European Journal of Cancer 193 (2023) 113291

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Computational intelligence analysis of high-risk neuroblastoma patient health records reveals time to maximum response as one of the most relevant factors for outcome prediction

Davide Chicco [a,b], Riccardo Haupt [c], Alberto Garaventa [d], Paolo Uva [e], Roberto Luksch [f,1], Davide Cangelosi [e,*,1]

[a] *Institute of Health Policy Management and Evaluation, University of Toronto, Toronto, Ontario, Canada*
[b] *Dipartimento di Informatica Sistemistica e Comunicazione, Università di Milano-Bicocca, Milan, Italy*
[c] *DOPO Clinic, Department of Hematology/Oncology, IRCCS Istituto Giannina Gaslini, Genoa, Italy*
[d] *Unità di Oncologia, IRCCS Istituto Giannina Gaslini, Genoa, Italy*
[e] *Unità di Bioinformatica Clinica, IRCCS Istituto Giannina Gaslini, Genoa, Italy*
[f] *S.C. Pediatria oncologica, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*

Received 1 February 2023; Received in revised form 24 July 2023; Accepted 9 August 2023
Available online 19 August 2023

KEYWORDS
Artificial intelligence;
Neuroblastoma;
Maximum response to first-line treatment;
Time to maximum response to first-line treatment;
Random forests;
Feature importance ranking

Abstract ***Objective:*** Seek new candidate prognostic markers for neuroblastoma outcome, relapse or progression.
Materials and methods: In this multicentre and retrospective study, Random Forests coupled with recursive feature elimination techniques were applied to electronic records (55 clinical features) of 3034 neuroblastoma patients. To assess model performance and feature importance, dataset was split into a training set (80%) and a test set (20%).
Results: In the test set, the mean Matthews correlation coefficient for the Random Forests models was greater than 0.46. Feature importance analysis revealed that, together with maximum response to first-line treatment (D_MAX_RESP), time to maximum response to first-line treatment (TIME_MAX_RESP.days) is a relevant predictor of both patients' outcome and relapse\progression. We showed the prognostic value of the max response to first-line treatment in clinically relevant subsets of high-, intermediate-, and low-risk patients for both overall and relapse-free survival (Log-rank p-value < 0.0001). In high-risk patients older than 18 months and stage 4 tumour achieving a complete response or very good partial response, patients who exhibited a D_MAX_RESP greater than 9 months showed a better

* *Correspondence to*: IRCCS Istituto Giannina Gaslini, Largo Gerolamo Gaslini 5, 16147 Genoa, Italy.
E-mail address: davidecangelosi@gaslini.org (D. Cangelosi).
[1] equal contribution

https://doi.org/10.1016/j.ejca.2023.113291

prognosis with respect to patients achieving D_MAX_RESP earlier than 9 months (overall survival): hazard ratio 3.3 95% confidence interval 1.8–5.9, Log-rank p-value p < 0.0001; relapse-free survival: 3.2 95%CI 1.8–5.6, Log-rank p-value p < 0.0001).

Conclusion: Our findings evidence the emerging role of the TIME_MAX_RESP.days in addition to the D_MAX_RESP as relevant predictors of outcome and relapse\progression in neuroblastoma with potential clinical impact on the management and treatment of patients.

1. Introduction

Peripheral neuroblastic tumours are a family of rare tumours of the sympathetic nervous system derived from the primitive neural crest, accounting for 7–10% of all tumours in children [1]. Neuroblastoma is the most common and clinically relevant peripheral neuroblastic tumour histologic variant [2]. Patients diagnosed with neuroblastoma are assigned to different risk groups, which drive the choice of the treatment protocol [3]. Although the advances in neuroblastoma treatment have improved neuroblastoma patient survival rates with respect to the past, the clinical and biological heterogeneity of neuroblastoma still poses serious challenges to physicians [4,5]. About 20% of patients with high-risk neuroblastoma early progress or are refractory to standard induction therapy, and 50% of patients who achieve remission later relapse [6–8]. Novel prognostic markers able to accurately predict patients' relapse\progression or deaths are fundamental to improve patients' stratification by tailoring treatment to patients' characteristics [9–11].

The ability of promising biomarkers to improve the actual patients' stratification system has already been demonstrated in preclinical studies [5,9,12–17] However, most of the proposed markers are still waiting for clinical validation to date [18].

The Registro Italiano dei Tumori Neuroblastici Periferici (RINB) is the official Italian register for collecting clinical data about patients with neuroblastoma treated at institutions of the Italian Association of Pediatric Hematology-Oncology (AIEOP) network [19]. Since its foundation in 1976, RINB database collects features about personal, clinical, biological, histological, and treatment data on patients with neuroblastoma, thereby becoming an essential resource for supporting clinical decisions and for providing data to research projects. Epidemiological studies based on selected data from the RINB database have already been published showing that the outcome of children with neuroblastoma has progressively improved over the years, stage 4s neuroblastoma is curable in nearly 90% of cases, the cure rate could be further increased through timely identification of patients at risk who might benefit from surgical techniques, infants diagnosed with stage 4s neuroblastoma who underwent upfront tumour resection had a better outcome than who did not, and infants who progressed to stage 4 did worse in relation to older age at progression and longer interval between diagnosis and progression [19–23]. However, a systematic analysis of the entire RINB database to identify new predictive markers is currently lacking.

Machine learning provides the theoretical basis and practical solutions to handle scientific problems including discovering predictive features of disease severity and patient survival from electronic medical records [9,24].

In the present study, we apply machine learning techniques to data in the RINB database for discovering new prognostic markers of the neuroblastoma outcome or relapse/progression with potential clinical impact on the management and treatment of patients.

2. Materials and methods

2.1. Patient characteristics, inclusion criteria and clinical parameters definition

A total of 3756 patients aged 0–36 years were diagnosed with neuroblastoma between January 1979 and December 2020 in 32 institutions of the Italian AIEOP network and registered in the RINB database. Tumour stage and maximum response to first-line treatment (D_MAX_RESP) were defined according to Italian Cooperative Group for Neuroblastoma (ICGNB) criteria before 1989 [25], and from 1989 according to the International Neuroblastoma Staging System (INSS) and the International Neuroblastoma Response Criteria system [26,27]. Patients were treated by the ICGNB protocols until 1994 [25,28–32], and by protocols of the International Society of Pediatric Oncology European Neuroblastoma Group (SIOPEN) thereafter [33–36].

All protocols were approved by the local institutional review board. Before being enroled and treated on their respective trials, patients or their guardians signed a consent form allowing the use of their clinical and nongenetic data for clinical research purposes.

Time to the maximum response to first-line treatment (TIME_MAX_RESP.days) was computed by subtracting the date of the D_MAX_RESP from the date of diagnosis. Only patients who deceased from tumour, and patients with available status at the clinical follow-up were included in the study. Furthermore, features

with 80% or greater of non-missing values were also included.

2.2. Study endpoints

The primary endpoint was the clinical status, alive or dead, referred to as the outcome. Overall survival (OS) was defined as the time (in years) from disease diagnosis to patient death or the last follow-up if the patient survived. The secondary endpoint was a binary variable expressing whether a patient experienced a relapse/progression. Relapse-free Survival (RFS) was defined as the time (in years) from disease diagnosis to relapse\progression or the last follow-up if no relapse\progression occurred.

2.3. Machine learning build-up and validation

To be analysed by machine learning techniques, the dataset required data filtering, harmonisation, and transformation steps whose effect was to exclude features providing redundant information, not sufficiently reliable, or unsuitable to be analysed with machine learning techniques. Missing values might have negative effects on the reliability of the analyses, therefore we excluded features whose number of missing values exceeded 20% of total values. Machine learning analysis was carried out on patient's status, alive or dead, with at least 3-year follow-up to improve the reliability of the reference variables used in the machine learning analysis.

