Published in final edited form as:

Eur J Nucl Med Mol Imaging. 2018 February ; 45(2): 292–305. doi:10.1007/s00259-017-3829-7.

Validation of the mIBG skeletal SIOPEN scoring method in two independent high-risk neuroblastoma populations: the SIOPEN/HR-NBL1 and COG-A3973 trials

Ruth Ladenstein^{1,2}, Bieke Lambert³, Ulrike Pötschger¹, Maria-Rita Castellani⁴, Valerie Lewington⁵, Zvi Bar-Sever⁶, Aurore Oudoux⁷, Anna Iliw ska⁸, Katerina Taborska⁹, Lorenzo Biassoni¹⁰, Gregory A. Yanik¹¹, Arlene Naranjo¹², Marguerite T. Parisi¹³, Barry L. Shulkin¹⁴, Helen Nadel¹⁵, Michael J. Gelfand¹⁶, Katherine K. Matthay¹⁷, Julie R. Park¹⁸, Susan G. Kreissman¹⁹, Dominique Valteau-Couanet²⁰, and Ariane Boubaker²¹

¹Children's Cancer Research Institute, Zimmermannplatz 10, 1090 Vienna, Austria ²Medical University, Department of Paediatrics, Vienna, Austria ³Radiology and Nuclear Medicine, Ghent University, Ghent, Belgium ⁴Nuclear Medicine Division, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy ⁵Guy's and St Thomas' NHS Foundation Trust, London, UK ⁶Schneider Children's Medical Center of Israel, Petah-Tikva, Israel ⁷Department of Nuclear Medicine Lille,

Correspondence to: Ruth Ladenstein.

Ruth Ladenstein and Bieke Lambert share first authorship

Part of data previously presented: ASCO Annual Meeting 2014, Chicago, USA, May 30–June 3; J Clin Oncol 32:5 s, 2014 (suppl; abstr 10,029); EANM Annual Meeting Congress 2014, Oct 20–23, Gothenburg, Sweden

Authors' contributions: Ruth Ladenstein and Bieke Lambert share first authorship.

Conception and design.

Ruth Ladenstein, Ulrike Pötschger, Valerie Lewington, Gregory Yanik, Katherine K. Matthay, Julie Park, Susan G. Kreissman, Ariane Boubaker.

Provision of study material or patients.

Ruth Ladenstein, Bieke Lambert, Valerie Lewington, Zvi Bar-Sever, Aurore Oudoux, Anna (X0015A)Iliw ska, Katerina Taborska, Lorenzo Biassoni, Gregory Yanik, Arlene Naranjo, Marguerite T. Parisi, Barry L. Shulkin, Helene Nadel, Michael J. Gelfand, Julie Park, Susan G. Kreissman, Dominique Valteau-Couanet, Ariane Boubaker.

Collection and assembly of data.

Ruth Ladenstein, Bieke Lambert, Maria-Rita Castellani, Zvi Bar-Sever, Aurore Oudoux, Anna (X0015A)Iliw ska, Katerina Taborska, Lorenzo Biassoni, Gregory Yanik, Arlene Naranjo, Marguerite T. Parisi, Barry L. Shulkin, Helene Nadel, Michael J. Gelfand, Julie Park, Susan G. Kreissman, Dominique Valteau-Couanet, Ariane Boubaker.

Data analysis and interpretation.

Ruth Ladenstein, Bieke Lambert, Ulrike Pötschger, Maria-Rita Castellani.

mIBG reviewers.

Bieke Lambert, Maria-Rita Castellani, Zvi Bar-Sever, Aurore Oudoux, Anna (X0015A)Iliw ska, Katerina Taborska, Lorenzo Biassoni, Ariane Boubaker.

Manuscript writing.

Ruth Ladenstein, Bieke Lambert, Ulrike Pötschger, Gregory Yanik, Ariane Boubaker.

Final approval of manuscript.

Ruth Ladenstein, Bieke Lambert, Ulrike Pötschger, Valerie Lewington, Maria-Rita Castellani, Zvi Bar-Sever, Aurore Oudoux, Anna Iliw ska, Katerina Taborska, Lorenzo Biassoni, Gregory Yanik, Arlene Naranjo, Marguerite T. Parisi, Barry L. Shulkin, Helene Nadel, Michael J. Gelfand, Katherine K. Matthay, Julie Park,

Compliance with ethical standards: Disclosure of potential conflict of interest: All authors declare that they have no conflict of interest.

Ethical approval: All procedures performed in the analysis of the studies HR-NBL1/SIOPEN and COG-A3973 involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Oscar Lambret Center, Lille, France ⁸Children's Memorial Health Institute, Warsaw, Poland ⁹Motol University Hospital, Prague, Czech Republic ¹⁰Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK ¹¹University of Michigan Medical Center, Ann Arbor, MI, USA ¹²Children's Oncology Group Statistics and Data Center, University of Florida, Gainesville, FL, USA ¹³Department of Radiology, University of Washington School of Medicine/ Seattle Children's Hospital, Seattle, WA, USA ¹⁴Department of Diagnostic Imaging, St. Jude Children's Research Hospital, Memphis, TN, USA ¹⁵Department of Radiology, BC Children's Hospital, Vancouver, BC, Canada ¹⁶Department of Radiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA ¹⁷Department of Pediatrics, University of California San Francisco School of Medicine, San Francisco, CA, USA ¹⁸Department of Pediatrics, University of Washington School of Medicine/ Seattle Children's Hospital, Seattle, WA, USA ¹⁹Department of Pediatrics, Duke University Medical Center, Durham, NC, USA ²⁰Pediatric and Adolescent Oncology, Gustave Roussy Institute, Université Paris-Sud, Villejuif, France ²¹Institute of Radiology, Clinique de La Source, Lausanne, Switzerland

Abstract

Background—Validation of the prognostic value of the SIOOPEN mIBG skeletal scoring system in two independent stage 4, mIBG avid, high-risk neuroblastoma populations.

