International Criteria for Diagnosis, Staging, and Response to Treatment in Patients With Neuroblastoma


Neuroblastoma is one of the most common tumors in childhood. However, it often has been difficult to compare clinical and laboratory studies of this disease due to a lack of uniform criteria for diagnosis, staging, and response. An international group of conferees addressed each of these issues and reached a consensus. Specific criteria for making a diagnosis of neuroblastoma are defined. A new neuroblastoma staging system is proposed that takes into account the most important elements of current but incompatible systems. Finally, criteria for response to treatment are standardized. The criteria proposed herein represent an international consensus of essentially every major pediatric oncology group or organization in the United States, Europe, and Japan. The staging system should be referred to as the International Neuroblastoma Staging System, and the response criteria as the International Neuroblastoma Response Criteria. Implementation of these criteria will greatly facilitate the comparison of clinical and laboratory studies by different groups and countries. Furthermore, these criteria should serve as a foundation on which future modifications or improvements can be based.


NEUROBLASTOMA, a tumor of the autonomic nervous system, is one of the most common tumors of childhood. It occurs in nine per million per year in children <15 years of age in the United States, and it accounts for about 8% of all pediatric cancers. However, there has been relatively little improvement in the long-term survival of most patients with this disease, despite dramatic improvements in the survival of patients with many other pediatric malignancies, notably acute lymphoblastic leukemia and Wilms' tumor. One of the obstacles to progress in this disease is the difficulty in comparing results of protocols from different centers or countries, largely because of a lack of uniform criteria for diagnosis, for staging, and for determining response to therapy.

Neuroblastoma is a "small, round, blue cell tumor," which generally arises in the adrenal medulla or sympathetic chain in a child. To confirm this diagnosis, some histologic evidence is generally required which demonstrates neural origin or differentiation by light microscopy, electron microscopy, or immunohistology. Alternatively, because of the frequency of involvement of the bone marrow in 50% to 60% of patients, some patients are considered to have neuroblastoma based on the presence of "compatible tumor cells" involving the bone marrow, accompanied by increased urinary catecholamine metabolites. A third set of criteria used by some is a "compatible" radiographic or scinti-
graphic appearance, with increased urinary catecholamine metabolites. Since all groups do not accept all methods for making a diagnosis of neuroblastoma, there exist some discrepancies between studies.

There are three major staging systems used for neuroblastoma throughout the world: (1) the system used by the Children’s Cancer Study Group (CCSG), and others; (2) the system used by St Jude Children’s Research Hospital (SJCRH) and the Pediatric Oncology Group (POG); and (3) systems based on the Tumor Node Metastasis (TNM) classification proposed by the International Union Against Cancer (UICC). Modifications of these systems are used by the Italian Cooperative Working Group, the Malignant Tumor Committee of the Japanese Society of Pediatric Surgeons, and others. In general, the various staging systems give comparable results in distinguishing low-stage, good-prognosis patients from high-stage, poor-prognosis patients. However, some of the differences between the staging systems are substantial, particularly as applied to individual patients, so the results of one group cannot be readily compared with another.

Points of disagreement include (1) the importance of resectability of the primary tumor; (2) the prognostic significance of tumor “crossing the midline”; (3) the prognostic importance of ipsilateral and/or contralateral lymph node involvement; and (4) the separation of patients (usually infants) with “stage IV-S” from other children with disseminated disease. Agreement on the definition of stage with regard to these and other issues would facilitate comparison of different studies.

A variety of terms have been used to report the response of neuroblastoma patients to a given treatment regimen: complete response (CR), very-good partial response (VGPR), good partial response (GPR), partial response (PR), mixed response (MR), stable disease (SD), and progressive disease (PD). Despite their general use, the same terms may have a different meaning when used by different groups. This is due in part to differences in the number and type of tests used for evaluation. In addition, the time at which response is evaluated often varies considerably.

A conference, sponsored by the William G. Forbeck Research Foundation, was held on November 10 to 11, 1986 to begin the process of standardizing definitions for diagnosis, staging, and treatment response. Individuals representing most of the major pediatric oncology groups in the world agreed in principle on definitions and drafted documents. These documents were then circulated to participants and others not in attendance. The proposals were presented and discussed at the Fourth International Symposium on Advances in Neuroblastoma Research in Philadelphia on May 14 to 16, 1987. This communication represents the conclusions reached by international correspondence following that meeting.

