Minimal residual disease monitoring in neuroblastoma

van Wezel, E.M.

Citation for published version (APA):
Chapter 7

Interim analysis of the prospective minimal residual disease monitoring study in high risk neuroblastoma

Esther van Wezel¹,⁶, Boris Decarolis², Janine Sutterheim¹,⁵, Lily Zappeij-Kannegieter¹, Mariette van Doornum¹, Jalenka van Wijk¹,⁵, Barbara Hero², Thorsten Simon², Frank Berthold², Max van Noesel⁴,⁵, Marta Fiocco³, Huib Caron⁶, Ellen van der Schoot¹, Godelieve Tytgat⁵,⁶

¹. Department of Experimental Immunohematology, Sanquin Research, Amsterdam The Netherlands and Landsteiner Laboratory of the AMC, University of Amsterdam, The Netherlands, 2. Children’s Hospital, University of Cologne, Pediatric Hematology and Oncology, Germany, 3. Department of Biostatistics, Leiden University Medical Center and Dutch Childhood Oncology Group The Hague, Netherlands, 4. Department of Pediatric Oncology, Sophia Children’s Hospital, Erasmus Medical Center, Rotterdam, The Netherlands, 5. Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands, 6. Department of Pediatric Oncology, Emma Children’s Hospital, Academical Medical Center, University of Amsterdam, The Netherlands
Abstract

Introduction
Real-time quantitative PCR (RQ-PCR) for the detection of minimal residual disease (MRD) has been demonstrated to be a predictor of survival in retrospective studies. Therefore, we performed a large prospective study in high risk neuroblastoma (NB) patients. We studied the prognostic value of the level of NB-specific mRNA at diagnosis and investigated the clinical significance of NB clearance from the BM during induction chemotherapy.

Methods
Serial BM samples from 223 high risk NB patients at diagnosis and at set time points during treatment were prospectively collected at Dutch and German centers between 2009 and 2015. qPCR was performed using five NB-specific markers, as previously described.

Results
High levels of neuroblastoma mRNA in BM at diagnosis were associated with a poor event-free survival (EFS) and overall survival (OS). Patients with a fast molecular response (BM clearance after 4 courses of therapy) did not differ significantly in terms of EFS and OS from patients without a fast response. Also after induction therapy no significant difference in outcome between patients with and without MRD positivity was observed.

Conclusion
In this interim analysis we show that high levels of BM infiltration at diagnosis are strongly associated with a poor outcome. In contrast to results found in other studies BM response after completion of induction chemotherapy was not associated with a poor outcome. However, in this interim analysis the follow-up time is still too short to draw definitive conclusions, moreover the group of patients with available serial BM samples was too small. Therefore, for the final analysis we need to wait until for many patients a follow-up time of 60 months is reached, and in the final analysis biological factors and treatment need to be correlated with the different response groups.
Introduction

Neuroblastoma is the most common extracranial solid tumor in children and has a broad spectrum of clinical behavior, ranging from spontaneous regression to incurable aggressive disease (1, 2). The presence of metastatic spread at diagnosis is the most important factor in determining outcome in patients with neuroblastoma (3-5). About 50% of patients present with bone marrow (BM) metastases at diagnosis. Neuroblastoma can be staged according to the International Neuroblastoma Staging System (INSS). Using the INSS system and biological factors neuroblastoma can be further stratified in low, intermediate or high risk tumors (6).

Despite intensive treatment, including high dose chemotherapy and autologous stem cell rescue (7-11), only approximately 40% of patients remain long-term free of disease (6). Residual neuroblastoma cells are thought to be the major cause of relapse and bone marrow is a frequent site of recurrent disease (12).

Morphology is the classical method for evaluation of BM infiltration (13). However, morphology has a limited sensitivity and is unable to detect small numbers of tumor cells in BM (14, 15). Immunocytoology is a more sensitive technique to detect MRD than morphology, has been standardized and is already implemented in several clinical protocols (16, 17). Real-time quantitative PCR (RQ-PCR) is a highly sensitive technique to detect small numbers of tumor cells in blood and BM (17, 18) and detection of neuroblastoma mRNA in BM can be predictive of outcome (19-21). However, the clinical significance of MRD detection remains controversial because of the retrospective character of most studies, the small numbers of patients studied and differences in methodology.