Results—The semi-quantitative SIOOPEN score evaluates skeletal meta-iodobenzylguanidine (mIBG) uptake on a 0–6 scale in 12 anatomical regions. Evaluable mIBG scans from 216 COG-A3973 and 341 SIOOPEN/HR-NBL1 trial patients were reviewed pre- and post-induction chemotherapy. The prognostic value of skeletal scores for 5-year event free survival (5 yr.-EFS) was tested in the source and validation cohorts. At diagnosis, both cohorts showed a gradual non-linear increase in risk with cumulative scores. Several approaches were explored to test the relationship between score and EFS. Ultimately, a cutoff score of 3 was the most useful predictor across trials. A SIOOPEN score 3 pre-induction was found in 15% SIOOPEN patients and in 22% of COG patients and increased post-induction to 60% in SIOOPEN patients and to 73% in COG patients. Baseline 5 yr.-EFS rates in the SIOOPEN/HR-NBL1 cohort for scores 3 were $47\% \pm 7\%$ versus $26\% \pm 3\%$ for higher scores at diagnosis ($p < 0.007$) and $36\% \pm 4\%$ versus $14\% \pm 4\%$ ($p < 0.001$) for scores obtained post-induction. The COG-A3973 showed 5 yr.-EFS rates for scores 3 of $51\% \pm 7\%$ versus $34\% \pm 4\%$ for higher scores ($p < 0.001$) at diagnosis and $43\% \pm 5\%$ versus $16\% \pm 6\%$ ($p = 0.004$) for post-induction scores. Hazard ratios (HR) significantly favoured patients with scores 3 after adjustment for age and MYCN-amplification. Optimal outcomes were recorded in patients who achieved complete skeletal response.

Conclusions—Validation in two independent cohorts confirms the prognostic value of the SIOOPEN skeletal score. In particular, patients with an absolute SIOOPEN score > 3 after induction have very poor outcomes and should be considered for alternative therapeutic strategies.

Keywords

High-risk neuroblastoma; mIBG; SIOOPEN score

Background

Neuroblastoma is the most common extra-cranial solid tumour in children. High-risk neuroblastoma (HR-NBL), defined by metastatic disease above the ages of 12–18 months [1, 2], amplification of the proto-oncogene MYCN (MYCNA) at any age, and unfavourable histopathologic features [3, 4] are associated with long-term survival rates of 40% [5–7]. Internationally agreed treatment options for HR-NBL include induction multi-agent chemotherapy, surgery, high dose chemotherapy (HDT) followed by autologous stem cell transplantation (ASCT), external beam radiotherapy, radionuclide therapy, differentiation therapies, and immunotherapy.

Accurate disease staging is essential to optimise treatment. Meta-iodobenzylguanidine (mIBG) is a norepinephrine analogue taken up by neuroblastoma cells. Iodine-123 (^{123}I) or Iodine-131 (^{131}I) labelled mIBG scintigraphy is recommended to assess disease extent in accordance with International Neuroblastoma Risk Group (INRG) guidance [8–10].

The International Society of Paediatric Oncology European Neuroblastoma Group (SIOPEN) is a network of 183 paediatric oncology treatment centres in 22 countries [11, 12]. The SIOPEN semi-quantitative mIBG skeletal scoring system has previously been shown to have excellent inter- and intra-observer reproducibility and to provide a reliable prognostic indicator in neuroblastoma [13]. The present study aimed to validate the SIOPEN scoring system as a prognostic response predictor in two independent trial populations.

Methods

Newly diagnosed stage 4 HR-NBL patients with mIBG-avid disease enrolled on the SIOPEN/HR-NBL1 [[ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT00030719)(NCT00030719) and EudraCT (2006–001489-17)] and the Children's Oncology Group (COG)-A3973 trial [[ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT00004181) NCT00004181] were the two populations used for retrospective validation of the SIOPEN semi-quantitative mIBG skeletal scoring system. Planar images acquired 24 h after ^{123}I -mIBG or ^{131}I -mIBG administration were collected and scored at diagnosis and at the end of induction chemotherapy. SPECT images were not used for scoring. Briefly, the SIOPEN method scores skeletal disease extent within 12 body segments on a 0 to 6 scale (Fig. 1) giving a maximum score of 72 [13]. The SIOPEN methodology showed a 95% concordance in a blinded review by six nuclear medicine physicians [13]. The mIBG scans from SIOPEN/HR-NBL1 and COG A3973 were reviewed independently by a team of eight SIOPEN Nuclear Medicine investigators, scoring in pairs by consensus the pre and post-induction scans of each patient from both patient cohorts. Poor quality (i.e. often ^{131}I -mIBG scans) and incomplete data sets (i.e. missing lateral skull views) were excluded. No clinical information was provided other than the primary tumour site. Soft tissues (ST) other than the primary tumour were scored in each body segment assigning the following scores: zero for no abnormality, one for a solitary, two for 2 and three for >2 lesions.

Patient flow and characteristics are shown in Annex 1 and Table 1. MIBG image data sets were judged to be of poor quality and/or incomplete in 72/413 (17.4%) patients of the SIOPEN/HRNBL1 trial and in 64/280 (22.9%) patients of the COG A3973 trial (Annex 1).

