Minimal residual disease detection and monitoring in children with neuroblastoma
Stutterheim, J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Minimal residual disease detection and monitoring in children with neuroblastoma
Chapter 8

Summary, conclusions and future directions
Summary, conclusions and future perspectives

PCR-based detection of minimal residual disease (MRD) in children with neuroblastoma (NB) at different time-points during follow-up might provide information on the effectiveness of treatment. Our aim was 1) to improve minimal residual disease (MRD) detection in neuroblastoma by identifying new MRD markers and 2) to study whether we can use MRD monitoring during treatment to identify patients who benefit from the current treatment and patients who need different, additional treatment. The ultimate goal of MRD detection in patients with NB is to stratify patients, using PCR-based MRD detection. The results and future directions for research are summarized and discussed below.

Summary and discussion of the results on PCR-based MRD detection in children with neuroblastoma

Optimal detection of MRD using real-time-quantitative (RQ)-PCR for mRNAs requires the identification of a target mRNA that is expressed in the target tumor cells but not in normal hematological cells. For the last two decades the first enzyme in the catecholamine synthesis pathway tyrosine hydroxylase (TH) and the enzyme for the synthesis of GD2, which is present on the cell membrane of most neuroblastomas, GD2 synthase (GD2S), have been used by several groups as MRD RQ-PCR targets. Because TH and GD2S are expressed in bone marrow (BM) and peripheral blood (PB) the interpretation of low positive MRD results are not unequivocal, i.e. these markers are not totally specific. Therefore, the search for other more specific targets for detection of minimal residual disease persisted. The ideal MRD marker is neuroblastoma specific with no expression in the normal hematological compartments such as BM, PB and peripheral blood stem cells (PBSC). During the course of this PhD project we and two other groups have tried to discover new NB markers by systematical approaches. Our group selected candidate MRD markers by comparing SAGE mRNA values of normal tissues with SAGE mRNA values of neuroblastoma tissues (chapter 3). From SAGE mRNA expression libraries, 28 genes were selected, which showed high expression in neuroblastoma tumors and little or no expression in normal tissues. By extensive RQ-PCR testing of NB tumors, control BM and control PB samples, the 6 most specific neuroblastoma markers were selected; PHOX2B, TH, DDC, DBH, CHRNA3, GAP43. The most promising gene was PHOX2B, which had no detectable expression in normal BM, PB, and PBSC.
In chapter 2, we tested PHOX2B in parallel to the currently most widely used MRD marker, TH, and showed that a sizable fraction of samples of patients with NB displayed TH levels that did not permit a clear-cut interpretation. The use of PHOX2B overcomes this problem, as a positive PHOX2B result clearly implies the presence of neuroblastoma mRNA. This represents a major advantage of PHOX2B compared to TH, GD2S, and other NB markers, which do have low expression in BM, PB, and/or PBSC samples of control individuals.

Two other groups also reported on identification of potential new markers by gene expression profiling, among which PHOX2B. Viprey et al.\textsuperscript{1} used the Affymetrix human U133 Plus 2.0 Array to compare gene expression in primary NB (n=32) and pooled PB from 24 healthy volunteers. They identified 240 differentially expressed genes of which eventually 11 neuronal genes were selected with no expression in hematological cells as determined by EST database analysis. Evaluation with RQ-PCR on PB samples (n=15) of healthy volunteers revealed that only PHOX2B and Doublecortin (DCX) were more sensitive and specific than TH. The third study was performed by Cheung et al.\textsuperscript{2} who carried out a gene expression array using Affymetrix human U95 gene chip on 48 stage 4 tumors and 9 remission marrows (from stage 4 patients with NB). They found 49 differential expressed genes with a better neuroblastoma-marrow ratio than TH. Using two NB cell lines (LAN1 and NMB7), they performed a sensitivity assay that showed that 11 of 49 genes also appeared to be more sensitive than TH. Next, the top 8 markers were tested for their ability to predict survival, and positivity of 6 these markers (CCND1, DDC, ISL1, PHOX2B, GABRB3, KIF1a) after two cycles of immunotherapy was highly prognostic for overall survival.