We applied Random Forests [37] with default parameters for binary classification on the general datasets and on the localised and metastatic subsets, both for outcome and relapse\progression. We employed the randomForest R software library version 4.7–1.1 with default parameters (500 trees to grow, number of variables randomly sampled as candidates at each split equal to square root of the number of features), on a computer running R version 4.1.3 through Miniconda.

We trained our model on a training set of 80% of randomly selected patients' data and applied the trained model on the remaining test set of 20% of patients' data. We repeated this procedure 100 times and measured prediction performance results with the Matthews correlation coefficient (MCC) [38,39].

We decided to use MCC performance metric because it is more informative than other metrics [40–42]. MCC ranges in [−1, +1] where +1 indicates perfect prediction and −1 indicates imperfect prediction.

We used a cutoff of 0.5 for the generation of the confusion matrix: we mapped values below 0.5 into 0s and mapped values greater or equal to 0.5 into1s. To assess feature importance, we employed Random Forests in a recursive feature elimination procedure [37], by removing a single variable and computing the binary classification, for a number of iterations equal to the number of features.

2.4. Statistical analysis

OS and RFS curves were plotted by the Kaplan-Meier method and were compared with the log-rank test. P-values lower than 0.005 are considered significant, as suggested by Benjamin and colleagues [43]. Patients with missing values were excluded from the Kaplan-Meier analysis. GraphPad Prism version 8.0 for Microsoft Windows, www.graphpad.com, was used to plot Kaplan-Meier curves and to compute log-rank p-values, hazard ratios(HR) and relative confidence intervals (CI). Logrank method was used to compute hazard ratios and relative confidence intervals. Treatment protocol was evaluated as a potential confounder. The lowest log-rank p-value or the highest HR, in case of identical log-rank p-value, was used to select the best cut-off value.

Additional details about methods are included in the Supplementary materials.

3. Results

3.1. Prediction performance and feature importance ranking using machine learning techniques

Analysis used the data of 3756 patients with neuroblastoma registered in the RINB database until 31st December 2020. Fig. 1 shows the schematic representation of the analysis workflow implemented in the present study.

Data filtering, harmonisation, and transformation steps are fundamental steps for preparing the data to be analysed by machine learning techniques and to ensure the validity and reliability of data analysis results [44]. The effect of data preprocessing was to exclude patient records or features providing redundant information, not sufficiently reliable, or unsuitable to be analysed with machine learning techniques. The complete preprocessing procedure with the number of patients and features excluded from subsequent analysis is reported in the Supplementary Fig. S1.

The filtering step excluded 11 (0.29%) patients who deceased for causes unrelated to neuroblastoma, 100 (2.66%) patients who deceased for toxicity or second tumour, and 611 (16.2%) patients with missing status at the clinical follow-up were excluded from subsequent analyses.

RINB dataset was originally constituted by 412 clinical and molecular parameters. The filtering step excluded 74 (17.9%) features providing redundant information, 273 (66.2%) features with number of missing values exceeding 20% of the total values, and 10 (2.4) features that cannot be analysed by ML methods. Data harmonisation and mapping were also carried out as

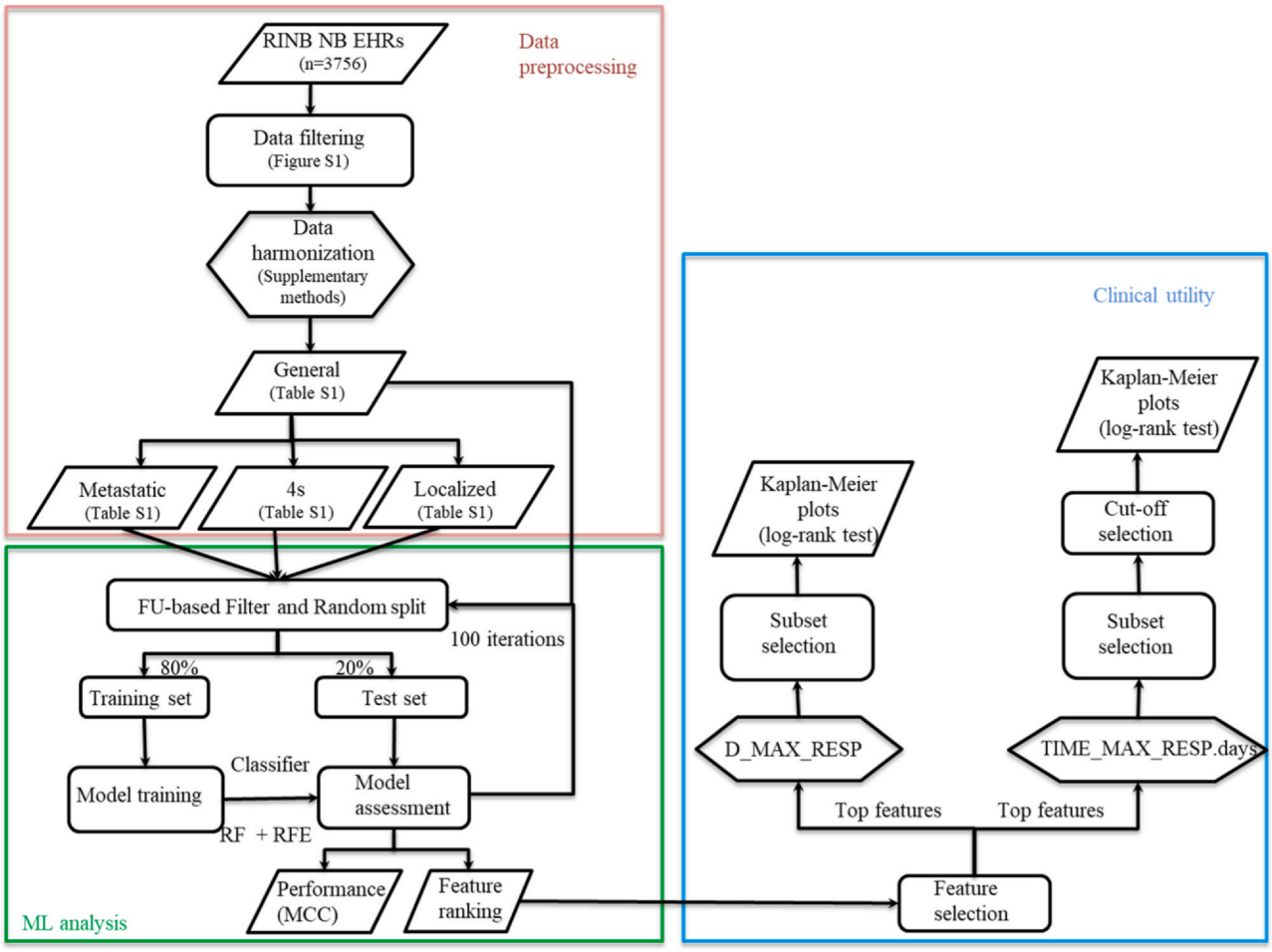

Fig. 1. Workflow of the analysis. Analysis is divided into three main sections: data preprocessing, machine learning analysis, and clinical utility assessment. In the preprocessing, 3756 records from the Registro Italiano dei Tumori Neuroblastici Periferici (RINB) database were filtered, new features were created and harmonised to be processed by machine learning techniques. At the end of this step, 3034 and 55 features constituted the general dataset. Metastatic, metastatic showing spontaneous maturation and regression, and localised subsets of patients from the general dataset were extracted to consider distinct disease peculiarities. Outcome and relapse\progression were used as endpoints for the four datasets. The eight resulting datasets were iteratively split into training and test sets, respectively. A minimum of 3 years of follow-up was set up to define alive patients' outcomes for machine learning analysis. Training data were used to build classifiers with Random Forest and recursive feature elimination techniques. Test data were used to assess prediction performance and rank features by relevance. The clinical utility of selected features was assessed in clinically relevant groups of patients by Kaplan-Meier method and log-rank test. The prognostic value of the selected features was assessed in selected subsets of patients. When necessary, distinct cut-off values were tested to find the most relevant one. Supplementary figures and tables were included between brackets to provide additional details about datasets.

