Combining TrkA Immunohistochemical-Score with Clinicopathologic Parameters in Neuroblastoma: An Influential Prognostic Nomogram

Juanqing Yue
Shanghai Jiaotong University School of Medicine Xinhua Hospital
Lei Cai
Shanghai Jiaotong University School of Medicine Xinhua Hospital
Ruifen Wang
Shanghai Jiaotong University School of Medicine Xinhua Hospital
Meng Qiao
Shanghai Jiaotong University School of Medicine Xinhua Hospital
Kezhou Wang
Shanghai Jiaotong University School of Medicine Xinhua Hospital
Wenwei Yu
Shanghai Jiaotong University School of Medicine Xinhua Hospital
Junlei Li
Shanghai Jiaotong University School of Medicine Xinhua Hospital
Meng Zhang
Shanghai Jiaotong University School of Medicine Xinhua Hospital
Wenbin Guan
Shanghai Jiaotong University School of Medicine Xinhua Hospital
lifeng Wang (wanglifeng@xinhuamed.com.cn)
Shanghai Jiaotong University School of Medicine Xinhua Hospital https://orcid.org/0000-0002-7118-0163

Primary research

Keywords: neuroblastoma, NTRK1, TrkA, immunohistochemistry, bioinformatics analysis, prognosis, Trk inhibitors, nomogram

Posted Date: January 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-143769/v1

Title Page

Combining TrkA Immunohistochemical-score with Clinicopathologic parameters in Neuroblastoma: An Influential Prognostic Nomogram

・Authors' names: Juanqing Yue†, Lei Cai†, Ruifen Wang, Meng Qiao, Kezhou Wang, Wenwei Yu, Junlei Li, Zhang Meng, Wenbin Guan*, Lifeng Wang*

・Authors'affiliations: Department of Pathology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

*** Correspondence:**

Wenbin Guan

guanwenbin@xinhuamed.com.

Lifeng Wang

wanglifeng@xinhuamed.com.

Abstract:

【Background】 Neuroblastoma (NB) is one of the most common solid tumors in children with varied clinical outcomes. Although there are some several risk stratification systems currently, their clinical applications are limited due to the testing conditions of different laboratory and the heavy financial burden on patients. TrkA is coded by *NTRK1*, belonging to tropomyosin receptor kinase family. We have observed that TrkA was differentially expressed in paraffin tissue sections of NB. The aim of this

study was to determine the immunohistochemical-score of TrkA as an independent prognostic factor for NB and establish a useful prognostic model for postoperative patients.

【 **Methods** 】 We systematically summarized the relationship between immunochemistry (IHC) score of TrkA and clinicopathological parameters in 86 NB cases. Fluorescence in situ hybridization (FISH) and qRT-PCR were used to detect *NTRK1* gene fusion. Furthermore, GSE96631, GSE16476, GSE49710 and GSE73537 datasets, originated from Gene Expression Omnibus (GEO), were analyzed to figure out the *NTRK1* related molecular characteristics by bioinformatics methods. And combined TrkA immunohistochemical-score with clinicopathologic parameters to construct a prognostic nomogram of overall survival (OS) for NB.

【**Result**】 In clinical samples and GEO database analyses, patients in the *NTRK1* / TrkA low expression group showed significantly poorer outcome than patients in high group. Multivariate cox regression analysis demonstrated *NTRK1* / TrkA as an independent prognostic factor for NB survival. Neither Fluorescence in situ hybridization nor qRT-PCR detected evidence for *NTRK1* gene fusion in clinical samples, indicating that differential expression in *NTRK1* / TrkA are caused by epigenetic changes. Bioinformatics analyses revealed that *MYC* target related pathway may play a critical role in low expression of TrkA, leading to unfavorable prognoses of NB.

【**Conclusion**】 The results of this study suggest that the IHC score of TrkA may be used as an independent predictor of postoperative OS of patients with NB. By combining the IHC score of TrkA and clinicopathological features, the proposed nomogram provides a feasible predictive tool for postoperative patients with NB. Simultaneously, this study also reveals that Trk inhibitors are not supposed to be taken in NB patients.

Keywords: neuroblastoma, *NTRK1*, TrkA, immunohistochemistry, bioinformatics analysis, prognosis, Trk inhibitors, nomogram

Number of Figures: 7

Number of Tables: 5

Number of Supporting Information:9

1.Introduction

In 1864, neuroblastoma (NB) was first described by Rudolf Virchow as a glioma tumor occurring in the abdominal cavity of children. Then it was named by James Homer Wright in 1910 as neurocytoma or neuroblastoma[1]. NB is one of the malignant solid tumors that often occur in children, with a median age of 18 months. Although typical lesions can also be seen in adults, these cases are extremely rare[2].

NB is a sympathetic nervous system embryogenic tumor originating from primitive neural crest cells, which can occur anywhere in the

sympathetic nervous system[3]. NB belongs to the neuroblastic tumor (NT) spectrum. According to different differentiation levels, NTs are categorized as ganglioneuroma (GN), intermediate ganglioneuroblastoma (GNB), and neuroblastoma (NB)[4]. NB is a type of complex disease showing different clinical, biological and genetic characteristics. Some patients may have spontaneous regression, while others may have disease progression or even death. Because of the trait of invasiveness and metastasis, NB accounts for nearly 15% of all childhood malignant tumor deaths[5]. Current risk stratification revealed that the age of initial diagnosis and the status of *MYCN* gene are the most powerful prognostic factor. Older age at initial diagnosis and *MYCN* amplification imply unfavorable outcome. Besides, many other factors matter, such as chromosomal deletion or acquisition[6], and DNA ploidy[7]. At present, there are mainly two kinds of risk classification systems in the world: COG (Children's Oncology Group） classification system[8] and INRG (International Neuroblastoma Risk Group) classification system[9,10]. The former introduced five risk factors to divide NB into three groups: low risk, intermediate risk and high risk. The five risk factors included INSS stage, age of initial diagnosis, *MYCN* amplification status, international pathological classification of NB and DNA ploidy. The INRG classification system divided NB into four groups: very low risk, low risk, intermediate risk and high risk. It took seven risk factors into consideration:

INRG stage, age, *MYCN* amplification status, international pathological classification of NB, tumor cell differentiation, DNA ploidy and chromosome 11q abnormality. However, their clinical applications are limited due to the limitations of different laboratory testing conditions and complexity of operation. As for some patients who cannot get available test, he systematic prediction of the risk of progrnosis is a formidable task. Therefore, it is necessary to develop new effective and convenient models related to prognosis of NB contributing to stratified risk management of patients and further appropriate therapeutic strategies.

Neurotrophic receptor tyrosine kinase（*NTRK*）gene family includes *NTRK1*, *NTRK2* or *NTRK3*, coding for three transmembrane high-affinity tyrosine-kinase receptors (TrkA, TrkB, and TrkC) for nerve growth factors respectively ,which is mainly involved in nervous system development[11,12]. Current studies have shown that *NTRK* was activated through multiple mechanisms in various malignant tumors, but *NTRK* fusion is the most definite driving factor for carcinogenesis and can cause overexpression of Trk protein. There are some approved Trk small molecule inhibitors, such as Entrectinib and Larotrectinib, can be used in patients with advanced solid tumors carrying *NTRK* fusion genes. Recently, some studies have found high expression of NGF/TrkA accompanied with malignant biological behavior in pancreatic cancer and breast cancer[13-15].

Clinically, we have observed the differential expression of TrkA protein in NB patients, but its clinical significance is not clear yet. Previous studies have found that the mRNA expression level of *NTRK1* was related to the poor prognosis of NB, high expression of mRNA for *NTRK1* tending to be with favorable progression of NB[16-19]. But there were few researches to verify the relationship between the expression level of TrkA, the protein product of *NTRK1* gene, and clinical features among NB's patients. In addition, the mechanisms of *NTRK1* differential expression level causing different clinical outcomes are still indefinite yet.

In this study, we combined clinical samples and public microarray datasets, and concentrated on exploring potential feasibility of TrkA immunohistochemical-score as an evaluation model for prediction of outcome in NB patients, aiming to provide a basis for prognostic and therapeutic guidance. Moreover, bioinformatics methods were used to analyze the possible mechanisms behind.

2.Material and Methods

2.1 Specimen Collection

The 86 clinical FFPE tissue samples diagnosed as NTs in this study were collected by the Department of Pathology of Xinhua Hospital Affiliated To Shanghai Jiao Tong University School Of Medicine from 2013 to 2020, including 52 cases of NB, 29 cases of GNB and 5 cases of GN, and information about patient's basic clinicopathological characteristics is

shown in **Table 1**. All patients underwent surgery in the hospital. Patients were diagnosed by pathologists according to World Health Organization published standardizations of diagnosis, and classified based on the International Neuroblastoma Staging System (INSS). Each specimen collected was consented to by both hospital and individual. This study was approved by the Medical Ethical Committee of Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine.

2.2 Microarry source

In this study, we downloaded the gene expression profiling data from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). A total of 311 series about human NB were retrieved from the database. Finally, we select four gene expression profiles (GSE73517, GSE16476, GSE49710 and GSE96631) with deliberation (**Table 2**), and detailed information is shown in **Tables S1-S4**.

2.3 Immunohistochemistry (IHC)

IHC staining was performed with an Anti-bodies against TrkA（EPR17341，Abcam，the United Kingdom）on an automated system (BOND MAX system, Leica, Germany) following operating instructions. A positive result is defined as staining of the cytoplasm, and/or cell membrane and/or nucleus. Inflammatory cells were viewed as internal controls.

