic analyses were provided by laboratories at the Children's Hospital of Los Angeles and the Center for Biomarker Discovery at Oregon Health and Science University.

Many thanks to my advisor and thesis chair, Jodi Lapidus, upon whose statistical instruction and guidance I relied heavily. I would also like to acknowledge the contributions of other members of my thesis committee; Motomi Mori, Stephen Roberts, and Srinivasa Nagalla.

Finally, I would like to express my gratitude to my family, my friends, and my husband, Kurt, for all of their support and encouragement throughout this project.

Abstract

Although great progress has been made recently in the treatment and cure of most types of pediatric cancer, similar success has eluded neuroblastoma. This malignancy is characterized by diverse outcomes including spontaneous regression, development into benign conditions, and progression to aggressive metastatic disease. The biologic mechanisms determining these presentations remain unknown, challenging those who seek accurate prognostic tools and successful therapies.

To increase understanding of neuroblastoma's underlying biology, we evaluated an existing dataset containing neurotrophin-associated gene expression levels and clinical prognostic assessment variables for 235 tumors. Genes studied included members of the Trk and GABA families of neurotrophin receptors. Through complementary statistical methodologies, including analysis of variance, logistic regression, and classification trees, we drew associations between gene expression levels and clinically derived risk groups. Gene expression levels of Trk A, p75, and GABA A delta emerged as potential predictors of neuroblastoma risk group. Further genetic studies are required for the development of an improved system of risk determination and the discovery of targeted treatments.

Background

Embryonic development is achieved through a complex and critical sequence of events. Early in this process, cells from the neural crest migrate throughout the body to differentiate into numerous specific cell types and structures. Many of these cells will become the peripheral nervous system and its subgroup, the sympathetic nervous system. Occasionally, a sympathetic neuron precursor will experience a genetic change that may lead to unregulated proliferation. These genetically abnormal cells may develop immediately into tumors, or lie dormant for varying lengths of time until a precipitating event triggers the onset of tumor development. This is the beginning of a neuroblastoma, a type of childhood malignancy characterized by puzzling presentations and diverse outcomes.

Epidemiology of neuroblastoma

Among children in the United States, cancer in all of its forms is the fourth leading cause of death, and the most common cause of disease-related death (1). Following leukemias, central nervous system tumors, and lymphomas, neuroblastoma is the fourth most prevalent pediatric malignancy, accounting for 7-10% of all childhood cancer (2,3). Approximately 650 new cases of neuroblastoma are diagnosed per year (4).

As children age, their risk of developing neuroblastoma decreases. In the United States, the annual incidence rate for neuroblastoma for children under the age of 15 years ranges from 8 to 12 per million (1). For infants of less than one year of age, however, the incidence rate is greater than 60 per million, making neuroblastoma the most common

infant malignancy. This rate was nearly twice that of leukemia, the next most common cancer (1). While neuroblastoma patients generally range in age from neonates to 15-year olds, the median age of existing neuroblastoma cases is 2 years (3).

The incidence of neuroblastoma is slightly higher in males than in females, and slightly higher in whites than blacks. Between 1975 and 1995, the average annual neuroblastoma incidence rates per million children under 15 years were 10.1 for white males, 9.6 for white females, 8.8 for black males, and 8.6 for black females (1). An insufficient number of U.S. neuroblastoma cases were identified in other races to determine incidence rates.

Etiology

Hereditary cases of neuroblastoma have been noted, but are extremely rare. Since the first discovery of familial neuroblastoma in 1945, there have only been approximately 40 family pedigrees described in which more than one case was found in first or second-degree relatives (19). Inheritance was determined to follow an autosomal dominant Mendelian model. According to this "two hit" model, tumors were predicted to arise as the result of two separate mutation events. In sporadic cases, these mutations were both somatic, whereas in familial cases one mutation was inherited and the other was somatic (20, 21).

Several case-control and population-based studies have been conducted to identify possible etiological risk factors for neuroblastoma development. Since neuroblastoma is a disease of infancy and young childhood, studies have focused primarily on parental and prenatal exposures. Several investigators reported significant

associations between maternal use of sex hormones prior to or during pregnancy and development of neuroblastoma (5,6,7,8). The results obtained from most etiologic studies, however, have been either null or contradictory. Due to the scarcity of neuroblastoma cases, studies are performed with few subjects, and lack the statistical power needed to detect significant differences. Conflicting results have been obtained in the studies of parental exposure to electromagnetic fields (9,10,11,12,13,14,15), maternal use of alcohol during pregnancy (5,6,16), and previous fetal loss (17,18).

Adding to this confusion are studies in which investigators have reported finding statistically significant predictors of neuroblastoma to be different between subsets of the disease. Dividing patients into groups based on age at diagnosis and risk status allowed different yet significant etiologic agents to be detected, including gestation duration, maternal alcohol consumption during pregnancy, and parental age (5,15). Such differences suggest that neuroblastoma actually represents more than one distinct disease.

Disease presentation and behavior

Primary neuroblastoma tumors may be found in a variety of anatomical locations. While approximately one half of neuroblastomas begin in the adrenal medulla, other possible sites of tumor origination include the pelvic ganglia and the paraspinal ganglia in the chest or abdomen (22). At the time of diagnosis, metastasis may have begun in as many as 70% of neuroblastomas (4). The range of primary tumor locations and metastatic sites results in a wide array of presenting symptoms, from abdominal discomfort and bowel irregularities to limping and periorbital hemorrhage (23).

Possible outcomes of neuroblastoma are diverse and puzzling. Tumors may undergo spontaneous regression, progression to metastatic disease, or development into benign conditions. Autopsies performed on infants expiring of non-neuroblastoma causes find small clusters of neuroblastoma cells in up to 1 of every 220 adrenal glands (24). This phenomenon suggests that most neuroblastomas that occur will spontaneously regress or mature without being clinically detected. A related group of benign tumors, the ganglioneuromas, are believed to have begun as neuroblastomas before maturing (24). Despite efforts to discover the biological explanation for these distinct outcomes, the pathology of neuroblastoma remains enigmatic, and its elucidation demands more research.

Clinical prognostic indicators

Although the exact biologic mechanisms that lead to neuroblastoma's diverse behavior are still unclear, several clinical variables have been found that have strong associations with disease outcome.

Age

In the early 1970's patient age was shown to be an important and independent predictor of neuroblastoma outcome (25). Patients diagnosed with neuroblastoma under the age of one year have a much more positive prognosis. The five-year survival for infants diagnosed at less than one year of age is approximately 83%. This percentage is very high compared to the five-year survival of children diagnosed between the ages of one and five years (55%), and those diagnosed over the age of five years (40%) (1).

Patient age at diagnosis and tumor stage are still considered the two most important prognostic indicators in clinical practice (26).

Stage

Prior to the late 1980's several staging systems had been developed worldwide to help classify and predict behavior of neuroblastoma tumors based on clinical features. Having different systems produced a lack of uniformity in prognosis and in assignment of a treatment regimen. As a result, investigators had difficulty comparing patients that were being treated in different institutions throughout the world, complicating the study of novel treatments. To resolve these difficulties, an international group formed to standardize the diagnosis, staging, and treatment response of neuroblastoma (27). This group developed and later revised the International Neuroblastoma Staging System (INSS) from a combination of factors taken from the three primary existing staging systems (28). The INSS is now the internationally accepted and employed method of staging neuroblastoma. This system classifies patients into one of five stages based on the location and infiltration of the tumor, the completeness of the tumor resection, and the extent of lymph node involvement and metastasis (27).

As shown in Table 1, prognoses by stage range from 1 (the most favorable) to 4 (the least favorable). Stage 4s is an unusual category that involves widespread dissemination at the time of diagnosis, but the prognosis is generally very good. These tumors, by definition, only occur in infants less than one year old (28). Since the inception of the INSS, several studies have tested and verified the system's utility in predicting disease outcome (29,30). Coupling age at diagnosis with INSS stage provides a superior estimate of disease behavior than considering either of these predictors alone.

Three-year remission for infants in stage 3 disease is 80% to 90%, whereas it is only 50% in children older than one year of age. These rates drop to 60% to 75% for stage 4 infants, and 15% for stage 4 children over 1 year (31). At any age at diagnosis, 75% to 90% of patients in stages 1, 2, and 4s are alive and event-free three years after their diagnoses (31).

Table 1: International Neuroblastoma Staging System criteria for classifying neuroblastoma by stage (28)

Stage	Definition
1	Localized tumor with complete gross excision, with or without microscopic residual disease; representative ipsilateral lymph nodes negative for tumor microscopically (nodes attached to and removed with the primary tumor may be positive).
2A	Localized tumor with incomplete gross excision; representative ipsilateral nonadherent lymph nodes negative for tumor microscopically.
2B	Localized tumor with or without complete gross excision, with ipsilateral nonadherent lymph nodes positive for tumor. Enlarged contralateral lymph nodes must be negative microscopically.
3	Unresectable unilateral tumor infiltrating across the midline, * with or without regional lymph node involvement; or localized unilateral tumor with contralateral regional lymph node involvement; or midline tumor with bilateral extension by infiltration (unresectable) or by lymph node involvement.
4	Any primary tumor with dissemination to distant lymph nodes, bone, bone marrow, liver, skin and/or other organs (except as defined for stage 4s).
4s	Localized primary tumor (as defined for stage 1, 2A or 2B), with dissemination limited to skin, liver, and/or bone marrow** limited to infants <1 year of age).

*The midline is defined as the vertebral column. Tumors originating on one side and crossing the midline must infiltrate to or beyond the opposite side of the vertebral column.
**Marrow involvement in stage 4s should be minimal, i.e., <10% of total nucleated cells identified as malignant on bone marrow biopsy or on marrow aspirate. More extensive marrow involvement would be considered to be stage 4. The MIBG scan (if performed) should be negative in the marrow.

Tumor histology

Neuroblastic tumors are composed of two populations of cells, including Schwann cells, responsible for myelin formation in the peripheral nervous system, and neuroblastic cells. Studies have shown that in normal development neuroblastic cells produce a chemotactic factor that attracts Schwann cells from surrounding tissues.

Schwann cells, in turn, produce factors responsible for stimulating neuroblast differentiation and ceasing neuroblast proliferation (32). Through this process some neuroblastomas mature, developing into benign conditions called ganglioneuromas. Others, however, do not.

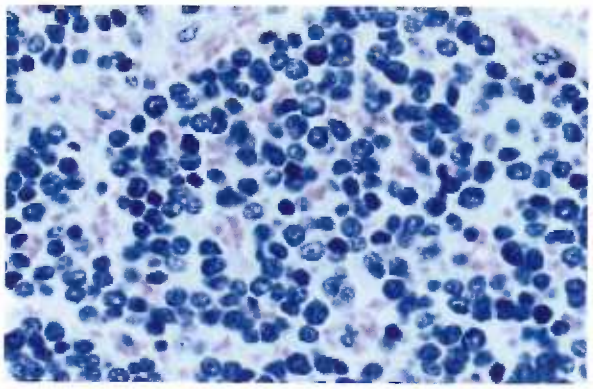
Techniques for classifying neuroblastic tumors were originally proposed by Shimada (33) and revised by the International Neuroblastoma Pathology Committee in 1999 to become the currently accepted and widely used International Neuroblastoma Pathology Classification (INPC) (34). The classification system determines tumor type based on the density of Schwann cells and histological appearance of neuroblastic cells within the tumors.

According to INPC criteria, neuroblastic tumors may be one of the four following types:

1. Neuroblastoma (Schwannian stroma-poor)

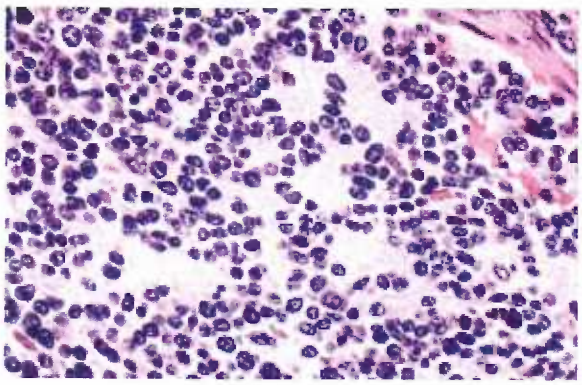
   These tumors consist of clusters of immature neuroblastic cells separated by few Schwann cells. Depending on the presence of various indicators of cellular maturity contained by the neuroblastic cells, tumors are subclassified into one of three groups (see Figures 1-3).

Figure 1: Neuroblastoma, undifferentiated subtype, as defined by the International Neuroblastoma Pathology Classification system (24)



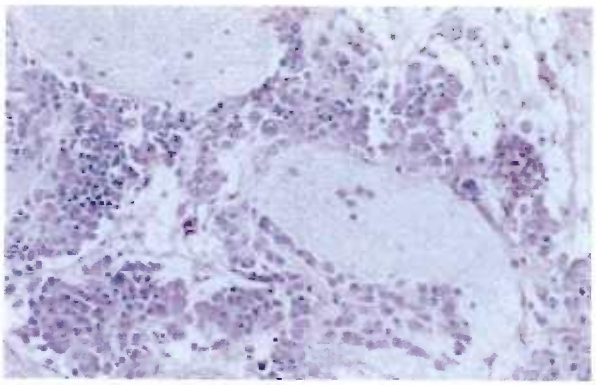
Neuroblasts are seen densely distributed with very little neuropil (the material present between nerve cells). The prognosis of this group is always unfavorable.

Figure 2: Neuroblastoma, poorly differentiating subtype, as defined by the International Neuroblastoma Pathology Classification system (24)



Undifferentiated neuroblasts are present within small amounts of neuropil.

Figure 3: Neuroblastoma, differentiating subtype, as defined by the International Neuroblastoma Pathology Classification system (24)

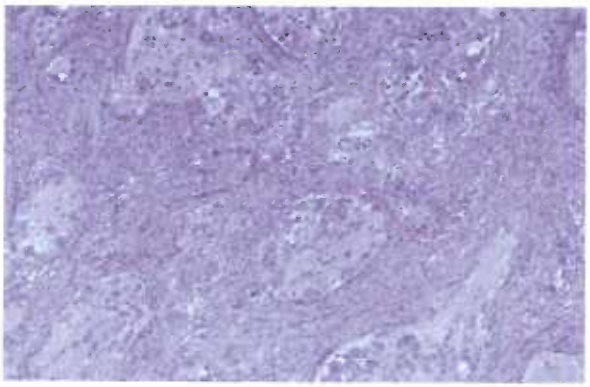


Differentiating neuroblasts are present within greater amounts of neuropil.

2. Ganglioneuroblastoma, intermixed (Schwannian stroma-rich)

Tumors of this type are in the process of maturation, but have not yet reached the full maturity of ganglioneuromas. They contain a mix of randomly distributed cells in various stages of maturation surrounded by a higher proportion of Schwann cells than were seen in neuroblastomas (see Figure 4). Regardless of the patient's age, the prognosis of such tumors is favorable.

Figure 4: Ganglioneuroblastoma, intermixed, as defined by the International Neuroblastoma Pathology Classification system (24)

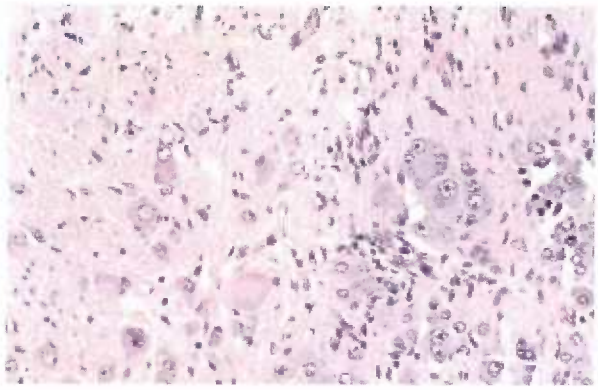


Different cell populations are present within large amounts of stroma.

3. Ganglioneuroma (Schwannian stroma-dominant)

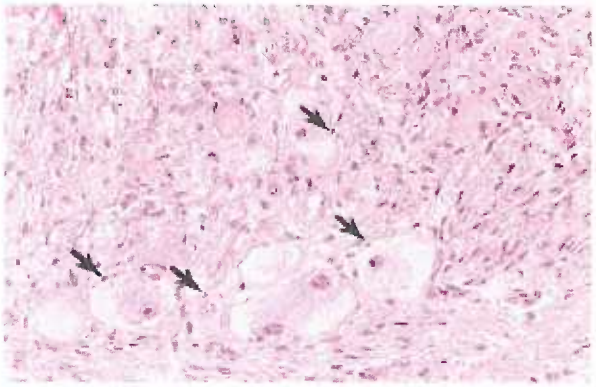
   The highest proportions of Schwann cells are seen in this prognostically favorable tumor class, which is further divided into two subgroups (see Figures 5-6).