In summary, all three studies used high through put analysis to identify new targets for MRD detection in neuroblastoma and in all three studies, the selection of new MRD markers was based on the differential expression between neuroblastoma tumor and normal hematological cells. Subsequently, different approaches were used by the groups: while our group and Viprey et al. chose the most specific markers, Cheung et al. selected markers on basis of sensitivity and survival prognostic tests. Even though approaches were different, all three groups identified PHOX2B as a very sensitive and specific maker. PHOX2B is now being evaluated in three large ongoing prospective studies in Europe (the combined study of the German pediatric oncology-hematology (GPOH) group and Dutch children oncology group (DCOG), the International Society of Pediatric Oncology European Neuroblastoma
Minimal residual disease detection and monitoring in children with neuroblastoma (SIOPEN) and the American study (Children Oncology Group (COG)).

Although PHOX2B seems to be an ideal marker, the question rose if MRD detection using only one marker is an optimal clinical strategy. Expression of different markers in primary neuroblastoma is very heterogeneous and it is unknown whether they are stably expressed during treatment. Therefore studying multiple markers might help to overcome the tumoral heterogeneity and thus increase sensitivity. In chapter 3, we showed that the use of a panel of markers indeed increases the sensitivity of MRD detection. Therefore, we advise to use a panel of markers to obtain maximal sensitivity. This panel should definitely include PHOX2B, because it is the only completely specific and most sensitive marker. Furthermore, it should include TH, as the most widely applied marker. In addition, for testing of different tissues, such as BM and PB, we have shown that the background of the markers differs between normal tissues, so different panels of markers could be applied for MRD testing in different hematological compartments (e.g., BM, PB, PBSC and/or CD34+). As an example, we include DBH for PB and PBSC testing, but not for BM testing.

Furthermore, we described in chapter 2 and 3, how we established cut-off levels for positivity for each PCR target to discriminate between clinically significant levels of tumor cell mRNA and levels in normal cells, since most markers show illegitimate expression in normal hematological cells. To be able to establish cut-off levels, we first defined absolute levels of expression of our panel of PCR targets in normal hematological tissues, such as BM, PB, PBSC and CD34 positive cells. Then we proposed to apply guidelines adapted from the European Study Group (ESG) on MRD detection in Acute Lymphoblastic Leukemia (ALL) to define the expression levels constituting for positive results. Thus, for each target and in each hematological tissue, different cut-off levels for MRD positivity were described. To minimize the chance on false positive results the cut-off levels were clearly apart from the background expression. Since PCR results will in the first phase be used to identify patients with NB who might benefit from more aggressive therapy, we aimed with our guidelines to prevent false positive results. Unfortunately, there is no international consensus on cut-off levels for positivity yet. In the ESIOP trial, a standardized method to perform RQ-PCR with TH as marker is described in which expression is normalized to β2 microglobulin and reported using the ΔΔCt method to correlate the result to a positive control (provided by the central reference laboratory at the University of Leeds). In the ESIOP study it was decided not to face the challenge of defining
Chapter 8 Summary, conclusions and future directions

what constitutes a positive and negative result on an individual test basis, but selection of the minimal clinically relevant level will be retrospectively defined by statistical analyses of the clinical results. In our opinion this is incorrect, since it is known that the expression of the PCR targets (TH) in primary tumors are highly heterogeneous between patients (up to 3 log) and furthermore it is unknown if expression levels of the PCR targets (TH) change during treatment. Thus a clinically relevant level can in theory be different for each patient. Therefore international consensus on what constitutes for a positive and negative result for different reference genes needs to be established, and for correct quantification of MRD the primary tumor is required.

For reliable PCR-based detection and quantification of MRD in NB it is important to know the stability of expression of PCR target genes. In chapter 4 we show in a comparative analysis of different pairs of neuroblastoma tumor and BM samples, that in general target expression is quite stable, however it might alter especially upon treatment and at relapse. When the primary tumor is available BM infiltration at diagnosis can be reliably quantified in most patients. Since variation between PCR targets is larger during treatment, relative quantification during treatment is more difficult than at diagnosis. Based on our results we propose to apply a panel of PCR targets. By using the median value of tumor cell infiltration, most (but not all) variations due to treatment and outgrowth of sub clones will be neutralized.