During the inclusion period of this study, Viprey et al (22) published the first prospective study investigating the clinical usefulness of the detection of neuroblastoma mRNA in PB and BM from stage 4 neuroblastoma patients. They demonstrated that high levels of neuroblastoma mRNA (mathematical methods were used to determine clinically significant mRNA levels) in PB or BM at diagnosis and in BM after completion of induction chemotherapy predicted for worse outcome. The magnitude of reduction in mRNA level between diagnosis and post induction therapy was not of predictive value.

The goal of our study was to investigate the prognostic value of the detection of a panel of 5 sensitive and specific mRNA markers (PHOX2B, DDC, TH, CHRNA3, GAP43) in BM from high risk neuroblastoma patients. Our group has recently demonstrated the superiority of this panel compared to the use of single markers (18). With the design of the study it was estimated that 50% of the children with high risk neuroblastoma would have no detectable tumor cells in their BM after 4 months. Using a power of 0.8 and alpha of 0.05 and a RR of 2, the sample size was calculated at 300 patients. In this interim analysis we studied the prognostic value of the level of mRNA at diagnosis and investigated the clinical significance of BM clearance during induction chemotherapy.
Results described in this chapter are from an interim analysis performed November 2014 for the Dutch patients (n=49) and March 2015 for the German patients (n=174). Final results will be analyzed after the inclusion period in 2016 and at the end of the follow up period in 2019.

Methods

Patients and samples

Patients were treated according to the similar HR protocols of the German NB2004 trial and the Dutch DCOG NB2009 trial. The HR treatment protocol consisted of two N8 courses (if randomized to, German patients), MIBG therapy (Dutch patients), alternating N5 and N6 courses (3 each) and high dose chemotherapy with autologous stem cell transplantation followed by retinoic acid for 12 months (Supplemental figure 1). Dutch patients diagnosed after 2012 (if eligible; patients in CR or VGPR) (n=20) received GD2 immunotherapy in the Children’s Hospital Philadelphia, USA (23).

Staging was done according to the International Neuroblastoma Staging System (INSS) (6, 13). For therapy stratification, the status of the MYCN oncogene was investigated using two different molecular techniques (FISH, Southern Blot or array-based comparative genomic hybridization) in the DCOG and GPOH reference Laboratories. The test result for each parameter was given according to the criteria of the European Neuroblastoma Pathology, Biology, and Bone Marrow Group (24). High-risk patients were defined as stage 4 over 1 year of age or all stages with MYCN amplification. Bone marrow (BM) involvement was assessed by two aspirates and two biopsies from bilateral sites, according to the International Neuroblastoma Risk Group (13). Written informed consent from parents or guardians was obtained for all patients. The study was approved by the Medical Research Ethics Committee of the Academic Medical Center (Amsterdam, the Netherlands) and the University of Cologne (Cologne, Germany).

BM samples (0.5 ml) at diagnosis and at set time points during induction chemotherapy (After 2x N8/ 2x MIBG therapy and after each cycle of N5 and N6) were taken from children diagnosed with high risk neuroblastoma between 2009 and 2014. RQ-PCR was performed in Amsterdam, the Netherlands and GD2 immunocytology was performed in Cologne, Germany. The GD2 immunocytology of BM samples was done according to the internationally standardized protocol described by Swerts et al. 2005 (25).

RNA extraction, reverse transcription and real-time quantitative PCR

RNA was isolated from PAX blood RNA tubes (Qiagen, Venlo, Netherlands) with the PAX blood RNA Kit (Qiagen). cDNA was synthesized from 2-3 μg of RNA, using 25 μmol/L
random hexamers (Invitrogen, Carlsbad, CA, USA), 1 mmol/L dNTPs (Promega, Madison, WI, USA) and 100U of MMLV transcriptase (Invitrogen, ThermoFisher, Waltham, USA), in a total reaction volume of 40 µl and incubated at 42°C for 45 minutes. Finally, the reverse transcriptase was inactivated by heating and the volume was diluted to 100 µl.

qPCR for paired-like homeobox 2B (PHOX2B), tyrosine hydroxylase (TH), dopa decarboxylase (DDC), cholinergic receptor alpha 3 (CHRNA3) and growth-associated protein 43 (GAP43) was performed using beta-glucoronidase (GUS) for normalization as described previously (18) on the Step-One-Plus (Applied Biosystems, Carlsbad, CA, USA). All qPCR experiments were carried out in triplicate and mean values were used. If more than one sample of a patient was available, samples were tested separately. In case of discrepant results, the positive sample was used in the survival analysis.