Figure 5: Ganglioneuroma, maturing subtype, as defined by the International Neuroblastoma Pathology Classification system (24)



While most of the cells seen in these tumors are fully mature, a few clusters of maturing neuroblastic cells in various stages may be seen.

Figure 6: Ganglioneuroma, mature subtype, as defined by the International Neuroblastoma Pathology Classification system (24)

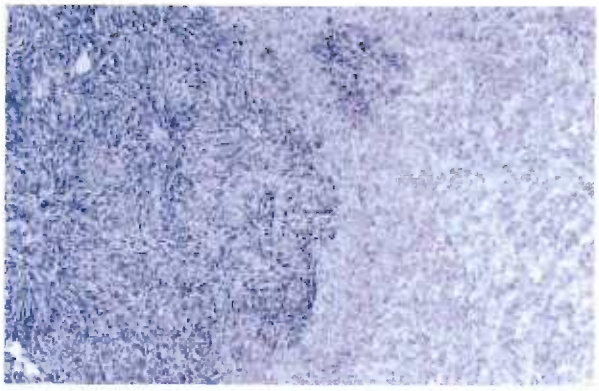


Arrows point to fully mature ganglion cells amid Schwannian stroma.

4. Ganglioneuroblastoma, nodular (composite, Schwannian stroma-rich/stroma-dominant and stroma-poor)

   This class of tumor is characterized by having one or more nodules of stroma-poor neuroblastoma contained within stroma-dominant ganglioneuroma or stroma-rich intermixed ganglioneuroblastoma tissue (see Figure 7). The stroma-poor nodules are aggressive and rapid to grow. As with undifferentiated neuroblastoma tumors, the prognosis of nodular ganglioneuroblastomas is always unfavorable (32).

Figure 7: Ganglioneuroblastoma, nodular, as defined by the International Neuroblastoma Pathology Classification system (24)



Stroma-poor neuroblastic cells on the right infiltrate the stroma-rich area to the left.

When assessing favorable versus unfavorable histology-based prognosis, the mitosis-karyorrhexis index (MKI) is also considered. This determinant is a measure of the proportion of tumor cells undergoing either of these processes. MKI is classified into low (<2% mitotic and karyorrhectic cells), intermediate (2%-4%), or high (>4%) (32). Since some immature neuroblastic tumors develop into nonmalignant conditions, tumor

class is combined with patient age at diagnosis and MKI to estimate prognosis (see Table 2).

Table 2: International Neuroblastoma Pathology Classification criteria for assessing favorable versus unfavorable histology (32)

Age	Favorable Histology Group	Unfavorable Histology Group
Any	Ganglioneuroma (Schwannian stroma-dominant) ♦ Maturing ♦ Mature	
Any	Ganglioneuroblastoma, intermixed (Schwannian stroma-rich)	Ganglioneuroblastoma, nodular (composite, Schwannian stroma-rich/stroma-dominant and stroma-poor)
Any		Neuroblastoma (Schwannian stroma-poor) ♦ Undifferentiated and any MKI
<1.5 Years	Neuroblastoma (Schwannian stroma-poor) ♦ Poorly differentiated and low or intermediate MKI* ♦ Differentiating and low or intermediate MKI*	Neuroblastoma (Schwannian stroma-poor) ♦ Poorly differentiated and high MKI* ♦ Differentiating and high MKI*
1.5 -5 Years	Neuroblastoma (Schwannian stroma-poor) ♦ Differentiating and low MKI*	Neuroblastoma (Schwannian stroma-poor) ♦ Poorly differentiated and any MKI ♦ Differentiating and intermediate or high MKI*
≥ 5 Years		Neuroblastoma (Schwannian stroma-poor) ♦ Any subtype and any MKI

*Low MKI: <2% mitotic and karyorrhectic cells, Intermediate MKI: 2%-4% mitotic and karyorrhectic cells, High MKI: >4% mitotic and karyorrhectic cells

N-myc amplification

Cancers often arise as the result of amplifications of genes in some way responsible for the control of cellular proliferation. One such gene, N-myc, is amplified in approximately 22% of neuroblastoma cases, making it the most common DNA amplification seen in this malignancy (35). Amplifications of N-myc are cytogenetically exhibited as homogeneously staining regions (HSRs) and double-minute chromatin

bodies (DMs) in tumor chromosome studies (22). It is likely, therefore, that N-myc amplification occurs by the replication-excision pathway, in which the gene sequence is over-replicated locally, and the resulting extra DNA is excised from the chromosome. These independent DNA fragments may be integrated into other chromosomes for further amplification (36).

N-myc amplification generally results in high gene expression levels, measured by N-myc mRNA and protein product levels. Increased expression of N-myc has been shown to stimulate tumorigenicity in mature cells and prevent apoptosis, or programmed cell death, in immature cells (36). As cells divide, they undergo four successive phases called G_1, S, G_2, and M. N-myc acts by shortening the G_1 phase of the cell-division cycle, thereby increasing the cell proliferation rate (36). Increased expression of N-myc does not necessarily require N-myc amplification. Tumors exhibiting greater than normal levels of N-myc mRNA and proteins without corresponding N-myc amplification may have irregularities in protein breakdown pathways (22). Therefore, while high N-myc expression levels are highly correlated with N-myc amplified tumors, these two measures are certainly not interchangeable.

Investigators have reported conflicting results regarding the prognostic significance of high expression levels of N-myc in non-amplified tumors (37,38). Nearly all N-myc amplified tumors, however, undergo rapid and aggressive disease progression, despite the status of other prognostic indicators. Between N-myc amplification status and N-myc expression, amplification is generally considered the more reliable prognostic indicator, and its determination is now routine in clinical risk assessment of neuroblastoma (35).

DNA index (ploidy)

DNA index (DI), or tumor "ploidy", refers to the overall number of chromosomes in the tumor cell. A DNA index of 1.0 reflects the 46 chromosomes expected in a normal diploid cell. Any DNA index value greater than 1.0 indicates that a state of hyperploidy exists. For example, a triploid cell will contain 69 chromosomes, and have a DI value of 1.5. Studies have found that in infants, near-triploid tumors seem to have the best outcomes, while diploid and tetraploid tumors correlate with poor prognosis and advanced stage (39,40,41). This relationship is lost, however, in patients diagnosed over 24 months of age. Therefore, DNA index is only a factor in assessing risk in infants.

1p loss and 17q gain

The most common chromosomal aberrations found in neuroblastoma tumors are gains of the long arm of chromosome 17 (17q), and deletions of the short arm of chromosome 1 (1p). 1p deletions correlate with other indicators of poor prognosis including N-myc amplification and DI near 1.0 (42). Although more than one deletion site have been found on 1p, deletions are generally seen in the distal end of the chromosome, suggesting that this may be the site of at least one tumor suppressor gene (43,44). 17q gains, identified in more than 90% of high-risk tumors studied, have been shown to be independent predictors of poor outcome in neuroblastoma (45). It is postulated that this site may contain a gene important for cell survival, and overexpression of this gene may lead to aggressive tumor growth. Often 1p loss accompanies 17q gain in a phenomenon called translocation. Translocation occurs when part of a chromosome breaks off and attaches to another chromosome. In 1:17 translocation part of the 17q chromosome replaces the missing area of 1p, doubly

increasing the risk. While initial reports are promising (35,46), the prognostic utility of these chromosomal abnormalities is still under investigation.

Risk groups

After standardizing the tumor staging system with the INSS, an international group of investigators identified the need to develop a standardized system of assigning neuroblastoma risk status to compare similar groups of patients being treated at different locations. Treatments would also be standardized based on this risk assessment. This group first proposed studying several biological variables for their prognostic significance when combined with age and INSS stage. Prior research in neuroblastoma biology provided the variables suggested for evaluation. The widely collected prognosticators of interest included indicators suggesting genetic abnormalities (N-myc amplification, DNA index, and individual chromosomal losses or gains), histopathological evaluation, and serum markers lactate dehydrogenase and ferritin (47). The serum markers were later discarded as potential prognostic indicators as they were determined to be more representative of tumor burden, or the amount of cancer present in the body (31).

Current methods of risk group determination used by the Children's Oncology Group (COG) depend on patient age, INSS stage, INPC histology summary, N-myc amplification status, and DNA index (in infants only). Risk group assignment is performed according to the criteria outlined in Table 3. Tumor stage is determined first, then patient age, N-myc amplification status, histology summary, and DNA index (where applicable) are examined successively to assign patients into low, intermediate, and high risk groups.

Table 3: Current Children's Oncology Group criteria for classifying neuroblastoma into risk groups (22)

Stage	Low Risk	Intermediate Risk	High Risk
1	All	None	None
2	Age < 1year, or Age 1-21 years and N-myc non-amplified, or Age 1-21 years and N-myc amplified with Favorable histology	None	Age 1-21 years and N-myc amplified with Unfavorable histology
3	None	Age < 1 year and N-myc non-amplified, or Age 1-21 years and N-myc non-amplified with Favorable histology	Age 0-21 years and N-myc amplified, or Age 1-21 years and N-myc non-amplified with Unfavorable histology
4	None	Age < 1 year and N-myc non-amplified	Age < 1 year and N-myc amplified, or Age 1-21 years
4s (all are <1 year)	N-myc non-amplified with Favorable histology and DNA index>1	N-myc non-amplified with Unfavorable histology, or N-myc non-amplified with Favorable histology and DNA index = 1	N-myc amplified

Tumor biology: neurotrophic factors and receptors

Neuroblastoma's diverse clinical manifestations and subgroup differences in prognostic indicators, genetic features, and etiologic agents suggests that this group of malignancies may encompass more than one biologically disparate disease state. Further comparison of the genetic activity of the different tumor types is a necessary step towards understanding the biological differences exhibited between the three risk groups.

Understanding the biology of neuroblastoma begins with understanding the process of normal nervous system development. Early in embryogenesis, neuron precursor cells travel from the neural crest to their target locations. Many more of these cells arrive than are needed. Circulating proteins known as neurotrophic factors at these

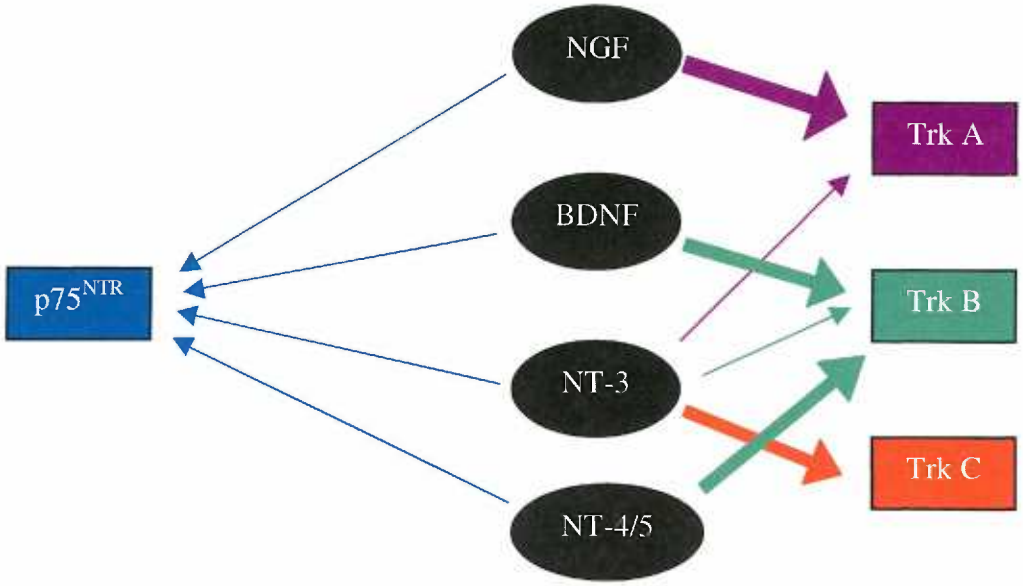
locations bind to the cells, enabling their survival and proliferation. Cells that are not able to bind adequate neurotrophic factor undergo apoptosis, or cell death. Expression of various cell-surface receptors enables the binding and subsequent action of neurotrophins.

Trk receptors

There are three tyrosine kinase receptors in the Trk family, each having different ligands and functions (see Figure 8). In normal cellular differentiation, expression levels of these receptors vary with embryonic stage. TrkA is highly expressed late in the differentiation process. This receptor binds nerve growth factor (NGF) and (less strongly) neurotrophin-3 (NT-3) (48). Nerve growth factor binding results in the differentiation of immature sympathetic neurons into mature ganglion cells. Failure of cells to bind NGF results in apoptosis (49). Recent studies have shown that neuroblastomas with high expression levels of TrkA have a much more favorable prognosis than those with low expressions of this receptor. High TrkA expression is correlated with younger age at diagnosis, lower stage, and non-amplified N-myc gene (22). High expressions of TrkB, however, are characteristic of tumors with a very poor prognosis (50). In normal cell differentiation, TrkB is highly expressed at early stages. Its ligands include brain-derived neurotrophic factor (BDNF), neurotrophin-4/5 (NT-4/5), and to a lesser extent NT-3 (48). Binding of these ligands promotes cell survival. The ligand NT-3 is also bound by the TrkC receptor. TrkC expression has not been studied as extensively as TrkA and TrkB in neuroblastoma, but preliminary data suggest that the relationship between TrkC and tumor prognosis follows the same pattern as that of TrkA (48). A final receptor noteworthy for its contribution to cellular differentiation is $p75^{NTR}$.

This receptor also binds ligands NGF, BDNF, NT-3, and NT-4/5. Its expression levels in neuroblastoma are unknown (48).

Figure 8: Patterns of neurotrophin binding by Trk family and p75 receptors



Neurotrophins are represented by black ovals, and corresponding receptors are shown in colored boxes. Block arrows indicate high-affinity binding. Narrow arrows indicate low-affinity binding

GABA receptors

The action of γ-aminobutyric acid (GABA) is dictated by the expression levels of three types of receptors, GABA A, GABA B, and GABA C. GABA A and GABA C exist as pentamers of various subunits. Identified GABA A subunits include six alpha subunits, three beta subunits, three gamma subunits, and one each of delta, epsilon, pi, and theta subunits. Complete receptors have been found to contain at least one alpha and one beta subunit, and one from the group including gamma, delta, epsilon, pi, or theta.

GABA C receptors consist of five rho subunits. Both GABA A and GABA C are ionotropic receptors, operating through control of ion channels. This is in contrast to GABA B, a metabotropic receptor that operates by means of enzymatic activation. The GABA B receptor consists of one R1 subunit and one R2 subunit associated with trimeric GTP-binding regulatory proteins (G proteins), which mediate the interactions between receptors and ligands (51).

GABA is most frequently thought of as an inhibitory neurotransmitter of the central nervous system. Recent research, however, has suggested further functions of GABA in neural development. In several laboratory studies, the addition of GABA to various neuronal cell populations in vitro was found to induce cellular differentiation, to affect DNA synthesis, and to regulate the migration of neurons (51). In another study, when researchers applied GABA to cultured peripheral neurons and neuroblastoma cells, neuron differentiation was stimulated and GABA receptor expression was changed, suggesting the role of GABA as a general neurotrophin (52). The potential role of GABA in the pathogenesis of neuroblastoma remains to be determined.

Summary

Neuroblastoma is a puzzling malignancy of infancy and early childhood characterized by diverse presentations and outcomes. The risk posed by this disease is currently assessed using a system combining several clinical prognosticators including age at diagnosis, tumor stage, histology summary, N-myc amplification status, and DNA index. Biologic mechanisms explaining the various outcomes are currently unknown, but may be revealed by differences in genetic expression levels between the tumors. Neurotrophin receptor genes are of particular interest. Since these genes are known

actors in normal neuronal development, it is likely that they also play a role in the pathogenesis of neuroblastoma.