**DIAGNOSIS**

The conferees and corresponding participants agreed on minimum criteria for establishing a diagnosis of neuroblastoma (Table 1). This definition excludes the combination of a compatible radiographic or scintigraphic appearance and increased urinary catecholamine metabolites. While allowing this definition may spare some patients an initial surgical procedure to confirm the diagnosis, a patient could potentially receive treatment that was not required, or alternatively could receive the wrong therapy. For example, some other tumors such as ganglioneuroma, pheochromocytoma, or peripheral neuroepithelioma could present in a similar manner.

In addition, biologic characterization of neuroblastoma cells is becoming increasingly im-

**Table 1. Diagnosis of Neuroblastoma**

A diagnosis of neuroblastoma is established if:

1. An unequivocal pathological diagnosis is made from tumor tissue by standard methods, including electron microscopy if necessary; or
2. Bone marrow contains unequivocal tumor cells (eg, syncytiotrophoblast) and urine contains increased urinary catecholamine metabolites (VMA and/or HVA >3 SD above the mean per mg creatinine for age*).

*Both the VMA and HVA should be measured. Normalization per milligram of creatinine makes a timed collection (such as a 24-hour collection, which is difficult in young children) unnecessary and avoids potential false negatives due to dilute urine. A test indicating comparable increase of serum catecholamines would suffice, assuming the appropriate controls and standardization were performed.
porate for diagnosis or prognostication, eg, histopathology, immunophenotype, N-myc gene copy number and/or expression, and DNA index.13-17 Some of these or other biologic properties are likely to supersede characteristics, such as crossing the midline or lymph node involvement, in future prognostic classifications. In order to perform the necessary biological studies, it is critical to obtain tumor tissue as part of the initial diagnostic workup.

A battery of monoclonal antibodies, recombinant DNA probes, and other markers are being developed that may allow increased specificity in making a diagnosis of neuroblastoma.1 Other “small, round, blue cell tumors of childhood,” such as rhabdomyosarcoma, Ewing's sarcoma, neuroepithelioma, and non-Hodgkin's lymphoma. It will be crucial to have tissue available for these studies as well. If surgery is needed to obtain tissue for diagnosis (Table 1), more definitive diagnostic and prognostic information may be obtained. This information may then allow the selection of a more appropriate treatment and may obviate the need for subsequent, extensive reevaluation or surgery.

STAGING

The proposed system is based on clinical, radiographic, and surgical evaluation of children with neuroblastoma. It is clear that primary and metastatic sites must be evaluated, but at issue is the number and type of tests used to determine the extent of disease. The proposed staging system uses components of previous systems (Table 2). Arabic numbers, rather than Roman numerals or letters of the alphabet, are used to distinguish this system from the two most widely used current systems for neuroblastoma.6

Stages 1 and 2

Stage 1 is similar to stages I and A in the two major staging systems.5,6 However, the greatest disagreement concerns the assignment of patients to the middle stages (II and III; B and C). This discrepancy is critical since some of these patients do extremely well, while others have tumors that are as aggressive as those with disseminated disease. A major controversy concerns the impact of ipsilateral v contralateral lymph node involvement on prognosis.5,18,19 Accordingly, it was decided to divide stage 2 into 2A (incompletely excised tumor), and 2B (ipsilateral lymph node involvement), so that these patients could be analyzed separately or together. Thus, it can be determined if the behavior of the patients with stage 2B more closely resembles stage 2A or stage 3. Localized, grossly resected tumors confined to the area of origin with adherent lymph nodes will be considered stage 1, even if the adherent lymph nodes are involved with tumor.

Stage 3

Most stage 3 tumors arise in the abdomen, since tumors crossing the midline by contiguous infiltration or by lymph node involvement are less common in the thorax. Infiltration is meant to indicate contiguous invasion of tissues across the midline, rather than a pedunculated tumor that hangs over the midline. There is no obvious prognostic significance to the different patterns of tumor spread by which a patient is considered to have stage 3 disease, so this stage has not been subcategorized. However, it is recommended that the specific findings by which a patient is categorized as stage 3 be recorded so that the different patterns of spread (by tumor infiltration, lymph node metastasis, or both) could be analyzed retrospectively.
Tumors arising in the midline have been a special problem. In general, these tumors arise from the organ of Zuckerkandl or from a sympathetic ganglion in the pelvis. Tumors arising in the adrenal gland, or in a sympathetic ganglion in the thorax or abdomen, by definition are not considered midline. Midline tumors without lymph node involvement or disseminated metastases would be classified as stage 1 or 2A, based on resectability. Those with unilateral lymph node involvement would be considered stage 2B, but those with bilateral lymph node involvement or grossly unresectable tumors extending on both sides of the midline (with or without lymph node involvement) would be considered stage 3.