Both trial designs are summarised in Fig. 2 and are fully described elsewhere [11, 12, 14, 15]. I-131-mIBG therapy was not part of treatment in either of the two protocols. Inclusion criteria differed slightly between trials primarily in the younger age groups. The SIOPEN/HR-NBL1 high-risk criteria included eligible patients as those with INSS stage 2 to 4, <12 months with MYCNA, or metastatic neuroblastoma >12 months, all up to an age < 21 years at diagnosis. COG-A3973 required at least one unfavourable biological feature (MYCNA, unfavourable pathology, and/or DNA index = 1) for stage 4 patients 12–18 months and MYCNA for stage 4 < 12 months. Both cooperative groups used intensive induction regimens [11, 14, 15] aimed for gross resection of the primary tumour (at different time points), applied HDT/ASCT local radiotherapy and subsequent maintenance therapy with isotretinoin. SIOPEN/HR-NBL1 patients underwent HDT/ASCT if in complete or partial response (CR/PR) with 3 mIBG avid sites and >50% reduction of mIBG avid osseous lesions while COG-A3973 patients in CR/PR post-induction were eligible to proceed to HDT/ASCT regardless of the number of residual mIBG-avid sites. Immunotherapy with anti-GD2 monoclonal antibody was only given to a minority of patients (Table 1) [16, 17].

Both trials were approved by their respective institutional review boards or national ethical lead committees, and written informed consent was obtained from patients (or legal guardians) prior to study entry according to national and institutional guidelines.

Statistical analysis

For event-free survival (EFS), time to event was defined as the time from the pre-induction mIBG scan until the time to relapse, progressive disease, secondary malignancy, or death. Patients alive without event were censored at the time they were last known alive. The time to event post-induction was calculated from the post-induction mIBG scan date. Only patients with positive skeletal baseline scans were included in the post-induction analysis. Patients with an event between the date of diagnosis and the date of post-induction mIBG scan were excluded from the post-induction analysis.

Survival analyses were performed using the methods of Kaplan and Meier [18], with standard errors per the methods of Greenwood [19]. EFS is presented as the rate $\pm$ standard error. The prognostic value of mIBG score upon survival was assessed using a log-rank test Cox proportional hazards (PH) model [19]. The latter was also used to adjust for MYCNA and age, where age was included as a linear predictor in the model.

Several approaches were employed to explore the relationship between score and EFS. First, restricted cubic splines [20] were used to estimate the non-linear relationship between skeletal score and hazard ratio (HR). Second, time-dependent receiver operating characteristic (ROC) curves for censored survival data as developed by Heagerty [21, 22] were used to investigate the prognostic value of the scoring system. The area under the ROC curve (AUC) was used to characterise the ability of the scoring system to distinguish between patients with and without event at 5 years. Third, an analysis was performed, using tertiles of the total SIOPEN score. Fourth, the prognostic value of the threshold score of 3 [12] was tested. Fifth, a grid search to identify an optimal cut point was conducted aligned with the Youden-Index approach previously used by the COG group [23, 24]. The Youden-

Index was based on estimates for sensitivity and specificity of the time-dependent ROC-method [22].

Finally, the percentage change from diagnostic to post-induction scans was calculated.

Analyses were performed in SAS9.3 (SAS Institute, Cary, NC, USA) and R [25]. *p*-values less than 0.05 were considered statistically significant.

Results

Cohort characteristics (Table 1)

In SIOPEN/HR-NBL1, 341 patients with mIBG-avid primary tumours at baseline were available for review (Table 1). MIBG scans from 307 patients were evaluable for review post-induction, of whom 68% underwent HDT/ASCT on-study, 17% received altered HDT/ASCT (phase II) off-protocol and 15% who did not undergo HDT/ASCT. For post-induction analysis, 28 patients without skeletal uptake at diagnosis were excluded. The COG-A3973 dataset comprised 216 baseline mIBG scans of stage 4 patients with mIBG-avid primary tumours, with 140 mIBG scans evaluable post-induction, of whom 89% underwent HDT/ASCT with CEM (Fig. 2).

The SIOPEN cohort had a higher proportion of patients age > 5 years, but a significantly lower number of patients with mIBG-negative skeletal uptake at baseline. The distribution of MYCNA and patients 18 months was comparable in the trials. In the analysis data set, only 20 patients in SIOPEN/HR-NBL1 and 24 in COG-A3973 had received immunotherapy [16, 17].

Five-year EFS from diagnosis for the two cohorts was 29% (SD $\pm$ 3%) for SIOPEN/HR-NBL1 patients and 38% (SD $\pm$ 3%) for COG-A3973 ($p = 0.165$). There was no difference between patients included in the current analysis, compared with the rest of the population treated on SIOPEN/HR-NBL1. For COG-A3973 no difference was found between all study participants compared with the analysed cohort (Annexs 1 and 2).

SIOPEN score validation analysis

SIOPEN skeletal mIBG score at diagnosis and 5 yr.-EFS (Fig. 5)—Median skeletal metastatic scores at diagnosis were identical in the COG and SIOPEN cohorts (Table 1; Fig. 3a). ROC analysis demonstrated a higher probability of events in patients with rising baseline mIBG scores in both cohorts: AUC 0.602 and 0.606 for COG and SIOPEN, respectively. The risk of an event compared with baseline score is shown in Fig. 3b. Both cohorts exhibited a nonlinear increase in risk of events with rising SIOPEN score. Flattening of the curve with higher scores was observed.

Grouping according to tertiles of the maximum score at diagnosis (Fig. 3c) found superior EFS in the lowest score group in both cohorts. An intermediate risk category was identified in the COG cohort, but not in SIOPEN patients.

Using the Youden-Index [23, 24], different SIOPEN score cut points were found with the two cohorts at diagnosis; 11 for SIOPEN/HR-NBL1 and 39 for COG-A3973 patients.

