

Radiogenomics prediction for MYCN amplification in Neuroblastoma: a hypothesis generating study

Angela Di Giannatale¹, Pierluigi Di Paolo¹, Davide Curione¹, Jacopo Lenkowicz², Antonio Napolitano¹, Aurelio Secinaro¹, Paolo Toma¹, Franco Locatelli¹, Aurora Castellano¹, and Luca Boldrini²

¹Bambino Gesù Pediatric Hospital

²Policlinico Universitario Agostino Gemelli

January 23, 2021

Abstract

Background: MYCN amplification represents a powerful prognostic factor in neuroblastoma (NB) and may occasionally account for intratumoral heterogeneity. Radiomics is an emerging field of advanced image analysis that aims to extract a large number of quantitative features from standard radiological images, providing valuable clinical information Procedure: In this retrospective study, we aimed to create a radiogenomics model by correlating computed tomography (CT) radiomics analysis with MYCN status and overall survival (OS). NB lesions were segmented on pre-therapy CT scans and radiomics features subsequently extracted using a dedicated library. Dimensionality reduction/features selection approaches were then used for features procession and logistic regression models have been developed for the considered outcome. Results: Seventy-eight patients were included in this study, 24 presented MYCN amplification. In total, 232 radiomics features were extracted. Eight features were selected through Boruta algorithm and 2 features were lastly chosen through Pearson correlation analysis: mean of voxel intensity histogram ($p=0.0082$) and zone size non-uniformity ($p=0.038$). Five-times repeated 3-fold cross-validation logistic regression models yielded an Area Under the Curve (AUC) value of 0.879 on the training and 0.865 on the testing set for MYCN. No statistical significant difference has been observed comparing radiomics predicted and actual OS data. Conclusions: CT based radiomics is able to predict MYCN amplification status and OS in NB, paving the way to the in depth analysis of imaging based biomarkers that could enhance outcomes prediction.

INTRODUCTION

Neuroblastoma (NB) is the most common extracranial solid tumor of childhood and arises in the developing sympathetic nervous system, resulting in tumors of the adrenal glands and/or sympathetic ganglia. NB exhibits different clinical presentations and outcomes, in function of tumor biology, ranging from extremely aggressive behavior to spontaneous regression during infancy [1]. Treatment decisions in NB are based on prognostic factors such as age at diagnosis, the tumor stage (according to the International Neuroblastoma Staging System (INSS) [2] or the International Neuroblastoma Risk Grouping Staging System (INRGSS) [3], the histology according to the International Neuroblastoma Pathology Classification (INPC) [4, 5], and molecular and cytogenetic characteristics [6]. The molecular profiling of NB is based on biopsies or surgical specimens following resection. However, a significant challenge for the accuracy of molecular profiling may arise because of tumor heterogeneity and partial sampling of lesions. MYCN amplification occurs in around 20% of NB and is par excellence associated with poor survival with treatment implications [7].

MYCN amplification occasionally shows intratumoral heterogeneity (hetMNA) and may depend on the analyzed amount of tumor material and on the applied techniques. More rarely, heterogeneity may occur spatially within the tumor and metastases or temporally, during disease evolution. Due to the unclear

biological and clinical impact of hetMNA, no specific therapeutic strategies exist for those patients [8, 9]. Standard radiological imaging can identify macroscopic evidence of hetMNA, providing useful information on disease status. Computed tomography (CT) and magnetic resonance imaging (MRI) are recommended for diagnosis and staging of NB, although there currently is no consensus about the optimal imaging modality for assessing local disease. Both techniques are routinely used, depending on availability and other factors. CT uses ionizing radiation and contrast medium, but it is characterized by very fast acquisitions thereby eliminating the need for sedation. MRI requires long acquisition times and is considered superior in patients with intraspinal extension [10].