**Data analysis**

A sample was scored positive if one out of five markers was above the threshold for positivity as has been described previously (18). In short, PHOX2B was scored positive if there was any amplification and the other markers were scored positive if the ΔCt of the sample was >3 Ct than the ΔCt found in control BM samples. To calculate the level of infiltration, the expression level of the RNA PCR targets were related to the expression level in neuroblastoma cell line IMR32, according to the following formula: $2^{\Delta\Delta Ct} (dCt_{IMR32} - dCt_{BM}) \times 100\%$.

A patient was scored as fast-responder if no neuroblastoma mRNA was detected after 4 courses of treatment and as slow-responder if qPCR results were positive after 4 courses of treatment and negative after induction chemotherapy. A patient was scored as non-responder if qPCR results were positive after completion of induction chemotherapy (also patients with missing samples after 4 courses were included in this analysis). A patient was scored as fast/slow responder if qPCR results were negative after induction therapy and no sample was tested after 4 courses of treatment (See **Supplemental table 1** for response scoring).

Survival rates were estimated by using Kaplan-Meier’s methodology. To assess the significant differences between the estimated survival curves the log-rank test has been used. Univariate and multivariate prognostic analyses were performed for OS and EFS by using the Cox proportional hazards regression model. All statistical analyses were performed by using SPSS version 21. Data collection and testing was performed in the Netherlands, the survival analysis was performed in Germany by a statistician not involved in this study. This was done because in the German cohort a clinical trial is ongoing (extra N8 chemotherapy versus standard therapy) and survival data could not be released without unblinding the clinical trial participants.
Results

Patient characteristics

Two hundred twenty three children with high risk disease were included in this study at time of interim analysis (184 stage 4, 22 localized and 17 unknown stage patients) (Table 1). Median patient age was 31.20 months and ranged from 0.2 – 295 months. Most patients were German patients (78%). Primary tumor MYCN status was available for more than 90% of children; of these, 43% had MYCN amplification.

Contribution of the different mRNA markers

From 217 patients 313 diagnosis BM samples were available (from 160 patients 1 sample was available, from 38 patients 2 samples were available and from 19 patients > 2 samples were available). For the marker analysis all samples were included in the analysis (for survival analysis only the sample with the highest level of infiltration was included (at diagnosis no discrepant results in terms of positivity were observed). At diagnosis 280 out of 313 samples (89%) showed positive results for one or more markers (Figure 1A). In most of these positive samples all markers were positive (84%). However,
Figure 1. Marker positivity a. contribution of the different markers to the positive samples. Each ellipse represents positive results of one marker at diagnosis and follow-up. b. Level of infiltration relative to IMR32 at different time point during induction chemotherapy.
in some samples only a single marker was positive (PHOX2B 5%; GAP43 0.4%). PHOX2B was positive in 279 out of 280 positive samples (99.6%). The quantitative data show that the mRNA levels can vary greatly between children at diagnosis (Figure 1B). As expected the mRNA level was higher at diagnosis than during or after induction chemotherapy (p<0.0001). During or after induction chemotherapy 729 BM samples from 223 patients were tested with all 5 mRNA markers (Figure 1A). In these follow up samples more differences in marker positivity were observed. In 159 out of 729 BM samples (22%) all markers were positive. PHOX2B positivity was observed most frequently (451 times out of 468 positive samples (96%)). In 17 of the positive follow-up samples (0.4%) PHOX2B was not detected. TH, GAP43 and CHRNA3 were all less informative (TH in 262/468 (56%), GAP43 in 257/468 (54.9%) and CHRNA3 in 238/486 (48.9%). DDC was the least informative marker and was found positive in only 177 out of 468 positive samples (38%) and as single marker in only one sample (0.2%).