reported in the Supplementary methods. Pre-processing steps reduced a dataset of 3034 records composed of 55 clinical features. To consider tumour subsets, we decided to perform machine learning analyses on the general dataset and on three subsets defined based on tumour stage, which included patients with metastatic (stage 4), localised (stage 1, 2, and 3), or patients with special metastatic disease associated with a favourable prognosis (stage 4s), respectively. The number of patients and relative percentage of each clinical feature in the general, metastatic, localised, and 4s datasets is summarised in Table S1. To simplify the legibility of the manuscript, we indicated the general, subset of metastatic, localised, or 4s with A, M, L, and 4s, respectively, followed by _OS or _RFS, if outcome or relapse\progression was used as endpoint. The eight datasets were: A_OS, A_RFS, M_OS, M_RFS, L_OS, L_RFS, 4s_OS, and 4s_RFS. Patients' outcome or relapse\progression were used as reference variables to perform supervised machine learning analysis. A detailed evaluation of the patient follow-up revealed that 535 (17.6%) alive patients had a follow-up lower than 3 years. Since the outcome of alive patients cannot be reliably determined, we decided to exclude patient records from the subsequent machine learning analysis. OS and RFS features were excluded from the machine learning analysis to avoid trivial conclusions. The resulting dataset included 2499 out of 3034 (82.4%) patients. The characteristics of

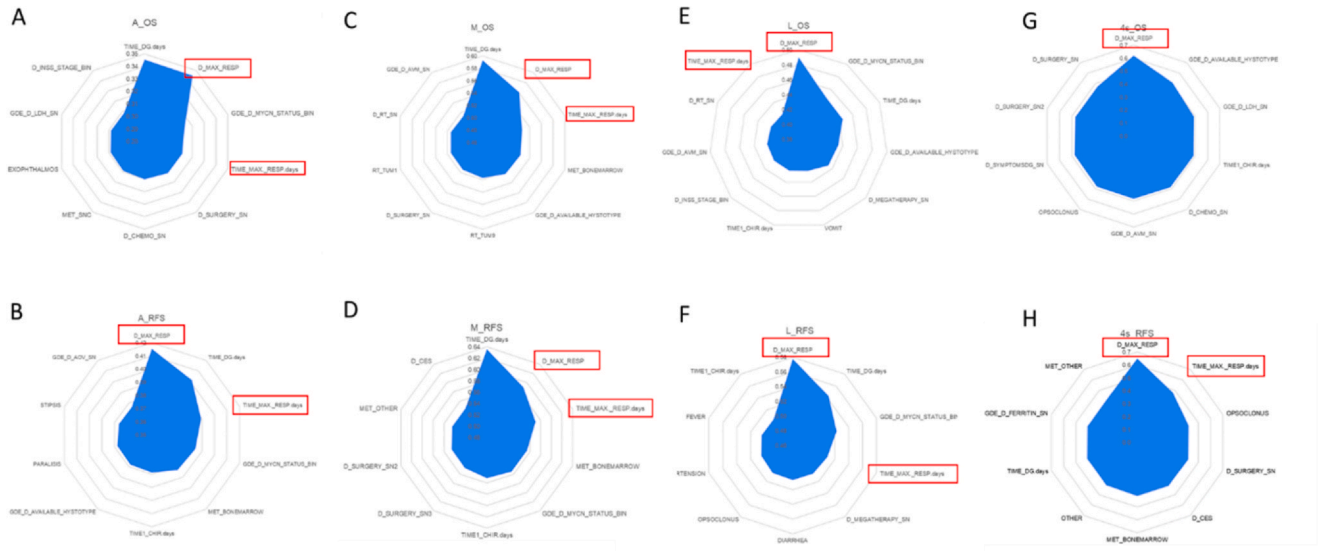

Fig. 2. Spider chart of the top 10 relevant features predicted by Random Forest in eight distinct datasets. Features are ordered by relevance in clockwise order. The most relevant appears on the top side of the chart. Relevance was calculated as 1 minus mean MCC obtained by a feature during the execution of recursive feature elimination procedure. Charts are relative to the analysis in the following datasets: A) A_OS, B) A_RFS, C) M_OS, D) M_RFS, E) L_OS, F) L_RFS, G) 4s_OS, and H) 4s_RFS. maximum response to first-line treatment (D_MAX_RESP) and time to the maximum response to first-line treatment (TIME_MAX_RESP.days) were evidenced by a red box.

patients analysed by machine learning techniques are reported in Table S2.

Each dataset was randomly split into two groups of 80% and 20% of patients for 100 consecutive iterations. In the test set, the resulting mean MCC for the Random Forests models from the highest to the lowest was: 0.697 for A_OS, 0.632 for 4s_RFS, 0.629 for A_RFS, 0.584 for L_OS, 0.556 for 4s_OS, 0.509 for L_RFS, 0.498 for M_OS, and 0.462 for M_RFS.

Features in each dataset were sorted by decreasing order of relevance and ranking is reported in the Table S3-S10. The top ten relevant clinical parameters in each dataset were displayed by spider charts for additional investigations (Fig. 2).

Our findings evidence that age at diagnosis (TIME_DG.days), D_MAX_RESP, TIME_MAX_-RESP.days, and MYCN status (GDE_D_-MYCN_STATUS_BIN) are the most relevant predictors of both patients' outcome and relapse\progression.

Furthermore, our findings suggest that D_MAX_RESP and TIME_MAX_RESP.days might be of clinical utility for neuroblastoma.

3.2. Assessing the clinical utility of the D_MAX_RESP for neuroblastoma

D_MAX_RESP was classified into six categories: complete response (CR), very good partial response (VGPR), partial response (PR), mixed response (MR), no response (NR), and progressive disease (PD).

To evaluate the clinical utility of D_MAX_RESP for neuroblastoma, survival analysis was carried out in all datasets (Fig. 3A-H). D_MAX_RESP significantly stratified patients in the general dataset, in the subset of patients with metastatic tumour, in the subset with localised tumour, as well as those with 4s tumour, thus highlighting the high prognostic value of D_MAX_RESP in neuroblastoma for both OS and RFS (Log-rank p-value < 0.0001, Fig. 3A-H). The prognostic value of D_MAX_-RESP was confirmed for OS and RFS in additional clinically relevant subgroups of patients defined by combination of established prognostic markers including the group of high-risk patients older than 18 months with stage 4 tumours (Log-rank p-value < 0.0001; Fig. 4A-B), low/intermediate-risk patients younger than 18 months with stage 4 tumours lacking Neuroblastoma-Derived V-Myc Avian Myelocytomatosis Viral Related Oncogene (MYCN) amplification (Log-rank p-value < 0.0001; Fig. 4C-D), intermediate-risk patients older than 18 months with stage 3 and not amplified MYCN tumours (Log-rank p-value < 0.0001; Fig. 4E-F), and low-risk patients with stage 1, 2, 4s tumours lacking MYCN amplification (Log-rank p-value < 0.0001; Fig. 4H-G).

The high percentage of deaths or relapses\progressions in the subsets of patients achieving a CR or VGPR suggested that D_MAX_RESP is insufficient to stratify these patients.