Applying the threshold of ≥ 3 mIBG-avid skeletal spots resulted in significant EFS differences in both trials (Fig. 3d). In the SIOPEN cohort, the 5 yr.-EFS was 47% ($\pm 7\%$) for children with baseline scores of ≥ 3 versus 26% ($\pm 3\%$) for higher scores ($p = 0.007$). The corresponding results in the COG-A3973 cohort were 5 yr.-EFS rates of 51% ($\pm 7\%$) and 34% ($\pm 4\%$), respectively ($p < 0.001$). After adjustment for age and MYCNA, the hazard ratios (HR) showed a higher risk for patients with scores of ≥ 3 ; specifically 1.74 fold for SIOPEN/HR-NBL1 (95% CI: 1.12–2.70; $p = 0.013$) and 1.73 for COG-A3973 (95% CI 1.05–2.86; $p = 0.032$).

Patients presenting with mIBG-avid disease at diagnosis, but without skeletal involvement (Table 1), i.e. a pre-induction score of zero, had a favourable prognosis with 5 yr.-EFS of 46% ($\pm 10\%$) for SIOPEN/HR-NBL1 and 52% ($\pm 8\%$) for COG-A3973.

SIOPEN skeletal mIBG score post- induction chemotherapy and 5 yr.-EFS

(Fig. 4)—Of 307 SIOPEN-patients evaluable for post-induction scoring, 283 had pre-induction scores ≥ 3 , which improved to a post-induction score of ≥ 3 in 57% (163/283 patients). Of 140 COG patients evaluable for post induction scoring, 133 had pre-induction scores ≥ 3 , which improved to a post-induction score of ≥ 3 in 71% (95/133 patients).

Figure 4a illustrates the efficacy of induction treatment on skeletal tumour burden in both cohorts. A significantly higher proportion of COG patients (56%) achieved complete skeletal remission compared with 37% of SIOPEN patients (Table 1) ($p = 0.001$).

The predictive value of the post-induction SIOPEN score in the ROC analysis was higher in both groups, with a higher AUC of 0.628 and 0.660 for the SIOPEN/HR-NBL1 and COG-A3973 datasets, respectively. Higher post-induction scores correlated with poorer outcome in both cohorts, but only very low post-induction score values predicted a favourable outcome. Again, a nonlinear relationship between score values and 5 yr.-EFS was observed (Fig. 4b).

Analysis based upon tertiles of the maximum score of 72 showed a very limited sample size at the highest tertile, but a significant difference between the other two tertiles in both trial populations (Fig. 4c).

The Youden-Index showed an optimal cut-point = 3 for SIOPEN/HR-NBL1, whereas the optimal cut-point was 0 for the COG-A3973 cohort.

The predefined score threshold of 3 was reliable in both cohorts: 5 yr.-EFS was 36% ($\pm 4\%$) for scores of ≥ 3 versus 14% ($\pm 4\%$) for higher scores ($p < 0.001$) in the SIOPEN/HR-NBL1 cohort and was 43% ($\pm 5\%$) vs. 16% ($\pm 6\%$) in the COG-A3973 cohort ($p = 0.004$) (Fig. 4d). All differences were maintained after adjustment for age and MYCNA. Patients with scores of ≥ 3 fared significantly worse as demonstrated by a HR of 1.53 (95% CI: 1.12–2.13; $p = 0.008$) for SIOPEN/HR-NBL1 and a HR of 1.92 (95% CI 1.09–3.44; $p = 0.023$) in the COG-A3973 population.

A significantly better outcome was found in both trial cohorts for patients with mIBG skeletal positivity at diagnosis, but subsequent complete clearing of skeletal lesions (post-

induction score 0) (Fig. 5). However, patients without complete clearing of skeletal disease post-induction showed no significant difference in EFS, based upon the percent change in SIOPEN score from diagnosis to post-induction scans for patients with both sets of mIBG scan readings.

SIOPEN soft tissue scores: The incidence of ST-lesions other than the primary tumour pre- and post-induction was low in both cohorts, and hence results were summarised as either negative or positive (Table 1) with pre-induction rates of <20% and post-induction incidence <10%. ST-scores showed a high correlation with skeletal scores ($p < 0.001$) in both trial cohorts. No significant (NS) influence of ST on outcome was detectable after adjustment for ST. In the SIOPEN group 5 yr.-EFS rates pre- and post-induction were for ST negative 29% ($\pm 3\%$) and 29% ($\pm 3\%$)(NS) and for ST positive ones 30% ($\pm 6\%$) and 15% ($\pm 7\%$)(NS). In the COG-A3973 5 yr.-EFS rates pre- and post-induction were for ST negative 39% ($\pm 4\%$) and 37% ($\pm 4\%$)(NS) and for ST positive 34% ($\pm 8\%$) and 12% ($\pm 12\%$)(NS).

Discussion

Two semiquantitative mIBG scoring systems have been developed to evaluate tumour burden in high-risk neuroblastoma, the SIOPEN and Curie scoring systems. An International Neuroblastoma Risk Group (INRG) task force led by Matthay et al. [10] has examined both methodologies as a potential prognostic marker for outcome determination. The SIOPEN scoring methodology is currently being used in SIOPEN high-risk neuroblastoma trials, with Curie scoring used in COG trials. We have shown here that the SIOPEN mIBG score is highly prognostic of outcome in two independent data sets, SIOPEN/HR-NBL1, and COG-A3973.

In a prior COG review of COG-A3973, COG investigators [24] were unable to identify a mIBG (Curie) score at diagnosis that correlated with outcome. In contrast, using diagnostic scans from the same data set (COG-A3973) we have now shown that a SIOPEN score ≤ 3 is associated with improved 5-yr. EFS, when compared SIOPEN scores >3 at diagnosis, with similar results found in the SIOPEN/HR-NBL1 cohort. Based on only a small population of 58 patients, a SIOPEN cut point of 4 at diagnosis was previously proposed by the Cologne Interscore Comparison Study [26]. Importantly, our results presented here are based on more than 500 intercontinental patients. One may question if the HR-NBL1/SIOPEN eligibility criteria for HDT/ASCT have influenced the cut of value of 3. However, the independent data set of the COG-A3973 trial allowing responding patients to proceed to HDT/ASCT regardless of the number of residual mIBG-avid sites confirmed the SIOPEN methodology. In the small, better prognosis subgroup with a SIOPEN score ≤ 3 at diagnosis, the early favourable signal for outcome may prevent choosing additional treatment intensification where severe toxicities may outweigh potential benefits. However, a post-induction SIOPEN score > 3 identifies an ultra-high risk group of patients in need of innovative and/or intensified treatment strategies.

