Comparison of $^{123}$I-Metaiodobenzylguanidine (MIBG) and $^{131}$I-MIBG Semi-Quantitative Scores in Predicting Survival in Patients With Stage 4 Neuroblastoma: A Report From the Children’s Oncology Group

Arlene Naranjo, PhD,*1 Marguerite T. Parisi, MD, MS Ed,2 Barry L. Shulkin, MD,3 Wendy B. London, PhD,4 Katherine K. Matthey, MD,5 Susan G. Kreissman, MD,6 and Gregory Y. Yanik, MD7

INTRODUCTION

Outcomes for patients with high-risk neuroblastoma remain poor, with 5-year survival rates of only 30–40% for patients with stage 4 disease who are >1 year of age [1,2]. Despite multimodal therapy, complete remission rates remain low, with 15% of patients progressing and another 33% achieving less than a complete response during induction [3].

Metaiodobenzylguanidine (MIBG) is an aralkylguanidine which is structurally similar to the neurotransmitter norepinephrine, and the ganglionic blocking agent guanethidine [4]. Scintigraphic studies in the early 1980s confirmed the effectiveness of MIBG in localization of both pheochromocytomas [5–8] and ganglionic tumors [9–12]. Until recently, when the Food and Drug Administration (FDA) approved $^{123}$I-MIBG for use in children and neuroblastomas concentrate MIBG [16–19]. $^{123}$I-MIBG scans are widely available in Europe [13,14]. Each has the capability to detect neuroblastoma within bone marrow, cortical bone, or soft tissue [15]. Approximately 90% of neuroblastomas concentrate MIBG [16–19]. $^{123}$I-MIBG scans have superior image quality, allow single photon emission computed tomography (SPECT) imaging, and deliver a lower radiation dose to the patient than $^{131}$I-MIBG scans. The sensitivity and specificity of MIBG labeled with either radioidine isotope has been high in identifying both primary and metastatic tumors, and MIBG uptake has been noted in patients with either low or advanced stage disease, favorable or unfavorable Shimada histologic patterns, amplified or unamplified MYCN gene levels, and elevated or normal catecholamine levels [16,20].

The role of MIBG uptake as a surrogate marker for overall response has been examined in many clinical trials [21–27]. Initial studies made use of the qualitative features of the MIBG scans, noting whether disease was detectable or not on MIBG scintigraphy. Semi-quantitative scoring methods have now been developed to predict the extent and severity of MIBG avid disease [21–23]. Using such criteria, the presence of MIBG avid disease immediately prior to myeloablative therapy has been associated with later relapse or disease progression [22–24,28].

In the process of attempting to determine an appropriate definition of an ultra high-risk patient subgroup that may benefit from alternative therapies in designing a new high-risk trial, it has proved necessary to compare $^{123}$I-MIBG and $^{131}$I-MIBG scans to ensure that survival results are not scan-type dependent. Since $^{123}$I-MIBG did not have FDA approval nor was widely available at the time Children’s Oncology Group (COG) study A3973 opened, patients had scans of both types. The goal of this analysis was not to determine whether $^{131}$I-MIBG could replace $^{123}$I-MIBG but to evaluate the possibility of combining results in a single analysis; $^{123}$I-MIBG should not be abandoned and remains preferable to $^{131}$I-MIBG for neuroblastoma imaging as they deliver less patient radiation yet have greater sensitivity in disease detection. Both $^{123}$I-MIBG and $^{131}$I-MIBG scans were used for disease assessments of neuroblastoma patients enrolled on Children’s Oncology Group (COG) high-risk study A3973. The hypothesis was that $^{123}$I-MIBG and $^{131}$I-MIBG scans were sufficiently similar for clinical purposes in terms of ability to predict survival.

Key words: COG A3973; Curie score
was necessary to obtain statistical evidence for combining both types in a single analysis.