3.3. Assessing the clinical utility of TIME_MAX_RESP.days for neuroblastoma

Survival analysis based on TIME_MAX_RESP.days feature was carried out in the subset of 502 high-risk

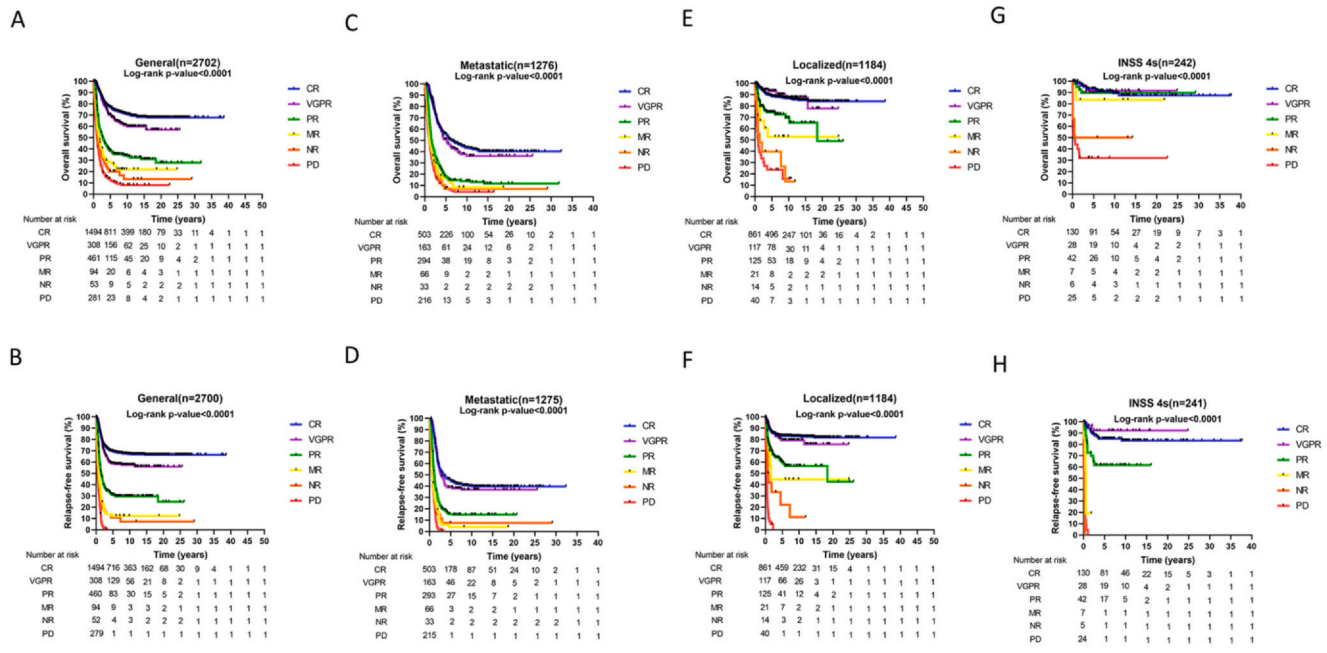

Fig. 3. Kaplan-Meier estimates of overall survival (OS) and relapse free survival (RFS) based on the D_MAX_RESP feature in the eight datasets. Plots are relative to: (A) A_OS, (B) A_RFS, (C) M_OS, (D) M_RFS, (E) L_OS, (F) L_RFS, (G) 4s_OS, and (H) 4s_RFS. Plots are entitled with a dataset name and number of patients. Survival curves for patients achieving complete response (CR) (blue), very good partial response (VGPR) (violet), partial response (PR) (green), mixed response (MR) (yellow), no response (NR) (orange), and progressive disease (PD) (red) were compared by log-rank test, respectively. Log-rank p-value is reported on top of each plot. When a log-rank p-value was lower than 0.0001, we indicated it as log-rank-p-value < 0.0001. A table reporting the number of patients at risk is shown below each plot.

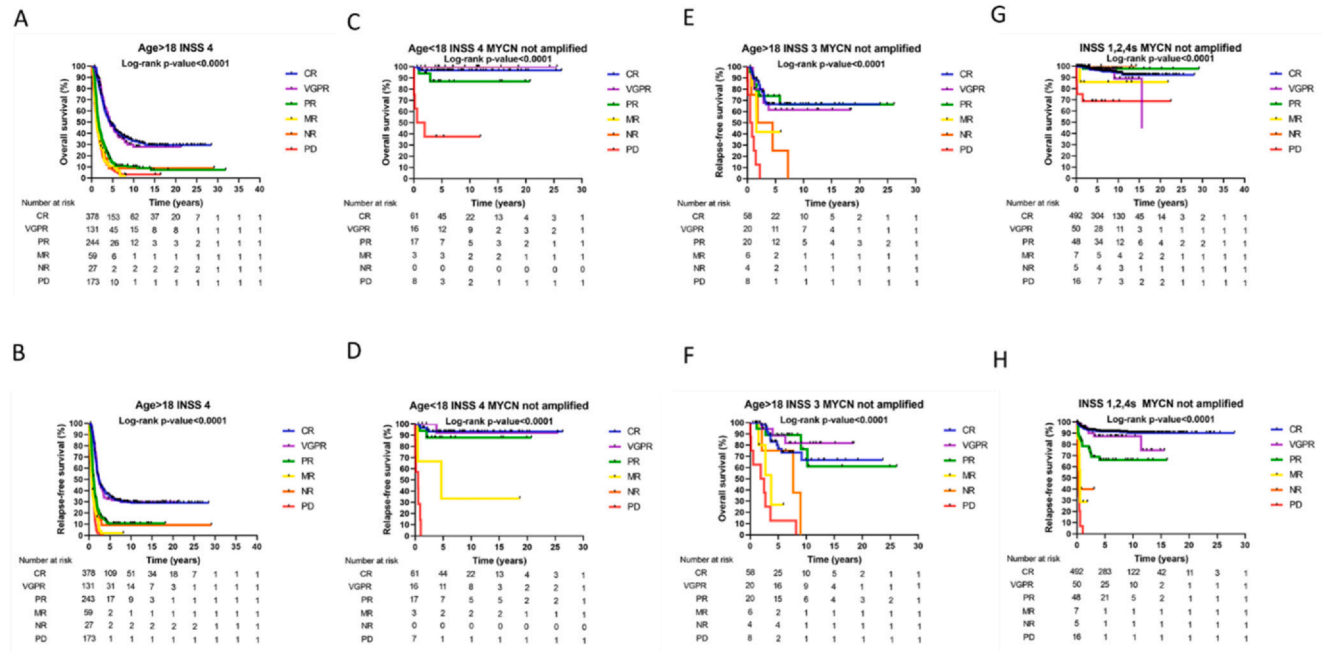

Fig. 4. Kaplan-Meier estimates of OS and RFS based on the D_MAX_RESP feature in clinically relevant groups of patients. Plots are relative to: (A) OS of high-risk patients, stage 4, and age > 18 months, (B) RFS of high-risk patients, stage 4, and age > 18 months, (C) OS of intermediate-risk patients, stage 4, age < 18 months, and with not amplified MYCN tumours, (D) RFS of intermediate-risk patients, stage 4, age < 18 months, and with not amplified MYCN tumours, (E) OS of intermediate-risk patients, age > 18 months, stage 3 with not amplified MYCN tumours, (F) RFS of intermediate-risk patients, age > 18 months, stage 3 with not amplified MYCN tumours, (H) OS of low-risk patients with stages 1, 2, 4s and no MYCN amplification tumour, and (G) RFS of low-risk patients with stages 1, 2, 4s and no MYCN amplification tumour. Plots are entitled with the characteristics of the patients' subset. Survival curves for patients achieving CR (blue), VGPR (violet), PR (green), MR (yellow), NR (orange), and PD (red) curves were compared by log-rank test, respectively. Log-rank p-value is reported on top of each plot. When a log-rank p-value was lower than 0.0001, we indicated it as log-rank-p-value < 0.0001. A table reporting the number of patients at risk is shown below each plot.