The results of IHC staining evaluated by 2 pathologists separately, were quantified by using the semiquantitative integration method based on the staining intensity and the positive rate (**Figure 1a**). Detailly, the scoring criterion for staining intensity was following: compared with the background color, colorless was 0 point, light yellow, yellow-brown and brown were counted as 1, 2, and 3 points respectively. The staining intensity was compared with the background color. The scoring standard for percentage of positive cells were following: <25%, 25%~50%, 51~75%, >75% were scored as 1, 2, 3, 4 points respectively. The final score = staining intensity score × positive rate score.

2.4 Fluorescence in situ hybridization (FISH) and qRT-PCR

FISH can detect large structural variants at the DNA level, such as abnormalities in the number and structure of chromosomes or genes. The *MYCN* copy number was identified by FISH （Vysis LSI *N-MYC* CEP 2 spectrum orange, Abbott, the United States). *MYCN* was considered to be amplified when the gene copy number was above 3 per haploid. All clinical cases were tested *MYCN* status. In this study, *NTRK1* fusion was detected by separation probe（*NTRK1* Break Apart Rearrangement Probe, F.01103. An Biping, China）and was only performed on the cases with strongly IHC staining (n=22). However, there is no clear standard for the positive threshold currently. In our study, of the 50 cells evaluated, more than 30% tumor cells detected separate signals, was viewed as fusion status of

the *NTRK1*. The interpretation of the above results was assessed by two pathologists independently.

qRT-PCR was performed to further verify *NTRK1* fusion. We selected NB samples from the above 22 cases which met the following testing conditions: 1) samples need to be obtained in recent two years; 2) the content of tumor cells in the tissue should exceed 30%. Total RNA extraction from FFPE specimens was performed using the kit (8.126001W008A, Amoy Dx, China) according to the manufacturer's instructions. The RNA of the tumor cells in the FFPE samples were extracted and reverse transcribed into cDNA, then combined with fluorescent probe technology to test the *NTRK* gene fusion qualitatively on a fluorescent quantitative PCR instrument (ABI 7500, Thermo Fisher, the United States). When the CT (cycle times) value of HEX (VIC) signal for housekeeping gene is $<$ 20, and simultaneously CT value of FAM signal for target gene is ≤25, the sample was thought to be detected as gene fusion for target gene.

2.5 Bioinformatics analysis of microarray datasets

To explore the function of *NTRK1* in the prognosis of NB, all NB cases in three datasets (GSE73517, GSE16476 and GSE49710) were divided into *NTRK1* high-expression group (NB-*NTRK1*^{high}) and *NTRK1* low-expression group (NB-*NTRK1*^{low}) respectively according to the the median value of *NTRK1* expression level. The Sanger box software was used to detect the

differentially expressed genes (DEGs) between two groups, and the adjusted P-value and |logFC| were calculated. When genes met the cutoff criteria, adjusted $P < 0.05$ and $|logFC| \geq 1.0$, those genes were considered as DEGs. Making a statistical analysis of each dataset and use the Venn diagram webtool (bioinformatics.psb.ugent.be/webtools/Venn/) to determine the intersecting part. Then Gene Ontology (GO) functional annotation analysis is a common strategy for extensive functional enrichment studies, which is divided into biological process (BP), cellular components (CC) and molecular functions (MF). Using the Database for Annotation, Visualization and Integrated Discovery (DAVID) tools (https://david.ncifcrf.gov/), we can achieve GO annotation analysis of DEGs. $P < 0.05$ and gene counts≥10 were considered statistically significant. Then we visualized the protein–protein interaction (PPI) network of DEGs by using Cytoscape software (www.cytoscape.org/). Finally, Gene Set Variation Analysis (GSVA) was carried out in NB-*NTRK1*^{high} and NB-*NTRK1*^{low} groups by the GSVA R package that also annotated gene set c2.cp.kegg.v7.0.symbols.gmt. Then the results of GSVA were intersected by R package （http://cran.r-project.org/web/packages/VennDiagram/）, and the venn diagram was drawn.

2.6 Statistical Methods

Data were analyzed with Graphpad Prism 8 software （http://www.graphpad.com/） and R software（version 3.6.3）. The nomogram was based on the results of the multivariate regression analysis, combining with pathological diagnosis, and compiled using the RMS package in R software. Student's two-tailed t test, Chi-square test and survival analysis were used in statistical analysis. When P value < 0.05 was considered statistically significant.

3.Results

3.1 The correlation between *NTRK1*/ TrkA and NB clinical parameters

TrkA protein expression levels were assessed in 86 NB clinical samples using IHC. TrkA protein were differentially expressed between clinical samples (**Figure 1b-c**). Moreover, correlation of IHC score of TrkA and clinical parameters were found out that IHC score of TrkA was strongly associated with age of initial diagnosis (≤18months/＞18months), *MYCN* status (amplification or not) and INSS stage. The results turn out that IHC score of TrkA trended to be higher in younger, *MYCN* not amplified, and low INSS stage (stage 1 and 2) or 4s patients. There was no significant difference in gender. Further, we compared the expression of TrkA score in the primary tumor and the metastasis in 32 of 86 clinical cases which the FFPE tissue of metastases can be obtained. It was found that there was no statistical difference in the IHC score of TrkA of tumor cells between metastasis and primary tissue（**Figure 1d**, P=0.5526）.

The microarray datasets (GSE73517, n=105; GSE16476, n=88; GSE49710, n=498) were obtained to further verify the relationship between mRNA expression level of *NTRK1* and clinical characteristics of NB. The mRNA expression level of *NTRK1* was higher in younger, *MYCN* not amplified and low INSS stage (stage 1 and 2) or 4s stage, but did not significantly correlate with gender (P>0.05) (**Figure 2a-d**), which was consistent with analyses derived from clinical samples. The result concluded from GSE96631 showed significant downregulated mRNA expression of *NTRK1* in NB patients compared with embryonic neural creat cells (ENCC)、adrenal cortex (AC) and adrenal marrow (AM) (P <0.0001, **Figure 2e**).

According to the the median value level of *NTRK1*, all NB cases in three datasets were divided into NB-*NTRK1*^{high} groups and NB-*NTRK1*^{low} groups respectively. The correlation analysis results （**Table 3**）revealed that the *NTRK1* expression level in NB patients significantly correlated with age of initial diagnosis, *MYCN* status and INSS stage (P<0.05).

3.2 *NTRK1*^{low} is associated with unfavorable outcome in NB patients

According to the IHC score and the median value (score=6) as the cut-off value, the 86 clinical samples were divided into high-expression (TrkA^{high}) and low-expression (TrkA^{low}) groups, and survival analyses were carried out. Besides, the two datasets (GSE16476 and GSE49710) which prognostic information were available were also analyzed for

survival. The result derived from clinical cohort showed that $TrkA^{low}$ group was significantly associated with shorter overall survival (OS) time, as well as the result got from public datasets, showing lower mRNA expression of *NTRK1* indicating shorter OS time (**Figure 3a**). In the multivariate Cox regression analysis of OS, the variables included were age of first diagnosis, sex, INSS stage, *MYCN* status and the IHC score of TrkA or mRNA expression level of *NTRK1* (**Figure 3b-d**). Through 86 clinical sample analysis, it presented that the prognosis of NB had a strong correlation with the TrkA score and the status of *MYCN* gene. In both GSE49710 and GSE16476, the prognosis was associated with age and *NTRK1* expression level, besides, in GSE49710, the results identified that the prognosis of NB patients is also significantly correlated with INSS stage and *MYCN* status

In order to assess the specificity and sensitivity of *NTRK1*/TrkA as a reliable predictive factor of OS, Receiver Operator Characteristic curve (ROC), Decision Curve Analysis（DCA）analyses and nomogram were further performed. The AUCs of the TrkA among clinical samples, GSE16476 and GSE49710 were 0.784, 0.878 and 0.815, respectively. DCA suggesting that TrkA has good clinical utility for prognostic evaluation (**Figure 3e-f**). Based on the results of the multivariate analysis of OS, combining pathological diagnosis, we established a prognostic nomogram of OS (**Figure 4**). Data analysis showed that TrkA score0 better than 3 in the nomogram, this may be due to the cases of score0 (n=4) and 3(n=1) is limited. Besides, score4 ranks between 6 and 8

283 in the nomogram may owing to insufficient follow-up time of some cases because the
284 time of inclusion of cases is different. However, it showed the trend that OS was
285 significantly increased with the increase in scores, the C-index (=0.95) manifested the
286 preferable effectiveness of model.

287 3.3 FISH and qRT-PCR

288 FISH techniques were performed in FFPE samples (**Figure 5a**). No
289 fusion status of *NTRK1* was found in above cases. PCR amplification curve
290 of positive control showed that the cycle threshold (CT) value of HEX (VIC)
291 signal for housekeeping gene is 12.5 (≤20), and CT value of FAM signal
292 for target gene is 13.43 (≤25), prompting the experimental conditions are
293 qualified (**Figure 5b**). PCR amplification curve of clinical sample showed
294 that the CT value of HEX (VIC) signal for housekeeping gene is 12.27
295 (≤20), and CT value of FAM signal for target gene is undetermined,
296 implying that no evidence involving *NTRK1* fusion was detected by qRT-PCR
297 (**Figure 5c**).