A number of studies examining the prognostic impact of MIBG scintigraphy have combined both $^{123}$I-MIBG and $^{131}$I-MIBG data in their analyses [23,24,26,29], whereas others have examined $^{123}$I-MIBG scans alone [21,22,30–32]. As most studies have contained significantly greater numbers of $^{123}$I-MIBG than $^{131}$I-MIBG scans, limited comparative data can be generated on the efficacy of one scintigraphic method over the other [23,24,29]. A large comparative study of 162 patients undergoing diagnostic imaging with either $^{131}$I-MIBG (99 patients) or $^{123}$I-MIBG (63 patients) found no response advantage with the use of either radiisotope [26].

We now compare $^{123}$I-MIBG and $^{131}$I-MIBG scans in predicting survival for over 300 stage 4 patients treated on COG A3973.

METHODS

Patient Selection and Treatment

Newly diagnosed high-risk patients with International Neuroblastoma Staging System (INSS) stage 4 neuroblastoma enrolled on COG protocol A3973 who had completed $^{123}$I-MIBG or $^{131}$I-MIBG scans at one or more of the following time points—diagnosis, post-induction, post-transplant, or post-biotherapy—were included in this analysis. To be eligible for COG A3973, patients with stage 4 disease had to be aged 30 years or younger at the time of initial diagnosis. If younger than 12 months, MYCN amplification (>10 copies) was required; if between 12 and 18 months of age, any unfavorable (MYCN amplification, unfavorable histology, and diploid) or unknown biologic feature was required. Normal renal, cardiac, hepatic, and hematopoietic function was required, as well as no prior systemic therapy. All patients had a signed written informed consent approved at the local Institutional Review Board.

Patients received six cycles of chemotherapy every 21 days, with Cycles 1, 2, 4, and 6 consisting of cyclophosphamide 4.2 g/m², doxorubicin 75 mg/m², and vincristine 2 mg/m², and Cycles 3 and 5 including cisplatin 200 mg/m² and etoposide 600 mg/m². Surgical resection of the primary tumor occurred after Cycle 5. Patients were randomized to either purged or unpurged autologous stem cell transplant after induction and then received radiation and 13-cis-retinoic acid biologic therapy.

MIBG Scans

Supersaturated potassium iodide was administered, generally 24 hr prior to the diagnostic MIBG dose and subsequently continued for 3–7 days following the dose, to reduce thyroid accumulation of free radiiodine. Either $^{123}$I-MIBG or $^{131}$I-MIBG was administered via intravenous injection over 90 sec (methodology discussed below). For determination of MIBG scores, 10 different sites were evaluated, including 9 skeletal sites (head, chest, T-spine, L-spine, pelvis, upper arms, lower arms, femurs, and lower legs) and an additional 10th site for soft tissue lesions. Skeletal sites were individually scored from 0 to 3 as follows:

0 = no MIBG involvement; 1 = one MIBG avid lesion present; 2 = greater than one MIBG avid lesion present; and 3 = MIBG avidity present in >50% of an individual site. Soft tissue lesions were scored: 0 = no MIBG involvement; 1 = one MIBG avid soft tissue lesion present; 2 = greater than one MIBG avid soft tissue lesion present; and 3 = MIBG avidity in a soft tissue lesion occupying >50% of the chest or abdomen. A patient’s Curie score at each time point was calculated as the sum of his/her scores over all individual sites. The maximum Curie score was 30. All MIBG scans were centrally reviewed by two nuclear medicine physicians, without knowledge of the original scan reports or other clinical or imaging information to validate the data.

$^{123}$I-MIBG scintigraphy. Planar images with or without tomographic images were acquired [20]. Typically, for planar imaging, overlapping anterior and posterior spot views imaging the entire body are obtained for 10 min each at 24 hr following the administration of $370$ MBq (10 mCi)/1.7 m² of body surface area (BSA) of $^{123}$I-MIBG. A large field of view dual-head gamma camera and low-energy collimators were recommended. Alternately, a low-speed, whole-body scan was acceptable. A single- or multiheaded camera with a low-energy collimator, rotated 360° with 120 projections at ~20 sec per stop, was used for SPECT imaging at 24 hr. Filtered back-projection with a Butterworth filter and a cut-off frequency of 0.2–0.5 was used to reconstruct the images.