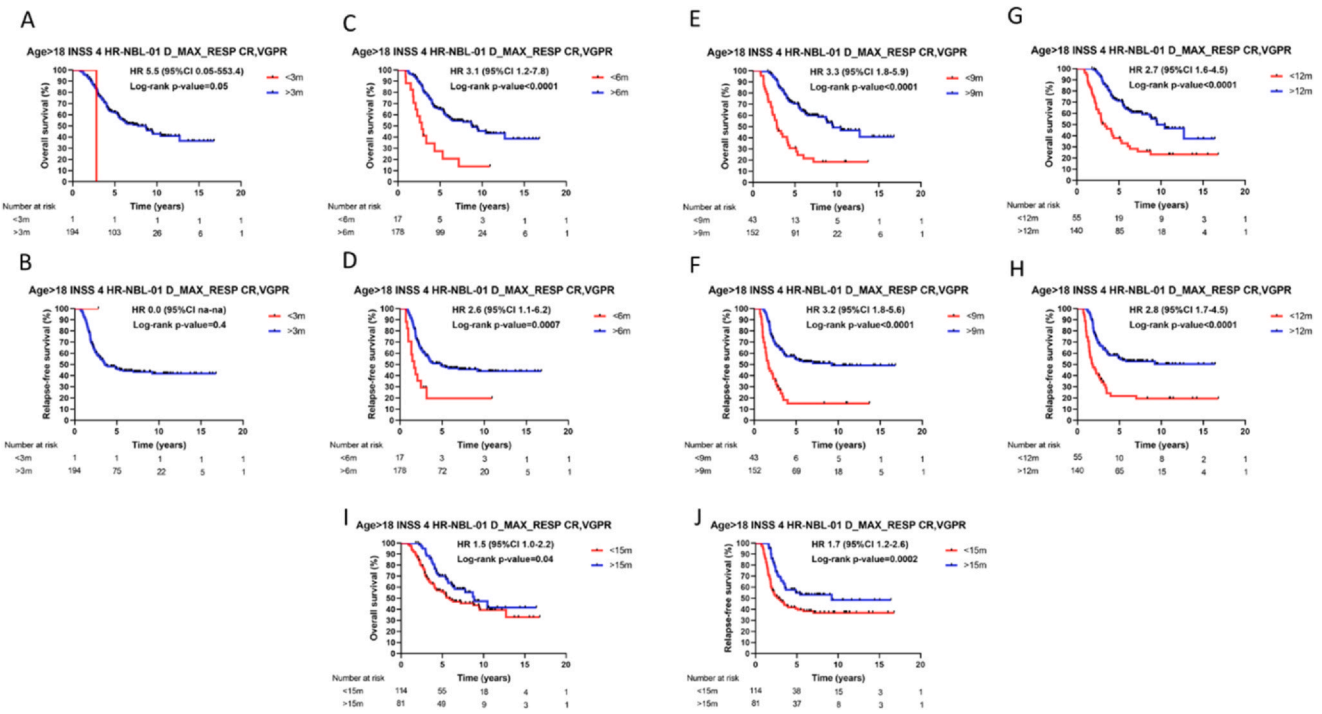

Fig. 5. Kaplan-Meier estimates for OS and RFS based on the TIME_MAX_RESP.days in the subset of high-risk patients treated with HR-NBL-01 protocol. Plots are relative to overall or RFS of patients older than 18 months with stage 4 tumour who achieved a CR or VGPR and treated with HR-NBL-01 protocol. Plots are entitled with the characteristics of the patients' subset. Cuts-off at 3 (A and B), 6 (C and D), 9 (E and F), 12 (G and H), and 15 (I and J) months were used to split patients into early or late response. Log-rank p-value and hazard ratio (HR) with 95% of confidence interval (95%CI) are reported on top of each plots. When a log-rank p-value was lower than 0.0001, we indicated it as log-rank-p-value < 0.0001. A table reporting the number of patients at risk is shown below each plot.

patients older than 18 months and stage 4 tumour who achieved a CR or VGPR. Cut-off values at 3, 6, 9, 12, and 15 months were evaluated to assess the prognostic value of TIME_MAX_RESP.days. The most significant stratification was achieved at 12 months for both OS and RFS (OS: HR 2.6 95%CI 2.0–3.2, Log-rank p-value $p < 0.0001$; RFS: 2.6 95%CI 2.1–3.2, Log-rank p-value $p < 0.0001$; Fig. S2A-S2J).

Analysis of the protocols administered to patients revealed a heterogeneous distribution, where HR-NBL-01, the most recent treatment protocol administered to manage high-risk patients in our dataset [45], was the most frequent (Fig. S3). Therefore, we decided to perform a new survival analysis using the data of 195 patients older than 18 months with INSS 4 tumour achieving a CR or VGPR treated with the HR-NBL-01 protocol and with defined TIME_MAX_RESP.days to confirm our findings in a homogenous subset of patients. A significant difference was found at 6, 9, and 12 months cut-offs for OS (Log-rank p-value < 0.005, Fig. 5C, E, G), and at 6, 9, 12, and 15 months cut-offs for RFS (Log-rank p-value < 0.005, Fig. 5D, F, H, J).

Interestingly, for 6, 9, and 12 months cut-offs the OS and RFS of patients who achieved an early CR or VGPR was significantly lower than that of patients who achieved the same types of response, but later, thus confirming our findings in the previous analysis (Log-rank p-value < 0.005; Fig. 5). Nine months was the cut-off that achieved the most significant stratification for both OS and RFS (OS: HR 3.3 95%CI 1.8–5.9, Log-rank p-value $p < 0.0001$; RFS: 3.2 95%CI 1.8–5.6, Log-rank p-value $p < 0.0001$; Fig. 5E-F).

HR-NBL-01 protocol treatment programme lasts approximately 14–16 months and includes induction chemotherapy, surgical resection of the primary tumour, high-dose chemotherapy with autologous stem cell rescue, radiation therapy, and finally immunotherapy plus differentiation therapy with isotretinoin [45]. During protocol administration patients are subjected at different time-points to evaluation of the status of the disease and eventually addressed to alternative treatments when a progression or relapse occurs. Patients achieving a satisfactory response after induction according to the SIOPEN criteria [36], proceed with the treatment programme. Patients who do not have a satisfactory response receive additional treatments [46]. Fig. S4 summarises the most common scenarios physicians might manage during periodical evaluations of patients' response to treatment.

Furthermore, we dissected the subset of patients treated with HR-NBL-01 protocol and assessed the differences between early or late responses in the subset of 171 patients, which get a remission after the end of the protocol and may be considered off-therapy. Again,