Study Objectives

The broad objective of this study is to add to the growing body of knowledge explaining the biologic mechanisms for the development of neuroblastoma's diverse outcomes. Specifically, the expression levels of several families of genes believed to be associated with the pathology of this malignancy will be studied. Of particular interest are genes associated with the receptors of neurotrophic factors. These factors all play critical roles in the normal differentiation of neural crest cells into mature sympathetic neurons.

A data set compiled jointly by the Children's Hospital of Los Angeles (CHLA) and Oregon Health and Science University will be examined with this objective in mind. CHLA investigators gathered tumor samples and information from 235 neuroblastoma patients diagnosed throughout the United States. The data set they compiled includes variables describing the clinical presentations of the patients' tumors, as well as gene expression levels of various neurotrophic factors. Variables to be analyzed in this study include expression levels of genes associated with GABA receptors (GABA A alpha 1-6, GABA A beta 1-3, GABA A gamma 2-3, GABA A delta, GABA A epsilon, GABA A pi, GABA A theta, GABA B, GABA C, GAD-1, and GAD-2), the Trk A and p75 receptors, and NGF. Unfortunately, expression levels of Trk B and Trk C were unavailable in this data set.

This study is a secondary analysis of these data. The primary analysis included a statistical examination of a subset of the GABA receptor family of genes, using time to disease progression/death as the outcome variable (53). This study will use clinical risk group as the outcome variable, and will examine the expression levels of neurotrophin receptor genes across these groups.

Specific aims

Specific aims for this study include:

1. Categorize patients into risk status groups according to INRG criteria
2. Examine the relationships between the gene expression levels in Trk and GABA families
3. Determine the associations between gene expression variables and risk status groups
4. Create models to describe risk group based on factors and/or individual gene expressions
5. Determine how well risk groups can be classified by using combinations of factors and individual gene expressions

Methods

Clinical methods

Study participants

Subjects enrolled in the study were patients at one of the many U.S. institutions partnered with the former Children's Cancer Group (CCG). The CCG has since joined with the Pediatric Oncology Group (POG) to become the Children's Oncology Group (COG) (see Appendix A for a complete list of partner institutions). All participating institutions obtained IRB approval for the study. A total of 235 neuroblastoma tumor samples were collected over sixteen years, from January 1980 through April 1996. Samples were collected from patients at the time of their diagnoses, prior to treatment with chemotherapy or other methods.

Variable collection

The tumor samples were evaluated for traditional prognostic markers including patient age at diagnosis, stage, Shimada histology profile, and N-myc amplification status at partner institutions. These data, as well as a sample of the tumor itself, were submitted to the Neuroblastoma Center at the Children's Center for Cancer and Blood Diseases in Children's Hospital, Los Angeles (CHLA). The patients' institutions periodically submitted further data including survival times and disease progression, and CHLA added variables documenting length of follow-up time and outcome to the data set.

Expression levels for various genetic markers were determined by RT-PCR at CHLA. These variables included Trk A, NGF, and p75. Tumor tissue was then submitted to OHSU for further genetic expression testing.

Molecular methods

At the Center for Biomarker Discovery at Oregon Health and Science University, Department of Pediatrics, the complimentary techniques of suppression subtractive hybridization and cDNA microarray assays were employed to find genes that were differentially expressed in neuroblastoma tumor samples compared to normal adrenal tissue. A pool of mRNA taken from tumors of all stages was tested with mRNA from normal adrenal glands, identifying 1600 genes preferentially expressed in neuroblastoma. These genes were used to create microarray slides to test expression levels in adrenal controls and tumor samples. GABA system gene expressions were determined to be significantly different between various samples and controls. RT-PCR was used to quantify expression levels in genes of interest. Gene expression levels were determined for several genes comprising the GABA family. GABA genes measured included 15 GABA A receptor subunits, GABA B and GABA C receptors, and metabolizing enzymes GAD1 and GAD2 (see reference 53 for complete methods of molecular analysis). RT-PCR results were added to the data compiled by CHLA.

Two variables were obtained for every gene examined. The first variable, called cycle threshold, represents the number of RT-PCR cycles required to produce a target quantity of cDNA. Low results for this variable correspond to high initial levels of mRNA in the sample, and therefore to high levels of gene expression. Samples with gene expression product (mRNA) absent achieve the maximum cutoff value of 40 for this variable, since no cDNA was produced. The second variable conveys relative numerical values for the gene expressions of the samples.

Statistical methods

Several methods of statistical analysis were used to examine the data from various perspectives while accomplishing the proposed specific aims of the study. The statistical methods used for each aim are described below. These analyses were performed using SPSS software unless otherwise noted (54).

Data examination and transformation (Aim 1)

Clinical variables

Clinical prognostic indicators were used to create the variables age at diagnosis (broken down into categories of greater than or less than one year), stage (1,2,3,4, or 4s), N-myc amplification status (amplified or non-amplified), and histology summary (favorable, unfavorable, or indeterminate).

Risk status variables

A new variable was created to assign cases to a category of high, intermediate, or low risk based on clinical variables as depicted in Table 3 (page 16). Frequencies of cases in each level were obtained, and missing cases were isolated and examined. Since missing cases were due to incomplete histology, DNA index, and N-myc status data, we decided to create two binary variables to divide risk status into low and intermediate vs. high risk, and low vs. intermediate and high risk. Creating these binary variables allowed us to include many of these missing cases into some analyses.

Gene expression variables

The ranges and medians of all continuous gene expression variables were determined, and their distributions were examined. To determine whether these values depicted meaningful expression or merely background noise, the cycle threshold

variables were examined. Cycle threshold values of less than 30 correspond to highly expressed genes. Values between 30 and 38 indicate a low level of expression, and values greater than 38 are considered undetectable. Variables with cycle threshold values greater than 38 for all observations were omitted from further analyses.

Histograms of expression level variables to be included in analyses were examined, and a $\log_{10}$ transformation was applied to all continuous variables to reduce skewness. Prior to taking the logs, variables containing negative and zero values were adjusted to permit log transformations of all values. Variables having a minimum value of zero had 0.01 added to each observation, and variables having negative expression values had a constant added to each observation in order that the new minimum value become 0.01. Distances between observations within variables were therefore preserved while permitting logs of every value to be obtained. Distributions and correlations (obtained by Spearman rank method) of log-transformed variables were examined.

Survival analysis

Disease progression-free survival times of each risk group were compared to gain insight into the utility of the newly created risk group variable. Kaplan-Meier analysis and log-rank testing was performed on the data examining time to disease progression stratified by patient risk group.

Principal components analysis (Aim 2)

PCA was performed to reduce the numerous gene expression variables into a few principal components calculated as linear combinations of original variables (factor scores). Creating uncorrelated factor scores also allowed us to avoid the problem of variable correlation in regression analyses. Initially, the analysis was performed using all

gene expression variables, and factors created with eigenvalues greater than one were saved. These underwent varimax rotation to eliminate inter-factor correlation. The analysis was rerun several times, omitting variables that either loaded equally into more than one factor, or did not load significantly into any. Based on the results of initial analyses the decision was made to remove Trk A and p75 from the PCA variable pool and take their average to create a new variable (referred to as Trk75). The factor scores from the final analysis were saved as new variables (referred to as Factor 1 and Factor 2) to include in future analyses. Internal reliability of factors was assessed by Cronbach's alpha. To get a sense of the clinical meaning behind the factors, the factor score variables were split into quartiles, and the mean number of genes expressed was computed for each quartile.

Examination of new predictor variables

We examined histograms of new variables including the factor scores created by principal components analysis and the variable created by taking the average of Trk A and p75. Boxplots of these variables with cases divided into risk groups were also examined to identify differences that might prove to be important in regression analysis and modeling of the risk variables.

Analysis of variance (Aim 3)

A series of analyses of variance were performed to determine whether there were significant differences in gene expression levels between the three risk groups. The new variables created by principal components analysis as well as individual gene expression variables were examined. Post hoc pairwise comparisons using Tukey's HSD method were conducted to determine which of the groups significantly differed.

Regression analysis (Aim 4)

To examine the possibility that different levels of risk behave as separate disease states characterized by the expression of different genes, regression analysis was performed using the two binary risk variables (low vs. intermediate/high risk, referred to as Risk 2, and low/intermediate vs. high risk, referred to as Risk 3) as dependent variables and combinations of gene expression variables as predictor variables.

Logistic regression

A series of logistic regression models were fit using Risk 2 and Risk 3 as the dependent variables and Factor 1, Factor 2, Trk75, and individual gene expressions as independent variables. Results from these analyses were considered when creating multiple logistic regression models to predict Risk 2 and Risk 3. Initially, only the factor scores and Trk75 were entered into these models, then individual component gene expression variables were substituted for factors in further analyses. The individual gene variables in the final models were selected based on model fit and correlation levels with other variables.

Evaluation of predictor linearity

Independent variables found to be significantly associated with risk level by logistic regression were further explored in generalized additive models using the R software program (55). Splines of significant predictors were examined to detect nonlinearity in their functions. Predictors with nonlinear splines were broken into categorical variables corresponding to regions of differential response. These variables were then re-entered into logistic regression analyses, and the improvement in model fit was determined.

Logistic regression using data subgroups

Regression models were fit using pairs of individual risk groups. The models developed compared low and intermediate risk groups, intermediate and high risk groups, and low and high risk groups. Independent variables used included gene expression levels that emerged as significant predictors in previous regression analyses.

Classification of intermediate risk group

The logistic regression model developed using high risk vs. low risk as the outcome variable was used to create an equation to determine fitted values for risk probability. Intermediate risk cases were entered into this equation to compute fitted values, and the distribution of these values was examined. This distribution was compared to those of the predicted values for the low and high risk groups.

Classification trees (Aim 5)

Salford Systems' CART methodology was used to create trees in which different subsets of predictor variables were used to classify risk into its three categories of low, intermediate, and high (56). Initially, predictor variables were limited to Factors 1 and 2, and Trk75. Results from this model were compared to a tree created to predict risk using a combination of individual gene expression variables. The preferred trees were selected from the groves of trees grown based on a combination of tree complexity and relative cost. Cross-validation misclassification rates from the final trees selected were examined and the classification powers of the trees were assessed.

Results

Each specific aim will be restated followed by the results of analyses performed to reach the aim.

Aim 1: Categorize patients into risk status groups according to stage, age at diagnosis, N-myc amplification status, and Shimada histology summary

The clinical variables have the distributions displayed in Table 4. Although most of the cases were N-myc non-amplified and histology favorable, overall the data showed diverse clinical presentation with all subgroups of variables being well represented.

Table 4: Frequency of neuroblastoma cases in subgroups of stage, age at diagnosis, N-myc status, and histology summary

Variable	Subgroups	Number of Cases
Stage	1	24
	2	50
	3	42
	4	96
	4s	23
Age at diagnosis	Less than1 year	117
	Greater than 1 year	118
N-myc status	Amplified	40
	Non-amplified	194
	Unknown	1
Histology summary	Favorable	121
	Indeterminate	5
	Unfavorable	56
	Unknown	53

Assignment of clinical variables into the risk status variable resulted in the distribution of cases displayed in Table 5. Overall, distribution of cases among the risk groups was fairly even. Most cases (39%) were classified as high risk, closely followed by low risk (35%). Slightly fewer fell into the intermediate risk group (26%).

Table 5: Frequency of neuroblastoma cases in each risk group stratified by subgroups of stage, age at diagnosis, N-myc status, and histology summary

		Risk Status		
		Low	**Intermediate**	**High**
Stage	1	24		
	2	48		2
	3		27	8
	4		27	68
	4s		1	2
Age at diagnosis	Less than 1 year	46	45	5
	Greater than 1 year	26	10	75
N-myc status	Amplified	1		39
	Non-amplified	71	55	41
Histology summary	Favorable	58	38	7
	Indeterminate			1
	Unfavorable	7	4	45
	Unknown	7	13	27
Total		**72**	**55**	**80**

The risk status variable is composed of 207 valid cases and 28 missing values. These missing cases are the result of unknown and indeterminate histology classifications, one missing N-myc amplification status, and unknown DNA index information. The breakdown of missing cases by clinical variables is shown in Table 6. Unfortunately, most of the 4s cases could not be included in this categorical variable as a result of unavailable DI data.

Most of these missing cases were included as useable data in several analyses, however, by the creation of the two binary variables Risk 2 (low and intermediate versus high risk) and Risk 3 (low versus intermediate and high risk). The distributions of these variables are shown in Table 7. The total number of cases classified is 227 for Risk 2 and 215 for Risk 3.

Table 6: Frequency of missing cases stratified by risk status determinants

Number of Cases	Stage	Age at Diagnosis	N-myc Status	Histology Summary	DNA Index	Possible Risk Groups
4	3	> 1 year	non-amplified	indeterminate	N/A	Intermediate or High
3	3	> 1 year	non-amplified	unknown	N/A	Intermediate or High
1	4	<1 year	unknown	unknown	N/A	Intermediate or High
2	4s	< 1 year	non-amplified	unknown	unknown	Low or Intermediate
18	4s	< 1 year	non-amplified	favorable	unknown	Low or Intermediate

Table 7: Frequency of cases in subgroups of dichotomous risk variables (Risk 2 and Risk 3)

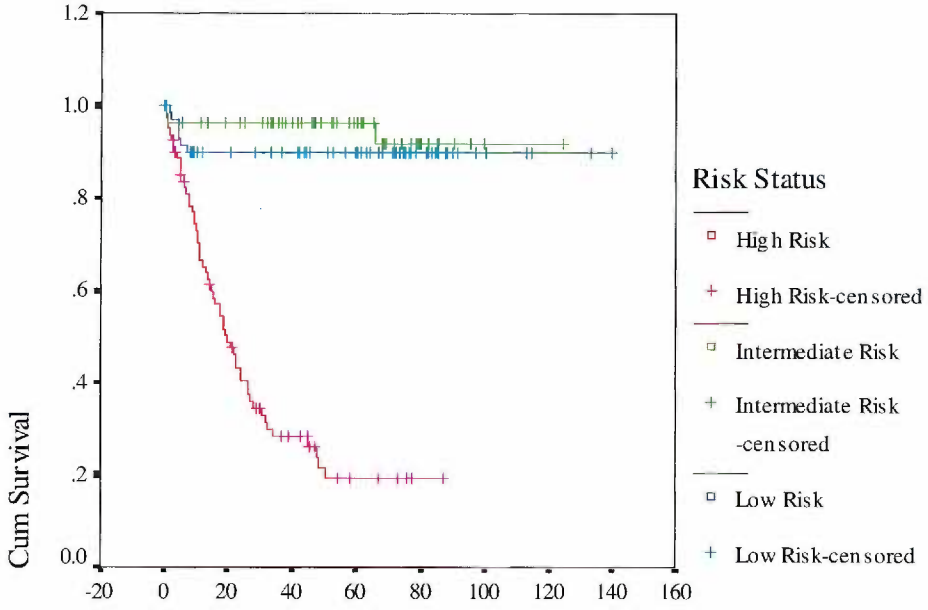
Low/Intermediate vs. High Risk	Low/Intermediate Risk	147
(Risk 2)	High Risk	80
	Missing	8
Low vs. Intermediate/High Risk	Low Risk	72
(Risk 3)	Intermediate/High Risk	143
	Missing	20

Mean survival times for each risk group and tests of significance by Kaplan-Meier are shown in Table 8. This table, along with the survival plot (see Figure 9), reveals significantly longer survival times of the low and intermediate risk groups when compared to the high risk group. Mean survival times do not differ significantly between the low and intermediate risk groups. In each of these two groups less than ten percent of the cases expired or had progressive disease, compared to more than two thirds of the high risk cases.