Yanik et al. [24] previously reported the prognostic significance of post-induction Curie scores in high risk neuroblastoma in the COG-A3973 data, with 3-yr. EFS of 44.9% vs. 15.4% for patients with Curie scores ≤ 2 vs. >2 post-induction. This result has now been

validated on the HR-NBL1/SIOPEN trial population [27]. For comparison, in the COG-A3973 cohort, a SIOPEN score > 3 post-induction was associated with inferior survival (5-yr. EFS 16%) when compared with a SIOPEN score ≤ 3 post-induction (5-yr. EFS 43%). In SIOPEN/HR-NBL-1 inferior survival was likewise associated with post-induction SIOPEN scores > 3 . Due to differences in scoring methodology between the two systems, a Curie score of 2 cannot be directly translated to a SIOPEN score of 3, but both reflect limited metastatic disease burden. Clearly, patients with a complete clearance of their skeletal disease to a SIOPEN score of zero had the best outcomes in both trial populations, similar to the prior study of Curie score in A3973. However, in patients with incomplete clearance of skeletal disease, no prognostic subgroups based on the percentages of response were detectable. Apparently depth of response is not a good discriminant of outcome prediction, as patients with extensive initial involvement may individually have a response $> 75\%$, but a SIOPEN score above 3. We found that the absolute score was a better predictor of outcome than the relative score [10].

Improved outcomes were previously reported for patients achieving a complete skeletal response following induction chemotherapy [24, 26, 28]. However, it is the cut-off score of > 3 post-induction that clearly discriminates an actionable decision point for altered innovative strategies in a very poor prognosis group. The higher rate of complete skeletal disease remission in the COG A3973 trial suggests that the COG A3973 induction regimen is more effective in clearing skeletal disease than the SIOPEN induction regimen, either due to the inclusion of anthracyclines in COG A3973 or related to the longer duration of induction therapy [6, 11, 12, 14, 29]. SIOPEN currently is addressing the induction question by randomising Rapid COJEC [11] against the modified N7 [29] regimen, but results are not yet complete. Future further treatment intensification may include the introduction of ch14.18/CHO antibody during induction or the early use of mIBG treatment. Upfront mIBG treatment as standard of care in a collaborative trial may be hampered by the logistics of performing such therapies in newly diagnosed patients since only a few treating centres are available in Europe and the US. However, tandem HDT as recently reported [30] raises hope that intensification of treatment intensities with mIBG treatment as part of a tandem approach may allow improved survival rates particularly in the ultra-high-risk neuroblastoma patients. These strategies are currently the subject of trials within both North American and European cooperative groups.

Though distinct differences in both the study design and patient demographics existed between the two trials (Table 1), the prognostic impact of SIOPEN scores in stage 4 high-risk neuroblastoma remained. By focusing on metastatic skeletal uptake only, the SIOPEN scoring method offers three advantages. First, it avoids the difficulties in assessing soft tissue disease from two dimensional planar images, particularly if SPECT and SPECT/CT techniques were not applied. The potential influence of soft tissue lesions other than the primary tumour was investigated. A low incidence and a high correlation with the skeletal score were observed in both populations without any independent prediction power, hence justifying not considering ST lesions. Second, different study guidelines may produce confounding results if the primary tumour site is not separated out from the overall score. In SIOPEN/HR-NBL1, mIBG scans were obtained preoperatively at the end of induction, to determine if surgery or further cycles of chemotherapy (TVD) [15] should be given. In

contrast, the post-induction scans in the COG-A3973 population were obtained after removal of the primary tumour. Third, the 12 body segments of the SIOPEN score allow a straightforward attribution of skeletal scores to anatomic regions, facilitating a rapid review process and providing information regarding laterality. Nonetheless, the prognostic significance and cut points of the two scoring methods are similar.

Various approaches were tested to identify a reliable and straightforward prognostic SIOPEN score cut-point. ROC-analysis yielded a higher AUC following induction chemotherapy, underlining that the prognostic impact of the SIOPEN score is stronger post-induction than at diagnosis. This is consistent with previous reports from the COG investigators using the Curie score. A gradual, non-linear trend towards an increased risk of event was observed with rising SIOPEN score. The latter explains why a single consistent cut point across time points and data sets was not identified when applying the Youden-Index. The Youden-Index methodology was not robust across the two trials in providing unique cut points for diagnosis and post-induction. Overall a flattening of risk in the higher score ranges emphasised that the risk is not proportionally increased with high SIOPEN scores. A favourable SIOPEN score risk discrimination occurred only at the lower skeletal score range and was in agreement with a previously clinically chosen cut point of three [31].

Conclusions

This analysis of two independent HR-NBL trials confirms the predictive value of the SIOPEN skeletal score method based on absolute scores at diagnosis and at the end of induction therapy. In particular, the post-induction skeletal score > 3 allows the identification of a very poor prognosis patient group requiring altered treatment strategies. Work is in progress to integrate these findings within a unified international consensus mIBG scoring method for global use in the paediatric oncology community based on the close collaborative work of COG and SIOPEN.

Acknowledgments

We thank Marek Nykiel, Ingrid Pribill, PhD and Claudia Zeiner-Koglin, MSc at the Children's Cancer Research Institute (CCRI) for their support.