Prognostic value of neuroblastoma mRNA levels in BM at diagnosis

High levels of neuroblastoma mRNA at diagnosis (infiltration level > 10% relative to IMR32) strongly predicted for EFS and OS in BM at diagnosis in univariate and multivariate analysis (including MYCN and stage) (Figure 2 and table 2). The patients without

**Figure 2.** Level of infiltration at diagnosis. Event free (a.) and overall survival (b.) for level of infiltration at diagnosis in 4 groups (no BM infiltration, <0.1%, between 0.1% and 10% and >10%)
BM infiltration did not differ significantly in terms of EFS and OS from children with BM infiltration. However, the group of patients without BM involvement was very small (n=12). Also when analyzing level of infiltration as a continuous variable, higher level of infiltration was significantly associated with a poor outcome (EFS: HR 1.01 p=0.007 and OS: HR 1.01 p=0.016).

**Clinical significance of BM clearance during and at the end of induction therapy**

To study the clinical significance of early and late BM clearance the BM response was analyzed at two time points; MRD negativity after 4 courses of induction chemotherapy was classified as fast response and positivity after 4 courses and negativity at the end of induction chemotherapy as slow response, whereas children with positive BM at that time point were scored as non-responders (Figure 3 and 4). The 12 children without detectable bone marrow infiltration at diagnosis were not included in this analysis. For 130 children the response after 4 courses could be assessed and for 116 children response at the end of induction therapy could be evaluated (for detailed response assessment see supplemental table 1). After 4 courses of induction therapy neuroblastoma mRNA was not detectable in in BM any more in 52 out of 130 children (fast-responders). At the end of induction chemotherapy neuroblastoma mRNA was still detected in 41 out of 116 patients (non-responders). In 12 patients neuroblastoma mRNA was detectable after 4 courses of chemotherapy and became undetectable at the end of induction chemotherapy (slow-responders).
Children with a fast molecular response did not differ significantly in terms of EFS and OS from children without a fast molecular response (5-y EFS of 29% versus 28%, log rank test p= 0.4; 5-y OS of 45% versus 55%, log rank test p= 0.98) (Figure 3). Patients with a molecular response (fast and slow responders) did not differ significantly in terms of EFS and OS from patients without a molecular response (No response versus response, EFS: HR 1.42 (0.85-2.37), p=0.185; OS: HR 1.33 (0.72-2.43), p=0.368) (Figure 4 and table 3).

There was a trend that slow response was associated with a better EFS and OS than fast response (EFS: HR 0.43, p=0.091 and OS: HR 0.34, p=0.10), but this was not statistically significant and only 12 patients showed a slow response.

Table 3. Cox regression response groups

<table>
<thead>
<tr>
<th></th>
<th>Event free survival</th>
<th>Overall survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P-value</td>
</tr>
<tr>
<td>No response versus Response</td>
<td>1.42 (0.85;2.37)</td>
<td>0.185</td>
</tr>
<tr>
<td>Slow response versus Fast response</td>
<td>0.43 (0.15;1.26)</td>
<td>0.091</td>
</tr>
</tbody>
</table>

Cox regression analysis for no response versus response and slow response versus fast response in EFS and OS. HR = Hazard ratio; CI= confidence interval.
Discussion

The interim analysis of this study demonstrates that high levels of neuroblastoma mRNA in BM at diagnosis (>10%) are associated with a poor EFS and OS. This is in line with the results found by Viprey and colleagues in a large prospective study and by Trager et al. in a retrospective study (22, 26). Our group has previously described in a retrospective study that the presence of BM metastases rather than the level of BM infiltration is prognostic. However, this was a retrospective study and only 39 stage 4 patients were studied. Furthermore, when looking into the data of our retrospective study in more detail, also in this study 25% of patients with more than 10% BM infiltration (n=16) were still alive whereas in the patients with less than 10% (n=23) 42% was still alive, so possibly if more patients would have been studied also a significant association between level of infiltration and outcome was observed. The results of the two prospective studies suggest that at diagnosis a group of patients with a very poor outcome can be identified. For these patients current treatment is failing in the far majority of patients and especially these patients might benefit from alternative (experimental) treatment strategies. In the final analysis of this study the levels of infiltration should be compared with immunocytology and other methods for disease detection in neuroblastoma, like