298 3.4 GO Functional enrichment analysis and GSVA analysis

299 To probe the function of *NTRK1* in the good prognosis of NB, we conducted a
300 bioinformatics analysis on the database data. With the criteria of FDR $<$
301 0.05 and $|logFC| \geq 1$, total of 1978 DEGs were identified from GSE73517,
302 including 1839 upregulated genes and 139 downregulated genes, while 407
303 DEGs were identified from GSE16476, 210 upregulated genes and 197
304 downregulated genes. And 833 DEGs were identified in GSE49710,

305 including 769 genes were upregulated, and 64 genes were downregulated
306 (**Table 4** and **Figure 6a**). All DEGs were selected out through comparing
307 NB-*NTRK1*^{high} and NB-*NTRK1*^{low} samples. Detailed information about
308 DEGs of each data set were included in **Tables S5-S7**. Subsequently, Venn
309 analysis was performed to get the intersection of the DEGs profiles (**Figure 6b**).
310 Finally, 151 DEGs were significantly differentially expressed among all three groups,
311 including 129 upregulated genes and 22 downregulated genes, and a summary of
312 intersection of the DEGs profiles is displayed in **Tables S8.** GO function for DEGs
313 were performed using the DAVID (**Table 5**). The results of GO analysis
314 indicated that DEGs were mainly enriched in BPs, including positive
315 regulation of GTPase activity and apoptotic process. CC analysis showed that the
316 DEGs were significantly enriched in plasma membrane and integral
317 component of plasma membrane. As for MF analysis, the DEGs were enriched
318 in calcium ion binding. Details about enriched GO terms of DEGs is shown in
319 **Tables S9.** Protein interactions among the DEGs were predicted with
320 STRING tools. Totally, 149 nodes and 135 edges were involved in the PPI
321 network (**Figure 6c**).

322 The function of *NTRK1* was further revealed by GSVA by comparing the
323 differential expression group (**Figure 7**). The most significant
324 HALLMARK gene set in the *NTRK1* low expression group was *MYC*
325 targets, which was the intersected part of venn diagram for the GSVA
326 analysis results of the three data sets. On the contrary, in the *NTRK1* high

expression group, the common HALLMARK gene set was hedgehog (Hh) signaling pathway.

4.Discussion

NB is one of the most common solid tumors in children, with highly variable symptoms and clinical process. Although some risk stratification strategies are available, their clinical applications remain limited owing to their complexity of operation and expensive testing costs.

There are many techniques in pathology, the immunohistochemistry technology has become a mature technology after more than 40 years of development. In the past 20 years, it has been widely used in medical biology research and diagnostic pathology practice. It has become an important method to identify antigen molecules such as proteins, peptides and carbohydrates in tissues and cells. Compared with complex and expensive molecular tests, immunohistochemical staining techniques are faster and more economical, moreover, it consumes less tissue. Most pathology laboratories of the hospital have launched immunohistochemical testing projects, which makes this test be available to nearly all patients.

This study is the first time that TrkA has been quantitative evaluated by immunohistochemical staining in a large sample of neuroblastoma and constructed a prognostic nomogram of OS for NB. IHC score of TrkA is easy to conduct that can be substituted for texting mRNA expression level of

348 *NTRK1*. Patients with higher IHC score of TrkA tend to have better
349 prognoses.

350 FISH and qRT-PCR were used to detect the *NTRK1* gene status in the
351 TrkA overexpression samples, but no fusion was found. Furthermore,
352 bioinformatics analyses revealed that *MYC* targets was the obvious
353 pathway that highly associated with the low expression of *NTRK1*/TrkA,
354 taking unignorable effects in poor prognoses of NB. The *MYC* (*c-MYC*) gene,
355 along with the *MYCN* gene, belongs to the *MYC* gene family. Their encoding protein
356 products, either MYCN protein or MYC protein, are known to make a complex with
357 MAX (MYC-associated factor X) protein, and the complex binds to target genes,
358 leading to activation of transcription, involving in cell proliferation, inhibition of
359 differentiation, apoptosis, and so on[20]. Recent studies proposed that *MYC*
360 family-driven NB, defined as augmented expression of MYC protein and/or
361 MYCN protein, is highly aggressive. And MYC protein overexpression has
362 been reported as a novel prognostic factor for NB[21-24]. Based on
363 previous studies and our results, TrkA may promote the cell proliferation,
364 migration and invasion of neuroblastoma through the *MYC* axis.

365 Contrastively, the Hh signaling pathway was the associated with the
366 high expression of *NTRK1*. Many studies have noticed the activating effect of Hh
367 signaling on autophagy. For instance, Hh antagonist cyclopamine inhibited
368 autophagy activation in NB cell line SHSY5Y, suggesting that Hh signaling mediated
369 the activation of autophagy in NB cell lines. Besides, some researchers observed that

370 in cell line C3H10T1/2, suppression of Hh signaling activation led to decreased
371 protein expression of ATG5, a key protein for autophagy activation[25]. Moreover,
372 some researchers believed that the Hh signaling pathway may promote
373 differentiation of NB. Above researches uncovered a probable explanation for
374 spontaneous regression of NB. In the light of these, we could conceive that the
375 activation of the Hh signaling pathway in NB is dramatically connected
376 with the high expression of *NTRK1*, which may play a crucial role in good
377 prognosis of NB, that was consistent with previous study[26,27]. Based on
378 previous research, we can know that the downstream pathways of *NTRK*
379 gene are PI3K, RAS/MAPK/ERK and PLC-γ pathway[28], which are
380 independent of either *MYC* targets or Hh signaling pathways. Therefore, the
381 above results mean that the different expression of TrkA among NB
382 patients was caused by epigenetic changes, instead of gene alterations.

383 Besides, analyzing about IHC score of TrkA between metastases and
384 primary tumors among clinical samples, no significant statistical
385 difference was found. Therefore, when the primary tumors were difficult to
386 be obtained clinically, metastases were supposed to be detected.

387 There are some highlights of our study. Firstly, we constructed a model
388 based on a biomarker associated with outcomes of NB. Besides, the
389 technology was easy to obtained by most laboratories and to some extent
390 can relieve financial pressure of patient's family. Secondly, the results of
391 pathological molecular testing suggest that *NTRK1* gene fusion was not

found in NB, and TrkA high expression may contribute a good outcome, so the usage of Trk inhibitors maybe inadequate for NB. Thirdly, bioinformatic analysis were performed to explore underlaying mechanism, providing a novel view for NB.

5. Conclusion

In this study, we proposed a useful predictor model easy to operate technically. The IHC score of TrkA in tumor cells can be used as a supplementary material, contributing to individualized treatment and optimizing the hierarchical management of children with NB.

6.Abbreviations

CI: Confidence interval

DEGs: Differentially expressed genes

GSEA: Gene set enrichment analysis

FFPE: Formalin-fixed and paraffin-embedded

FISH: Fluorescence in situ hybridization

GEO: Gene Expression Omnibus

IHC: Immunohistochemical

TrkA: Tyrosine kinase A

NB: neuroblastoma

NTRK1: neurotrophic receptor tyrosine kinase1

PPI: Protein–protein interaction

ROC: Receiver operating characteristic

STRING: Search Tool for the Retrieval of Interacting Genes Database

7.Declarations

7.1 Ethics approval and consent to participate

Each specimen collected was consented to by both the Medical Ethical Committee of Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine and individual and every patient's parents.

7.2 Consent for publication

The authors consent for publication.

7.3 Availability of data and materials

The datasets analyzed in this study were obtained from Gene Expression Omnibus (GEO) database for free. The accession numbers are mentioned in the article (GEO accession number: GSE96631, GSE73517, GSE16476 and GSE49710).

7.4 Competing interests

The authors declare that they have no competing interests.

7.5 Funding

This study is sponsored by the Joined Medicine-Engineering Fund of Shanghai Jiao Tong University (No. YG2016QN71).

7.6 Author' Contributions

JY and LC performed the research study, analyzed the data, and wrote the paper. RW, MQ, KW, and WY supported during performing of the

experiments, and contributed essential reagents and laboratory equipment. JL and ZM collected patient data. WG and LW conceived and designed the study. All authors contributed to the article and approved the submitted version.

Corresponding author

Correspondence to Wenbin Guan & Lifeng Wang

7.7 Acknowledgments

None.

7.8 Authors' information

Juanqing Yue and Lei Cai contributed equally to this work.

Affiliations: Department of Pathology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

Juanqing Yue, Lei Cai, Ruifen Wang, Meng Qiao, Kezhou Wang, Wenwei Yu, Junlei Li, Zhang Meng, Wenbin Guan & Lifeng Wang

8.References:

1. **Wright JH.** Neurocytoma or neuroblastoma, a kind of tumor not generally recognized. *The Journal of experimental medicine*. 1910; 12: 556-61.

2. **Suzuki M, Kushner BH, Kramer K, Basu EM, Roberts SS, Hammond WJ, LaQuaglia MP, Wolden SL, Cheung N-KV, Modak S.** Treatment and outcome of adult-onset neuroblastoma. *International journal of cancer*. 2018; 143: 1249-58.

3. **Nakagawara A, Li Y, Izumi H, Muramori K, Inada H, Nishi M.** Neuroblastoma. *Japanese*

Journal of Clinical Oncology. 2018; 48: 214-41.

4. **Xie Y, Xu H, Fang F, Li Z, Zhou H, Pan J, Guo W, Zhu X, Wang J, Wu Y.** A 3-Protein Expression Signature of Neuroblastoma for Outcome Prediction. *Am J Surg Pathol*. 2018; 42: 1027-35.