Susan G. Kreissman, Dominique Valteau-Couanet, Ariane Boubaker. Funding This study was funded by the European Commission Community Research, Fifth Framework Program, Quality of Life and Management of living Resources: EC Grant No. QLRI-CT-2002-01768 (11. 2002–10. 2005) [<https://www.siopen-r-net.org>]; by the National Cancer Institute Paediatric and Adolescent Solid Tumour Steering Committee and received the following travel funding: Alex's Lemonade Stand Foundation, Ben Towne Foundation, Children's Neuroblastoma Cancer Foundation.

Annex **Annex 1: Consort Diagram**

Annex 2

	Patients	Events	5-yrs. EFS	p-value
Included	216	138	0.39±0.03	0.870
Not included	270	166	0.39±0.03	.

a) Comparison of event-free survival (EFS) of patients on the SIOPEN trial and in the analysis mIBG review data set.

	Patients	Events	5-yrs. EFS	p-value
Included	341	237	0.29±0.03	0.170
Not included	1187	769	0.34±0.01	.

b) Comparison of event-free survival (EFS) of patients on the COG A3973 trial and in the analysis mIBG review data set.

References

1. London WB, Castleberry RP, Matthay KK, Look AT, Seeger RC, Shimada H, et al. Evidence for an age cutoff greater than 365 days for neuroblastoma risk group stratification in the Children's oncology group. *J Clin Oncol*. 2005; 23:6459–65. [PubMed: 16116153]
2. Moroz V, Machin D, Faldum A, Hero B, Iehara T, Mosseri V, et al. Changes over three decades in outcome and the prognostic influence of age-at-diagnosis in young patients with neuroblastoma: a report from the international neuroblastoma risk group project. *Eur J Cancer*. 2011; 47:561–71. [PubMed: 21112770]
3. Seeger RC, Brodeur GM, Sather H, Dalton A, Siegel SE, Wong KY, et al. Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *N Engl J Med*. 1985; 313:1111–6. [PubMed: 4047115]
4. Shimada H, Stram DO, Chatten J, Joshi VV, Hachitanda Y, Brodeur GM, et al. Identification of subsets of neuroblastomas by combined histopathologic and N-myc analysis. *J Natl Cancer Inst*. 1995; 87:1470–6. [PubMed: 7674334]
5. Simon T, Hero B, Faldum A, Handgretinger R, Schrappe M, Klingebiel T, et al. Long term outcome of high-risk neuroblastoma patients after immunotherapy with antibody ch14.18 Or oral metronomic chemotherapy. *BMC Cancer*. 2011; 11:21. [PubMed: 21244693]
6. Valteau-Couanet D, Le Deley MC, Bergeron C, Ducassou S, Michon J, Rubie H, et al. Long-term results of the combination of the N7 induction chemotherapy and the busulfan-melphalan high dose chemotherapy. *Pediatr Blood Cancer*. 2014; 61:977–81. [PubMed: 23970413]
7. Matthay KK, Reynolds CP, Seeger RC, Shimada H, Adkins ES, Haas-Kogan D, et al. Long-term results for children with high-risk neuroblastoma treated on a randomized trial of myeloablative

therapy followed by 13-cis-retinoic acid: a children's oncology group study. *J Clin Oncol.* 2009; 27:1007–13. [PubMed: 19171716]

8. Brodeur GM, Pritchard J, Berthold F, Carlsen NL, Castel V, Castelberry RP, et al. Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *J Clin Oncol.* 1993; 11(8):1466–77. [PubMed: 8336186]
9. Monclair T, Brodeur GM, Ambros PF, Brisse HJ, Cecchetto G, Holmes K, et al. The international neuroblastoma risk group (INRG) staging system: an INRG task force report. *J Clin Oncol.* 2009; 27:298–303. [PubMed: 19047290]
10. Matthay KK, Shulkin B, Ladenstein R, Michon J, Giammarile F, Lewington V, et al. Criteria for evaluation of disease extent by (123)I-metaiodobenzylguanidine scans in neuroblastoma: a report for the international neuroblastoma risk group (INRG) task force. *Br J Cancer.* 2010; 102:1319–26. [PubMed: 20424613]
11. Ladenstein R, Valteau-Couanet D, Brock P, Yaniv I, Castel V, Laureys G, et al. Randomized trial of prophylactic granulocyte colony-stimulating factor during rapid COJEC induction in pediatric patients with high-risk neuroblastoma: the European HR-NBL1/ SIOOPEN study. *J Clin Oncol.* 2010; 28:3516–24. [PubMed: 20567002]
12. Ladenstein R, Pötschger U, Pearson ADJ, Brock P, Luksch R, Castel V, et al. Busulfan and melphalan versus carboplatin, etoposide, and melphalan as high-dose chemotherapy for high-risk neuroblastoma (HR-NBL1/SIOOPEN): an international, randomised, multi-arm, open-label, phase 3 trial. *Lancet Oncol.* 2017; 18:500–14. [PubMed: 28259608]
13. Lewington V, Lambert B, Poetschger U, Sever ZB, Giammarile F, McEwan AJB, et al. 123I-mIBG scintigraphy in neuroblastoma: development of a SIOOPEN semi-quantitative reporting, method by an international panel. *Eur J Nucl Med Mol Imaging.* 2017; 44:234–41. [PubMed: 27663238]
14. Kreissman SG, Seeger RC, Matthay KK, London WB, Spoto R, Grupp SA, et al. Purged versus non-purged peripheral blood stem-cell transplantation for high-risk neuroblastoma (COG A3973): a randomised phase 3 trial. *Lancet Oncol.* 2013; 14:999–1008. [PubMed: 23890779]
15. Garaventa A, Luksch R, Biasotti S, Severi G, Pizzitola MR, Viscardi E, et al. A phase II study of topotecan with vincristine and doxorubicin in children with recurrent/refractory neuroblastoma. *Cancer.* 2003; 98:2488–94. [PubMed: 14635085]
16. Ladenstein R, Weixler S, Baykan B, Bleeke M, Kunert R, Katinger D, et al. Ch14.18 Antibody produced in CHO cells in relapsed or refractory stage 4 neuroblastoma patients: a SIOOPEN phase 1 study. *MABs.* 2013; 5:801–9. [PubMed: 23924804]
17. Yu AL, Gilman AL, Ozkaynak MF, London WB, Kreissman SG, Chen HX, et al. Anti-GD2 antibody with GM-CSF, interleukin-2, and isotretinoin for neuroblastoma. *N Engl J Med.* 2010; 363:1324–34. [PubMed: 20879881]
18. Kaplan E, Meier P. Nonparametric Estimation from Incomplete Observations. *J Am Stat Assoc.* 1958; 53(282):457–81.
19. Marubini, E., Valsecchi, MG. *Analysing survival data from clinical trials and observational studies.* New York: Wiley; 2004.
20. Heinzl H, Kaider A. Gaining more flexibility in cox proportional hazards regression models with cubic spline functions. *Comput Methods Prog Biomed.* 1997; 54:201–8.
21. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics.* 2005; 61:92–105. [PubMed: 15737082]
22. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics.* 2000; 56:337–44. [PubMed: 10877287]
23. YODEN WJ. Index for rating diagnostic tests. *Cancer.* 1950:32–5. [PubMed: 15405679]
24. Yanik GA, Parisi MT, Shulkin BL, Naranjo A, Kreissman SG, London WB, et al. Semiquantitative mIBG scoring as a prognostic indicator in patients with stage 4 neuroblastoma: a report from the Children's oncology group. *J Nucl Med.* 2013; 54:541–8. [PubMed: 23440556]
25. Heagerty PJ. and packaging by PS-C. survivalROC: Time-dependent ROC curve estimation from censored survival data. 2013. R package version 1.0.3. 2013
26. Decarolis B, Schneider C, Hero B, Simon T, Volland R, Roels F, et al. Iodine-123 metaiodobenzylguanidine scintigraphy scoring allows prediction of outcome in patients with stage