$^{131}$I-MIBG scintigraphy. $^{131}$I-MIBG scintigraphy is typically performed 24–48 hr and occasionally 72–96 hr after administering $18.5–37$ MBq (0.5–1.0 mCi)/1.7 m² BSA of $^{131}$I-MIBG. Anterior and posterior views of the head, neck, chest, abdomen, and pelvis are obtained for 100,000 counts or 20 min [20]. Alternately, a low-speed, whole-body scan would also be adequate. The use of a dual-head gamma camera with a large field of view and high-energy collimators is again preferred. SPECT imaging is not performed with $^{131}$I-MIBG scintigraphy. MIBG loses activity over time due to radioactive decay; thus anatomic features become less recognizable at later imaging times. However, tumors often become more evident over time, due to a reduction in surrounding background levels.

Statistical Analysis

Differences in median Curie scores between the two MIBG scan types were tested with the two-sided Wilcoxon Rank-Sum test at each time point. The relationship between $^{123}$I-MIBG and $^{131}$I-MIBG scan types and standard prognostic factors age (<18 months vs. ≥18 months), MYCN status (not amplified vs. amplified), ploidy (hyperdiploid vs. diploid), and Shimada histology (favorable vs. unfavorable) were examined via Fisher’s exact test.

Time to event for $^{123}$I-MIBG versus $^{131}$I-MIBG was compared using log-rank test comparisons of event-free survival (EFS) and overall survival (OS) at each time point. The relationship between $^{123}$I-MIBG and $^{131}$I-MIBG scan types and standard prognostic factors age (<18 months vs. ≥18 months), MYCN status (not amplified vs. amplified), ploidy (hyperdiploid vs. diploid), and Shimada histology (favorable vs. unfavorable) were examined via Fisher’s exact test.

Time to event for $^{123}$I-MIBG versus $^{131}$I-MIBG was compared using log-rank test comparisons of event-free survival (EFS) and overall survival (OS) at each time point. EFS and OS were calculated using the method of Kaplan and Meier [33] with standard errors per Peto et al. [34]. For EFS, time to event was calculated from diagnosis until the first occurrence of an event (relapse, progressive disease, secondary malignancy, or death) or until the time of last contact if no event occurred. For OS, time to event was calculated from diagnosis until the time of death or last contact if the patient did not die. At each post-diagnosis time point, time to event was calculated from the date of the MIBG scan. Patients who had an event between the diagnosis and post-diagnosis MIBG scan (post-induction, post-transplant, and post-biotherapy) went off-study and hence were not included in the analysis for that particular time point.

Cox proportional hazards (PH) models with the Efron method of handling tied event times were used to determine the prognostic strength for survival of the Curie score and scan type in the
TABLE I. Curie Score Summary Statistics and Tests by Time Point and Scan Type

<table>
<thead>
<tr>
<th>Time point</th>
<th>123I-MIBG</th>
<th>131I-MIBG</th>
<th>Wilcoxon Rank-Sum P-value</th>
<th>Log-rank P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Median (range)</td>
<td>N Median (range)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td>218 11 (0–30)</td>
<td>89 8 (0–30)</td>
<td>0.1217</td>
<td>0.9582 0.8283</td>
</tr>
<tr>
<td>Post-induction</td>
<td>226 0 (0–28)</td>
<td>76 0 (0–20)</td>
<td>0.6024</td>
<td>0.3501 0.5377</td>
</tr>
<tr>
<td>Post-transplant</td>
<td>167 0 (0–24)</td>
<td>62 0 (0–5)</td>
<td>0.1154</td>
<td>0.5333 0.6366</td>
</tr>
<tr>
<td>Post-biotherapy</td>
<td>91 0 (0–26)</td>
<td>23 0 (0–5)</td>
<td>0.5353</td>
<td>0.6243 0.2642</td>
</tr>
</tbody>
</table>

MIBG, metaiodobenzylguanidine; N, number of patients; EFS, event-free survival; OS, overall survival. *Time to event for the log-rank test calculated from the date of the MIBG scan.

RESULTS

Patient Characteristics and MIBG Scans

For the 350 stage 4 patients with at least one MIBG scan, the median age at diagnosis was 3.1 years (range 6.8 months to 29 years). MYCN amplification was observed in 41% of tumors analyzed. Diploid tumors were more prevalent (57%) and the majority of patients had unfavorable histology (97%). No statistically significant difference in age, MYCN amplification, ploidy, and histology distribution was found between the 123I-MIBG and 131I-MIBG cohort of patients at any of the time points.