Stage 4

Patients are categorized as having stage 4 disease on the basis of disseminated disease involving distant lymph nodes, bone, bone marrow, liver, and/or other organs (except as defined in stage 4S). This definition of stage 4 is essentially identical to stages IV and D in the two major staging systems. However, there is some evidence that patients who have stage 4 on the basis of distant lymph node, liver, or marrow involvement (excluding 4S) do better than those who have stage 4 on the basis of cortical bone involvement. This is especially true for patients <2 years old. Since this may be an important distinction in terms of prognosis and choice of therapy, it is recommended that the criteria by which patients are considered to have disseminated disease be recorded for all stage 4 patients.

Stage 4S

Stage IV-S (or D-S) has been retained as 4S based on the favorable outcome generally experienced with these patients and because of recent biologic evidence distinguishing these patients from infants with conventional stage IV disease, such as serum ferritin, DNA index, and N-myc copy number. However, the issue of bone marrow involvement remains a problem, since a "positive" marrow can mean anything from a few small clumps of cells in an otherwise normal marrow to total marrow replacement. Most felt that the latter case would more appropriately belong in stage 4, but sampling makes it difficult to make an objective assessment of the extent of marrow involvement. Therefore, patients with any extent of marrow involvement that otherwise fit the 4S category will be staged as 4S.

The issue of placing an upper age limit of 1 year on stage 4S was raised. Some felt that the unique behavior of tumors in these patients was generally restricted to the first 6 to 12 months of life. Others felt that the rare patient over 1 year of age who met the criteria for stage 4S might be overtreated if assigned to stage 4. The number of true 4S patients over 1 year of age should be very small, and it is not clear if their tumors behave in a nonaggressive manner similar to their younger counterparts, or in an aggressive manner typical of disseminated disease in older children. Since there are no age restrictions on other stages, it was decided not to include an age restriction for this group of patients.

Minimum Testing

The minimum testing necessary to define these stages is presented in Table 3. Certainly, the more tests that are performed, the greater the likelihood of finding disseminated disease. Uniformity with respect to minimum testing should improve the comparability of studies, and the tests recommended can be performed in

<table>
<thead>
<tr>
<th>Tumor site</th>
<th>Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>CT, ultrasound, or MRI with 3D measurements</td>
</tr>
<tr>
<td>Metastases</td>
<td>Bilateral posterior iliac bone marrow aspirates and core biopsies (four adequate specimens necessary to exclude tumor)</td>
</tr>
</tbody>
</table>
|            | Bone radiographs and scintigraphy by 
|            | $^{99m}$Tc-diphosphonate, $^{123}$I-MIBG  |
|            | or $^{123}$I-MIBG  |
|            | Abdominal and liver imaging by CT, ultrasound, or MRI |
|            | Chest radiograph (AP and lateral) and chest CT |
| Markers    | Urinary catecholamine metabolites (VMA and HVA) |

NOTE. For evaluation of bone metastases, $^{99m}$Tc-diphosphonate scintigraphy is recommended for all patients and is essential if MIBG scintigraphy is negative in bone.

Abbreviations: CT, computed tomography; MRI, magnetic resonance imaging; MIBG, meta-iodobenzylguanidine.
most medical centers. This is an area that is continually evolving, and changes will undoubtedly be recommended in the future to reflect progress.

The specificity and sensitivity of meta-iodobenzylguanidine (MIBG) scintigraphy for evaluation of bone and soft tissue involvement by neuroblastoma was discussed. Unfortunately, MIBG scintigraphy is not readily available in most medical centers in the United States and in some other countries. It is likely that once the MIBG technique is more widely available and more experience is gained, it will be included in the routine evaluation of patients with suspected or proven neuroblastoma. Until that time, 99mTc-diphosphonate scintigraphy will remain the standard for evaluation of bone disease, and computed tomography (CT) and/or magnetic resonance imaging (MRI) the standard for evaluation of soft tissue disease.

An important area of difference concerns the assessment of bone marrow disease by marrow aspirates and/or biopsy. Some centers do only a single marrow aspirate, whereas others routinely do nine aspirates and four biopsies under general anesthesia. Recent studies suggest that marrow biopsies add substantially to the detection of marrow involvement by tumor. Also, it is clear that obtaining more samples will increase the likelihood of detecting marrow involvement. However, the clinical relevance of marrow involvement detected with additional studies is unclear if it is not detected in the first few studies. A compromise was reached to use two marrow aspirates and two biopsies, ie, one aspirate and one biopsy from each of the posterior iliac crests. Only a single study positive for tumor is required to document bone marrow involvement, but all four studies are required to rule it out.