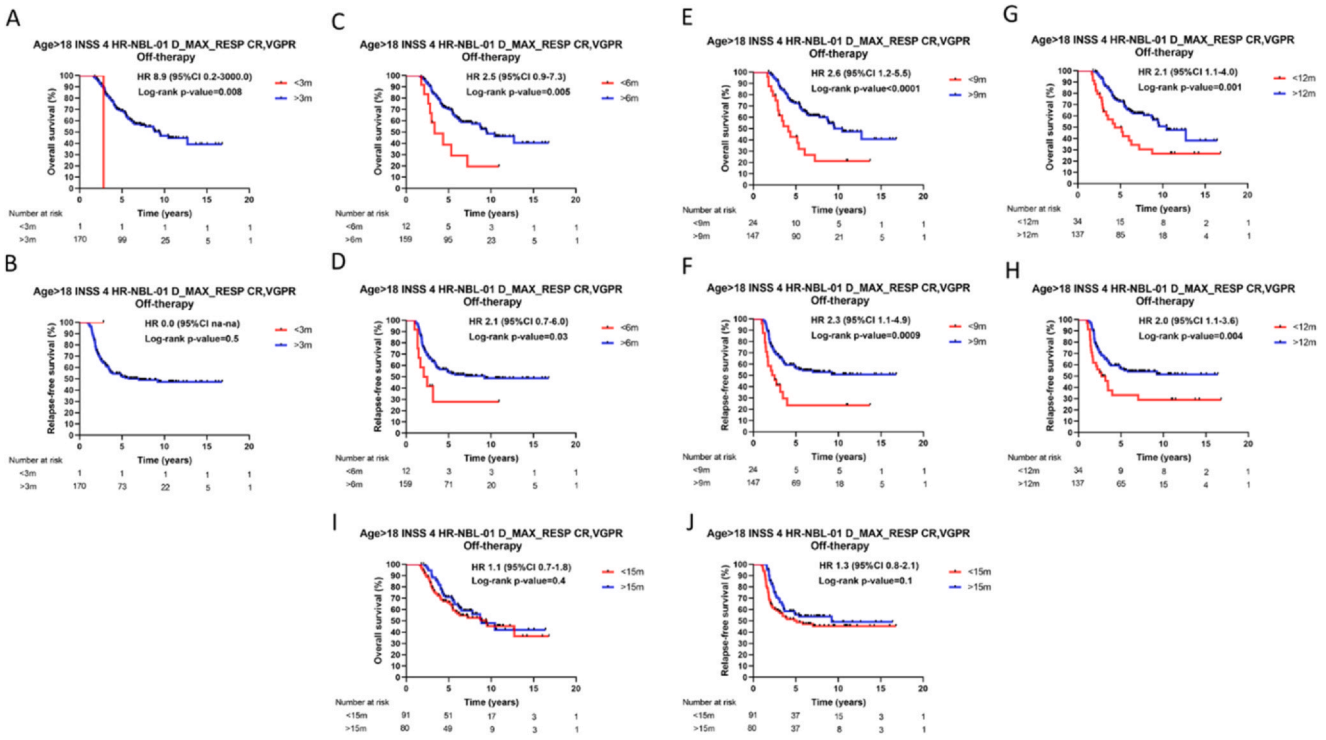

Fig. 6. Kaplan-Meier estimates for OS and RFS based on the TIME_MAX_RESP.days in the subset of high-risk off-therapy patients treated with HR-NBL-01 protocol. Plots are relative to OS or RFS of patients older than 18 months with stage 4 tumour who achieved a CR or VGPR, treated with HR-NBL-01 protocol and who completed the entire treatment protocol (Off-therapy). Plots are entitled with the characteristics of the patients in the sub-population and log-rank p-value. Cut-off of at 3 (A and B), 6 (C and D), 9 (E and F), 12 (G and H), and 15 (I and J) months were used to split patients into early or late responses. Log-rank p-value and HR with 95% of 95%CI are reported on top of each plots. When a log-rank p-value was lower than 0.0001, we indicated it as log-rank-p-value < 0.0001. A table reporting the number of patients at risk is shown below each plot.

as above described for the subset of patients with stage 4 tumour treated with the HR-NBL-01 protocol and achieving a CR or VGPR, a significant stratification of patients was found with 6, 9, and 12 months cut-offs for OS (Log-rank p-value < 0.005, Fig. 6C, E, G), but only for cut-offs of 9 and 12 months stratification was significant for RFS (Log-rank p-value < 0.005, Fig. 6F, H).

Interestingly, OS and RFS of patients with early CR or VGPR response were significantly lower than that of those who achieved a late CR or VGPR response (Log-rank p-value < 0.005; Fig. 6). Nine months was the cut-off that achieved the most significant stratification for both OS and RFS (OS: HR 2.6 95%CI 1.2–5.5, Log-rank p-value $p < 0.0001$; RFS: 2.3 95%CI 1.1–4.9, Log-rank p-value $p = 0.0009$; Fig. 6E-F).

These findings indicate that TIME_MAX_RESP.days could additionally stratify clinically relevant subsets of patients even if patients achieved a CR or VGPR to first-line treatment.

4. Discussion

In the present study, we provide robust evidence of the independent predictive and prognostic value of the time to D_MAX_RESP in addition to the maximum response for neuroblastoma OS and RFS. These findings were obtained from the analysis of 3756 records of patients diagnosed with neuroblastoma and registered in the RINB database, which includes both established features, such as age and non-established features, such as patients' symptoms at diagnosis. Published studies have integrated clinical and molecular markers into classification models able to predict the survival of neuroblastoma patients [47,48]. However, our dataset represents the largest compendium of features ever reported for neuroblastoma. Although the data registered in the RINB database has been analysed in part on other published reports [7,19,21,36,46–49], our study is the first exploiting the predictive power of machine learning techniques to the entire set of records stored in the RINB database. Machine learning analysis was based on patient status alive or dead with at least 3 years of follow-up. This procedure is in accordance with previously published studies [9,50,51].

Feature ranking by relevance was instrumental to show that D_MAX_RESP and TIME_MAX_RESP.days are among the top 4 relevant factors for the prediction of the outcome in the general or metastatic dataset and the relapse or progression in the general, metastatic, localised, or 4s dataset.

Survival analysis on high–risk patients older than 18 months and stage 4 tumour achieving a CR or VGPR pointed out that there exists a significantly lower OS and RFS for patients who obtain an early maximum response to therapy with respect to those with a later maximum response to therapy for either the good responders or the worse responders. Different cut-off values at 3, 6, 9, 12, and 15 months were evaluated, and all showed a significant stratification ability demonstrating that TIME_MAX_RESP.days is prognostic for high-risk patients. The prognostic value of the TIME_MAX_RESP.days has never been reported in the literature, thereby our findings show that it is a novel prognostic factor for neuroblastoma and provide the first indication of the clinical utility of this feature in neuroblastoma.

Clinical practice in Europe establishes that patients with high-risk neuroblastoma achieving CR/PR response in the first few months of treatment receive the treatment expected from the HR-NBL-01 protocol moving to consolidation phase with high-dose chemotherapy followed by maintenance with cis-retinoic acid and immunotherapy [7,36]. This treatment philosophy is pursued in the ongoing HR-NBL-02 clinical trial (EudraCT N°: 2019–001068–31, ClinicalTrials.gov Identifier: NCT04221035). Otherwise, patients who did not achieve an optimal response (CR \VGPR) received additional treatments [46]. To take this practice into account, we additionally refined our analysis focusing on the subset of patients older than 18 months with stage 4 tumour treated with HR-NBL-01 protocol who achieved a CR or VGPR. We demonstrate that the OS and RFS of patients who exhibit a CR or VGPR to treatment after 9 or 12 months from the starting date of therapy is significantly higher with respect to those who exhibit the same maximum response, but earlier than 9 or 12 months. Furthermore, we identify the 9 months as the cut-off that achieves the largest separation of both OS and RFS. Our findings are demonstrated in the subset of patients treated with HR-NBL-01 protocol and in the subset getting a remission after the end of the protocol. Both analyses were necessary to additionally investigate the clinical utility of TIME_MAX_RESP.days by analysing as homogenous as possible subsets of patients, thereby excluding the potential confounding effect of the treatment protocol or avoiding that stratification might be due to the occurrence of disease progression\relapse during treatment. A limitation of the analysis, albeit unavoidable, is the exact time at which maximum response is achieved. In fact, the transition from one state of the disease to another is a dynamic phenomenon throughout the treatment programme and patients are not under continuous evaluation. For this reason, the maximum response is linked to the date of examinations at different time points as scheduled during the treatment plan.

Our findings countertrend the clinical practice and this may be justified with the fact that patients with later responses receive more prolonged treatments than patients with a rapid response in the first months of induction therapy. These results could be related to biological characteristics of early responding patients, i.e. the presence of high proliferation rate that confer high response during the induction phase. In practice, the results of the present study encourage clinicians not to give up on the goal of definitive disease control after an induction therapy that has not led to a remission considered satisfactory, pending response signals to the induction therapy itself. Furthermore, it could be speculated that the prolongation of time of treatment would be beneficial for both patients who have a slow responding disease and those with a rapid disease remission. In fact, in the last 2 decades the prolongation of the treatment duration in high-risk neuroblastoma, with the addition of a maintenance phase with retinoic acid, and further addition of immunotherapy has been demonstrated to be as efficacious [49,52]. Further attempts to prolong the treatment duration, such as the addiction to the maintenance of prolonged treatment with Difluoromethylornithine could be beneficial and deserve evaluation, balancing on one hand the need to improve therapeutic results and on the other hand to preserve patients' quality of life [53].

Our study demonstrates the ability of computational intelligence applied to large compendium of data from neuroblastoma patients to discover new features with potential clinical impact on the management and treatment of patients.