5. **Ward E, DeSantis C, Robbins A, Kohler B, Jemal A.** Childhood and adolescent cancer statistics, 2014. *CA: a cancer journal for clinicians*. 2014; 64: 83-103.

6. **Brodeur GM, Fong CT, Morita M, Griffith R, Hayes FA, Seeger RC.** Molecular analysis and clinical significance of N-myc amplification and chromosome 1p monosomy in human neuroblastomas. *Progress in clinical and biological research*. 1988; 271: 3-15.

7. **Look AT, Hayes FA, Nitschke R, McWilliams NB, Green AA.** Cellular DNA content as a predictor of response to chemotherapy in infants with unresectable neuroblastoma. *N Engl J Med*. 1984; 311: 231-5.

8. **Strother DR, London WB, Schmidt ML, Brodeur GM, Shimada H, Thorner P, Collins MH, Tagge E, Adkins S, Reynolds CP, Murray K, Lavey RS, Matthay KK, Castleberry R, Maris JM, Cohn SL.** Outcome after surgery alone or with restricted use of chemotherapy for patients with low-risk neuroblastoma: results of Children's Oncology Group study P9641. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2012; 30: 1842-8.

9. **Maris JM.** Recent advances in neuroblastoma. *N Engl J Med*. 2010; 362: 2202-11.

10. **Cohn SL, Pearson AD, London WB, Monclair T, Ambros PF, Brodeur GM, Faldum A, Hero B, Iehara T, Machin D, Mosseri V, Simon T, Garaventa A, Castel V, Matthay KK.** The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. *Journal of clinical oncology : official journal of the American Society of Clinical*

Oncology. 2009; 27: 289-97.

11. **Nakagawara A.** Trk receptor tyrosine kinases: A bridge between cancer and neural development. *Cancer Letters*. 2001; 169: 107-14.

12. **Drilon A, Siena S, Ou S-HI, Patel M, Ahn MJ, Lee J, Bauer TM, Farago AF, Wheler JJ, Liu SV, Doebele R, Giannetta L, Cerea G, Marrapese G, Schirru M, Amatu A, Bencardino K, Palmeri L, Sartore-Bianchi A, Vanzulli A, Cresta S, Damian S, Duca M, Ardini E, Li G, Christiansen J, Kowalski K, Johnson AD, Patel R, Luo D, Chow-Maneval E, Hornby Z, Multani PS, Shaw AT, De Braud FG.** Safety and Antitumor Activity of the Multitargeted Pan-TRK, ROS1, and ALK Inhibitor Entrectinib: Combined Results from Two Phase I Trials (ALKA-372-001 and STARTRK-1). *Cancer Discov*. 2017; 7: 400-9.

13. **Marchesi F, Piemonti L, Mantovani A, Allavena P.** Molecular mechanisms of perineural invasion, a forgotten pathway of dissemination and metastasis. *Cytokine & growth factor reviews*. 2010; 21: 77-82.

14. **Liebig C, Ayala G, Wilks JA, Berger DH, Albo D.** Perineural invasion in cancer: a review of the literature. *Cancer*. 2009; 115: 3379-91.

15. **Romon R, Adriaenssens E, Lagadec C, Germain E, Hondermarck H, Le Bourhis X.** Nerve growth factor promotes breast cancer angiogenesis by activating multiple pathways. *Mol Cancer*. 2010; 9: 157-.

16. **Suzuki T, Bogenmann E, Shimada H, Stram D, Seeger RC.** Lack of high-affinity nerve growth factor receptors in aggressive neuroblastomas. *J Natl Cancer Inst*. 1993; 85: 377-84.

17. **Kogner P, Barbany G, Dominici C, Castello MA, Raschellá G, Persson H.** Coexpression of messenger RNA for TRK protooncogene and low affinity nerve growth factor receptor in

neuroblastoma with favorable prognosis. *Cancer Res*. 1993; 53: 2044-50.

18. **Borrello MG, Bongarzone I, Pierotti MA, Luksch R, Gasparini M, Collini P, Pilotti S, Rizzetti MG, Mondellini P, De Bernardi B, 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: 540-5.

19. **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. *N Engl J Med*. 1993; 328: 847-54.

20. **Meyer N, Penn LZ.** Reflecting on 25 years with MYC. *Nature reviews Cancer*. 2008; 8: 976-90.

21. **Ikegaki N, Shimada H, Fox AM, Regan PL, Jacobs JR, Hicks SL, Rappaport EF, Tang XX.** Transient treatment with epigenetic modifiers yields stable neuroblastoma stem cells resembling aggressive large-cell neuroblastomas. *Proceedings of the National Academy of Sciences of the United States of America*. 2013; 110: 6097-102.

22. **Tornóczky T, Kálmán E, Kajtár PG, Nyári T, Pearson AD, Tweddle DA, Board J, Shimada H.** Large cell neuroblastoma: a distinct phenotype of neuroblastoma with aggressive clinical behavior. *Cancer*. 2004; 100: 390-7.

23. **Wang LL, Teshiba R, Ikegaki N, Tang XX, Naranjo A, London WB, Hogarty MD, Gastier-Foster JM, Look AT, Park JR, Maris JM, Cohn SL, Seeger RC, Asgharzadeh S, Shimada H.** Augmented expression of MYC and/or MYCN protein defines highly aggressive MYC-driven neuroblastoma: a Children's Oncology Group study. *Br J Cancer*. 2015; 113: 57-63.

24. **Wang LL, Suganuma R, Ikegaki N, Tang X, Naranjo A, McGrady P, London WB, Hogarty MD, Gastier-Foster JM, Look AT, Park JR, Maris JM, Cohn SL, Seeger RC, Shimada H.**

Neuroblastoma of undifferentiated subtype, prognostic significance of prominent nucleolar formation, and MYC/MYCN protein expression: a report from the Children's Oncology Group. *Cancer*. 2013; 119: 3718-26.

25. **Milla LA, González-Ramírez CN, Palma V.** Sonic Hedgehog in cancer stem cells: a novel link with autophagy. *Biological research*. 2012; 45: 223-30.

26. **Souzaki R, Tajiri T, Souzaki M, Kinoshita Y, Tanaka S, Kohashi K, Oda Y, Katano M, Taguchi T.** Hedgehog signaling pathway in neuroblastoma differentiation. *Journal of pediatric surgery*. 2010; 45: 2299-304.

27. **Oue T, Yoneda A, Uehara S, Yamanaka H, Fukuzawa M.** Increased expression of the hedgehog signaling pathway in pediatric solid malignancies. *Journal of pediatric surgery*. 2010; 45: 387-92.

28. **Amatu A, Sartore-Bianchi A, Siena S.** NTRK gene fusions as novel targets of cancer therapy across multiple tumour types. *ESMO Open*. 2016; 1: e000023.

Table 1 Patients' clinicopathological characteristics in our clinical cohort (N = 86)

characteristics	Alive (n=66)	Dead (n=20)	Total (n=86)
Age at initial dignosis			
≤18m	20(30.3%)	2(10%)	22(25.6%)
＞18m	46(69.7%)	18(90%)	64(74.4%)
Gender			
Female	24(36.4%)	5(25%)	29(33.7%)

Male	42(63.6%)	15(75%)	57(66.3%)
Pathological dignosis			
Neuroblastoma	39(59.1%)	13(65%)	52(60.5%)
Ganglioneuroblastoma	23(34.8%)	6(30%)	29(33.7%)
Ganglioneuroma	4(6.1%)	1(5%)	5(5.8%)
N-MYC status			
Amplified	7(10.6%)	14(70%)	21(24.4%)
Not amplified	59(89.4%)	6(30%)	65(75.6%)
Losses of 1p			
Yes	3(4.5%)	5(25%)	8(9.3%)
No	35(53.1%)	6(30%)	41(47.7%)
NA	28(42.4%)	9(45%)	37(43.0%)
Losses of 11q			
Yes	4(6.1%)	1(5%)	5(5.8%)
No	13(19.7%)	3(15%)	16(18.6%)
NA	49(74.2%)	16(80%)	65(75.6%)
INSS stage			
1	15(22.7%)	0(0%)	15(17.4%)
2	14(21.2%)	0(0%)	14(16.3%)
3	10(15.2%)	2(10%)	12(14.0%)
4	25(37.9%)	18(90%)	43(50.0%)
4s	2(3.0%)	0(0%)	2(2.3%)

539

540

541

542

543 **Table 2**. Basic information of four microarray datasets from GEO database

Dataset ID	Platform	Sample	Total number
GSE16476	GPL570	NB	88
GSE49710	GPL16876	NB	498
GSE73517	GPL16876	NB	105
GSE96631	GPL22166 and GPL22167	NB; AM; AC; ENCC	75 (64 NB; 4 AM; 3 AC; 4 ENCC)

544 **Abbreviation:** AC, adrenal cortex; AM, adrenal marrow; ENCC, embryonic neural

545 crest cells; GEO, Gene Expression Omnibus database; NB, neuroblastoma;

546

547

548

549

550 **Table 3**. Correlation between *NTRK1* expression and clinical parameters in 3 GEO datasets

	GSE16476			**GSE49710**			**GSE73517**		
variables	***NTRK1*^{low} (n=44)**	***NTRK1*^{high} (n=44)**	***P***	***NTRK1*^{low} (n=248)**	***NTRK1*^{high} (n=250)**	***P***	***NTRK1*^{low} (n=248)**	***NTRK1*^{high} (n=250)**	***P***

Age at first diagnosis (month)			<0.001*			<0.001*		<0.001*
≤18	15	33		94	206	18	44	
> 18	29	11		154	44	34	9	
Gender			1.00			0.939		NA
Male	27	26		142	145	NA		
Female	17	18		106	105			
INSS			<0.001*			<0.001*		<0.001*
1	1	7		31	90	1	9	
2	3	12		22	56	1	8	
3	5	8		26	37	4	6	
4	31	9		149	34	42	14	
4S	4	8		20	33	4	16	
***MYCN*_amp**			<0.001*			<0.001*		<0.001*
Yes	15	1		91	1	32	1	
No	29	43		154	247	20	52	

Note: * indicates P value ＜0.05.