Methods for detection of marrow involvement by tumor are in evolution. Most centers rely on standard cytology of marrow smears and histology of marrow biopsies, whereas others are using "neuroblastoma-specific" immunocytoLOGY. The latter approach is more sensitive in detecting occult disease. However, until techniques and reagents become standardized, the determination of bone marrow involvement will continue to rely on standard cytology and histopathology.

The assessment of lymph node involvement has been a problem, particularly for surgeons. It is not always clear which lymph node groups need assessment if none are enlarged. In addition, the distinction between adherent, adjacent, draining, and "distant" nodes may be difficult. Any enlarged lymph nodes in the operative field should be biopsied to document the presence or absence of tumor involvement, and other identifiable nodes also should be biopsied. Each lymph node that is biopsied should be labeled so that its relationship to the primary tumor is apparent. However, it is not necessary to biopsy nodes in the contralateral thorax to rule out tumor involvement if none are enlarged, since the likelihood of distant lymph node spread across the midline in the thorax is unlikely. A standardized approach to lymph node biopsy and assessment is being formulated by surgery and pathology representatives.

The presence or absence of liver involvement should be determined at the time of surgery for patients with primary tumors in the abdomen. Otherwise, the diagnostic imaging studies are adequate to assess liver involvement in individuals with primary tumors in the neck or thorax. Other sites of suspected tumor involvement should be biopsied if possible, especially if the findings would alter the stage of the patient.

**RESPONSE TO TREATMENT**

Criteria for determining the response to therapy vary considerably from one institution to another, as well as the time after diagnosis at which the assessment of response is made. The same tests that are used for determining extent of disease (Table 3) should be used to assess response of primary and metastatic sites to treatment. In responding patients, it is not necessary to repeat all tests if they were negative on initial evaluation. However, patients who relapse should be thoroughly reassessed for extent of disease as if they were newly diagnosed patients.

Table 4 lists criteria to determine response to therapy. It is important to note that a given level of overall response involves thorough assessment of both primary and metastatic sites. For example, a CR requires that the primary and all metastatic sites fulfill CR criteria. A CR in metastatic sites and a PR in the primary tumor.
Table 4. Definitions of Response to Treatment

<table>
<thead>
<tr>
<th>Response</th>
<th>Primary</th>
<th>Metastases</th>
<th>Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CR</td>
<td>No tumor</td>
<td>No tumor (chest, abdomen, liver, bone, bone marrow, nodes, etc)</td>
<td>HVA/VMA normal</td>
</tr>
<tr>
<td>2. VGPR</td>
<td>Reduction &gt;90% but &lt;100%</td>
<td>No tumor (as above except bone); no new bone lesions, all preexisting lesions improved</td>
<td>HVA/VMA decreased &gt;90%</td>
</tr>
<tr>
<td>3. PR</td>
<td>Reduction 50%-90%</td>
<td>No new lesions; 50%-90% reduction in measurable sites; 0 to 1 bone marrow samples with tumor; bone lesions same as VGPR improved</td>
<td>HVA/VMA decreased 50% to 90%</td>
</tr>
<tr>
<td>4. MR</td>
<td>No new lesions; &gt;50% reduction of any measurable lesion (primary or metastases) with &lt;50% reduction in any other; &lt;25% increase in any existing lesion*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. NR</td>
<td>No new lesions; &lt;50% reduction but &lt;25% increase in any existing lesion*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. PD</td>
<td>Any new lesion; increase of any measurable lesion by &gt;25%; previous negative marrow positive for tumor</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE. Evaluations of primary and metastatic disease as outlined in Table 3. Abbreviation: NR, no response.

*Quantitative assessment does not apply to marrow disease.

would be considered a PR overall. Evaluations for response in newly diagnosed patients are recommended at the end of induction (usually 3 to 4 months), at the end of maintenance (usually 8 to 12 months), before and after surgical procedures, before marrow transplantation, and as indicated clinically. If surgery is performed as part of the evaluation after treatment, it will be important to report the response achieved both before and after surgery, so the relative contribution of chemotherapy and of surgery can be assessed.

Although examples are given for responses in liver and thorax, involvement of other organs or sites should be evaluated in a similar manner. Physical examination may be informative for assessment of lymph node involvement or skin metastases, and this should also be included in any evaluation of response to treatment. Three-dimensional measurements should be possible for primary tumors and many metastatic lesions based on CT or other diagnostic imaging modalities. The response criteria are expressed as percent decrease in size, and by convention are based on the product of the greatest two diameters measurable, not on volume.