The retrospective nature of RINB dataset limited our control over the data collection and the database entry procedures. RINB records have been collected over 30 years and manual curation of the data introduced a heterogeneous vocabulary and many missing values. For these reasons, many features included in the original RINB dataset were not sufficiently reliable such as features with a large number of missing values, or were not suitable to be analysed by machine learning techniques. Therefore, new features such as the time to max response to first-line treatment were created and included in the dataset and diverse data filtering and harmonisation steps were necessary to grant the validity of our statistical and computational analyses. A prospective study will be designed in the future to validate our findings in a more specific and controlled setting thus reducing any potential source of bias.

Funding

This work was supported by Fondazione Italiana per la Lotta al Neuroblastoma ONLUS and by the Italian Ministry of Health with "2023 Ricerca Corrente" funds - ID n. RRC-2023-23683432. This study was also

funded by the European Union – Next Generation EU programme, in the context of The National Recovery and Resilience Plan, Investment Partenariato Esteso PE8 "Conseguenze e sfide dell'invecchiamento'", Project Age-It (Ageing Well in an Ageing Society). This work was also partially supported by Ministero dell'Università e della Ricerca of Italy under the "Dipartimenti di Eccellenza 2023-2027" ReGAInS grant assigned to Dipartimento di Informatica Sistemistica e Comunicazione at Università di Milano-Bicocca.

Role of funding source

Funding sources were not involved in the study design, in the collection, analysis and interpretation of data, in the writing of the report; and in the decision to submit the article for publication.

CRediT authorship contribution statement

Davide Chicco: Formal analysis, Methodology, Software, Validation, Writing – review & editing. **Riccardo Haupt:** Data curation, Resources, Writing – review & editing. **Alberto Garaventa:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Paolo Uva:** Conceptualization, Investigation, Visualization, Writing – review & editing. **Roberto Luksch:** Supervision, Roles/Writing – original draft, Writing – review & editing. **Davide Cangelosi:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Roles/Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgements

We would like to thank Giovanni Erminio, for providing the Registro Italiano dei Tumori Neuroblastici Periferici (RINB) data file, detailed documentation and for helping with delighted discussions, Dr. Stefano Parodi, for his thoughtful suggestions on the statistical analysis, Barbara Galleni for helping with the RINB data, and Martina Fragola for helping with Association of Pediatric Hematology-Oncology (AIEOP) centres data. We thank AIEOP centres for their contribution of patients' data on the RINB database.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejca.2023.113291.

References

[1] Matthay KK, Maris JM, Schleiermacher G, et al. Neuroblastoma. Nat Rev Dis Prim 2016;2:1–21. https://doi.org/10.1038/nrdp.2016.78.
[2] Speleman F, Park JR, Henderson TO. Neuroblastoma: a tough nut to crack. Am Soc Clin Oncol Educ Book Am Soc Clin Oncol Annu Meet 2016;35:e548–57. https://doi.org/10.1200/EDBK_159169.
[3] Irwin MS, Naranjo A, Zhang FF, et al. Revised neuroblastoma risk classification system: a report from the children's oncology group. J Clin Oncol 2021;39:3229–41. https://doi.org/10.1200/JCO.21.00278.
[4] Tolbert VP, Matthay KK. Neuroblastoma: clinical and biological approach to risk stratification and treatment. Cell Tissue Res 2018;372:195–209. https://doi.org/10.1007/S00441-018-2821-2.
[5] Lerone M, Ognibene M, Pezzolo A, et al. Molecular genetics in neuroblastoma prognosis. Children 2021;8:456. https://doi.org/10.3390/CHILDREN8060456.
[6] Chung C, Boterberg T, Lucas J, et al. Neuroblastoma. Pediatr Blood Cancer 2021;68:e28473. https://doi.org/10.1002/PBC.28473.
[7] Garaventa A, Poetschger U, Valteau-Couanet D, et al. Randomized trial of two induction therapy regimens for high-risk neuroblastoma: HR-NBL1.5 International Society of Pediatric Oncology European Neuroblastoma Group Study. J Clin Oncol 2021;39:2552–63. https://doi.org/10.1200/JCO.20.03144.
[8] Qiu B, Matthay KK. Advancing therapy for neuroblastoma. 2022 198 Nat Rev Clin Oncol 2022;19:515–33. https://doi.org/10.1038/s41571-022-00643-z.
[9] Cangelosi D, Morini M, Zanardi N, et al. Hypoxia predicts poor prognosis in neuroblastoma patients and associates with biological mechanisms involved in telomerase activation and tumor microenvironment reprogramming. Cancers 2020;12:1–45. https://doi.org/10.3390/CANCERS12092343.
[10] Shawraba F, Hammoud H, Mrad Y, et al. Biomarkers in neuroblastoma: an insight into their potential diagnostic and prognostic utilities. Curr Treat Options Oncol 2021;22:102. https://doi.org/10.1007/S11864-021-00898-1.
[11] Liu Q, Wang Z, Jiang Y, et al. Single-cell landscape analysis reveals distinct regression trajectories and novel prognostic biomarkers in primary neuroblastoma. Genes Dis 2022;9:1624–38. https://doi.org/10.1016/J.GENDIS.2021.12.020.
[12] Grasso S, Cangelosi D, Chapelle J, et al. The SRCIN1/p140Cap adaptor protein negatively regulates the aggressiveness of neuroblastoma. Cell Death Differ 2020;27:790–807. https://doi.org/10.1038/S41418-019-0386-6.
[13] Garbati P, Barbieri R, Cangelosi D, et al. MCM2 and carbonic anhydrase 9 are novel potential targets for neuroblastoma pharmacological treatment. Biomedicines 2020;8:1–19. https://doi.org/10.3390/BIOMEDICINES8110471.
[14] Bettinsoli P, Ferrari-Toninelli G, Bonini SA, et al. Favorable prognostic role of tropomodulins in neuroblastoma. Oncotarget 2018;9:27092–103. https://doi.org/10.18632/ONCOTARGET.25491.
[15] Ognibene M, Pagnan G, Marimpietri D, et al. CHL1 gene acts as a tumor suppressor in human neuroblastoma. Oncotarget 2018;9:25903–21. https://doi.org/10.18632/ONCOTARGET.25403.
[16] Morini M, Cangelosi D, Segalerba D, et al. Exosomal microRNAs from longitudinal liquid biopsies for the prediction of response to induction chemotherapy in high-risk neuroblastoma patients: a proof of concept SIOPEN study. Cancers 2019;11:1476. https://doi.org/10.3390/CANCERS11101476.
[17] Dondero A, Morini M, Cangelosi D, et al. Multiparametric flow cytometry highlights B7-H3 as a novel diagnostic/therapeutic target in GD2neg/low neuroblastoma variants. J Immunother cancer 2021;9:e002293. https://doi.org/10.1136/JITC-2020-002293.
[18] Bernauer C, Man YKS, Chisholm JC, et al. Hypoxia and its therapeutic possibilities in paediatric cancers. Br J Cancer 2021;124:539–51. https://doi.org/10.1038/S41416-020-01107-W.
[19] Haupt R, Garaventa A, Gambini C, et al. Improved survival of children with neuroblastoma between 1979 and 2005: a report of the Italian Neuroblastoma Registry. J Clin Oncol 2010;28:2331–8. https://doi.org/10.1200/JCO.2009.24.8351.