Radiomics is an emerging field of advanced image analysis that aims to extract a large number of quantitative features from standard radiological images, providing valuable clinical information [11]. To date radiomics is mainly used in oncology to predict diagnosis, survival, progression of disease and gene mutation (i.e. radiogenomics) [12-15]. Since radiomics has the potential for predicting molecular characteristics, it can potentially be used for determination of mutational or amplification status of specific genes, as suggested by promising preliminary experiences also in the specific field. The purpose of this study was to evaluate if radiomics features extracted from standard staging CT scans could help in predicting MYCN amplification status in patients affected by NB.

METHODS

Patients selection and data collection

Patients affected by NB treated from March 2005 to March 2019 at the Bambino Gesù Children Hospital of Rome, Italy, were retrospectively enrolled. The patient's disease was staged according to age at the time of diagnosis (< or [?] 18 months), INSS [2] and the INRGSS [3]. The amplification of MYCN gene was determined by fluorescent in situ hybridization (FISH) [25]. Clinical and pathological data of all cases were retrieved from the medical records available in the institutional archives. The inclusion criteria were defined as follows: (i) patients affected by NB; (ii) patients who underwent MYCN gene determination; (iii) patients with at least 6 months of follow up (iv) patients images had to be stored in Dicom format.

Ethics Statement

This study was conducted in accordance with the declaration of Helsinki and approved by the local institutional review board. All patients guardians signed a consent form allowing the use of their data for clinical research purposes.

Radiomics analysis

The available staging CT imaging data were manually retrieved from the institutional Picture archiving and communication system (PACS) system and anonymized according to study protocol before being analyzed by the Fondazione Policlinico Universitario "A. Gemelli" Radiomics lab of Rome, Italy. Images were reviewed by two experienced pediatric radiologists in consensus (PLD and DC) for quality check and NB were manually contoured in their entire extension, defining a single region of interest (ROI) per lesion on Eclipse Treatment Planning System (Varian Medical Systems, Palo Alto, CA, USA). In figure 1 is shown an example of NB lesion segmentation. All exams were acquired according to standard CT protocols after intravenous administration of iodinated contrast medium [26]. Equilibrium contrast phase was preferred for image segmentation as the enhancement obtained better delineates the main radiological features of NB [26, 27]. A subsequent independent check of the segmentations has been provided by a radiation oncologist belonging to the radiomics analysis team, in order to ensure ROI integrity and anatomical coherence of the segmented volumes. Radiomics features extraction was then performed using the MODDICOM R library, according to the standard recommended by the image biomarker standardisation initiative (IBSI), [16, 28].

Statistical analysis and predictive model set up

Radiomics features can be strongly correlated with each other, so that clusters of similarly distributed ones could be identified. The reduction of features' space dimensionality allows to reduce the probability of over-

fitting the data. Radiomics features were normalized according to Z-score and feature selection was performed via Boruta algorithm with 3-fold cross validation repeated 5 times. . Features, which were Boruta-selected at least 7 times out of 15, were kept for the next analysis phase. These selected features were then tested for reciprocal pairwise correlation with Pearson correlation test, and only those with a correlation value below 0.8 were kept as the final set of features. Finally, logistic regression models were trained with the selected features in 3-fold cross-validation repeated 5 times with the same dataset splits used in the feature selection step, with the aim to predict MYCN amplification status. Area Under the Curve (AUC) values and classification matrix statistics at prediction 0.3 cut-off (positive case prevalence) were averaged among the cross-validation results to obtain an estimate of the model out-of-sample performance. Furthermore, the identified radiomics features were also pooled with patients follow up information, in order to verify their OS prediction capability.

RESULTS

Patients Clinical Characteristics

This study included 81 patients diagnosed with NB between March 2005 and March 2019. Three patients have been discarded from the analysis due to lack of data regarding MYCN status. Of the remaining 78, 50 were male and 28 female, with a median age at diagnosis of 28 months (range 14 days-199 months). Table 1 summarizes the characteristics of patients. Patients were divided in three categories according to MYCN gene amplification: amplified (24), not amplified (43) and gain (11). Both a three (amplified versus not amplified versus gain) and a two variables (amplified plus gain versus not amplified) analysis were performed in order to test model's performance. At the time of the analysis, mean OS was 47 months (range: 17-171).