Table 8: Kaplan-Meier and log-rank results for analysis of time to disease progression/death stratified by risk group

Risk Group	Mean Survival Time	95% Confidence Interval	Number of Events	Number Censored	Percent Censored
Low Risk	125.8	(116.0, 135.6)	7	65	90.3
Intermediate Risk	117.1	(109.0, 125.2)	3	52	94.6
High Risk	31.3	(24.2, 38.4)	57	23	28.8
	Statistic	**df**	**p-value**		
Log Rank	101.4	2	<0.001		

Figure 9: Kaplan-Meier survival curve for analysis of time to disease progression/death stratified by risk group



Time (days) to disease progression/death

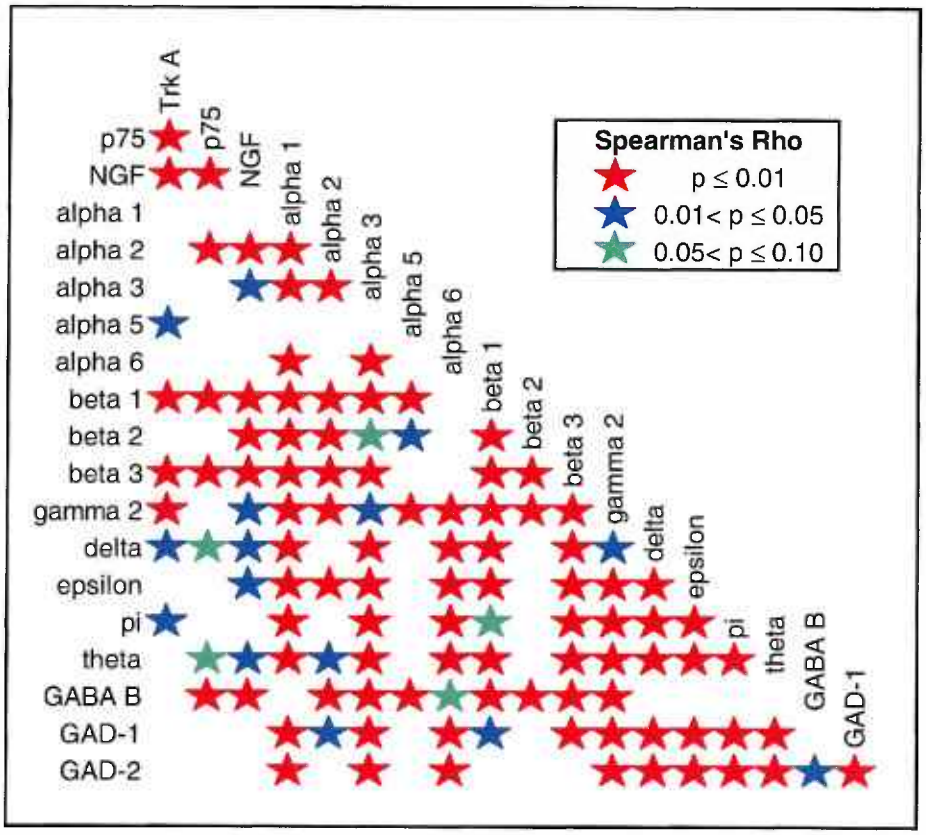
Aim 2: Examine the relationships between the gene expressions in Trk and GABA families

Gene expression quantification variables

Genes with expression detected by RT-PCR include GABA A (subgroups alpha 1-3 and 5-6, beta 1-3, gamma 2, delta, epsilon, pi, and theta) GABA B, GAD 1 and 2, TrkA, NGF, and p75. Genes with cycle threshold values greater than 38 for all observations include GABA A alpha 4, GABA A gamma 3, and GABA C. These three variables were excluded from further analyses, as cycle thresholds greater than 38 indicate that no gene expression was detected.

Distributions of gene expression variables were found to be consistently and strongly skewed to the right with most values close to zero. The $\log_{10}$ transformations made the majority of the variables' distributions more closely resemble a normal distribution. Correlations were determined conservatively using the nonparametric Spearman rank method. Many pairs of variables exhibit significant correlation. In Figure 10, significantly correlated pairs of variables are indicated using stars color-coded by ranges of p-values. Approximate absolute values of Spearman's correlation coefficient are between 0.108 and 0.131 for green stars, 0.132 and 0.167 for blue stars, and greater than 0.168 for red stars (see Appendix B for a complete table of correlations).

Figure 10: Correlations between gene expression variables displayed by p-value ranges of Spearman's rho



Principal components analysis

An analysis was performed in which three principal components were created. GABA A alpha 5 failed to load significantly into a factor, and NGF loaded equally into two. Neither of these variables was included in the final factor scores. The only variables to contribute significantly to the third factor were Trk A and p75 (see Table 9). Both these variables are missing data for several cases, and were not included in the final factor scores.

Table 9: Relative contributions of gene expression variables to components produced by principal components analysis

	Component		
	1	2	3
GABA A epsilon	0.907		
GAD-1	0.904		
GABA A alpha 1	0.884		
GABA A pi	0.873		
GABA A theta	0.859		
GABA A alpha 6	0.747		
GAD-2	0.690		
GABA A delta	0.663		
GABA A alpha 3	0.610		
GABA B		0.772	
GABA A alpha 2		0.771	
GABA A beta 1		0.705	
GABA A beta 3		0.704	
GABA A beta 2		0.699	
GABA A gamma 2		0.490	
Trk A			0.889
p75			0.793

Extraction method: Principal Component Analysis
Rotation method: Varimax with Kaiser Normalization
GABA A alpha 5 and NGF omitted
Cells <0.25 suppressed

A PCA that produced two factors was the preferred solution. This solution explained 60.8% of the variance of the 15 variables (Table 10). Coefficients were determined for each variable that reflected the relative contribution of the variable to the factor. These coefficients were then multiplied by the gene expression values for each of the cases. Factor scores were calculated by taking the sum of all of these coefficient-variable products. The scores from both factors were saved as new variables called Factor 1 and Factor 2 in the data set. In addition, the two variables (Trk A and p75) that contributed highly to the third factor obtained from the second analysis (see Table 9) were averaged to create a new variable (Trk75).

Table 10: Relative contributions of gene expression variables to components produced by preferred principal components analysis

	Component	
	1	2
GABA A epsilon	0.911	
GAD-1	0.908	
GABA A pi	0.882	
GABA A alpha 1	0.878	
GABA A theta	0.859	
GABA A alpha 6	0.754	
GAD-2	0.716	
GABA A delta	0.689	
GABA A alpha 3	0.582	
GABA A beta 1		0.763
GABA A alpha 2		0.763
GABA B		0.746
GABA A beta 3		0.737
GABA A beta 2		0.689
GABA A gamma 2		0.533

Extraction method: Principal Component Analysis
Rotation method: Varimax with Kaiser Normalization
GABA A alpha 5, NGF, Trk A and p75 omitted
Cells <0.25 suppressed

The Cronbach's alpha (a value that evaluates internal consistency among variables) was 0.9110 and 0.7671 for Factor 1 and Factor 2, respectively. There were no variables in either factor whose exclusion would result in a sizeable increase in alpha. Most of the variables, had they been removed one by one, would have resulted in decreasing the alpha value.

New variables: Factors 1 and 2 and Trk75

As shown in Table 10, variables contributing most strongly to Factor 1 include GABA A epsilon, GAD-1, GABA A pi, GABA A alpha 1, GABA A theta, GABA A alpha 6, GAD-2, GABA A delta, and GABA A alpha 3. Factor 1 values ranged from –2.59 to 3.05. The lowest quartile of Factor 1 scores had an average of 5.7 of these variables expressed. The second, third, and highest quartiles had means of 7.8, 8.4, and 8.8 variables expressed, respectively.

Most of the weight in the second factor was contributed by the variables GABA A beta 1, GABA A alpha 2, GABA B, GABA A beta 3, GABA A beta 2, and GABA A gamma 2. Values of this variable ranged from a low of –2.02 to a high of 3.26. The first quartile showed an average of 2.6 of these variables expressed. The second and third quartiles had means of 4.0 and 4.7 variables, and the highest quartile had an average of 5.6 variables expressed (see Table 11).

Table 11: Average number of highly contributing gene expression variables expressed in quartiles of each PCA factor

	Average # genes expressed				Total # genes included in factor score
	Quartile				
	1st	2nd	3rd	4th	
Factor 1	5.7	7.8	8.4	8.8	9
Factor 2	2.6	4.0	4.7	5.6	6

Factors 1 and 2 had distributions close to normal. The new variable created by taking the average of Trk A and p75 (Trk75) exhibited a slight right skew. Boxplots of Factor 1 and Trk75 revealed lower median values in the low risk groups with a linear increase through the high risk group. A comparison of the boxplots depicting Factor 2 values for each risk group, however, showed nonlinear results, with the intermediate risk group displaying the lowest median value.

Aim 3: Determine the associations between gene expression variables and risk status groups

Results of analysis of variance for each gene expression variable and composite variable are shown in Table 12. Five of the gene expression levels differed significantly ($p<0.05$) between the three risk groups; Trk A, p75, GABA A delta, NGF, and Trk75 composite score. The low risk group is significantly different from the high risk group for all five variables. For Trk A, p75, and Trk75 composite, the intermediate risk group significantly differed from the high risk group, but the low and intermediate groups did not differ significantly. In contrast, GABA A delta low and intermediate groups were significantly different, whereas intermediate and high risk groups were not.

Table 12: Risk group mean values and ANOVA results for gene expression variables and factor scores

Variable	Risk Group Means			F-score	p-value	Significantly different subgroups*	
	Low	Int.	High				
Individual genes							
Trk A	5.31	5.47	4.38	34.399	<0.001	Low/High**	Int/High**
p75	3.08	3.09	2.24	15.429	<0.001	Low/High**	Int/High**
GABA A delta	0.22	-0.21	-0.35	10.557	<0.001	Low/Int***	Low/High**
NGF	0.88	0.61	0.38	3.382	0.036	Low/High****	
GABA A gamma2	1.70	1.22	1.46	2.612	0.076		
GABA A theta	0.12	-0.07	-0.04	2.297	0.103		
GABA B	-0.80	-1.10	-0.64	2.185	0.115		
GABA A alpha 2	-0.11	-0.59	-0.29	2.099	0.125		
GABA A epsilon	-0.70	-0.93	-0.85	2.037	0.133		
GABA A alpha 1	1.21	0.94	1.08	1.559	0.213		
GABA A beta 1	0.54	0.33	0.19	1.544	0.216		
GABA A alpha 6	-1.44	-1.64	-1.59	1.537	0.217		
GABA A alpha 5	-1.47	-0.93	-1.30	1.372	0.256		
GABA A beta 2	-1.88	-1.93	-1.88	1.235	0.293		
GAD-1	0.69	0.37	0.66	1.202	0.303		
GABA A pi	0.62	0.62	0.50	1.121	0.328		
GABA A beta 3	1.09	0.90	1.04	0.688	0.504		
GAD-2	-1.26	-1.27	-1.22	0.328	0.721		
GABA A alpha 3	-0.21	-0.23	-0.20	0.010	0.990		
Composite variables							
Trk75	4.29	4.30	3.30	32.096	<0.001	Low/High**	Int/High**
Factor 1	0.15	-0.09	-0.07	1.897	0.153		
Factor 2	0.21	-0.14	0.11	1.442	0.239		

*Multiple comparisons by Tukey's HSD
** $p<0.001$
*** $p<0.01$
**** $p<0.05$

Aim 4: Create a model to describe risk group based on factors and/or individual gene expressions

Regression analysis

Risk 2: Low and Intermediate risk vs. High risk

Several individual genes achieved statistical significance ($p<0.05$) when regressed one at a time onto the Risk 2 variable. These variables include Trk A, p75, NGF, GABA A delta, and GABA B. Considering the composite variables (Factors 1 and 2 and Trk75), only Trk75 proved to be an independent significant predictor of Risk 2. All gene expression variables with the exception of GABA B gave odds ratios less than 1.0, indicating an inverse relationship between gene expression levels and risk group. However, it appears that when low and intermediate risk groups are pooled, higher GABA B expression is associated with increased risk.

In a multiple logistic regression model considering composite variables only as predictors, Factor 2 and Trk75 proved to be significant, whereas Factor 1 did not. Model statistics included a deviance value of 188.459, and a Cox and Snell's R^2 of 0.285. Only GABA B expression and Trk75 were significant in analyses using individual gene expression variables to predict Risk 2. This second model appeared to be a slightly better fit for the data than was the first, reducing the deviance to 185.160 and increasing the R^2 value to 0.298. The results of univariate and multiple logistic regression analyses for the Risk 2 variable are summarized in Table 13.

Table 13: Summary of logistic regression models using Risk 2 (low and intermediate risk vs. high risk) as the dependent variable.

Univariate analyses				
	Wald Chi Square	**p-value**	**Odds Ratio**	**95% Confidence Interval for OR**
Individual genes				
TrkA	43.153	<0.001	0.280	(0.192, 0.410)
p75	20.589	<0.001	0.389	(0.259, 0.585)
NGF	4.068	0.044	0.749	(0.565, 0.992)
GABA A alpha 1	0.002	0.963	1.007	(0.742,1.368)
GABA A alpha 2	0.674	0.412	1.091	(0.887, 1.341)
GABA A alpha 3	0.006	0.940	1.010	(0.788, 1.294)
GABA A alpha 5	0.165	0.685	0.969	(0.833, 1.128)
GABA A alpha 6	0.861	0.354	0.835	(0.570, 1.223)
GABA A beta 1	1.327	0.249	0.876	(0.700, 1.097)
GABA A beta 2	1.302	0.254	2.210	(0.566, 8.636)
GABA A beta 3	0.281	0.596	1.085	(0.803, 1.466)
GABA A gamma 2	0.011	0.915	1.013	(0.801, 1.281)
GABA A delta	9.212	0.002	0.566	(0.392, 0.817)
GABA A epsilon	0.161	0.688	0.921	(0.616, 1.377)
GABA A pi	2.545	0.111	0.663	(0.401, 1.098)
GABA A theta	0.569	0.451	0.836	(0.525, 1.331)
GABA B	5.174	0.023	1.295	(1.036, 1.617)
GAD-1	0.546	0.460	1.081	(0.879,1.331)
GAD-2	0.799	0.371	1.431	(0.652, 3.140)
Composite variables				
Factor 1	0.726	0.394	0.887	(0.673, 1.169)
Factor 2	0.996	0.318	1.148	(0.875, 1.507)
Trk75	36.724	<0.001	0.208	(0.125, 0.346)
Multivariate models				
	Wald Chi Square	**p-value**	**Odds Ratio**	**95% Confidence Interval for OR**
Model 1				
Trk75	38.652	<0.001	0.178	(0.103, 0.306)
Factor 2	6.162	0.013	1.610	(1.105, 2.346)
Model 2				
Trk75	37.847	<0.001	0.188	(0.110, 0.320)
GABA B	9.148	0.002	1.603	(1.181, 2.176)

Highlighted results: $p<0.05$
OR = odds ratio

Risk 3: Low risk vs. Intermediate and High risk

Regressing individual gene expression variables onto Risk 3 in univariate analyses yielded more significant predictors than obtained in the previous analyses using Risk 2 as the outcome variable. As seen for Risk 2, Trk A, p75, NGF, and GABA A delta all proved to be individually predictive of Risk 3. In addition, GABA A gamma 2 and GABA A theta showed significance. GABA B, however, was no longer predictive of outcome when intermediate risk was grouped with high risk instead of low risk. Of the composite variables, only Trk75 was a significant predictor in univariate analyses. Odds ratios generated by these analyses are all less than one. It appears, therefore, that these gene expression levels decrease as risk level increases.

In multivariate models built using Risk 3 as the dependent variable and Factors 1 and 2 and Trk75 as independent variables, the latter two proved to be significant predictors. Deviance and Cox and Snell's R^2 values for this model were 178.51 and 0.126, respectively. Using individual gene expressions instead of factors yielded a similar model that used Trk75 and GABA A delta expression to predict Risk 3. This model appeared to improve slightly upon the first, reducing the deviance to 172.70 and increasing R^2 to 0.155. Expression variables GABA A alpha 3 and GABA A alpha 6 were also significant in models using individual gene expressions. These variables were too highly correlated to be included together in a model, however, and neither was significant when considered independently. Results of these analyses are summarized in Table 14.

Table 14: Summary of logistic regression models using Risk 3 (low risk vs. intermediate and high risk) as the dependent variable.