GSE73517 datasets do not provide information about gender of patients.

Abbreviation: *MYCN*_amp indicates *MYCN* gene amplification； NA, not available.

Table4. Statistics of DEGs from three microarray databasets

Dataset ID	Upregulated genes	Downregulated genes	Total number
GSE16476	210	197	407
GSE49710	769	64	833
GSE73517	1839	139	1978

Table5. Significantly enriched GO terms of DEGs

Category	**Term**	**Description**	**Gene Count**	***P*-value**
BP_Term	GO:0043547	positive regulation of GTPase activity	10	0.034343373
BP_Term	GO:0006915	apoptotic process	10	0.035005858
CC_Term	GO:0005886	plasma membrane	58	1.54E-06
CC_Term	GO:0005887	integral component of plasma membrane	25	2.11E-04
MF_Term	GO:0005509	calcium ion binding	16	4.13E-04

Abbreviations: BP, biological process; CC, cellular component; DEG, differentially expressed gene; GO, Gene Ontology; MF, molecular function.

8.Figure legends:

Figure 1. Analysis of the correlation between IHC scores of TrkA and clinical parameters in clinical cohort. **a**, Immunohistochemistrical semiquantitative integration model. Immunohistochemistrical (IHC) score was based on the degree of staining

intensity and percentage. The final IHC score = staining intensity score × positive rate score. **b**, TrkA was stained by anit-TrkA by IHC technique in the 86 NB clinical samples, showing different staining status. **c**, Analysis of the correlation between IHC scores of TrkA and clinical parameters in 86 clinical samples, including age of intial diagnosis (≤18 month;＞18 month), *MYCN* status (amplified;not amplified), gender, and INSS stage. **d**, TrkA score between primary and metastatic tumor tissue of 32 cases.

NOTE:*P < .05, **P < .01, ***P < .001, ****P＜0.001, ns, no statistical difference (P>0.05). IHC, immunohistochemistry; NB, neuroblastoma.

Figure 2. Associations between *NTRK1* expression and clinical paraments in three public datasets (GSE16476, GSE49710 and GSE73517). **a**, *NTRK1* expression and age of initial diagnosis(≤18 month;＞18 month). **b**, *NTRK1* expression and *MYCN* status (amplified; not amplified). **c**, *NTRK1* expression and gender. **d**, *NTRK1* expression among different INSS stage. **e**, *NTRK1* mRNA expression in NB (n=64)、embryonic neural crest cells (n=4)、adrenal cortex (n=3) and adrenal marrow (n=4)TrkA expression in GSE96631.

Note: *P < .05, **P < .01, ***P < .001, ****P＜0.001, ns, no statistical difference (P>0.05).

Figure 3. Survival analyses, time-dependent ROC curves and DCA curves for the prognostic model in clinical cohort and GEO cohort. **a**, Kaplan–Meier overall

survival (OS) analyses for *NTRK1*/TrkA expression in86 clinical NB samples(P=0.0001), 88 patients of GSE16476 （P＜0.0001）and 498 patients of GSE49710 (P＜0.0001). **b-d**, Multivariate cox regression analysis of risk factors associated with overall survival (OS) times of clinical samples, GSE16476 and GSE49710 respectively. **e**, Survival prediction power of the *NTRK1*/TrkA signature was determined by ROC analysis. The AUC values of clinical samples, GSE16476 and GSE49710 were calculated separately from the ROC curve. **f**, Decision Curve Analysis of clinical samples, GSE16476 and GSE49710.

Note: *MYCN*-amp, *MYCN* gene amplification; CI, confidence interval.

Figure 4. Nomogram to predict the probabilities of 1, 2 and 3-year survival probability. Points are assigned for pathological diagnosis, TrkA score, Tumor INSS stage, age of initial diagnosis and *MYCN* status by drawing a line upward from the corresponding values to the "Points" line. Draw an upward vertical line to the "Points" bar to calculate points. Based on the sum, draw a downward vertical line from the "Total Points" line to calculate 1, 2 and 3-year survival probability. Internal validation using the bootstrap method showed that the C-index for the model was 0.95.

Figure 5. Detection for *NTRK1* gene fusion. **a**, Fluorescence in-situ hybridization (FISH) of *NTRK1* "break-apart" probe. Separation signal was detected within 15% of tumor cells which indicated there was no fusion in *NTRK1*. **b**, qRT-PCR amplification

plot of positive control. The CT value of HEX (VIC) signal for internal control was 12.5 (≤20), and CT value of FAM signal for target gene was 13.43 (≤25). **c**, qRT-PCR amplification plot of clinical sample. The CT value of HEX (VIC) signal was 12.27 (≤20), and CT value of FAM signal for target gene was below detection limitation.

Figure 6. Identification of differentially expressed genes (DEGs) between NB-*NTRK1*^{high} and NB-*NTRK1*^{low} groups in three datasets (GSE16476, GSE49710 and GSE73517). **a**, Volcano plot showing the DEGs in three datasets. The red points represent upregulated genes screened based on the criteria of P＜0.05 and logFC≥ 1.The blue points represent downregulation of the expression of genes screened based on the criteria of P＜0.05 and logFC≤-1. The black points represent genes with no significant difference. **b**, Venn diagram of upregulated and downregulated DEGs common to all three GEO datasets. **c**, PPI network constructed with the DEGs, green nodes are upregulated genes, and yellow nodes are downregulated genes.

Figure 7. Gene set variation analysis (GSVA) for the *NTRK1* in three GEO datasets. **a-c**, The heatmaps of differentially expressed pathways for *NTRK1* through the calculation of GSVA in GSE16476, GSE49710, GSE373517 respectively. **d**, Venn diagram of pathways to three GEO datasets.

Figures

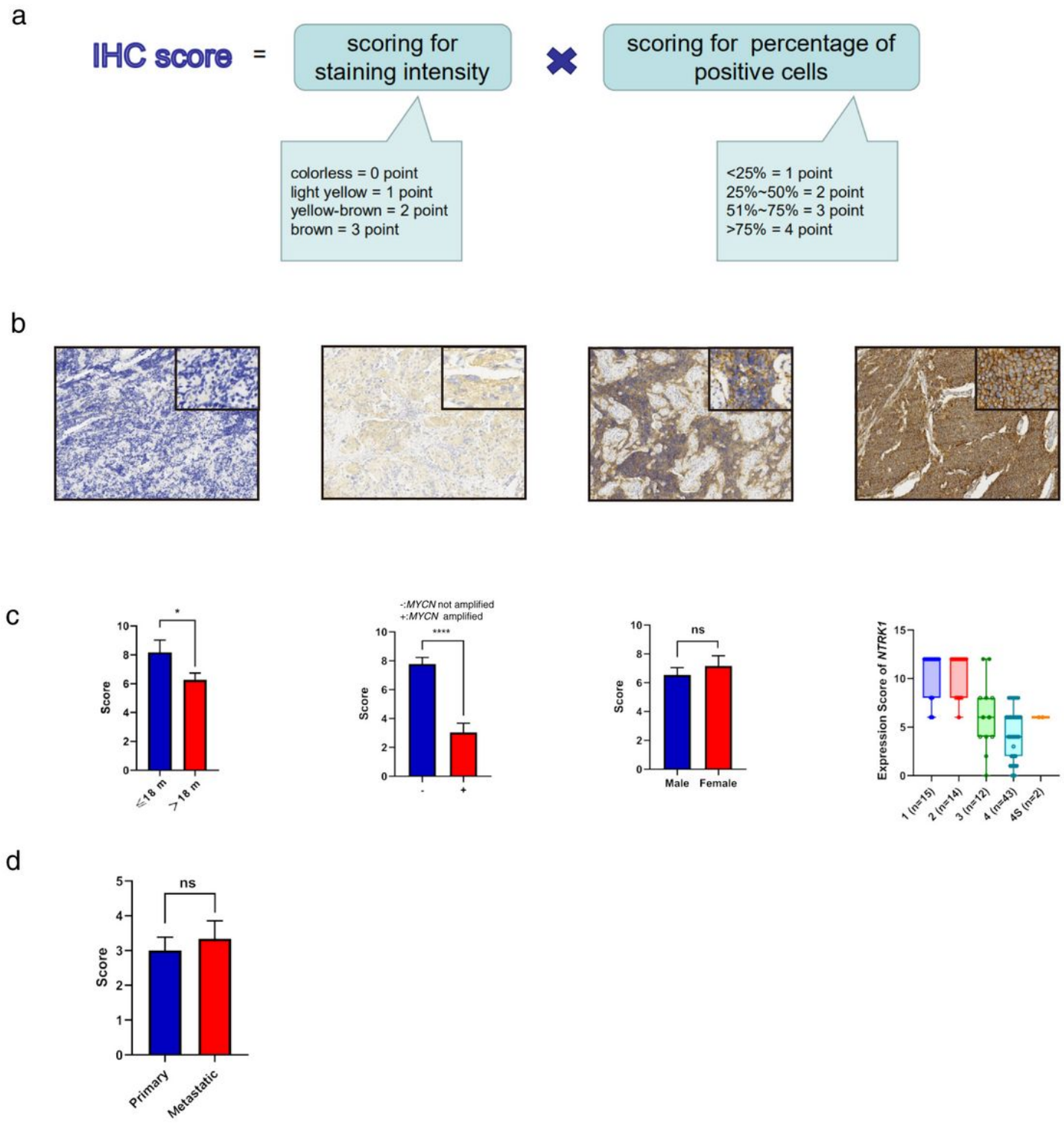

Figure 1

Analysis of the correlation between IHC scores of TrkA and clinical parameters in clinical cohort. a, Immunohistochemistrical semiquantitative integration model. Immunohistochemistrical (IHC) score was based on the degree of staining intensity and percentage. The final IHC score = staining intensity score ×