Radiomics features selection and modeling

All CT scans, derived from 78 patients, were considered technically adequate for the purpose of the analysis and therefore included in the study. In the figure 2 is shown the workflow of the study. Images were acquired on two different scanners (SOMATOM Plus 4 before 2011, SOMATOM Definition Flash after 2011, Siemens Healthineers, Erlangen, Germany). Mean pixel spacing was 0.44 (0.42-0.45) while mean slice thickness 2.6 mm (1-5 mm). In total, 232 radiomics features have been firstly extracted, belonging to the following feature classes: 20 statistical features (grey-level histogram); 14 morphological features; 100 texture features GLCM (grey level co-occurrence matrix); 66 texture features GLRLM (grey level run length matrix); 32 texture features GLSZM (grey level size zone matrix). After the Boruta selection procedure, 8 features were selected and addressed to the further step (Figure 3). After Pearson correlation analysis, 2 features were lastly retained: F_stat.mean (mean of voxel intensity histogram) and F_szm.2.5D.zsnu (zone size non-uniformity, computed in the 2.5D version). According to IBSI definition, F_stat.mean is a morphological intensity-based statistical feature that describes the distribution of the grey levels within the considered ROI. F_szm.2.5D.zsnu is a textural feature that assesses the distribution of zone counts over the different zone sizes. The uniformity of the zone sizes is low when the zone counts are distributed equally along zone sizes [16].

Model Prediction

The five-times repeated 3-fold cross-validation logistic regression models with the 2 selected radiomics features yielded a mean AUC value of 0.879 (SD=0.04) on the training and 0.865 (SD=0.08) on the testing set (prediction cut off = 0.3) for MYCN amplification status prediction (Figure 4A). After the first binary outcome identification (i.e. MYCN amplified versus not amplified), a following analysis considering also the additional "gain" mutational status was ran. Interestingly out of the 11 gain patients, nine were located under the amplification prediction cutoff threshold of 0.3. Figure 4B discloses the gain patients in the general model. This may support the existence of a common radiomics pattern between non amplified and gain patients, confirming the similar clinical behavior within the two classes. In order to test the ability of the two selected radiomics features to predict survival outcomes, actual OS data were plotted with radiomics based predictions and relative Kaplan-Meier curves were designed. Log-rank tests have been done between observed and predicted amplified and not amplified patients curves ($p=0.003$ and 0.05 respectively),

while no statistical significant difference has been observed between the observed and actual curves (p =not significant), (Figure 5).

DISCUSSION

Several somatic genetic alterations, that provide critical prognostic information, have been described in NB. Among those, MYCN amplification is a crucial prognostic factor, being associated to advanced tumor stage and poor outcome, and is central to the risk stratification systems. According to the recommendations of the INRG Biology Committee, the methods accepted for detecting MYCN amplification are represented by FISH, Polymerase Chain Reaction (PCR), Array comparative genomic hybridization (aCGH) and Multiple Ligation dependent Probe Amplification (MLPA) on tumor tissue [17]. In clinical practice, the availability of tumor material is sometimes challenging even if biopsy remains the gold standard for the diagnosis of NB. Several minimally invasive surgical approaches have been adopted, including samples from fine-needle aspirates, but they result to be burdened by a significant risk of hetMNA, giving clinicians only partial information that could be misleading in the choice of the most appropriate therapeutic approach. Furthermore, MYCN amplification may rarely show hetMNA itself, making the sampling and following diagnosis even more challenging. Recently, other experimental methods to detect MYCN amplification by using cell-free DNA in serum and plasma have been developed in patient affected by NB [18, 19]. However, these approaches have not been validated yet in the current clinical practice. In this context, radiomics and radiogenomics, may play a significant role thanks to the quantitative noninvasive and repeatable analysis of standard clinical imaging that encompasses the whole tumor volume, taking into account the hetMNA. Radiomics features may be associated with specific molecular pathways or mutations, offering reliable predictive tools for clinicians in breast cancer [20], glioma [14], lung cancer [15, 21] and even in healthy tissue [22]. Furthermore, CT represent the most common modality used to diagnose and stage NB and for this reason the images are easily accessible even for retrospective analysis. Brisse and colleagues correlate semantic features to the genomic profile of NB, showing a significant relationships among the sympathetic origin of the tumor, its genomic profile and the outcome [23]. In our preliminary study we aimed to develop a radiomics single phase CT based predictive model as a convenient and reliable biomarker for MYCN amplification status in NB. The proposed radiomics signature was set up to differentiate among MYCN amplified and MYCN wild types tumors, representing a promising tool for prognostic prediction in a particularly innovative scenario. Wu and colleagues have recently published a similar model, achieving similar results in MYCN prediction using a combined clinical-radiomics modeling approach. Despite the similar structure of the two experiences, our model focuses only on one CT sequence, reducing the foreseen workload for a real-world application of our approach in clinical practice and is based on a lower number of radiomics features (2 versus 7), reducing the computational burden and offering a more immediate workflow [24].