Univariate analyses				
	Wald Chi Square	**p-value**	**Odds Ratio**	**95% Confidence Interval for OR**
Individual genes				
TrkA	11.675	0.001	0.549	(0.389, 0.774)
p75	7.707	0.005	0.550	(0.360, 0.839)
NGF	4.964	0.026	0.648	(0.442, 0.949)
GABA A alpha 1	2.262	0.133	0.774	(0.555, 1.081)
GABA A alpha 2	3.179	0.075	0.819	(0.657, 1.020)
GABA A alpha 3	0.003	0.957	0.993	(0.761, 1.294)
GABA A alpha 5	1.606	0.205	1.119	(0.940, 1.333)
GABA A alpha 6	3.508	0.061	0.680	(0.454, 1.018)
GABA A beta 1	3.214	0.073	0.804	(0.633, 1.021)
GABA A beta 2	0.855	0.355	0.521	(0.131, 2.074)
GABA A beta 3	0.673	0.412	0.873	(0.632, 1.207)
GABA A gamma 2	4.828	0.028	0.752	(0.583, 0.970)
GABA A delta	16.655	<0.001	0.465	(0.322, 0.672)
GABA A epsilon	3.755	0.053	0.652	(0.423, 1.005)
GABA A pi	1.031	0.310	0.769	(0.464, 1.277)
GABA A theta	4.507	0.034	0.590	(0.363, 0.960)
GABA B	0.148	0.700	0.956	(0.761, 1.201)
GAD-1	0.810	0.368	0.902	(0.719, 1.130)
GAD-2	0.069	0.793	1.114	(0.497, 2.496)
Composite variables				
Factor 1	3.023	0.082	0.768	(0.571, 1.034)
Factor 2	2.493	0.114	0.796	(0.599, 1.057)
Trk75	13.532	<0.001	0.376	(0.224, 0.633)
Multivariate models				
	Wald Chi Square	**p-value**	**Odds Ratio**	**95% Confidence Interval for OR**
Model 1				
Trk75	14.210	<0.001	0.355	(0.207, 0.608)
Factor 1	5.273	0.022	0.620	(0.413, 0.932)
Model 2				
Trk75	11.821	0.001	0.388	(0.226, 0.666)
GABA A delta	10.488	0.001	0.458	(0.285, 0.734)

Highlighted results: $p<0.05$
OR = odds ratio

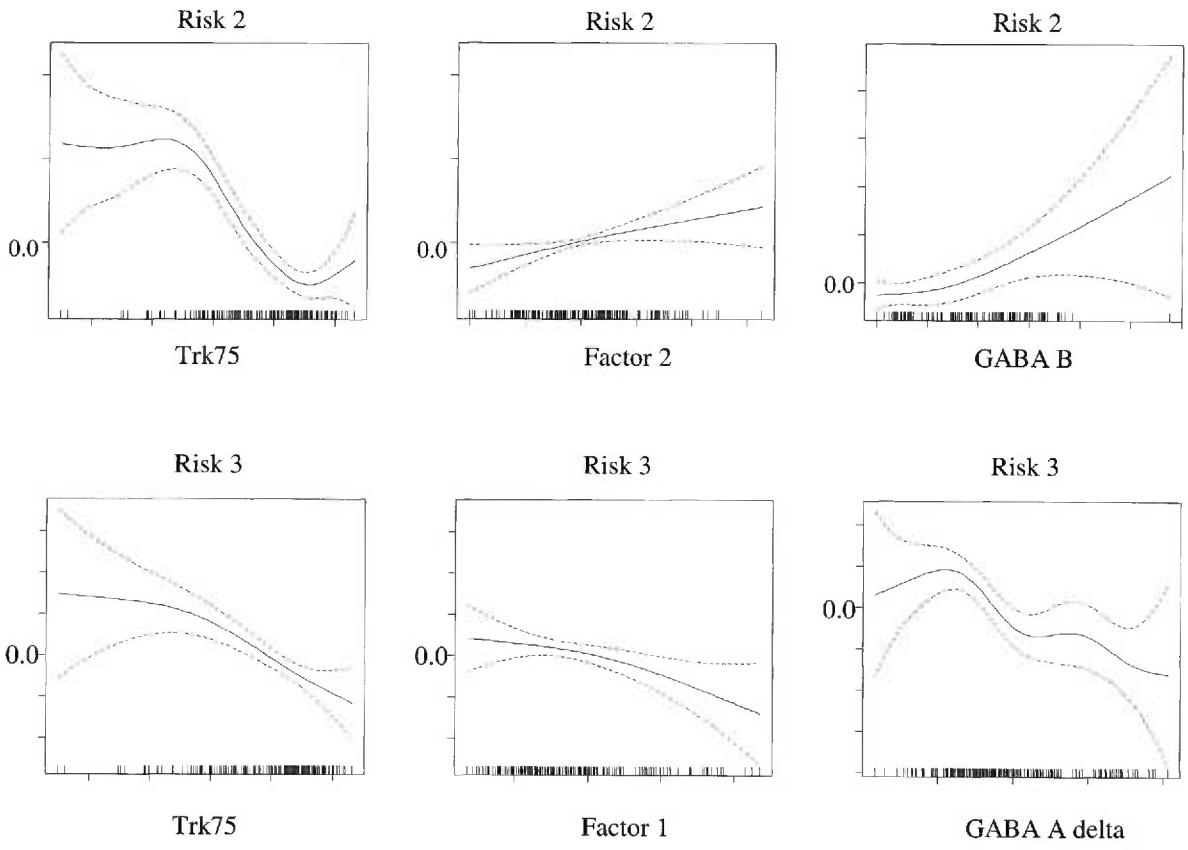
Evaluation of predictor linearity

Spline plots created to evaluate the linearity of independent variables significant in logistic models predicting Risk 2 and Risk 3 yielded mixed results. In Risk 2 models, spline plots reveal some apparent nonlinearity in Trk75, whereas plots for Factor 2 and GABA B do not deviate substantially from linearity. In Risk 3 models, spline plots of both Trk75 and Factor 1 appear fairly linear within the bulk of their observations. GABA A delta's plot exhibits some nonlinearity, however there appears to be an overall downward trend within most of the observations (Figure 11).

Evaluation of regression models using coding variables

Regression models using coding variables created to divide the most nonlinear variables (Trk75 for Risk 2 and GABA A delta for Risk 3) into ranges of values exhibiting uniform activity did not improve upon original regression models. The level of non-linearity exhibited by these variables was not sufficient to remove them from logistic regression models, so original models were retained.

Figure 11: Spline plots for independent variables significant in logistic regression models predicting Risk 2 and Risk 3



Plots depict values of each observation of the various predictor variables along the x-axes by their corresponding spline functions along the y-axes for outcome variables Risk 2 and Risk 3. Spline functions with negative slopes are indicative of predictor variables having a protective effect whereas spline functions with positive slopes are indicative of predictor variables conveying risk. At spline function values of 0.0 relative risk is null.

Logistic regression: two group comparisons

Binary logistic regressions run on subsets of data isolating low and intermediate risk, intermediate and high risk, and low and high risk groups produced the models detailed in Table 15.

Table 15: Summary of best regression model results: gene expression variables regressed onto binary risk outcome variables (low vs. intermediate risk, intermediate vs. high risk, and low vs. high risk)

Logistic regression: individual group data					
Dependent variable	**Predictor variables**	**Wald Chi Square**	**p-value**	**Odds Ratio**	**95% Confidence Interval for OR**
Low vs. Intermediate risk	GABA A delta	7.954	0.005	0.503	(0.313, 0.811)
Intermediate vs. High risk	Trk75	23.541	<0.001	0.168	(0.082, 0.345)
	GABA B	6.194	<0.013	1.671	(1.115, 2.503)
Low vs. High risk	Trk75	20.981	<0.001	0.237	(0.128, 0.439)
	GABA A delta	7.365	0.007	0.448	(0.251, 0.800)

These models further illuminate the relationships between the groups indicated by ANOVA and previous regression analyses. Values of odds ratios for GABA A delta in models predicting low versus intermediate risk and low versus high risk are 0.503 and 0.448, respectively. These values suggest that high GABA A delta levels may be predictive of low risk disease, and that levels progressively decrease as risk increases. High levels of Trk75 appear to be predictive of high risk disease (odds ratios 0.168 and 0.237 for intermediate versus high risk and low versus high risk), and high levels of GABA B may distinguish the high risk from the intermediate risk group (odds ratio 1.671).

Prediction of intermediate class

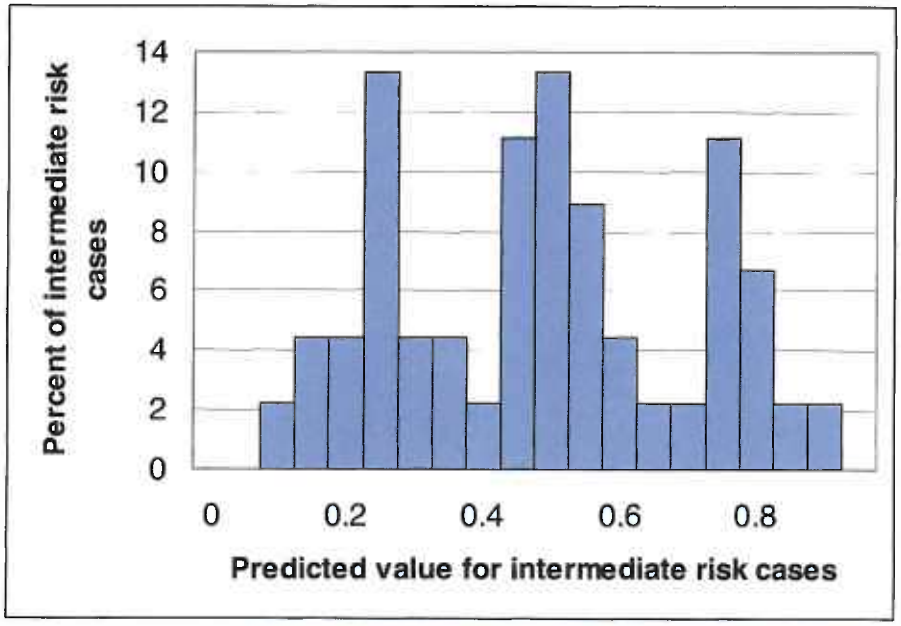
Regression analysis of gene expression variables onto the binary outcome variable low vs. high risk group produced the following equation for determining predicted probability:

$$\hat{\pi}(x) = \frac{e^{5.943 + 1.439 * Trk75 - 0.802 * GABAAdelta}}{1 + e^{5.943 - 1.439 * Trk75 - 0.802 * GABAAdelta}}$$

Prediction of values for intermediate risk group subjects produced a three-peaked distribution (see Figure 12). The first peak covers the area under which most of the values for low risk patients fell. Values in the third peak were similar to the bulk of values obtained for the high risk individuals. The middle peak, whose values fell directly between the other two, might represent individuals whose outcomes are truly more severe than low risk patients but better than the high risk group.

Since no other data set was available with which to classify patients according to Trk75 and GABA A delta values, cross-validation with intermediate risk patients not included in the regression analysis was necessary. Thus, the histogram depicted in Figure 12 was reflects the values obtained only on a small subset of the available data.

Figure 12: Histogram of predicted values for intermediate risk group individuals obtained by the regression model of GABA A delta and Trk75 expressions regressed onto low vs. high risk groups

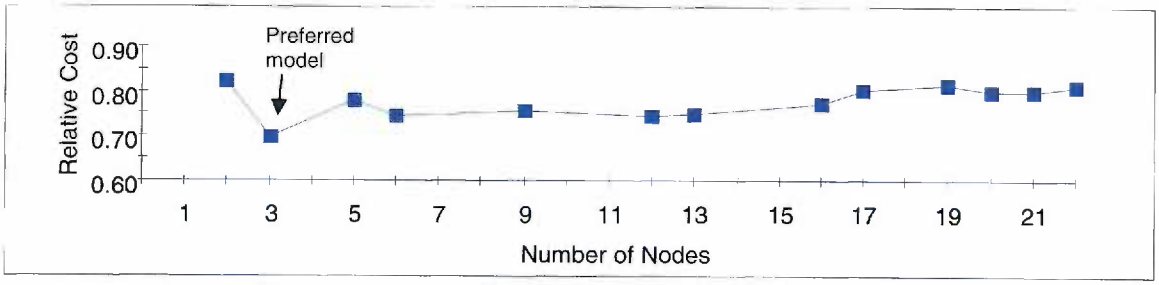


Aim 5: Determine how well risk groups can be classified by using combinations of factors and individual risk variables

Classification Trees

When Trk 75 and Factors 1 and 2 were used to classify risk group, a series of trees were generated with the error curve shown in Figure 13. The preferred tree chosen from this series has a relative cost of 0.699, and only 3 terminal nodes. This model corresponded to the lowest relative cost achieved in the series.

Figure 13: Error curve for series of classification trees using Factors 1 and 2 and Trk75 to classify risk



Splitters and resulting nodes of the preferred model are depicted in Figure 14. The first node created in this tree is comprised of 54 cases, mostly high risk, determined by the first splitter, Trk75 $\leq$ 3.56. The remainder of the data is then divided into two nodes based on Factor 1 scores of greater than or less than 0.822. The second node is composed of 99 predominantly intermediate risk cases, and the third node consists of 14 cases, most of which are low risk. The second node contains much more misclassification than the other two. Overall misclassification rates for this tree were very high for the low risk group, and fairly high for the high risk group (Table 16). Most

misclassification was due to the large proportion of low and high risk individuals in node 2.

Figure 14: Preferred classification tree produced by using Factors 1 and 2 and Trk75 to classify risk group

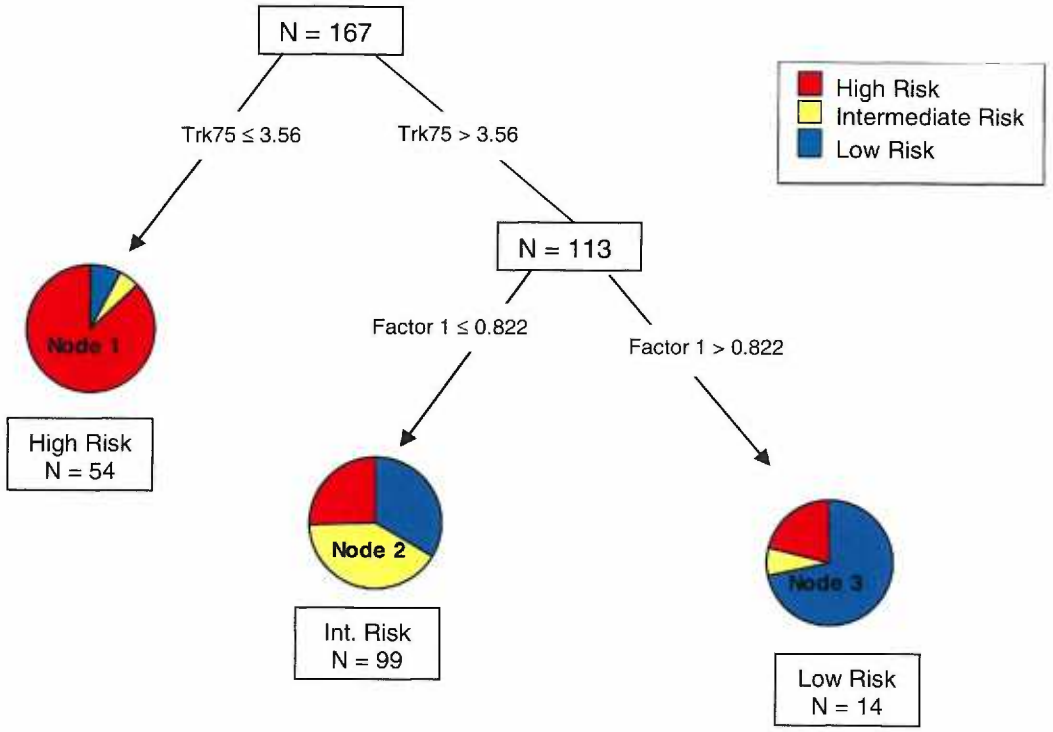


Table 16: Misclassification rates for classification trees using factors 1 and 2 and Trk75 to classify risk

Class	% Misclassified	
	Learn	Cross-validation
Low	78.72	76.60
Intermediate	8.89	17.78
High	37.33	45.33

The error curve in Figure 15 depicts a series of trees created using all gene expression variables. Choosing these variables as predictors enabled the inclusion of cases for which values of Trk75 are missing. The preferred tree selected from this grove has a cost of 0.793. This model was selected instead of the one having the lowest cost (0.753) because that tree is excessively complex, having many more terminal nodes, each of which contained few cases.