positive rate score. b, TrkA was stained by anit-TrkA by IHC technique in the 86 NB clinical samples, showing different staining status. c, Analysis of the correlation between IHC scores of TrkA and clinical parameters in 86 clinical samples, including age of intial diagnosis (≤18 month;>18 month), MYCN status (amplified;not amplified), gender, and INSS stage. d, TrkA score between primary and metastatic tumor tissue of 32 cases. NOTE:*P < .05, **P < .01, ***P < .001, ****P<0.001, ns, no statistical difference (P>0.05). IHC, immunohistochemistry; NB, neuroblastoma.

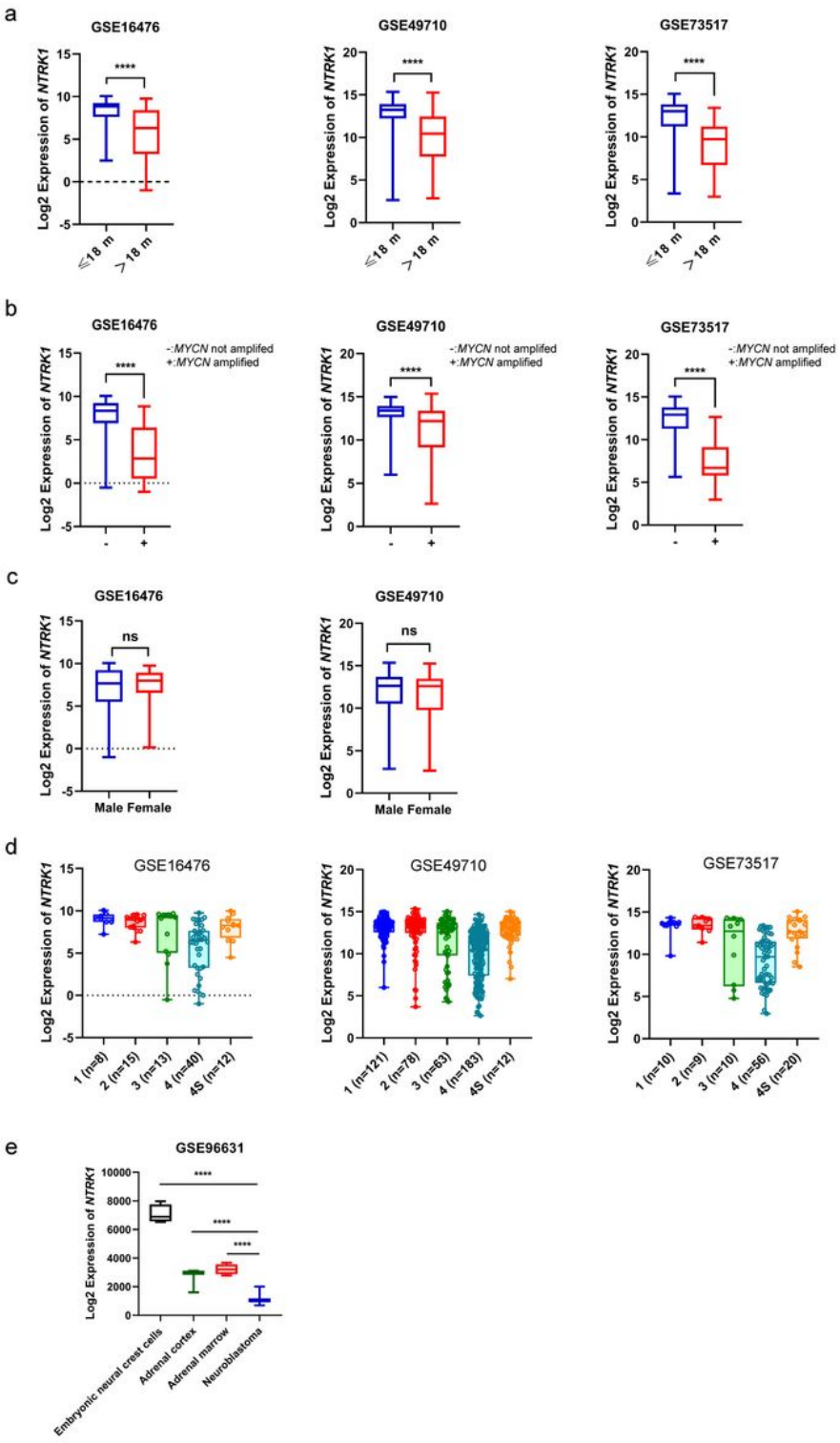

Figure 2

Associations between NTRK1 expression and clinical paraments in three public datasets (GSE16476, GSE49710 and GSE73517). a, NTRK1 expression and age of initial diagnosis(≤18 month;>18 month). b, NTRK1 expression and MYCN status (amplified; not amplified). c, NTRK1 expression and gender. d, NTRK1 expression among different INSS stage. e, NTRK1 mRNA expression in NB (n=64) ， embryonic neural crest cells (n=4) ，adrenal cortex (n=3) and adrenal marrow (n=4)TrkA expression in GSE96631. Note: *P < .05, **P < .01, ***P < .001, ****P<0.001, ns, no statistical difference (P>0.05).