A total of 926 MIBG scans were reviewed, with 683 scans (74%) utilizing 123I-MIBG and 243 scans (26%) using the 131I-MIBG radiolabel for diagnostic imaging. In 8% (28/350) of the stage 4 cases, patients switched from one scan modality to the other during the course of treatment. The remaining 92% of patients underwent MIBG scintigraphy with the same MIBG radioisotope throughout their therapy course or only had one scan performed. Of these, 243/350 (69%) patients had only 123I-MIBG scans and 79/350 (23%) had only 131I-MIBG scans performed. There were no differences in median Curie scores between scan type modalities (123I-MIBG vs. 131I-MIBG) at any of the time points (Table I).

Survival Analysis

The overall 5-year EFS and OS rates were 33.4 ± 3.6% and 45.6 ± 4.0%, respectively. None of the EFS and OS comparisons for 123I-MIBG versus 131I-MIBG at any time point were statistically significant (Table I). The 5-year EFS for the 243 patients with only 123I-MIBG scans and the 79 patients with only 131I-MIBG scans were 34.1 ± 4.6% and 39.9 ± 7.1%, respectively (log-rank P-value = 0.4945), and the 5-year OS rates were 46.5 ± 5.1% and 51.8 ± 7.5%, respectively (log-rank P-value = 0.5409).

At post-induction, 112/302 (37%) patients had a Curie score above the median value of 0. These 112 patients had a median Curie score of 3. Patients with Curie score above 0 were found to have statistically significantly worse EFS (log-rank P-value = 0.0009; Fig. 1). A Cox model for EFS indicated that Curie score and MYCN status were predictive of survival at post-induction, with Curie scores > 0 corresponding to an increase in the risk of event of 1.598 and patients with MYCN amplified tumors having significantly higher risk of event but the effect decreasing over time (Table II). There was no difference in EFS between 123I-MIBG and 131I-MIBG scan types. The PH assumption was violated for MYCN status so it was included in the Cox model as a time-varying covariate with a 1-day difference as the time scale. Curie score did not maintain independent statistical significance for OS.

DISCUSSION

Historically, the 123I-MIBG scan has been determined to be more sensitive than the 131I-MIBG scan. In other words, 123I-MIBG shows abnormal areas of uptake more clearly than 131I-MIBG and often shows sites of disease not visualized with 131I-MIBG. This may or may not result in a higher Curie score in certain patients.

123I-MIBG is clearly the agent of choice for the scintigraphic depiction of neuroblastoma. The imaging characteristics of the 123I-MIBG label are superior to those of the 131I-MIBG isotope. 123I-MIBG has a lower energy photon more suitable for modern...
gamma cameras and for acquisition of SPECT images. It has a shorter physical half-life (13.2 hr vs. 8.05 days) and does not have substantial beta emission, which adds to the radiation burden but does not contribute to imaging. These characteristics allow about 10-fold higher doses of 123I-MIBG to be administered for the same radiation burden as 131I-MIBG. Thus, image quality is superior and normal and abnormal features much more readily identified [20].

In this study, survival analyses did not identify a difference by scan type at post-induction. For future survival analyses using MIBG Curie score at post-induction as a possible component in the definition of an ultra high-risk group, the particular isotope of the radioiodine label may be ignored since the survival predictive ability is similar for 123I-MIBG and 131I-MIBG. A multivariate analysis with other clinical factors including age, MYCN amplification, and ploidy showed that Curie score at post-induction is an independent prognostic marker of EFS. When 123I-MIBG and 131I-MIBG were compared directly, Curie score was similarly predictive of outcome for the two scan types. For future studies or in meta-analyses, this statistical analysis will allow us to combine Curie score data at post-induction from both isotopes in assessing survival.

ACKNOWLEDGMENT

We would like to thank Fran Laurie and staff at the Quality Assurance Review Center (QARC) for their tremendous support of this project.

REFERENCES