Documentation of MR, SD, and PD with initial or subsequent therapy may be important. First, these responses document failure with a given therapy, since patients who do not achieve at least a PR for their overall response to treatment have little chance of cure. These criteria also will be useful for evaluating responses to phase I or II agents. Second, some patients who achieve a MR or SD based on a failure of the primary tumor or a metastatic lesion to decrease substantially in size may have a benign ganglioneuroma or a fibrotic, calcified mass. If careful histological evaluation of multiple different sites fails to demonstrate viable neuroblastoma cells, these tumors might then be considered to have achieved a CR. Cases in which CR is defined in this way should be identified accordingly.

It is difficult to give criteria for the quantitative assessment of response of bone and bone marrow disease. It is not currently possible to derive a quantitative scintigraphic value for a bone lesion. Moreover, scintigraphic changes may lag behind histologic changes by weeks or months. It would be useful to have a quantitative scintigraphic value for a bone lesion at diagnosis and at reevaluation, similar to Hounsfield units in CT. However, currently scintigraphy is used qualitatively to determine if a lesion has disappeared, improved, or progressed, or if a new lesion has developed.

With respect to marrow involvement, a decrease from a packed marrow to a single clump on a single aspirate is clearly a meaningful re-
response, although the marrow would still be viewed as "positive." Since standard histologic evaluation of marrow aspirates and biopsies is imprecise and subject to sampling artifacts, it is difficult to be more quantitative than we have been in Table 4. Perhaps in the future, immunocytoLOGY will provide a more objective means of quantitating marrow involvement that would be more precise and possibly obviate the need for multiple marrow sampling.

Tumor marker response, in the form of urinary catecholamine metabolites, is included in the evaluation of response to treatment. However, there are additional markers that may reflect response in some patients, such as serum ferritin, neuron-specific enolase (NSE), and the ganglioside GD2. If these or other markers of neuroblastoma disease activity become standardized and more generally available, they would allow additional criteria by which to evaluate response to treatment. However, due to limitations of standardization and availability, ferritin, NSE, and GD2 are not included in this version of the international response criteria.

**DISCUSSION**

An international consensus has been reached in neuroblastoma regarding criteria for diagnosis, staging, and response to treatment. Although these criteria do not contain any radical concepts or major changes from previous systems, they do embody the most important features in a system that is comprehensive and rigorously defined. More importantly, the use of these standardized criteria should facilitate comparisons between clinical and laboratory studies from different institutions and different countries. As such, this represents a major milestone in international cooperation.

These criteria are a foundation upon which future modifications and improvements can be built. It is likely that advances in immunodiagnosis, molecular biology, diagnostic imaging, and other areas will result in substantial changes in the above criteria. An increasing number of clinical and laboratory tests are becoming available that appear to improve diagnosis, staging, and assessment of disease status in patients with neuroblastoma.

Indeed, new diagnostic and prognostic tools have been developed in recent years, including monoclonal antibodies and recombinant DNA probes. Although any new tests will have to be evaluated prospectively, it is possible to assess their use retrospectively if samples and clinical data are available. Accordingly, the conference recommended that serum, tumor tissue, bone marrow and, if possible, some normal cells or tissue be saved in institutional, regional, or national repositories to permit this. While such a recommendation does require some expense, labor, and record keeping, it should prove to be an invaluable resource for assessing the use of new tests.

Ultimately, it may be possible to use a combination of clinical and biological criteria to divide patients into three groups: those who require no further therapy following surgery; those who are curable with conventional therapy; and those who would fail conventional therapy and who should receive more aggressive or experimental approaches (e.g., high-dose chemo- or chemoradiotherapy with bone marrow transplantation). International implementation and assessment of these approaches will be facilitated by the framework provided here, as well as by the network of communication and cooperation that has been established in order to generate this consensus.

**ACKNOWLEDGMENT**

We wish to acknowledge the invaluable contributions of Jennifer and George Forbeck and the William Guy Forbeck Research Foundation, Hilton Head, SC, for sponsoring the initial meeting and supporting this international effort. The writing committee would like to thank the selfless efforts of all contributors from around the world who were willing to put aside personal, regional, and national interests and work in a true spirit of cooperation for the sake of an international agreement on these very important issues. We also wish to thank the following individuals for helpful suggestions in the preparation of this manuscript: Julie Blatt, Christopher Fryer, Shinmaku Imashuku, Michio Kaneko, Ahira Naka-gawara, Mark Nesbit, Ruprecht Nitschke, Teresa Vietti, and K.Y. Wong.

**REFERENCES**

2. Voute PA: Neuroblastoma, in Sutow WW, Fernbach
INTERNATIONAL CRITERIA FOR NEUROBLASTOMA

DJ, Vietti TJ: Clinical Pediatric Oncology. St Louis, Mosby, 1984, pp 559-587


11. Tsuchida Y: Personal communication, 1988