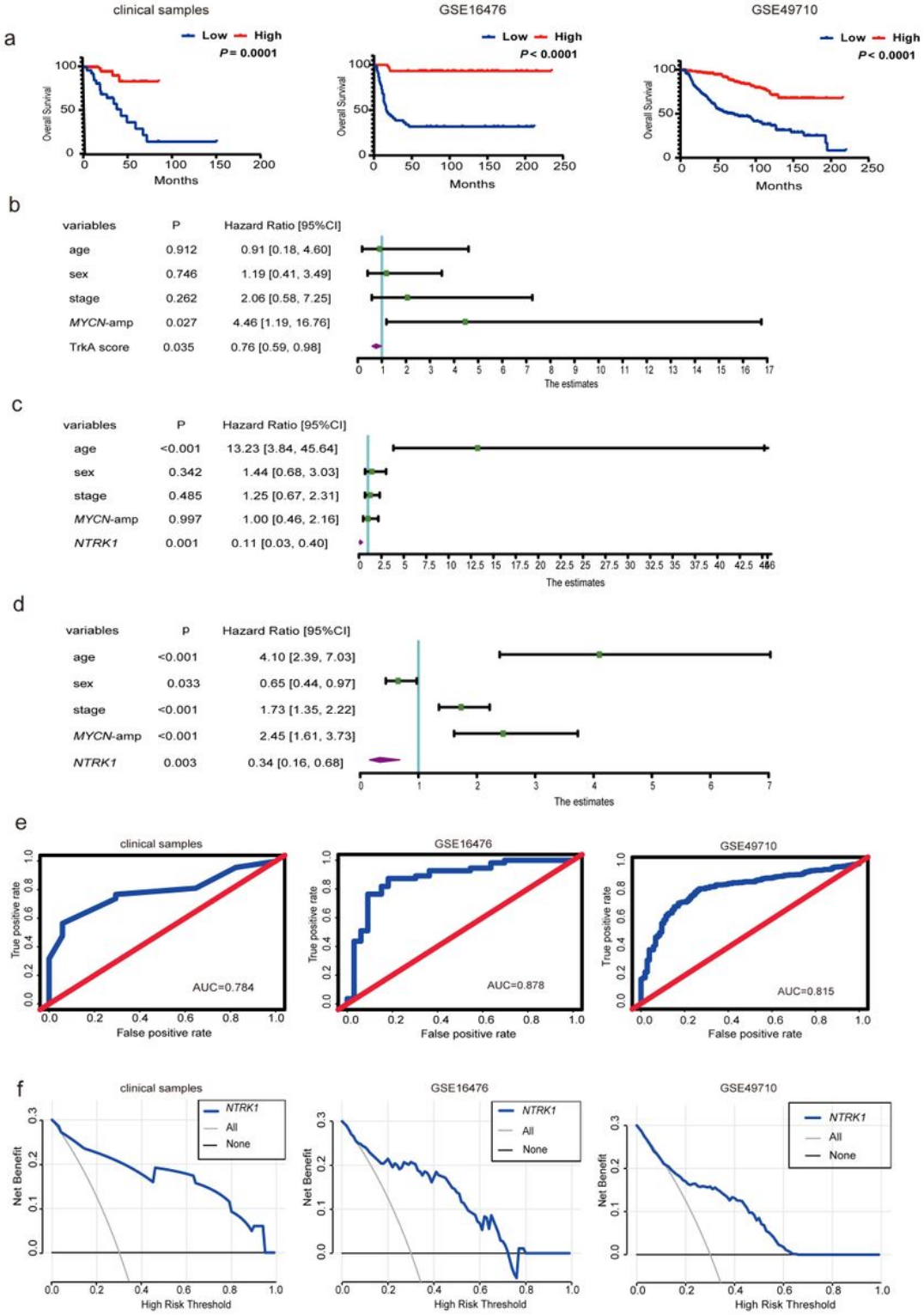

Figure 3

Survival analyses, time-dependent ROC curves and DCA curves for the prognostic model in clinical cohort and GEO cohort. a, Kaplan–Meier overall survival (OS) analyses for NTRK1/TrkA expression in86 clinical NB samples(P=0.0001), 88 patients of GSE16476 (P0.0001)and 498 patients of GSE49710 (P0.0001). b-d, Multivariate cox regression analysis of risk factors associated with overall survival (OS) times of clinical samples, GSE16476 and GSE49710 respectively. e, Survival prediction power of the NTRK1/TrkA signature was determined by ROC analysis. The AUC values of clinical samples, GSE16476 and GSE49710 were calculated separately from the ROC curve. f, Decision Curve Analysis of clinical samples, GSE16476 and GSE49710. Note: MYCN-amp, MYCN gene amplification; CI, confidence interval.

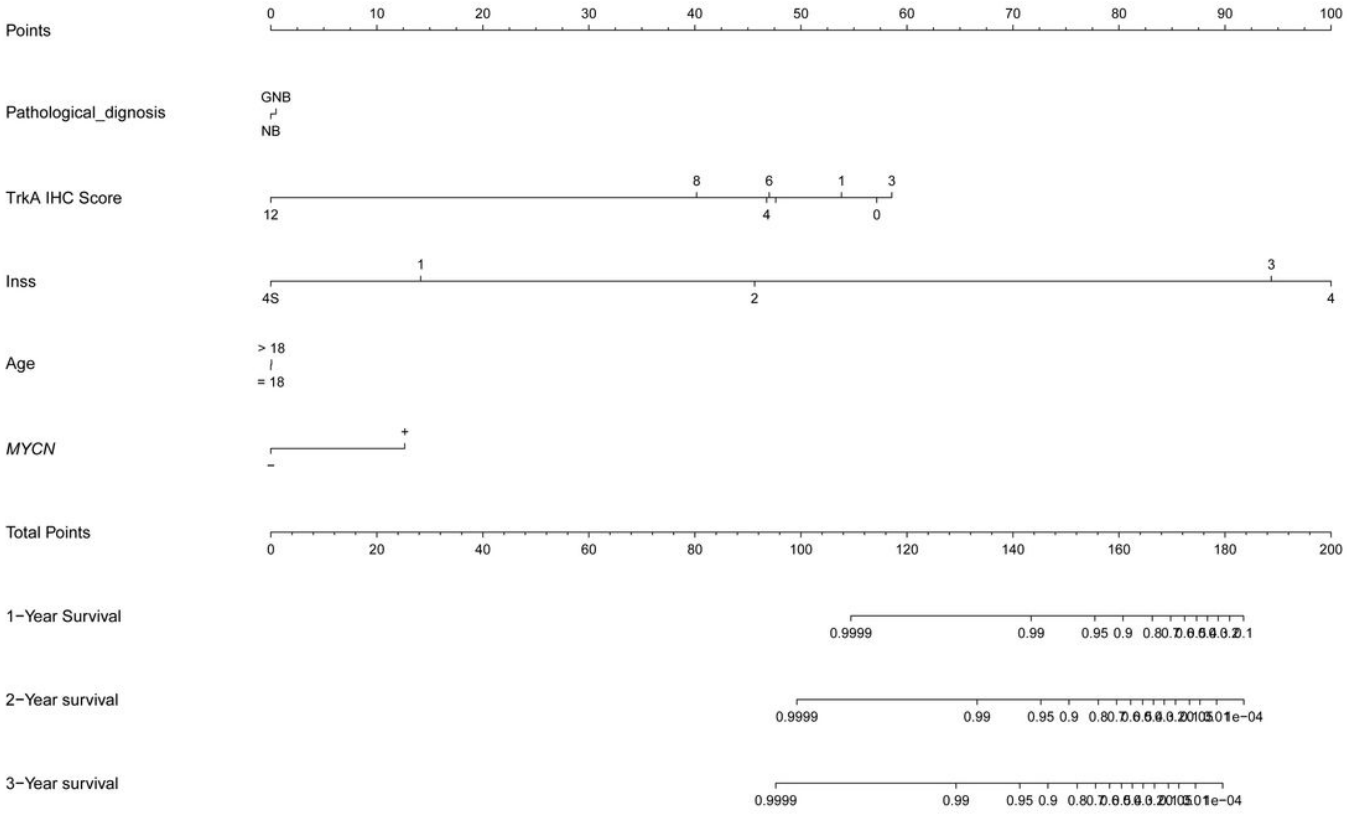

Figure 4

Nomogram to predict the probabilities of 1, 2 and 3-year survival probability. Points are assigned for pathological diagnosis, TrkA score, Tumor INSS stage, age of initial diagnosis and MYCN status by drawing a line upward from the corresponding values to the “Points” line. Draw an upward vertical line to the “Points” bar to calculate points. Based on the sum, draw a downward vertical line from the “Total Points” line to calculate 1, 2 and 3-year survival probability. Internal validation using the bootstrap method showed that the C-index for the model was 0.95.

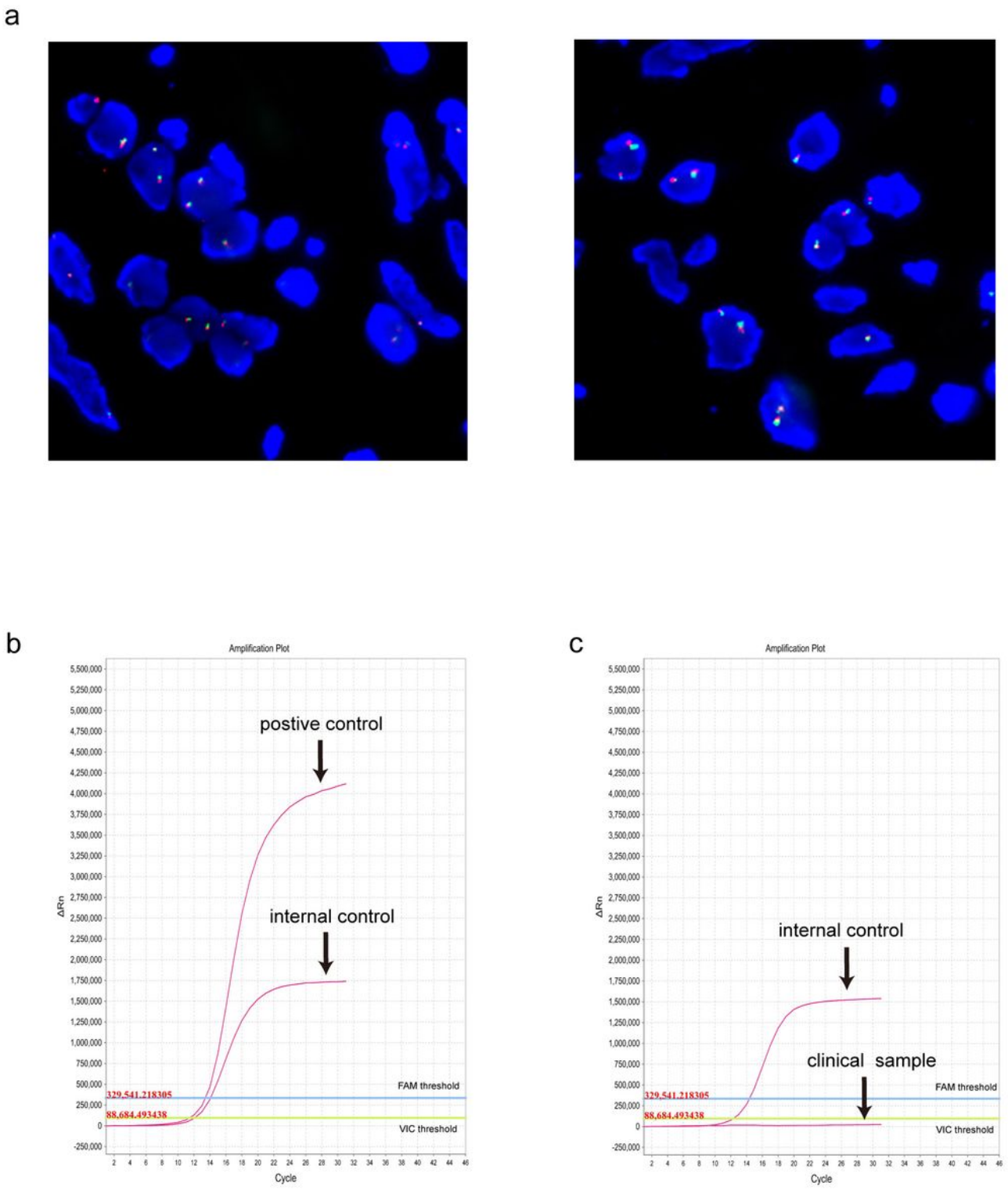

Figure 5

Detection for NTRK1 gene fusion. a, Fluorescence in-situ hybridization (FISH) of NTRK1 “break-apart” probe. Separation signal was detected within 15% of tumor cells which indicated there was no fusion in NTRK1. b, qRT-PCR amplification plot of positive control. The CT value of HEX (VIC) signal for internal control was 12.5 (≤20), and CT value of FAM signal for target gene was 13.43 (≤25). c, qRT-PCR

amplification plot of clinical sample. The CT value of HEX (VIC) signal was 12.27 (≤20), and CT value of FAM signal for target gene was below detection limitation.