Figure 4. Survival curves for EFS (a.) and OS (b.) for BM response groups during induction chemotherapy. In blue: fast responders; in green: slow responders; in yellow: non responders; in purple: fast or slow responders
levels of urinary catecholemines and MIBG score (27). Unfortunately at time of interim analysis these data were not available for all patients.

At time of interim analysis we had BM samples from 223 high risk neuroblastoma patients. For most patients we received a BM sample at diagnosis. However, during induction chemotherapy we did not receive BM samples for all patients at all time points. For only 130 patients BM samples after 4 courses of chemotherapy (58%) were received and at the end of induction the response could be assessed for only 116 children (52%). In the final analysis of this study we should compare clinical data and factors that may influence response in the group with and without missing samples. It is possible that the patients with missing samples may differ in response from the patients without missing samples, which may lead to bias. The goal of our study was to include 300 HR patients in 60 months. We nearly reached the goal of 300 patients, however since the response could only be assessed if we had samples available after 4 courses and at the end of induction therapy we should continue including patients to be able to assess the clinical significance of the molecular response patterns. Furthermore, the monitoring of serial sampling should be improved.

Retrospective studies have shown the clinical significance of MRD monitoring during treatment (19, 21) and Viprey and colleagues demonstrated that detection of mRNA in BM after completion of induction therapy in a prospective study is associated with a poor EFS and OS, however the level of mRNA at diagnosis was the strongest outcome predictor and neuroblastoma mRNA detection at end of induction did not have additive value to the predictive power of the level at diagnosis. In this interim analysis the detection of neuroblastoma mRNA after 4 courses of chemotherapy and after completion of induction chemotherapy was not associated with an unfavorable outcome. This might be caused by the relatively short follow-up time. In the study by Viprey et al. there is no difference in survival curves during the first 1.5 years after diagnosis and at the moment of interim analysis of our study the median follow-up time was still relatively short. Also in the study by Viprey et al. the effect of MRD level at the end of induction therapy was only moderate and lower than was anticipated when we designed our study. As mentioned in the introduction it was expected that 50% of patients would not have detectable BM disease after 4 months of treatment and that patients with BM disease had a 2 x higher risk of relapse than patients without detectable BM disease. After 4 courses of treatment (3-4 months) 78/130 (60%) still had detectable BM disease. Furthermore, in the fast response group 49% of patients had an event, whereas 62% had an event in the patients without fast response (however, follow-up time was still relatively short). Therefore, the effect of fast BM clearance on survival seems less than expected and we need to increase our sample size and follow-up time.

Viprey et al. have used mathematical methods to calculate below which cut points for a specific marker the mRNA level is associated with better survival compared to the
patients with a level above this cut point. Because for each marker the cut points were calculated separately, in this study a patient can be below the cut point for one marker and above the cut point for another marker. In our study we have used thresholds for positivity based on background expression in control BM samples and samples were scored positive or negative, and for this scoring the results of all markers were combined. Therefore, the difference in analysis and interpretation of the MRD results between our study and the study by Viprey et al. might have led to differences in results. Furthermore, it will be difficult to compare our results since the cut points depend on marker, time point and survival of the cohort studied. Since the patients in our cohort were treated differently it is possible that these cut points will not be the same. Also these cut points may not be reproducible in subsequent clinical trials with changes in treatment protocol. Response may be influenced by biological factors or differences in treatment, but in this interim analysis we were unable to study this because the trial participants could not be unblinded. Kushner and colleagues (28) suggested that a major response to upfront chemotherapy is crucial for a good outcome in patients with MYCNA, but not for patients without MYCNA. Therefore, in the final analysis of this study associations between BM molecular response and biological factors should be assessed. Viprey et al. found a significant association between EFS and OS and qPCR positivity after induction, even when correcting for MYCN status. However, mRNA values at diagnosis were the strongest outcome predictor.