- 4 neuroblastoma: results of the cologne interscore comparison study. *J Clin Oncol.* 2013; 31:944–51. [PubMed: 23341514]
27. Yanik, GA., Parisi, MT., Naranjo, A., Nadel, H., Gelfand, MJ., Park, JR., et al. Validation of post-induction Curie scores in high risk neuroblastoma. A Children's Oncology Group (COG) and SIOPEX group report on SIOPEX/HR-NBL1. *J Nucl Med.* 2017. <https://doi.org/10.2967/jnumed.117.195883>
 28. Hartmann O, Valteau-Couanet D, Vassal G, Lapierre V, Brugières L, Delgado R, et al. Prognostic factors in metastatic neuroblastoma in patients over 1 year of age treated with high-dose chemotherapy and stem cell transplantation: a multivariate analysis in 218 patients treated in a single institution. *Bone Marrow Transplant.* 1999; 23:789–95. [PubMed: 10231141]
 29. Kushner BH, Kramer K, LaQuaglia MP, Modak S, Yataghene K, Cheung NKV. Reduction from seven to five cycles of intensive induction chemotherapy in children with high-risk neuroblastoma. *J Clin Oncol.* 2004; 22:4888–92. [PubMed: 15611504]
 30. Park JR, Kreissman SG, London WB, Naranjo A, Cohn SL, Hogarty MD, et al. A phase III randomized clinical trial (RCT) of tandem myeloablative autologous stem cell transplant (ASCT) using peripheral blood stem cell (PBSC) as consolidation therapy for high-risk neuroblastoma (HR-NB): A Children's Oncology Group (COG) study. *J Clin Oncol Am Soc Clin Oncol.* 2016; 34(suppl; abstr LBA3):LBA3–LBA3.
 31. Ladenstein R, Philip T, Lasset C, Hartmann O, Garaventa A, Pinkerton R, et al. Multivariate analysis of risk factors in stage 4 neuroblastoma patients over the age of one year treated with megatherapy and stem-cell transplantation: a report from the European bone marrow transplantation solid tumor registry. *J Clin Oncol.* 1998; 16:953–65. [PubMed: 9508178]

Abbreviations

ASCT	Autologous stem cell transplantation
AUC	Area under the curve
CADO	Cyclophosphamide, Adriamycin, Vincristine
CBDCA	Carboplatin
CDDP	Cisplatin
CEM	Carboplatin-Etoposide-Melphalan
CI	Confidence interval
COG	Children's Oncology Group
COJEC	Rapid, platinum-containing induction schedule (CBDCA, CDDP, CYC, VCR, VP16)
CR/PR	Complete remission/partial remission
CT	Computed tomography
CYC	Cyclophosphamide
DNA	Deoxyribonucleic acid
EFS	Event free survival
HDT	High dose chemotherapy

HR	Hazard ratios
HR-NBL	High Risk Neuroblastoma
INRG	International Neuroblastoma Risk Group
mIBG	Metaiodobenzylguanidine
MYCN	Proto-oncogene
PH	Proportional hazard
ROC	Receiver operating characteristic
SIOPEN	Société International d'Oncologie Pédiatrique European Neuroblastoma
SPECT	Single-photon emission computed tomography
TVD	Topotecan, Vincristine, Doxorubicin
VCR	Vincristine
VP16	Etoposide

Fig. 1.