We showed that using our panel of neuroblastoma specific targets, MRD could be reliably detected. However the clinical relevance of MRD detection using this panel of markers still had to be established. In chapter 5 the clinical relevance of MRD detection at two time points during treatment was studied in a retrospective study;
1) an early molecular response 3 months after diagnosis and
2) a late molecular response after induction chemotherapy.
In our cohort of patients with metastatic high risk disease, molecular BM remission was observed in 11 of 38 (29%) patients at three months after diagnosis, which was associated with favorable outcome (5-y-OS 62% ± 15.0 versus 19% ± 8;p=0.009). In bivariate analysis, this early molecular response remained significantly corrected with survival when corrected for conventional clinical parameters, such as INRC response, residual metastatic MIBG uptake and presence of urinary catecholamine metabolites. It seems that already after a few cycles of induction therapy, the patients that respond to therapy and thus survive can be recognized. It is likely that there is a critical time point at which clearing of the BM from detectable transcripts
correlates to a better outcome and this has to be established in future prospective studies in uniformly treated patients.

In the same chapter (5), we observed, that after completion of induction chemotherapy, BM of 12 of 29 (41%) patients was still MRD positive, which was significantly associated with poor outcome (5-y-OS 0% versus 52% ± 12; p<0.001). These observations were in line with those of Cheung et al. Altogether these studies show that patients with residual disease detected before high dose chemotherapy, even at very low levels, are prone to have progressive disease. This could be the result of either slow-responding or subclinical progressive disease (PD), which has important implications for future therapy design. Although slow responders may benefit from continued chemotherapy, PD demands, at the very least, change in therapeutic strategy. To investigate the difference between slow-responding and PD, MRD in BM and PB should be monitored at designated time points during induction therapy, before high dose chemotherapy and stem cell rescue and during consolidation therapy, in prospective studies.

In chapter 6 we investigated the incidence and impact of autologous graft contamination. In order to increase the sample size we collaborated with German centers and managed to collect the largest samples size described until now. The Dutch and German patients were treated according to different treatment protocols and different hematological tissues were used as source for autologous stem cells, but the different patient groups had comparable survival. We therefore performed analysis of the combined cohort and showed that, the presence of neuroblastoma mRNA in stem cell harvests seems to be associated with unfavorable outcome. When we only analyzed the patients of whom autologous harvest were reinfused, the correlation between MRD positive harvest and survival lost its significance. Therefore, we hypothesized that the autologous stem cell harvest reflects remission status and that metastatic remission status would be correlated with survival. Indeed we show that MRD detection in BM at time of harvest was most significantly associated with survival. These results are in line with the results on fast molecular response (chapter 5). However the results need to be verified in a larger, more homogeneously treated cohort.

Since the expression of RNA markers varies greatly between tumors (chapter 3) and are not always stably expressed during treatment (chapter 4), we studied the possibility of MRD detection using a DNA marker in chapter 7. In this study we show that methylated RASSF1a can be used as a DNA MRD RQ-PCR marker for MRD detection in patients with NB. The
applicability of methylated RASSF1a as an MRD marker was compared to RQ-PCR results of our panel of RNA markers, and showed to be of additional value. Furthermore, by measuring the methylated RASSf1a fraction and the unmethylated RASSF1a fraction, MRD in BM could be more accurately quantified. Thus, methylated RASSF1a is a novel sensitive and specific DNA RQ-PCR marker for MRD detection in BM cells of patients with neuroblastoma and can be included, as marker in prospective MRD studies of neuroblastoma, for accurate quantification of MRD levels.