In conclusion, this preliminary study suggests that standard post-contrast CT logistic radiogenomics classifier may help to identify MYCN amplification status, which may successfully integrate the traditional invasive MYCN status testing and provide survival predictions for patients presenting amplification. The limitations of this study are represented by the relatively small sample size, which cannot reflect the overall MYCN status population, and the retrospective and single-institution nature. Further prospective multicenter validation studies are necessary to confirm the generalizability of the observed results and to externally validate the proposed model.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest regarding the publication of this article.

ACKNOWLEDGEMENTS

Dr. Angela Di Giannatale research is supported by a grant from the Ministero della Salute (GR-2016-02364088), The Authors acknowledge the support of the Radiomics Working Group of the Italian Alliance Against Cancer (<https://www.alleanzacontroilcancro.it/>).

REFERENCES

1. Matthay KK, Maris JM, Schleiermacher G, Nakagawara A, Mackall CL, Diller L, Weiss WA. Neuroblastoma. *Nat Rev Dis Primers*. 2016; 2: 1-21.
2. Brodeur GM, Pritchard J, Berthold F, Carlsen NL, Castel V, Castelberry RP, De Bernardi B, Evans AE, Favrot M, Hedborg F, Kaneko M, Kemshead J, Lampert F, Lee REJ, Look AT, Pearson ADJ, Philip T, Roald B, Sawada T, Seeger RC, Tsuchida Y, Voute PA. Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *J Clin Oncol*. 1993; 11: 1466-77.
3. Monclair T, Brodeur GM, Ambros PF, Brisse HJ, Cecchetto G, Holmes K, Kaneko M, London WB, Matthay KK, Nuchtern JG, von Schweinitz D, Simon T, Cohn SL, Pearson AD. The International Neuroblastoma Risk Group (INRG) staging system: an INRG Task Force report. *INRG Task Force. J Clin Oncol*. 2009; 27: 298-303.
4. 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: 349-63.
5. Peuchmaur M, d'Amore ESG, Joshi VV, Hata J, Roald B, Dehner LP, Gerbing RB, Stram DO, Lukens JN, Matthay KK, Shimada H. Revision of the International Neuroblastoma Pathology Classification: confirmation of favorable and unfavorable prognostic subsets in ganglioneuroblastoma, nodular. *Cancer*. 2003; 98: 2274-81.
6. 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. *J Clin Oncol*. 2009; 27: 289-97.
7. Seeger RC, Brodeur GM, Sather H, Dalton A, Siegel SE, Wong KY, Hammond D. Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *N Engl J Med*. 1985; 313: 1111-6.
8. Berbegall AP, Bogen D, Potschger U, Beiske K, Bown N, Combaret V, Defferrari R, Jeison M, Mazzocco K, Varesio L, Vicha A, Ash S, Castel V, Coze C, Ladenstein R, Owens C, Papadakis V, Ruud E, Amann G, Sementa AR, Navarro S, Ambros PF, Noguera R, Ambros IM. Heterogeneous MYCN amplification in neuroblastoma: a SIOP Europe Neuroblastoma Study. *Br J Cancer*. 2018; 118: 1502-1512.
9. Campbell K, Naranjo A, Hibbitts E, Gastier-Foster JM, Bagatell R, Irwin MS, Shimada H, Hogarty M, Park JR, DuBois SG. Association of heterogeneous MYCN amplification with clinical features, biological characteristics and outcomes in neuroblastoma: A report from the Children's Oncology Group. *Eur J Cancer*. 2020; 133: 112-119.
10. Brisse HJ, McCarville MB, Granata C, Krug KB, Wootton-Gorges SL, Kanegawa K, Giammarile F, Schmidt M, Shulkin BL, Matthay KK, Lewington VJ, Sarnacki S, Hero B, Kaneko M, London WB, Pearson AD, Cohn SL, Monclair T. Guidelines for imaging and staging of neuroblastic tumors: consensus report from the International Neuroblastoma Risk Group Project. *Radiology*. 2011; 261:243-57.
11. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016; 278: 563-77.
12. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebers F, Rietbergen MM, Leemans CR, Dekker A, Quackenbush J, Gillies RJ, Lambin P. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014; 5: 4006.
13. Tu W, Sun G, Fan L, Wang Y, Xia Y, Guan Y, Li Q, Zhang D, Liu S, Li Z. Radiomics signature: A potential and incremental predictor for EGFR mutation status in NSCLC patients, comparison with CT morphology. *Lung Cancer*. 2019; 132: 28-35.
14. Kim M, Jung SY, Park JE, Jo Y, Park SY, Nam SJ, Kim JH, Kim HS. Diffusion- and perfusion-weighted MRI radiomics model may predict isocitrate dehydrogenase (IDH) mutation and tumor aggressiveness in diffuse lower grade glioma. *Eur Radiol*. 2020; 30: 2142-2151.
15. Ma DN, Gao XY, Dan YB, Zhang AN, Wang WJ, Yang G, Zhu HZ. Evaluating Solid Lung Adenocarcinoma Anaplastic Lymphoma Kinase Gene Rearrangement Using Noninvasive Radiomics Biomarkers.