[20] Luksch R, Castellani MR, Collini P, et al. Neuroblastoma (peripheral neuroblastic tumours). Crit Rev Oncol Hematol 2016;107:163–81. https://doi.org/10.1016/J.CRITREVONC.2016.10.001.
[21] De Bernardi B, Di Cataldo A, Garaventa A, et al. Stage 4 s neuroblastoma: features, management and outcome of 268 cases from the Italian Neuroblastoma Registry. Ital J Pediatr 2019;45:8https://doi.org/10.1186/S13052-018-0599-1.
[22] Avanzini S, Buffoni I, Gigliotti AR, et al. Resection of primary tumor in stage 4S neuroblastoma: a second study by the Italian Neuroblastoma Group. Pediatr Surg Int 2021;37:37–47. https://doi.org/10.1007/S00383-020-04766-1.
[23] Parodi S, Sorrentino S, Cataldo ADi, et al. Metastatic progression in infants diagnosed with stage 4S neuroblastoma. A study of the Italian Neuroblastoma Registry. Pediatr Blood Cancer 2021;68:e28904. https://doi.org/10.1002/pbc.28904.
[24] Chicco D, Oneto L. Data analytics and clinical feature ranking of medical records of patients with sepsis. BioData Min 2021;14:12. https://doi.org/10.1186/S13040-021-00235-0.
[25] De Bernardi B, Carli M, Casale F, et al. Standard-dose and high-dose peptichemio and cisplatin in children with disseminated poor-risk neuroblastoma: two studies by the Italian Cooperative Group for Neuroblastoma. J Clin Oncol 1992;10:1870–8. https://doi.org/10.1200/JCO.1992.10.12.1870.
[26] Brodeur GM, Seeger RC, Barrett A, et al. International criteria for diagnosis, staging, and response to treatment in patients with neuroblastoma. J Clin Oncol 1988;6:1874–81. https://doi.org/10.1200/JCO.1988.6.12.1874.
[27] Brodeur GM, Pritchard J, Berthold F, et al. Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. J Clin Oncol 1993;11:1466–77. https://doi.org/10.1200/JCO.1993.11.8.1466.
[28] De Bernardi B, Pianca C, Boni L, et al. Disseminated neuroblastoma (stage IV and IV-S) in the first year of life outcome related to age and stage. doi:⟨10.1002/1097–0142⟩.
[29] De Bernardi B, Nicolas B, Boni L, et al. Disseminated neuroblastoma in children older than one year at diagnosis: comparable results with three consecutive high-dose protocols adopted by the Italian Co-Operative Group for Neuroblastoma. J Clin Oncol 2003;21:1592–601. https://doi.org/10.1200/JCO.2003.05.191.
[30] De Bernardi B, Conte M, Mancini A, et al. Localized resectable neuroblastoma: results of the second study of the Italian Cooperative Group for Neuroblastoma. J Clin Oncol 1995;13:884–93. https://doi.org/10.1200/JCO.1995.13.4.884.
[31] Garaventa A, De Bernardi B, Pianca C, et al. Localized but unresectable neuroblastoma: treatment and outcome of 145 cases. Italian Cooperative Group for Neuroblastoma. J Clin Oncol 1993;11:1770–9. https://doi.org/10.1200/JCO.1993.11.9.1770.
[32] Garaventa A, Boni L, Lo Piccolo MS, et al. Localized unresectable neuroblastoma: results of treatment based on clinical prognostic factors. Ann Oncol Off J Eur Soc Med Oncol 2002;13:956–64. https://doi.org/10.1093/ANNONC/MDF165.
[33] De Bernardi B, Mosseri V, Rubie H, et al. Treatment of localised resectable neuroblastoma. Results of the LNESG1 study by the SIOP Europe Neuroblastoma Group. Br J Cancer 2008;99:1027. https://doi.org/10.1038/SJ.BJC.6604640.
[34] Canete A, Gerrard M, Rubie H, et al. Poor survival for infants with MYCN-amplified metastatic neuroblastoma despite intensified treatment: the International Society of Paediatric Oncology European Neuroblastoma Experience. J Clin Oncol 2009;27:1014–9. https://doi.org/10.1200/JCO.2007.14.5839.
[35] De Bernardi B, Gerrard M, Boni L, et al. Excellent outcome with reduced treatment for infants with disseminated neuroblastoma without MYCN gene amplification. J Clin Oncol 2009;27:1034–40. https://doi.org/10.1200/JCO.2008.17.5877.
[36] Ladenstein R, Pötschger U, Valteau-Couanet D, et al. Interleukin 2 with anti-GD2 antibody ch14.18/CHO (dinutuximab beta) in patients with high-risk neuroblastoma (HR-NBL1/SIOPEN): a multicentre, randomised, phase 3 trial. Lancet Oncol 2018;19:1617–29. https://doi.org/10.1016/S1470-2045(18)30578-3.
[37] Breiman L. Random Forests. 2001 451 Mach Learn 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.
[38] Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim Biophys Acta Protein Struct 1975;405:442–51. https://doi.org/10.1016/0005-2795(75)90109-9.
[39] Chicco D, Jurman G. The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. BioData Min 2023;16:1–23. https://doi.org/10.1186/S13040-023-00322-4/FIGURES/11.
[40] Chicco D, Warrens MJ, Jurman G. The Matthews correlation coefficient (MCC) is more informative than Cohen's kappa and brier score in binary classification assessment. IEEE Access 2021;9:78368–81. https://doi.org/10.1109/ACCESS.2021.3084050.
[41] Chicco D, Starovoitov V, Jurman G. The benefits of the Matthews correlation coefficient (MCC) over the diagnostic odds ratio (DOR) in binary classification assessment. IEEE Access 2021;9:47112–24. https://doi.org/10.1109/ACCESS.2021.3068614.
[42] Chicco D, Tötsch N, Jurman G. The Matthews correlation coefficient (Mcc) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. BioData Min 2021;14:1–22. https://doi.org/10.1186/S13040-021-00244-Z/TABLES/5.
[43] Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. 2017 21 Nat Hum Behav 2017;2:6–10. https://doi.org/10.1038/s41562-017-0189-z.
[44] Fan C, Chen M, Wang X, et al. A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. Front Energy Res 2021;9:652801. https://doi.org/10.3389/FENRG.2021.652801/BIBTEX.
[45] Smith V, Foster J. High-risk neuroblastoma treatment review. Children 2018;5:114. https://doi.org/10.3390/CHILDREN5090114.
[46] Amoroso L, Erminio G, Makin G, et al. Topotecan-vincristine-doxorubicin in stage 4 high-risk neuroblastoma patients failing to achieve a complete metastatic response to rapid COJEC: a SIOPEN study. Cancer Res Treat 2018;50:148. https://doi.org/10.4143/CRT.2016.511.
[47] Moreno L, Guo D, Irwin MS, et al. A nomogram of clinical and biologic factors to predict survival in children newly diagnosed with high-risk neuroblastoma: an International Neuroblastoma Risk Group project. Pediatr Blood Cancer 2021;68:e28794. https://doi.org/10.1002/PBC.28794.
[48] Cohn SL, Pearson ADJ, London WB, et al. The International Neuroblastoma Risk Group (INRG) classification system: an INRG task force report. J Clin Oncol 2009;27:289. https://doi.org/10.1200/JCO.2008.16.6785.
[49] Ladenstein R, Pötschgerulrike.poetschger@ccri.at U, Valteau-couanet D, et al. Investigation of the role of dinutuximab beta-based immunotherapy in the SIOPEN high-risk neuroblastoma 1 trial (HR-NBL1). Cancers 2020;12:309. https://doi.org/10.3390/CANCERS12020309.
[50] Wei JS, Greer BT, Westermann F, et al. Prediction of clinical outcome using gene expression profiling and artificial neural networks for patients with neuroblastoma. Cancer Res 2004;64:6883–91. https://doi.org/10.1158/0008-5472.CAN-04-0695.
[51] Schramm A, Schulte JH, Klein-Hitpass L, et al. Prediction of clinical outcome and biological characterization of neuroblastoma by expression profiling. 2005 2453 Oncogene 2005;24:7902–12. https://doi.org/10.1038/sj.onc.1208936.
[52] Yu C, Liu Y, Miao Z, et al. Retinoic acid enhances the generation of hematopoietic progenitors from human embryonic stem cell-derived hemato-vascular precursors. Blood 2010;116:4786–94. https://doi.org/10.1182/BLOOD-2010-01-263335.
[53] Sholler GLS, Ferguson W, Bergendahl G, et al. Maintenance DFMO increases survival in high risk neuroblastoma. Sci Rep 2018;8:14445. https://doi.org/10.1038/S41598-018-32659-W.