Figure 15: Error curve for series of classification trees using individual gene expression level variables to classify risk

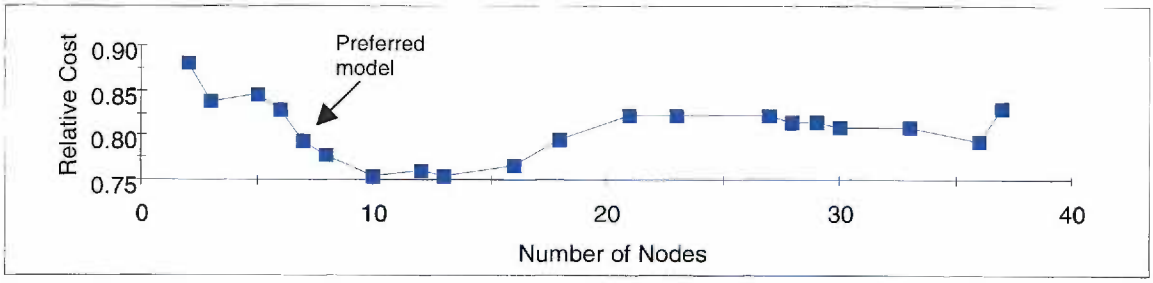
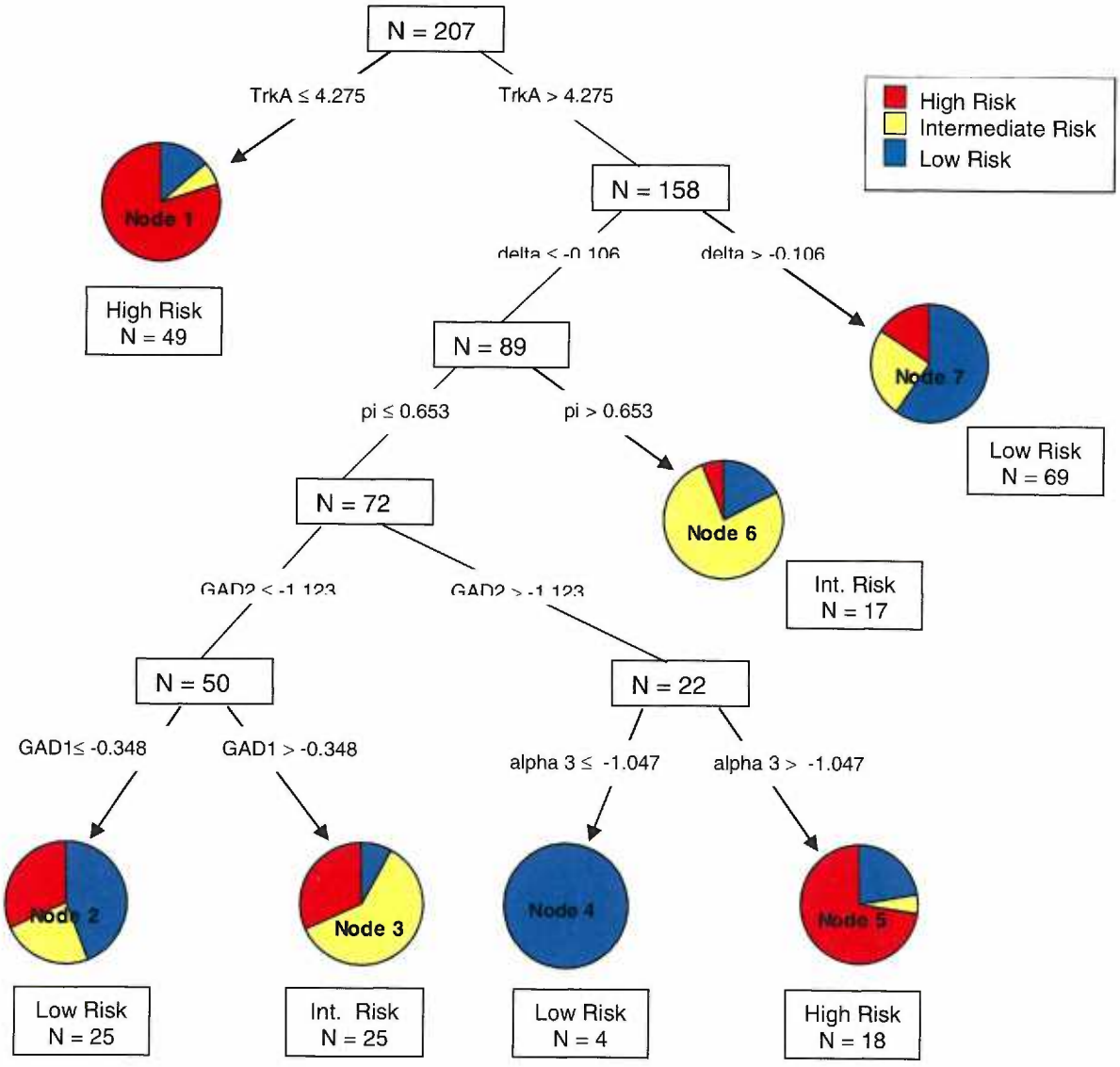


Figure 16 shows splitters and terminal node details for the preferred model. Cases with TrkA values less than 4.275 comprise the first node. This node resembles the first node seen in the previous model, consisting mainly of high risk cases. The second splitter used is a GABA A delta value of –0.106. This split creates node 7, largely consisting of low risk individuals. A GABA A pi value of 0.653 is then used to isolate node 6, a small group of mostly intermediate cases. The remainder of the data is split using the GAD-2 value of –1.123 to create two subsets, but no nodes. The first of these subsets is split using GAD-1 of –0.348 to create nodes 2 and 3. Node 2, containing much misclassification, is composed of 25 cases, a slight majority of which are low risk. Node 3 contains somewhat less misclassification, most of the 25 cases being intermediate risk.

The second subset of the GAD-1 split is further separated based on a GABA A alpha 3 value –1.047. This split creates node 4, a small group of low risk cases, and node 5, containing a predominance of high risk cases.

Figure 16: Preferred classification tree produced by using individual gene expression variables to classify risk group



Misclassification rates for low and high risk groups were lower than those seen in the previous model, however the rate for the intermediate group increased. Rates for all three classes remain much higher than desirable (Table 17).

Table 17: Misclassification rates for classification trees using individual gene expression level variable to classify risk

Class	% Misclassified	
	Learn	**Cross-Validation**
Low	22.22	65.28
Intermediate	49.09	54.55
High	35.00	38.75

Discussion

Several types of analyses were conducted to fulfill the aims set forth in this study. The combination of these methodologies was used to examine the data from several perspectives, and to expose complexities in the relationships between the variables. Analyses of variance suggested key variables that might be useful as predictor variables in further analyses. Regression analyses identified relative levels of risk for single variables predicting risk group, as well as basic interactions between groups of these predictors. Classification trees were grown to elucidate complex interactions between variables that are unable to be determined by the strict linear relationships imposed upon the data by logistic regression. These interactions are discovered when data are examined, split into subgroups using cut points of predictor variables, and re-examined, exposing relationships that would require a combination of several regression models to achieve.

While examining the results of these methodologies, it is important to remember that the utility of these gene expression variables comes from their relative, not absolute, value. RT PCR studies evaluate relative gene expression levels by quantifying replication cycle time required to produce target quantities of cDNA from original mRNA levels of samples subjected to identical conditions. These cycle threshold values may vary from run to run based on many factors including cycling conditions and substrate amounts. Small variations in laboratory protocol, therefore, may result in different absolute values for an individual sample, but should not change relative values when comparing all samples. Thus, it is of little interest to examine the individual odds

ratio provided by the regression of a gene expression variable onto risk, or to examine the precise value of a variable used in growing a classification tree. It is, however, of interest to examine the relative values of these odds ratios, and to determine those genes conveying risk as opposed to those providing an apparent protective effect.

When the relationships between the Trk and GABA genes were explored, three groups of co-expressed genes emerged. Factors comprised of GABA genes were of little utility in distinguishing risk levels in further analyses. In the ANOVA, logistic regression and CART models developed to clarify relationships between risk groups and neurotrophin gene expression levels, individual expression variables proved more valuable as predictors than did these factors. The variable that averaged the p75 and Trk A expression levels (Trk75), however, proved to be a significant predictor of high risk in the regression analyses.

Several individual GABA gene expression variables surfaced during the analyses as potentially important predictors of risk group. GABA A delta was exposed as the second most valuable predictor of class in the CART model, and an important predictor of low risk in regression analysis. Results obtained for GABA B were puzzling. Significant in a regression model distinguishing the intermediate from the high risk group, this gene was the only single gene expression variable whose value did not exhibit an inverse relationship to risk. Studies of GABA's role in neuronal differentiation have discovered that the types of receptors present in young cells vary as they mature (51). Considering this, our results support the suggestion that disease severity may depend in part on the stage of cellular differentiation at the time of malignancy development.

Our finding that decreased Trk A expression predicts high risk group is in agreement with the results of other studies (39,57). High risk tumors generally express high levels of Trk B that binds BDNF and NT-4/5, promoting cell survival. The favorable prognoses experienced by tumors expressing high levels of Trk A result either from the binding of adequate NGF to stimulate cell maturation, or inadequate NGF, causing cell death. Our results provide further support for the proposal that Trk A levels be used routinely in the clinical assessment of disease prognosis.

Our analyses also exposed a strong correlation between p75 and Trk A. P75 was found to be a significant predictor of risk group independently of Trk A, but the predictive power of the composite variable Trk75 was stronger than that of either of its component variables alone. The action of p75 as reported by other neuroblastoma studies, however, has not been so consistently demonstrated as that of Trk A. A recent study reported the results of an in vitro analysis of two different neuroblastoma cell lines both expressing p75, and neither expressing Trk A. The addition of NGF to the cell lines had the opposite effects of increasing proliferation in one clone, and decreasing proliferation in the other (58). Another study unearthed a possible new role for p75 as an inducer of cellular death (48). Surprisingly, the p75 gene has been mapped to 17q 21-22 (59), and thus is likely present in more than one copy in tumors exhibiting 17q gains. As discussed earlier, 17q gains have been identified in a high percentage of tumors having poor outcomes. Based on the strong correlation of Trk A and p75 seen in our study, it would be expected that these tumors would have more favorable outcomes, however the reverse has been shown consistently (42). This diverse set of results suggests that further study of the role of p75 in neuroblastoma is warranted.

A few of our results reinforce the conclusions reached by Roberts et al in the primary examination of these data (53). This initial set of analyses focused on the relationship between GABA A receptor components and patient outcome as measured by time to disease progression or death. The authors found that higher levels of GABA A delta correlated with longer survival and remission times in subsets of cases. Their finding is strengthened by the results of our analyses. High expression levels of this subunit were found to be predictive of low risk group. As mentioned by the authors of the primary analysis, chromosomal mapping of the GABA A delta gene further reinforces this conclusion. This gene was mapped to 1p36.3, an area of chromosome 1 commonly deleted in aggressive and rapidly growing tumors (60).

The contributors to the primary analysis also isolated two factors by principal components analysis. These factors were very similar in variable composition to those created by our analyses. Factor 1 of our study, however, also contained GAD-1 and GAD-2, and Factor 2 contained GABA B. In the previous study, high values of the factor that the authors termed GAP (analogous to our Factor 2) were found to be predictive of longer survival times. Our attempts to distinguish risk groups using Factor 2 did not yield complementary results. Although Factor 2 was found to be significant in regression analysis, higher levels of this factor actually predicted high risk group. Potentially the inclusion of GABA B with its peculiar properties into Factor 2 led to this disparity in results.

Overall, results from this study complement those of the primary analysis by examining these data from a different perspective. The use of a different outcome variable avoided some problems, but introduced others. Early in our analyses survival

times by risk group were examined. Significant differences in survival time were exhibited between the low and high risk groups, and between the intermediate and high risk groups. Survival times for the low and intermediate risk groups, however, were very similar. This phenomenon may be attributed to the moderately intensive and successful treatment given to intermediate risk patients (26). Since no treatment data were available, examination of differences in gene expression levels based solely on time to outcome would fail to distinguish differences between low and intermediate risk individuals. All individuals with favorable outcomes were considered together, regardless of how intensive the treatments were to achieve these outcomes.

While our analyses enabled comparison of low and intermediate risk groups, other problems were introduced into these analyses by using clinically derived risk groups as the outcome variable. The current definitions of these risk groups were developed as the International Neuroblastoma Risk Groups (INRG) in response to the need for an improved system of prognostication that would include newly discovered predictors of outcome in addition to age and INSS stage (47). The Children's Oncology Group adopted this risk stratification system to standardize treatment recommendations. Although the currently accepted determinants of risk group include age, stage, histopathology summary, DI, and N-myc amplification status, this classification system is still a work in progress, and new prognostic indicators are under evaluation for their utility in improving this system of risk prediction.

While the use of the neuroblastoma risk stratification system is currently the most accurate system known to assess prognosis, it is certainly not a perfect predictor of true disease outcome. Indeed, cases have been reported that were classified as high risk by

these criteria, but were treated successfully using only basic surgical methods recommended for low risk cases (61). Considering the heavy reliance of this system on exact age at diagnosis, an indicator that may be dependent on many non-biological factors, it is easy to see how discrepancies between prognosis and actual outcome may be introduced.

Since the radically different treatment recommendations of COG depend on this system of risk stratification, the implications of inaccurate prediction of outcome are drastic over- or under-treatment of patients. The management of intermediate risk patients is especially problematic. Clinical trials are underway in COG-partnered institutions nationwide to evaluate less intensive treatment options in intermediate risk patients, and to study other molecular characteristics of this group's tumors. Goals of these studies include a refinement of prognostication criteria and determination of appropriate treatment regimes.

Classification and prediction of intermediate risk in our analyses proved troublesome. To explore the nature of neurotrophin gene expression levels of this group, a regression equation was generated using low and high risk cases, and used to obtain predicted values for the intermediate group cases. The distribution of predicted values seen for the intermediate group suggested a lack of homogeneity in this group with respect to gene expression predictors. Perhaps our intermediate group consists of an amalgam of cases with varying levels of true severity.

Limitations

Our study was somewhat limited by missing data. Various cases were missing values for histology classifications, and others lacked gene expression levels of Trk A, p75 or NGF. Most 4s cases had to be omitted from analyses using the three-categoried risk status variable as the outcome due to missing DNA index data that prevented allocation of patients into the risk groups. These missing data decreased the useable sample size and power to detect significant differences. We were able to include many of these cases, however, by creating dichotomous risk variables for use in some analyses.

It is unlikely that much misclassification was introduced by the assessment of clinical prognosticators by the study's various constituent institutions. The determinations of N-myc amplification status and INSS stage are standardized. Tumor histology is subject to slightly more individual judgment, however all questionable cases were reviewed by Dr. Shimada at the Children's Hospital of Los Angeles.

Perhaps the most significant limitation hampering our study was the possible, but unknown, discrepancy between assigned risk group and actual risk posed by the malignancies. As mentioned earlier, the currently used COG risk group stratification system is a decidedly imperfect method with demonstrated shortcomings. While cases may have been accurately classified into risk group according to this system, they may actually have been misclassified in terms of true risk. This "misclassification" would have weakened associations between gene expression levels and outcome.

Assigning individuals to risk groups based on clinical prognosticators is also problematic in that it fails to allow the examination of the changes in gene expression occurring over the course of disease. It is likely that within the subgroups of low,

moderate, and high severity, fluctuations in gene expressions occur as the tumors grow, recede, or mature. Locking individuals into risk groups does not enable the study of these dynamics. To examine the changes in gene expression over time we should abandon the use of clinical prognosticators, and focus solely on gene expression levels and clinical outcomes adjusted for treatment.

Indications for future studies

Considering the disagreement between the current "gold standard" of neuroblastoma tumor risk assessment by the COG grouping system and true risk posed by tumors, a broad goal of future studies should continue to be improving disease prognostication. New technologies continue to increase the accuracy and ease in which gene expression levels can be determined. Ideally, studies should employ these technologies to develop models of tumor behavior based solely on neurotrophin gene expression levels to predict outcome in newly diagnosed cases. To develop these models it will be necessary to determine the patterns and changes in gene expression levels over time within the various strata of disease severity. The exclusion of currently used clinical prognosticators from future gene studies will be an important consideration in creating these models.