- Onco Targets Ther. 2020; 13: 6927–6935.
16. Zwanenburg A, Vallieres M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, Bogowicz M, Boldrini L, Buvat I, Cook GJR, Davatzikos C, Depeursinge A, Desserot MC, Dinapoli N, Dinh CV, EcheGARAY S, El Naqa I, Fedorov AY, Gatta R, Gillies RJ, Goh V, Gotz M, Guckenberger M, Ha SM, Hatt M, Isensee F, Lambin P, Leger S, Leijenaar RTH, Lenkiewicz J, Lippert F, Losnegard A, Maier-Hein KH, Morin O, Muller H, Napel S, Nioche C, Orlhac F, Pati S, Pfaehler EAG, Rahmim A, Rao AUK, Scherer J, Siddique MM, Sijtsema NM, Socarras Fernandez J, Spezi E, Steenbakkers RJHM, Tanadini-Lang S, Thorwarth D, Troost EGC, Upadhaya T, Valentini V, van Dijk LV, van Griethuysen J, van Velden FHP, Whybra P, Richter C, Lock S. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020; 295: 328-338.
 17. Ambros PF, Ambros IM, Brodeur GM, Haber M, Khan J, Nakagawara A, Schleiermacher G, Speleman F, Spitz R, London WB, Cohn SL, Pearson ADJ, Maris JM. International consensus for neuroblastoma molecular diagnostics: report from the International Neuroblastoma Risk Group (INRG) Biology Committee. *Br J Cancer*. 2009; 100: 1471–1482.
 18. Iehara T, Yagyu S, Gotoh T, Ouchi K, Yoshida H, Miyachi M, Kikuchi K, Sugimoto T, Hosoi H. A prospective evaluation of liquid biopsy for detecting MYCN amplification in neuroblastoma patients. *Jpn J Clin Oncol*. 2019; 49: 743-748.
 19. Lodrini M, Sprussel A, Astrahantseff K, Tiburtius D, Korschak R, Lode HN, Fischer M, Keilholz U, Eggert A, Deubzer HE. Using droplet digital PCR to analyze MYCN and ALK copy number in plasma from patients with neuroblastoma. *Oncotarget*. 2017; 8: 85234-85251.
 20. Li H, Zhu Y, Burnside ES, Huang E, Drukker K, Hoadley KA, Fan C, Conzen SD, Zuley M, Net JM, Sutton E, Whitman GJ, Morris E, Perou CM, Ji Y, Giger ML. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancersubtypes in the TCGA/TCIA data set. *NPJ breast cancer*. 2016; 2: 16012.
 21. Li H, Zhang R, Wang S, Fang M, Zhu Y, Hu Z, Dong D, Shi J, Tian J. CT-Based Radiomic Signature as a Prognostic Factor in Stage IV ALK-Positive Non-small-cell Lung Cancer Treated With TKI Crizotinib: A Proof-of-Concept Study. *Front Oncol*. 2020; 10: 57.