**Future directions**

**Prospective study**

All results of the retrospective studies as described in chapter 5 and 6 need to be and will be validated in a large multicenter prospective study. In order to reach enough power, DCOG NBL 2009 has joined forces with GPOH NB 2004. The backbone of the Dutch high-risk and GPOH high-risk protocol are the same, but in the GPOH there is a randomization between two additional course of induction chemotherapy (2 times N8) or no additional chemotherapy (standard arm), while in the Dutch protocol, patients will be treated with 2 courses of radiolabeled MIBG therapy before induction chemotherapy. At designated time points at diagnosis and during therapy BM and PB samples are taken for MRD analysis (figure 1). For molecular analysis material is collected in (special) PAX tubes, in which RNA is well stabilized. BM samples of all patients are tested by PCR in our lab as well as by GD2-immunocytology in the lab of our German collaborators (Cologne) for tumor invasion. This prospective study started in January 2009 and will keep including patients until January 2014. Until now, 20 DCOG and 60 GPOH patients with high risk disease have been included. In addition, we are also setting up collaboration with Turkish national NB studies (Ankara) TPOG NB 2009 to add a third validation cohort. The backbone of TPOH high risk protocol is the same as for DCOG and GPOH, but there is no additional treatment before induction chemotherapy, so this protocol is comparable to the standard arm of GPOH2004.
The goal of this prospective study is to improve risk stratification and treatment allocation for all stages of patients with NB by real-time quantitative PCR-minimal residual disease testing using the panel of RQ-PCR targets. In this study we will investigate
1) if MRD monitoring can improve BM staging at diagnosis (neuroblastoma all stages),
2) if MRD monitoring at different time points can predict survival in high risk patients,
3) if detection of MRD in autologous stem cell harvests predicts survival,
4) if MRD monitoring after completing therapy can predict early relapse and
5) if MRD monitoring in peripheral blood (PB) can add information compared to bone marrow (BM). In addition, all BM RQ-PCR results will be compared to GD2 immunocytology results.

We hope that this prospective study, together with the two other large prospective MRD studies (of SIOPEN and COG) will give definite conclusion on how PCR based MRD detection should be incorporated in future treatment protocols.
Isolation and characterization of disseminated and circulating tumor cells

Both immunocytology and RQ-PCR are sensitive methods to detect disseminated tumor cells (DTC) in the BM or circulating tumor cells (CTC) in the PB. These techniques can quantify the amount of tumor cells present in BM or PB, however they do not give information on molecular and functional characterization of disseminated tumor cells (DTC) and circulating tumor cells (CTC). Since DTCs and CTCs are the major source of progression and relapse, these cells are also of major interest for further functional studies. With the future introduction of new targetted anti-cancer drugs, new markers to determine response are needed. Future research will therefore also focus on the use of circulating tumor cells for monitoring the biological response within tumor cells in early clinical trials in children with cancer with novel compounds, such as ALK inhibitors. With the use of 3-colour FACS staining (CD45- (pacific blue)/ CD56+ (PE)/ GD2+ (PE-Cy5)) we are already able to isolate single NB cells from BM. However, for best isolation of DTCs and CTCs, this technique has to be compared to other techniques such as immuno-magnetic enrichment and the use of micropore filters. Then the isolated NB cells can be used to study the presence of the desired targets and biological efficacy of novel compounds during early clinical trials by measuring the level of target inhibition, concomittant downstream pathway modulation and the extent of the desired biological effect (e.g. cell cycle arrest etc.). In the ongoing prospective study DTCs and CTCs can be sorted from the BM and PB, respectively, for these purposes.

In bullet points; the implications for clinical practice and future research:

• The clinical relevance of MRD testing is being investigated in a prospective study.

• We have shown that patients which clear their BM early during treatment have a better prognosis, so MRD should be measured already early during treatment, after 2 courses of nuclear therapy or chemotherapy.

• Patients that have not cleared their BM after induction therapy have a very poor survival; therefore MRD should be measured after induction chemotherapy and before high dose chemotherapy.

• To evaluate the response of high dose chemotherapy MRD should also be measured after high dose chemotherapy in BM and PB. In this way it
can be established if high-dose chemotherapy can rescue the MRD positive patients, or that high-dose chemotherapy should be reserved as consolidation for patients that are MRD negative.

- All MRD tests should be measured in BM and PB with a panel of neuroblastoma specific RNA targets

- The panel of targets should at least include PHOX2B and TH, and others such as DDC, CHRNA3, DBH and GAP43 could be added.

- Because a DNA marker is more stable than mRNA markers, MRD should also be measured with a tumor specific DNA aberration; methylated RASSF1a, and compared to RNA targets in a prospective study.

- MRD should be measured in the autologous stem cell graft. If the autologous graft is selected for CD34, the positive, negative and the waste fraction should also be measured.

- If new experimental treatments are given, such as GD2 immunotherapy, the response of these treatments should also be evaluated by MRD testing in BM and PB.

- Finally: in the new protocol, response to treatment monitored by RQ-PCR should be included as standard response evaluation to determine which patient is responding and which patient can benefit from extra treatment or new drugs.
References