A prospective study involving COG-partnered institutions nationwide should be initiated. As with our study, tumor tissue should be excised and submitted for genetic studies to determine expression levels of neurotrophin genes. These genes should include the GABA family genes we examined, Trk A, B, and C, and p75. Vigilant follow-up of patients will be a critical step to ensuring the study's success. Remission times and dates

of disease progression and death will be necessary to create a measure of outcome. Complete and detailed treatment information should also be collected for outcome adjustment. In addition, tumors not completely excised at the time of diagnosis could continue to have aspirates taken over time. These additional samples might enable examination of the changes in gene expression levels as tumors progress either treated or untreated. To minimize bias that might be introduced by researchers' knowledge of assigned risk group, no clinical prognostication variables should be included into the data set. Since outcome requires adjusting by treatment, and risk group generally decides the treatment option, it may not be possible to truly "blind" researchers to traditional risk assessment. Nonetheless, efforts should be taken to avoid bias by original risk assignment, and to focus on drawing associations between neurotrophin gene expression levels and outcome. Ideally, these associations could be used to develop models of the continuums of various gene expression levels for malignancies corresponding to varying levels of severity.

Significance

Further knowledge of the role of neurotrophins and their receptors in the diverse clinical courses exhibited by neuroblastoma could lead to the discovery of novel prognosticators and an improved system of risk stratification. Consequently, more accurate predictions of disease outcome would decrease under- and over-treatment of neuroblastoma, thereby lowering mortality rates and decreasing unnecessary morbidity. In addition, an understanding of the biology underlying neuroblastoma may lead to specific, targeted approaches to therapy, further decreasing treatment-associated

morbidity. Many drugs are currently prescribed (the benzodiazepines, for instance), that function through interactions with the GABA receptors. Advances in the understanding of GABA involvement in neuroblastoma may reveal clinical utility of similar drugs in the treatment of subsets of this disease.

Conclusion

Through various statistical analyses we identified associations between the expression levels of a few neurotrophin receptor genes and neuroblastoma risk group. These genes included members of the GABA and Trk families. Although our results suggest that these gene expression levels may be used to predict tumor outcome, extensive future studies are needed to explore changes in gene expression over time within subsets of tumors, and to distinguish between different groups of malignancies leading to distinct outcomes. Implications of further understanding the behavior of neurotrophin-associated genes in neuroblastomas include an improved system of prognostication, targeted therapies, and decreased morbidity and mortality.

References

1. Ries LAG, Smith MA, Gurney JG, Linet M, Tamra T, Young JL, Bruin GR (eds). Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975-1995, National Cancer Institute, SEER Program. NIH Pub. No. 99-4649. Bethesda, MD, 1999.

2. Gurney JG, Davis S, Severson RK, Fang JY, Ross JA, Robison LL. Trends in cancer incidence among children in the U.S. Cancer 1996; 78: 532-541.

3. Izbicki T, Mazur J, Izbicki E. Epidemiology and etiology of neuroblastoma: an overview. Anticancer Research 2003; 23: 755-760.

4. Cancer reference information. American Cancer Society. Available at: http://www.cancer.org/docroot/CRI/CRI_0.asp

5. Schuz J, Kaletsch U, Meinert R, Kaatsch P, Spix C, Michaelis J. Risk factors for neuroblastoma at different stages of disease. Results from a population-based case-control study in Germany. Journal of Clinical Epidemiology 2001; 57: 702-709.

6. Kramer S, Ward E, Meadows AT, Malone KE. Medical and drug risk factors associated with neuroblastoma: a case-control study. Journal of the National Cancer Institute 1987; 78: 797-804.

7. Michalek AM, Buck GM, Nasca PC, Freedman AN, Baptiste MS, Mahoney MC. Gravid health status, medication use, and risk of neuroblastoma. American Journal of Epidemiology 1996; 143: 996-1001.

8. Mandel M, Toren A, Rechavi G, Dor J, Ben-Bassat I, Neumann Y. Hormonal treatment in pregnancy: a possible risk factor for neuroblastoma. Medical & Pediatric Oncology 1994; 23: 133-135.

9. Wilkins JR 3rd, Hundley VD. Paternal occupational exposure to electromagnetic fields and neuroblastoma in offspring. American Journal of Epidemiology 1990; 131: 995-1008.

10. Bunin GR, Ward E, Kramer S, Rhee CA, Meadows AT. Neuroblastoma and parental occupation. American Journal of Epidemiology 1990; 131: 776-780.

11. Spitz MR, Johnson CC. Neuroblastoma and paternal occupation. A case-control analysis. American Journal of Epidemiology 1985; 121: 924-929.

12. De Roos AJ, Teschke K, Savitz DA, Poole C, Grufferman S, Pollock BH, Olshan AF. Parental occupational exposures to electromagnetic fields and radiation and the incidence of neuroblastoma in offspring. Epidemiology 2001; 12: 508-517.

13. Kerr MA, Nasca PC, Mundt KA, Michalek AM, Baptiste MS, Mahoney MC. Parental occupational exposures and risk of neuroblastoma: a case-control study (United States). Cancer Causes & Control 2000; 11: 635-643.

14. Olshan AF, De Roos AJ, Teschke K, Neglia JP, Stram DO, Pollock BH, Castleberry RP. Neuroblastoma and parental occupation. Cancer Causes & Control 1999; 10: 539-549.

15. Carlsen NL. Neuroblastomas presenting in the first year of life: epidemiological differences from those presenting at older ages. Cancer Detection & Prevention 1996; 20 (3): 251-261.

16. Yang Q, Olshan AF, Bondy ML, Shah NR, Pollock BH, Seeger RC, Look AT, Cohn SL. Parental smoking and alcohol consumption and risk of neuroblastoma. Cancer Epidemiology, Biomarkers & Prevention 2000; 9: 967-972.

17. Neglia JP, Smithson WA, Gunderson P, King FL, Singher LJ, Robison LL. Prenatal and perinatal risk factors for neuroblastoma. A case-control study. Cancer 1988; 61: 2202-2206.

18. Hamrick SE, Olshan AF, Neglia JP, Pollock BH. Association of pregnancy history and birth characteristics with neuroblastoma: a report from the Children's Cancer Group and the Pediatric Oncology Group. Paediatric and Perinatal Epidemiology 2001; 15: 328-337.

19. Dodge HJ, Benner MC. Neuroblastoma of the adrenal medulla in siblings. Rocky Mt Med 1945; 42: 35-38.

20. Knudson AG, Strong LC. Mutation and cancer: neuroblastoma and pheochromocytoma. Amer J Hum Genet 1972; 24: 514-532.

21. Knudson AG. Mutation and cancer: statistical study of retinoblastoma. Proc Nat Acad Sci 1971; 68: 820-823.

22. Brodeur GM. Neuroblastoma: Biological insights into a clinical enigma. Nature Reviews 2003; 3: 203-216.

23. Castleberry RP. Biology and treatment of neuroblastoma. Pediatric Clinics of North America 1997; 44: 919-937.

24. Shimada H, Ambros IM, Dehner LP, Hata J, Joshi VV, Roald B. Terminology and morphologic criteria of neuroblastic tumors. Recommendations by the International Neuroblastoma Pathology Committee. Cancer 1999; 86: 349-363.

25. Breslow N, McCann B. Statistical estimation of prognosis of children with neuroblastoma. Cancer Res 1971; 31: 2098-2101.

26. Brodeur GM, Maris JM. Neuroblastoma. Principles and practice of pediatric oncology. Fourth edition. Lippincott Williams & Wilkins, Philadelphia PA, 2002: 896-937.

27. Brodeur GM, Seeger RC, Barrett A, Berthold F, Castleberry RP, D'Angio G et al. International criteria for diagnosis, staging and response to treatment in patients with neuroblastoma. Journal of Clinical Oncology 1988; 6: 1874-1881.

28. Brodeur GM, Pritchard J, Berthold F, Carlsen NLT, Castel V, Castleberry RP et al. Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. Journal of Clinical Oncology 1993; 11: 1466-1477.

29. Castel V, Garcia-Miguel P, Canete A, Melero C, Navajas A, Ruiz-Jimenez JI, Navarro S, Badal MD. Prospective evaluation of the International Neuroblastoma Staging System (INSS) and the International Neuroblastoma Response Criteria (INRC) in a multicentre setting. European Journal of Cancer 1999; 35: 606-611.

30. Ikeda H, Iehara T, Tsuchida Y, Kaneko M, Hata J, Naito H et al. Experience with International Neuroblastoma Staging System and Pathology Classification. British Journal of Cancer 2002; 86: 1110-1116.

31. Castleberry RP. Risk groups in loco-regional neuroblastoma. Medical and Pediatric Oncology 1998; 31: 253.

32. Shimada H, Roald B. Histology: tumors of the neuroblastoma group. Neuroblastoma. Elsevier Science, The Netherlands, 2000: 341-354.

33. Shimada H, Chatten J, Newton WA, Sachs N, Hamoudi AB, Chiba T, et al. Histopathologic prognostic factors in neuroblastic tumors; definition of subtypes of ganglioneuroblastoma and an age-linked classification of neuroblastomas. J Natl Cancer Inst 1984; 73: 405-416.

34. Shimada H, Inge MA, Dehner LP, Hata J, Joshi VV, Roald B, Stram DO et al. The International Neuroblastoma Pathology Classification (the Shimada System). Cancer, 1999; 86: 364-372.

35. Brodeur GM, Ambros PF. Genetic and biological markers of prognosis in neuroblastoma. Neuroblastoma. Elsevier Science, The Netherlands, 2000: 355-369.

36. Schwab M. MYCN amplification in neuroblastoma. Neuroblastoma. Elsevier Science, The Netherlands, 2000: 75-83.

37. Bordow SB, Norris MD, Haber PS, Marshall GM, Haber M. Prognostic significance of MYCN oncogene expression in childhood neuroblastoma. J Clin Oncol 1998; 16: 3286-3294.

38. Chan HSL, Gallie BL, DeBoer G, Haddad G, Ikegaki N, Dimitroulakos J et al. MYCN protein expression as a predictor of neuroblastoma prognosis. Clin Cancer Res 1997; 3: 1699-1706.

39. Ladenstein R, Ambros IM, Potschgr U,Amann G, Urban C, Fink FM et al. The prognostic significance of DNA di/tetraploidy in neuroblastoma. Med Pediatr Oncol 1998; 32: 189.

40. Kaneko Y, Kanda N, Maseki N, Sakurai M, Tsuchida Y, Takeda T. Different karyotypic patterns in early and advanced stage neuroblastomas. Cancer Res 1987; 47: 311-318.

41. Christiansen H, Lampert F. Tumour karyotype discriminates between good and bad prognostic outcome in neuroblastoma. Br J Cancer 1988; 57: 121-126.

42. Kaneko Y, Cohn SL. Ploidy and cytogenetics of neuroblastoma. Neuroblastoma. Elsevier Science, The Netherlands, 2000: 41-56.

43. White PS, Versteeg R. Allelic loss and neuroblastoma suppressor genes. Neuroblastoma. Elsevier Science, The Netherlands, 2000: 57-74.

44. Maris JM, Matthay KK. Molecular biology of neuroblastoma. J Clin Oncol 1999; 17: 2264-2279.

45. Meddeb M, Danglot G, Chudoba I, Venuat A, Bernard J, Avet-Loiseau H et al. Additional copies of a 25 Mb chromosomal region originating from 17q23.1-17qter are present in 90% of high-grade neuroblastomas. Genes Chromosomes Cancer 1996; 17: 156-165.

46. Pearson ADJ, Philip T. Prognosis of low-risk and high-risk neuroblastoma. Neuroblastoma. Elsevier Science, The Netherlands, 2000: 551-560.

47. Castleberry RP, Pritchard J, Ambros P, Berthold F, Brodeur GM, Castel V et al. The International Neuroblastoma Risk Groups (INRG): a preliminary report. European Journal of Cancer 1997; 33 (12): 2113-2116.

48. Nakagawara A, Kogner P. Expression and function of Trk and its related genes in human neuroblastoma. Neuroblastoma. Elsevier Science, The Netherlands, 2000: 147-157.

49. Kaplan, DR, Martin-Zanca D, Parada LF. Tyrosine phosphorylation and tyrosine kinase activity of the trk proto-oncogene product induced by NGF. Nature 1991; 350: 358-360.

50. Nakagawara A, Arima-Nakagawara M, Scavarda NJ, Azar CG, Cantor AB, Brodeur GM. Association between high levels of expression of the Trk gene and favorable outcome in human neuroblastoma. NEJM 1993; 328: 847-854.

51. Owens DF, Kriegstein AR. Is there more to GABA than synaptic inhibition? Nature Reviews 2002; 3: 715-727.

52. Spoerri PE, Wolff JR. Effect of GABA-administration on murine neuroblastoma cells in culture. I. Increased membrane dynamics and formation of specialized contacts. Cell Tissue Res 1981; 218: 567-579.

53. Roberts SS, Mori M, Pattee P, Lapidus J, Mathews R, O'Malley J et al. GABAergic system gene expression predicts clinical outcome in patients with neuroblastoma. J Clin Oncol 2004; 22: 4127-4134.

54. SPSS for Windows, Version 10.0. Information available at http://www.spss.com

55. R, Version 1.8.1. Available at http://www.r-project.org

56. Salford Systems' CART, Version 5.0. Available at http://www.salfordsystems.com/cart.php

57. Borrello MG, Bongarzone I, Pieriotti MA, Luksch R, Gasparini M, Collini P et al. Trk and ret proto-oncogene expression in human neuroblastoma specimens: high frequency of trk expression in non-advanced stages. Int J Cancer 1993; 54: 540-545.

58. Evangelopoulos ME, Weiss J, Kruttgen A. Neurotrophin effects on neuroblastoma cells: correlation with Trk and p75NTR expression and influence of Trk receptor bodies. J Neuro-Oncol 2004; 66: 101-110.

59. NCBI map viewer. Available at: http://www.ncbi.nlm.nih.gov/mapview

60. Emberger W, Windpassinger C, Petek E, Kroisel PM, Wagner K. Assignment of the human GABA A receptor delta-subunit gene (GABRD) to chromosome band 1p36.3 distal to marker NIB 1364 by radiation hybrid mapping. Cytogenet Cell Genet 2000; 89: 281-282.

61. Kushner BH, LaQuaglia MP, Kramer K, Cheung NV. Radically different treatment recommendations for newly diagnosed neuroblastoma: pitfalls in assessment of risk. J Ped Hem/ Oncol 2004; 26: 35-39.