 22. Nero C, Ciccarone F, Boldrini L, Lenkiewicz J, Paris I, Capoluongo ED, Testa AC, Fagotti A, Valentini V, Scambia G. Germline BRCA 1-2 status prediction through ovarian ultrasound images radiogenomics: a hypothesis generating study (PROBE study). *Sci Rep*. 2020; 10: 16511.
 23. Brisse HJ, Blanc T, Schleiermacher G, Mosseri V, Philippe-Chomette P, Janoueix-Lerosey I, Pierron G, Lapouble E, Peuchmaur M, Freneaux P, Galmiche L, Algret N, Peycelon M, Michon J, Delattre O, Sarnacki S. Radiogenomics of neuroblastomas: Relationships between imaging phenotypes, tumor genomic profile and survival. *PLoS One*. 2017; 12: e0185190.
 24. Wu H, Wu C, Zheng H, Wang L, Guan W, Duan S, Wang D. Radiogenomics of neuroblastoma in pediatric patients: CT-based radiomics signature in predicting MYCN amplification. *Eur Radiol*. 2020; doi: 10.1007/s00330-020-07246-1. Online ahead of print.
 25. Ambros PF, Ambros IM. Pathology and biology guidelines for resectable and unresectable neuroblastic tumors and bone marrow examination guidelines. *Medical and pediatric oncology*. 2001; 37: 492-504.
 26. Nievelstein RA, van Dam IM, van der Molen AJ. Multidetector CT in children: current concepts and dose reduction strategies. *Pediatr Radiol*. 2010; 40: 1324-1344.
 27. Papaioannou G, McHugh K. Neuroblastoma in childhood: review and radiological findings. *Cancer Imaging*. 2005; 5: 116-127.
 28. Dinapoli N, Alitto AR, Vallati M, Gatta R, Autorino R, Boldrini L, Damiani A, Valentini V. Moddicom: a complete and easily accessible library for prognostic evaluations relying on image features. *Annu Int Conf IEEE Eng Med Biol Soc*. 2015: 771-4.

FIGURE LEGENDS

Figure 1 . Abdominal NB lesion segmented on CT image

Figure 2 Workflow of the study showing the methodology used and radiomics data analysis.

Figure 3 . Radiomics features identification and selection procedure with Boruta selection algorithm.

Figure 4 . A) MYCN mutational status for binary prediction performance. Red dots represent non amplified patients, blue dots amplified ones. B) MYCN mutational status for three tiers prediction performance. Red dots represent non amplified patients, blue dots amplified ones and green dots gain patients.

Figure 5 . Kaplan-Meier curves for radiomics based survival prediction and observed data

Hosted file

Table 1.pdf available at <https://authorea.com/users/391329/articles/505473-radiogenomics-prediction-for-mycn-amplification-in-neuroblastoma-a-hypothesis-generating-study>