Definition of scores for the extension of the skeletal involvement with illustrative planar I-123-mIBG scans (arrows) of the various positive scores (left). The mIBG uptake over 12 body segments (right) is scored from 0 to 6 points per segment depending on the disease extension [13]

SIOPEN-HR-NBL-1 (Ref: *J Clin Oncol* 2010 Jul 20;28(21):3516-3524) **COG A3973** (Ref: *Lancet Oncol* 2013 Sep;14(10):999-1008) **Fig.2.**

Schematic overview of therapy on SIOPEN/HR-NBL-1 and COG A3973. Rapid COJEC [cycles given every 10 days; cycles 1 and 5: Vincristine, Carboplatin, Etoposide; cycles 2, 4, 6, and 8: Vincristine, Cisplatin; cycles 3 and 7: Vincristine, Etoposide, Cyclophosphamide]; Optional: TVD [cycles repeat every 3 weeks; Topotecan, Vincristine, Doxorubicin]; Modified N7 [cycles repeat every 3 weeks; cycles 1, 2, 4, and 6: cyclophosphamide, doxorubicin, vincristine; cycles 3 and 5: cisplatin and etoposide] mIBG scan; HDT: high dose therapy [BuMel: Busulfan-Melphalan; CEM: Carboplatin-Etoposide-Melphalan]; XRT, radiotherapy

Fig. 3.

Event-free survival (EFS) by SIOPEN at diagnosis (mIBG1). **a** Histogram of pre-induction mIBG score values at diagnosis; **b** Relationship between the hazards of an event (logHR) and the baseline score using restricted cubic splines (solid line estimated relationship; dotted line 95% confidence intervals); **c** EFS according to tertiles of the maximum score at diagnosis for SIOPEN/HRNBL1 (events/ patients for score 0–23: 76/131; score 24–48: 95/126; score > 48: 66/84) and COG–A3973 (events/ patients for: score 0–23: 51/92; score 24–48: 48/74; score > 48: 39/50). **d** EFS with a threshold of ≥ 3 mIBG positive skeletal spots at diagnosis

for SIOPEN/HRNBL1 (events/patients for: score 0–3: 26/52; score > 3: 211/289) and COG–A3973 (events/patients for score 0–3:26/49; score > 3: 112/ 167)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 4.

Event-free survival (EFS) by SIOPEN post-induction (mIBG2). **a** Histogram of pre-induction mIBG score values at diagnosis; **b** Relationship between the hazards of an event (logHR) and the baseline score using restricted cubic splines (solid line estimated relationship; dotted line 95% confidence intervals); **c** EFS according to tertiles of the maximum score at diagnosis for post-induction scores of SIOPEN/HRNBL1 (events/patients for: score 0–23: 177/259; score 24–48: 34/41; score > 48: 6/ 7 and COG–A3973 (events/patients for score 0–23: 74/122; score 24–48: 15/16; score > 48: 2/2). **d** EFS with a

threshold of 3 mIBG positive skeletal spots for SIOPEN/HRNBL1 (events/patients for score 0–3: 116/185; score>3: 101/122) and for COG–A3973 (events/patients for score 0–3: 59/102; score > 3: 32/38)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 5.

Event-free survival (EFS) by post-induction mIBG response. **a:** EFS of patients mIBG skeletal positive at diagnosis; post-induction score 0 vs. >0 for SIOPEN/HR-NBL-1 and COG A3973 trial cohorts. **b:** EFS of patients mIBG skeletal positive at diagnosis with scores >0 post-induction according to the percent change in SIOPEN mIBG score from diagnosis to post-induction mIBG scans

Table 1
Patient Characteristics at Diagnosis¹, Treatments given and Baseline/Post-induction Scintigraphic Findings in the Skeleton

	SIOPEN/HR-NBL1		COG-A3973		p-value
	n	%	n	%	
All	341		216		
Period	Numbers at baseline		2001–2006		
Sex	Years of accrual		2002–2010		
	Male	61%	121	56%	0.252
	Female	39%	95	44%	
Age	Median years (range)	2.92 (0.23–16.2)	3.18 (0.32–29.1)		
Age Groups	< 1 year	4%	8	4%	0.216
	1–1.5 years	9%	19	9%	
	1.5–5 years	67%	162	75%	
	> 5 years	19%	27	13%	
MYCN	Not amplified	60%	109	61%	0.923
	Amplified	40%	69	39%	
	Missing		38		
Baseline Skeletal mIBG Findings	Negative	8%	38	18%	0.001
	Positive	92%	178	82%	
	Metastatic load				0.407
	Low [1–23]	38%	92	43%	
	Intermediate (24–48)	37%	74	34%	
	High (48–72)	25%	50	23%	
	Median score		33		
	- Minimum		0		
	- Maximum		70		
Baseline Soft Tissue mIBG Findings (excluding Primary Tumour Site)	Negative	80%	180	85%	
	Positive	20%	32	15%	
	Not evaluable		29		
End of Induction			307 pts.	140pts	
Induction (Fig. 2)	COJEC	54%	166	–	–
	Additional TVD or CADO	42%	128	–	–

	SIOPEN/HR-NBL1		COG-A3973		p-value
	n	%	n	%	
High Dose Chemotherapy (Fig. 2)	Other additional CHT	13	4%	–	
	BuMel	144	47%	–	
	CEM	63	21%	124	89%–
	Other HDT	52	17%	–	
	Not given	48	15%	–	
	Ch14.18mAB/CHO	20	7%	–	
Immunotherapy	Ch14.18mAB/SP02	–	–	24	17%
	Negative	115	37%	79	56%
Post-induction Skeletal mIBG Findings	Positive	192	63%	61	44%
	SIOPEN score ranges				0.402
	Low (1–23)	259	84%	122	87%
	Intermediate (24–48)	41	13%	16	11%
	High (48–72)	7	2%	2	1%
	Median score	2		0	
	- Minimum	0		0	
	- Maximum	62		53	
	Negative	240	91%	132	94%
	Positive	25	9%	8	6%
Post-Induction Soft Tissue mIBG Findings (excluding Primary Tumour Site)	Not evaluable	42		0	