Appendix A

Children's Oncology Group (COG) affiliated United States institutions

	State	City	Institution
1	AL	Birmingham	University of Alabama
2	AL	Mobile	University of South Alabama
3	AR	Little Rock	University of Arkansas
4	AZ	Phoenix	Phoenix Childrens Hospital
5	AZ	Tucson	University of Arizona Health Sciences Center
6	CA	Downey	Southern California Permanente Medical Group
7	CA	Duarte	City of Hope National Medical Center
8	CA	Loma Linda	Loma Linda University Medical Center
9	CA	Long Beach	Miller Children's Hospital/Harbor-UCLA
10	CA	Los Angeles	Cedars-Sinai Medical Center
11	CA	Los Angeles	Childrens Hospital Los Angeles
12	CA	Los Angeles	UCLA School of Medicine
13	CA	Madera	Children's Hospital Central California
14	CA	Oakland	Childrens Hospital Oakland
15	CA	Orange	Childrens Hospital of Orange County
16	CA	Orange	University of California, Irvine
17	CA	Palo Alto	Stanford University Medical Center
18	CA	Sacramento	Kaiser Permanente Medical Group, Inc., Northern CA
19	CA	Sacramento	University of California, Davis
20	CA	San Diego	Children's Hospital San Diego
21	CA	San Francisco	UCSF School of Medicine
22	CA	San Francisco	Sutter Medical Center, Sacramento
23	CA	Santa Barbara	Santa Barbara Cottage Children's Hospital
24	CO	Colorado Springs	Presbyterian/St Lukes Medical Center and CHOA
25	CO	Denver	The Children's Hospital - Denver, CO
26	CT	Hartford	Connecticut Children's Medical Center
27	CT	New Haven	Yale University School of Medicine
28	DC	Washington	Children's National Medical Center - D.C.
29	DC	Washington	Georgetown University Medical Center
30	DE	Wilmington	Christiana Care Health Services/A.I. duPont Inst.
31	FL	Ft. Lauderdale	Broward General Medical Center
32	FL	Ft. Myers	The Children's Hospital of Southwest Florida Lee Memorial Health Sy
33	FL	Gainesville	University of Florida
34	FL	Hollywood	Joe DiMaggio Children's Hospital at Memorial
35	FL	Jacksonville	Nemours Children's Clinic-Jacksonville
36	FL	Miami	Baptist Hospital of Miami
37	FL	Miami	Miami Children's Hospital
38	FL	Miami	University of Miami School of Medicine
39	FL	Orlando	Florida Hospital Cancer Institute
40	FL	Orlando	Nemours Children's Clinic-Orlando

Children's Oncology Group (COG) affiliated United States institutions

	State	City	Institution
41	FL	Pensacola	Sacred Heart Hospital
42	FL	St. Petersburg	All Children's Hospital
43	FL	Tampa	Tampa Children's Hospital
44	FL	West Palm Beach	St. Mary's Hospital
45	GA	Atlanta	Children's Healthcare of Atlanta at Scottish Rite
46	GA	Augusta	Medical College of Georgia Childrens Medical Ctr
47	GA	Savannah	Memorial Medical Center/Backus Children's Hospital
48	HI	Honolulu	Cancer Research Center of Hawaii
49	IA	Des Moines	Raymond Blank Children's Hospital
50	IA	Iowa City	University of Iowa Hospitals & Clinics
51	ID	Boise	Mountain States Tumor Institute
52	IL	Chicago	Children's Memorial Medical Center at Chicago
53	IL	Chicago	Rush-Presbyterian St. Luke's Medical Center
54	IL	Chicago	University of Chicago Medical Center
55	IL	Chicago	University of Illinois
56	IL	Maywood	Loyola University Medical Center
57	IL	Oak Lawn	Advocate Hope Children's Hospital
58	IL	Park Ridge	Lutheran General Childrens Medical Center
59	IL	Peoria	St. Jude Midwest Affiliate
60	IL	Springfield	Southern Illinois University School of Medicine
61	IN	Indianapolis	Indiana University - Riley Childrens Hospital
62	IN	Indianapolis	St. Vincent Children's Hospital - Indiana
63	KS	Kansas City	University of Kansas Medical Center
64	KS	Wichita	Via Christi Regional Medical Center
65	KS	Wichita	Wesley Medical Center
66	KS	Wichita	Wichita CCOP
67	KY	Lexington	A.B. Chandler Medical Ctr - University of Kentucky
68	KY	Louisville	Kosair Childrens Hospital
69	LA	New Orleans	Children's of New Orleans/LSUMC CCOP
70	LA	New Orleans	Ochsner Clinic
71	LA	New Orleans	Tulane University Medical Center
72	MA	Boston	Boston Floating Hospital for Infants & Children
73	MA	Boston	Dana-Farber Cancer Institute and Children's Hospital, Boston
74	MA	Boston	Massachusetts General Hospital
75	MA	Springfield	Baystate Medical Center
76	MA	Worcester	University of Massachusetts Medical School
77	MD	Baltimore	Johns Hopkins Hospital
78	MD	Baltimore	Sinai Hospital of Baltimore
79	MD	Baltimore	University of Maryland at Baltimore
80	MD	Bethesda	National Cancer Institute - Pediatric Branch
81	ME	Bangor	Eastern Maine Medical Center
82	ME	Scarborough	Maine Children's Cancer Program
83	MI	Ann Arbor	C.S. Mott Children's Hospital
84	MI	Detroit	Children's Hospital of Michigan
85	MI	East Lansing	Michigan State University

Children's Oncology Group (COG) affiliated United States institutions

	State	City	Institution
86	MI	Flint	Hurley Medical Center
87	MI	Grand Rapids	DeVos Children's Hospital
88	MI	Gross Pointe Woods	St John Hospital and Medical Center
89	MI	Kalamazoo	Kalamazoo Center for Medical Studies
90	MI	Royal Oak	William Beaumont Hospital
91	MN	Minneapolis	Childrens Hospital & Clinics
92	MN	Minneapolis	University of Minnesota Cancer Center
93	MN	Rochester	Mayo Clinic and Foundation
94	MO	Columbia	University of Missouri - Columbia
95	MO	Kansas City	The Childrens Mercy Hospital
96	MO	St. Louis	Cardinal Glennon Children's Hospital
97	MO	St. Louis	Washington University Medical Center
98	MS	Jackson	University of Mississippi Medical Center Children's Hospital
99	MS	Keesler AFB	Madigan Army Medical Center (USOC)
100	MS	Keesler AFB	Tripler Army Medical Center (USOC)
101	MS	Keesler AFB	United States Air Force Med Ctr, Keesler AT (USOC)
102	MS	Keesler AFB	Walter Reed Army Medical Center (USOC)
103	MS	Keesler AFB	Naval Medical Center/Portsmouth (USOC)
104	NC	Asheville	Mission Hospitals
105	NC	Chapel Hill	University of North Carolina at Chapel Hill
106	NC	Charlotte	Carolinas Medical Center
107	NC	Charlotte	Presbyterian Hospital
108	NC	Durham	Duke University Medical Center
109	NC	Greenville	East Carolina University School of Medicine
110	NC	Winston-Salem	Wake Forest University School of Medicine
111	ND	Fargo	MeritCare Hospital
112	NE	Omaha	Childrens Memorial Hospital of Omaha
113	NE	Omaha	University of Nebraska Medical Center
114	NH	Lebanon	Dartmouth-Hitchcock Medical Center
115	NJ	Hackensack	Hackensack University Medical Center
116	NJ	Livingston	Saint Barnabas Medical Center
117	NJ	Morristown	Atlantic Health System
118	NJ	New Brunswick	University of Medicine and Dentistry of New Jersey
119	NJ	New Brunswick	Saint Peter's University Hospital
120	NJ	Newark	Newark Beth Israel Medical Center
121	NJ	Paterson	St. Joseph's Hospital and Medical Center
122	NM	Albuquerque	University of New Mexico School of Medicine
123	NV	Las Vegas	Sunrise Childrens Hospital, Sunrise Hosp & Med Ctr
124	NY	Albany	Albany Medical Center
125	NY	Bronx	Montefiore Medical Center
126	NY	Brooklyn	Brookdale Hospital Medical Center
127	NY	Brooklyn	Brooklyn Hospital Center
128	NY	Brooklyn	Maimonides Medical Center
129	NY	Brooklyn	SUNY Health Science Center at Brooklyn
130	NY	Buffalo	Roswell Park Cancer Institute

Children's Oncology Group (COG) affiliated United States institutions

	State	City	Institution
131	NY	Mineola	Winthrop University Hospital
132	NY	New Hyde Park	Schneider Children's Hospital
133	NY	New York	Columbia Presbyterian College of Phys & Surgeons
134	NY	New York	Memorial Sloan Kettering Cancer Center
135	NY	New York	Mount Sinai Medical Center
136	NY	New York	New York Hospital-Cornell Univ Medical Center
137	NY	New York	New York University Medical Center
138	NY	Rochester	University of Rochester Medical Center
139	NY	Stony Brook	State University of New York at Stony Brook
140	NY	Syracuse	State University of New York Upstate Medical University
141	NY	Valhalla	New York Medical College
142	OH	Akron	Childrens Hospital Medical Center-Akron, Ohio
143	OH	Cincinnati	Childrens Hospital Medical Center Cincinnati
144	OH	Cleveland	Rainbow Babies and Childrens Hospital
145	OH	Cleveland	The Children's Hospital at The Cleveland Clinic
146	OH	Columbus	Childrens Hospital of Columbus
147	OH	Dayton	Children's Medical Center Dayton
148	OH	Toledo	Mercy Children's Hospital
149	OH	Toledo	Toledo Children's Hospital
150	OH	Youngstown	Western Reserve Care System - Tod Childrens Hosp
151	OK	Oklahoma City	University of Oklahoma Health Sciences Center
152	OK	Tulsa	Warren Clinic, Inc
153	OR	Portland	Doernbecher Childrens Hospital - Oregon HSU
154	OR	Portland	Emanuel Hospital-Health Center
155	PA	Danville	Geisinger Medical Center
156	PA	Hershey	Penn State Children's Hospital, Hershey Med Ctr
157	PA	Philadelphia	Childrens Hospital of Philadelphia
158	PA	Philadelphia	St. Christopher's Hospital for Children
159	PA	Pittsburgh	Children's Hospital of Pittsburgh
160	PR	Santurce	San Jorge Children's Hospital
161	RI	Providence	Rhode Island Hospital
162	SC	Charleston	Medical University of South Carolina
163	SC	Columbia	South Carolina Cancer Center
164	SC	Greenville	Children's Hospital of the Greenville Hospital System
165	SD	Sioux Falls	Sioux Valley Children's Specialty Clinics
166	TN	Chattanooga	T.C. Thompson Children's Hospital
167	TN	Johnson City	East Tennessee State University
168	TN	Knoxville	East Tennessee Childrens Hospital
169	TN	Memphis	St. Jude Children's Research Hospital Memphis
170	TN	Nashville	Vanderbilt Children's Hospital
171	TX	Amarillo	Texas Tech UHSC - Amarillo
172	TX	Austin	Children's Hospital of Austin
173	TX	Corpus Christi	Driscoll Children's Hospital
174	TX	Dallas	North Texas Hosp for Children at Med City Dallas
175	TX	Dallas	University of Texas Southwestern Medical School

Children's Oncology Group (COG) affiliated United States institutions

	State	City	Institution
176	TX	Fort Worth	Cook Children's Medical Center
177	TX	Galveston	University of Texas Medical Branch
178	TX	Houston	M.D. Anderson Cancer Center
179	TX	Houston	Texas Children's Cancer Center at Baylor College of Medicine
180	TX	Lackland AFB	San Antonio Military Pediatric Cancer & Blood Disorders Center
181	TX	Lubbock	Children's Hem/Onc Team @ Covenant Children's Hosp
182	TX	San Antonio	Southwest Texas Methodist Hospital
183	TX	San Antonio	University of Texas Health Science Center at San Antonio
184	TX	Temple	Scott & White Memorial Hospital
185	UT	Salt Lake City	Primary Childrens Medical Center
186	VA	Charlottesville	University of Virginia Health Sciences Center
187	VA	Fairfax	Inova Fairfax Hospital
188	VA	Norfolk	Childrens Hospital-King's Daughters
189	VA	Richmond	Virginia Commonwealth Univ Health System - MCV
190	VA	Roanoke	Carilion Medical Center for Children at Roanoke Community Hospital
191	VT	Burlington	University of Vermont College of Medicine
192	WA	Seattle	Children's Hospital and Regional Medical Center
193	WA	Spokane	Sacred Heart Children's Hospital
194	WA	Tacoma	Mary Bridge Hospital
195	WI	Green Bay	St. Vincent Hospital
196	WI	La Crosse	Gundersen Lutheran
197	WI	Madison	University of Wisconsin - Childrens Hosp Madison
198	WI	Marshfield	Marshfield Clinic
199	WI	Milwaukee	Midwest Children's Cancer Center
200	WV	Charleston	West Virginia University HSC - Charleston
201	WV	Huntington	Cabell Huntington Hospital
202	WV	Morgantown	West Virginia University HSC - Morgantown

Appendix B

Correlations of gene expression variables by Spearman's rho

	Log Trk A	Log p75	Log NGF	Log GABA A alpha 1	Log GABA A alpha 2	Log GABA A alpha 3	Log GABA A alpha 5
Log Trk A	1.00	0.54	0.29	0.02	0.03	0.00	-0.14
Log p75	0.54	1.00	0.54	-0.03	0.21	0.03	-0.02
Log NGF	0.29	0.54	1.00	0.05	0.41	0.16	0.04
Log GABA A alpha 1	0.02	-0.03	0.05	1.00	0.22	0.57	-0.03
Log GABA A alpha 2	0.03	0.21	0.41	0.22	1.00	0.20	-0.02
Log GABA A alpha 3	0.00	0.03	0.16	0.57	0.20	1.00	0.10
Log GABA A alpha 5	0.16	-0.02	0.04	-0.03	-0.02	0.10	1.00
Log GABA A alpha 6	0.06	-0.06	-0.05	0.65	0.11	0.51	-0.10
Log GABA A beta 1	0.26	0.28	0.38	0.17	0.57	0.28	0.17
Log GABA A beta 2	0.01	0.05	0.28	0.23	0.45	0.12	0.13
Log GABA A beta 3	0.30	0.22	0.34	0.25	0.43	0.30	0.11
Log GABA A gamma 2	0.20	0.06	0.16	0.34	0.34	0.13	-0.24
Log GABA A delta	0.17	0.12	0.15	0.55	0.07	0.37	-0.10
Log GABA A epsilon	0.03	0.08	0.17	0.78	0.18	0.59	-0.03
Log GABA A pi	0.16	-0.04	0.04	0.76	0.06	0.46	0.00
Log GABA A theta	0.07	0.13	0.18	0.74	0.14	0.51	0.05
Log GABA B	0.02	0.20	0.43	-0.05	0.44	0.22	0.18
Log GAD-1	-0.08	-0.08	-0.01	0.81	0.15	0.57	-0.01
Log GAD-2	0.02	-0.10	-0.08	0.75	0.01	0.49	-0.03

Correlations of gene expression variables by Spearman's rho

	Log GABA A alpha 6	Log GABA A beta 1	Log GABA A beta 2	Log GABA A beta 3	Log GABA A gamma 2	Log GABA A delta	Log GABA A epsilon
Log Trk A	0.06	0.26	0.01	0.30	0.20	0.17	0.03
Log p75	-0.06	0.28	0.05	0.22	0.06	0.12	0.08
Log NGF	-0.05	0.38	0.28	0.34	0.16	0.15	0.17
Log GABA A alpha 1	0.65	0.17	0.23	0.25	0.34	0.55	0.78
Log GABA A alpha 2	0.11	0.57	0.45	0.43	0.34	0.07	0.18
Log GABA A alpha 3	0.51	0.28	0.12	0.30	0.13	0.37	0.59
Log GABA A alpha 5	-0.10	0.17	0.13	0.11	-0.24	-0.10	-0.03
Log GABA A alpha 6	1.00	0.06	0.01	0.11	0.20	0.37	0.54
Log GABA A beta 1	0.06	1.00	0.38	0.58	0.23	0.21	0.22
Log GABA A beta 2	0.01	0.38	1.00	0.37	0.34	0.08	0.04
Log GABA A beta 3	0.11	0.58	0.37	1.00	0.34	0.27	0.23
Log GABA A gamma 2	0.20	0.23	0.34	0.34	1.00	0.15	0.19
Log GABA A delta	0.37	0.21	0.08	0.27	0.15	1.00	0.68
Log GABA A epsilon	0.54	0.22	0.04	0.23	0.19	0.68	1.00
Log GABA A pi	0.62	0.11	0.02	0.27	0.18	0.58	0.76
Log GABA A theta	0.45	0.25	0.06	0.26	0.17	0.70	0.85
Log GABA B	-0.11	0.50	0.37	0.55	0.20	0.03	0.05
Log GAD-1	0.68	0.14	0.01	0.17	0.22	0.49	0.81
Log GAD-2	0.66	0.06	-0.06	0.08	0.21	0.46	0.74

Correlations of gene expression variables by Spearman's rho

	Log GABA A pi	Log GABA A theta	Log GABA B	Log GAD-1	Log GAD-2
Log Trk A	0.16	0.07	0.02	-0.08	0.02
Log p75	-0.04	0.13	0.20	-0.08	-0.10
Log NGF	0.04	0.18	0.43	-0.01	-0.08
Log GABA A alpha 1	0.76	0.74	-0.05	0.81	0.75
Log GABA A alpha 2	0.06	0.14	0.44	0.15	0.01
Log GABA A alpha 3	0.46	0.51	0.22	0.57	0.49
Log GABA A alpha 5	0.00	0.05	0.18	-0.01	-0.03
Log GABA A alpha 6	0.62	0.45	-0.11	0.68	0.66
Log GABA A beta 1	0.11	0.25	0.50	0.14	0.06
Log GABA A beta 2	0.02	0.06	0.37	0.01	-0.06
Log GABA A beta 3	0.27	0.26	0.55	0.17	0.08
Log GABA A gamma 2	0.18	0.17	0.20	0.22	0.21
Log GABA A delta	0.58	0.70	0.03	0.49	0.46
Log GABA A epsilon	0.76	0.85	0.05	0.81	0.74
Log GABA A pi	1.00	0.70	-0.08	0.78	0.77
Log GABA A theta	0.70	1.00	0.04	0.71	0.66
Log GABA B	-0.08	0.04	1.00	-0.06	-0.15
Log GAD-1	0.78	0.71	-0.06	1.00	0.86
Log GAD-2	0.77	0.66	-0.15	0.86	1.00