Furthermore, it should be emphasized that the patients in this study did not receive the same treatment. Children received either MIBG therapy followed by the standard treatment and from 2012 immunotherapy (if eligible, n=20) (Dutch patients), 2x N8 chemotherapy followed by the standard treatment or only the standard treatment (German patients). Because these different treatment protocols may influence the kinetics of the BM response as well as the outcome of the disease, it is difficult to assess the predictive value of BM clearance without correcting for treatment. For this interim analysis we were not able to unblind the trial participants.

Over the past decade PHOX2B, TH, GD2S, CCND1, DCX, ELAVL4, CHRNA3, DDC and GAP43 have been described as RNA MRD markers in neuroblastoma (18, 29-32) . Since the expression of these genes can change during treatment and at relapse and tumors are regarded to be heterogeneous, the use of a panel of markers is more reliable than using single markers (18, 33, 34). Several groups have used different panels of markers for MRD detection and it is unknown which marker panel is optimal (18, 22, 35, 36). In our study we have used a panel consisting of 5 markers. PHOX2B was the most informative marker and DDC was the least informative marker and was the only positive marker in one follow-up sample (0.2% of the positive samples). Therefore, in future studies DDC can be excluded as marker and the use of PHOX2B, TH, CHRNA3 and GAP43 will identify more than 99% of the positive samples. This is in line with the previous results of our
group, where only TH and CHRNA3 contributed to the additional value of the panel in follow up samples compared to PHOX2B alone (18). In the final analysis of the prospective study, the predictive power of the individual markers should be assessed.

In conclusion, in this interim analysis we show that the presence of high levels of BM infiltration at diagnosis is associated with a poor outcome. In contrast to results found in retrospective studies and in the study by Viprey and colleagues, BM response after completion of induction chemotherapy was not associated with a poor outcome. However, our follow-up time was still relatively short. Although we have nearly reached our goal of including 300 high risk patients, response could only be assessed in a much smaller group of children. Therefore, in the future more patients need to be included. In the final analysis of this study we need to compare BM results at diagnosis with other methods for disease detection, correct for biological factors and treatment in the different response groups and finally the predictive power of the individual markers should be assessed.
Reference List


15. Cheung IY, Barber D, Cheung NK. Detection of microscopic neuroblastoma in marrow by histology, immunocytology, and reverse transcription-PCR of multiple molecular markers. Clinical


35. Cheung IY, Feng Y, Cheung NK. Early negative minimal residual disease in bone marrow after immunotherapy is less predictive of late or non-marrow relapse among patients with high-risk stage 4 neuroblastoma. Pediatric blood & cancer. 2013;60(7):E32-4.
## Supplemental data

### Supplemental table 1. Response evaluation

<table>
<thead>
<tr>
<th>MRD after 4 courses</th>
<th>MRD at end of induction therapy</th>
<th>Response group</th>
<th>Number of pts</th>
</tr>
</thead>
<tbody>
<tr>
<td>neg</td>
<td>neg</td>
<td>fast</td>
<td>29</td>
</tr>
<tr>
<td>neg</td>
<td>missing</td>
<td>fast</td>
<td>23</td>
</tr>
<tr>
<td>pos</td>
<td>neg</td>
<td>slow</td>
<td>12</td>
</tr>
<tr>
<td>pos</td>
<td>pos</td>
<td>non</td>
<td>30</td>
</tr>
<tr>
<td>missing</td>
<td>pos</td>
<td>non</td>
<td>11</td>
</tr>
<tr>
<td>missing</td>
<td>neg</td>
<td>fast/slow</td>
<td>11</td>
</tr>
</tbody>
</table>

Neg = negative; pos = positive

### Supplemental figure 1. Different treatment arms and timeline in GPOH and DCOG protocols

116

6 weeks 6 weeks 6 weeks 6 weeks

6 weeks 6 weeks 6 weeks 6 weeks

6 weeks 6 weeks 6 weeks

Supplemental figure 1. Different treatment arms and timeline in GPOH and DCOG protocols