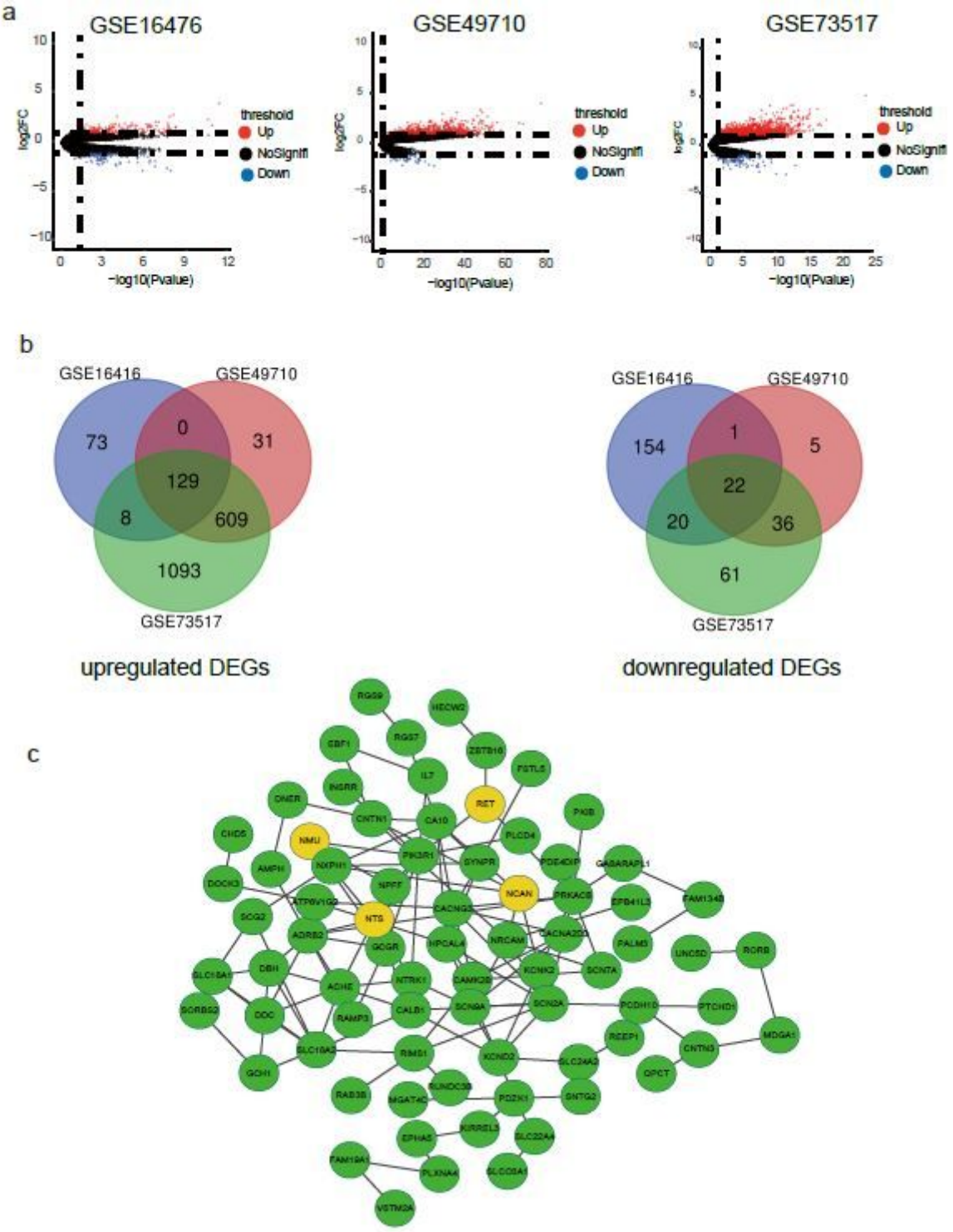

Figure 6

Identification of differentially expressed genes (DEGs) between NB-NTRK1high and NB-NTRK1low groups in three datasets (GSE16476, GSE49710 and GSE73517). a, Volcano plot showing the DEGs in three datasets. The red points represent upregulated genes screened based on the criteria of P□0.05 and logFC≥ 1.The blue points represent downregulation of the expression of genes screened based on the

criteria of P□0.05 and logFC≤-1. The black points represent genes with no significant difference. b, Venn diagram of upregulated and downregulated DEGs common to all three GEO datasets. c, PPI network constructed with the DEGs, green nodes are upregulated genes, and yellow nodes are downregulated genes.

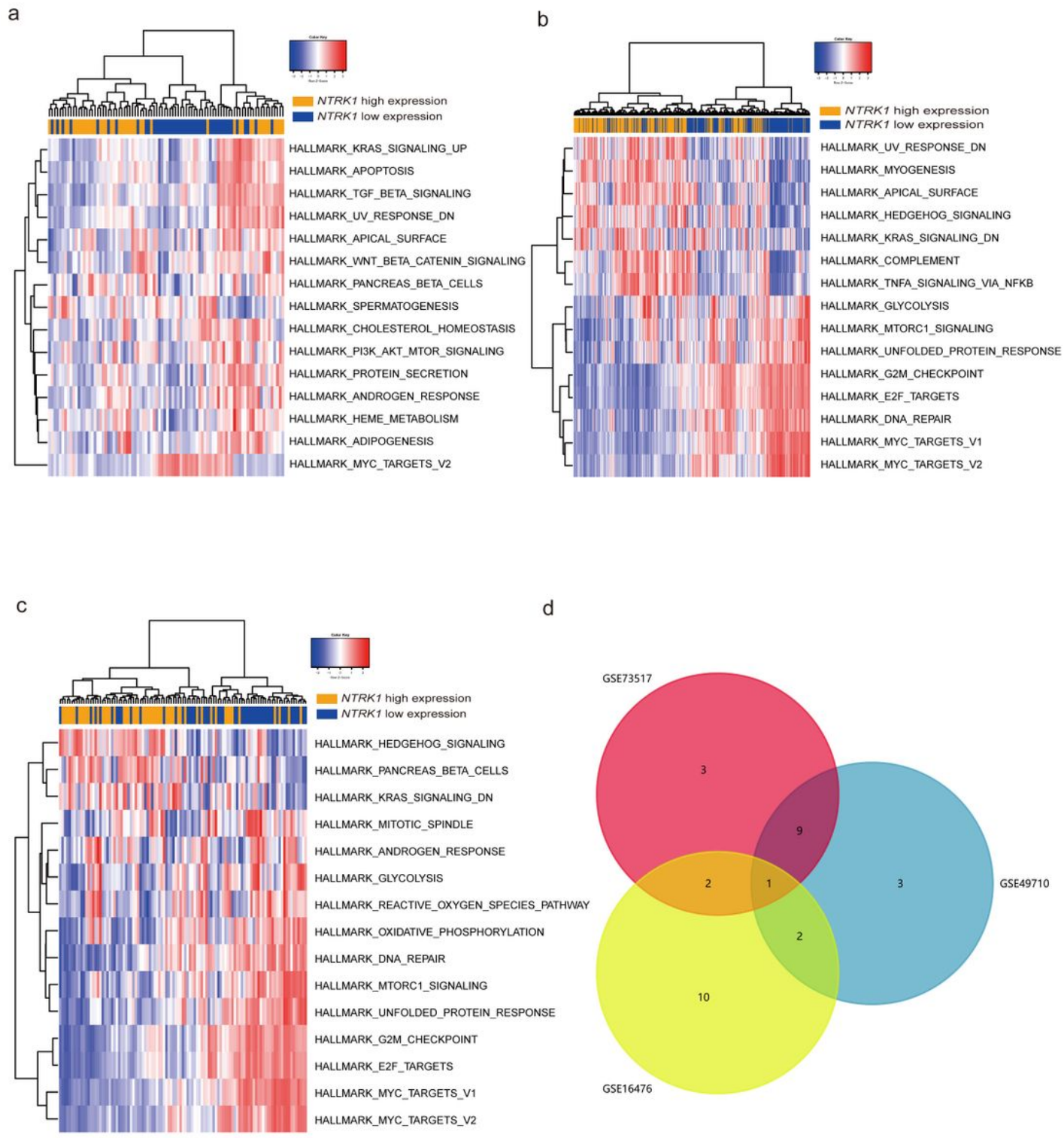

Figure 7

Gene set variation analysis (GSVA) for the NTRK1 in three GEO datasets a-c, The heatmaps of differentially expressed pathways for NTRK1 through the calculation of GSVA in GSE16476, GSE49710, GSE373517 respectively. d, Venn diagram of pathways to three GEO datasets.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xls
- TableS2.xls
- TableS3.xls
- TableS4.xls
- TableS5.xlsx
- TableS6.xlsx
- TableS7.xlsx
- TableS8.xls
- TableS9.xls